COLLABORATIVE ACTIVITY-BASED RIDE SHARING: LINKING HUMAN MOVEMENT, SOCIAL NETWORK, AND PLATIAL SEMANTICS FOR FUTURE TRANSPORTATION

YAO LI WANG

As a partial fulfillment to the degree of Doctor of Philosophy
PhD candidate
Infrastructure Engineering
The University of Melbourne
February 2018
Yaoli Wang: Collaborative activity-based ridesharing: Linking human movement, social network, and platial semantics for future transportation
© February 2018

SUPERVISORS:
Prof Stephan Winter
Dr Martin Tomko
Dr Nicole Ronald

LOCATION:
Melbourne, Australia
Dedicated to
the love of truth and innovation,
the love of family and friends,
and the love of sweetness and toughness of life.

— Yaoli, February 2018
ABSTRACT

This thesis presents a collaborative activity space model and the implemented prototype for ridesharing under the umbrella of this model. The ridesharing prototype accommodates socio-psychological barriers and travel flexibilities in multi-person travel behaviours. Despite the positiveness of ridesharing, such as reducing private car usage for social and environmental sakes, people find it uncomfortable or inconvenient to share rides. The two issues underneath are social preference and travel flexibility. If people are prioritised to match with those they prefer, and if an efficient strategy exists to reduce detour cost or the probability of no-ride, commitment to ridesharing might be increased. Current software or algorithms on the market, however, mostly overlook the socio-psychological aspects of ridesharing, and thus cannot better exploit the full potential of this new travel mode.

This work provides the solutions incorporating social network preference and travel flexibility from three perspectives:

1) A method called Social Network based Ridesharing allows travellers to prioritise their closer acquaintances as ride partners while still considering local strangers of low detour cost.

2) The Activity-based Ridesharing Algorithm involves travel aims and functions of places to provide alternative destination choices.

3) A combination of the social network based and the activity-based methods, called Collaborative Activity-based Ridesharing, is suggested where social network preference is not only satisfied but also used as a space search heuristic for fast retrieval of alternative destinations and potential ride partners.

The implementation of the model is experimented with agent-based simulation in a real study area with travel survey datasets. The outcomes justify that the proposed collaborative activity model and algorithms are capable of significantly increasing match rates and reducing socio-psychological obstacles for ridesharing.
Most ideas and figures have appeared previously in the following peer-reviewed publications:


Further papers are currently under peer-review:

- **Wang Y., Kutadinata R., Winer S. (under review).** The evolutionary interaction between taxi-sharing behaviours and social networks.


- **Wang Y., Bock F., Koetsier C., Winter S., Sester M. (under review).** Introducing ridesharing into parking problems.

Also, a non-reviewed magazine article presents the main idea of the thesis:


“Don’t ask yourself what the world needs.
Ask yourself what makes you come alive, and go do that,
because what the world needs is people who have come alive.”
— Howard Thurman

ACKNOWLEDGEMENTS

It is not easy to write this section and I believe (or hope) no one else has written it this way like I did. Three and half years, I have so many memories – sweet and bitter, laugh and tears, rejoice and stress. So I’m somehow making this section an autobiography.

In September 2014, I stepped on my second journey in my life to a new continent. This was going to be the third country I would live in. I was excited and curious. I was expecting some changes, but not sure how much until today – when I look back, I am surprised by how much I’ve changed during the 3.5 years.

I’ve been lucky always in terms of my relationship with my supervisors. It wouldn’t be better than sharing not only a similar research interest but also hobbies with my supervisors.

My major supervisor Prof Stephan Winter is such a cautious scientist with high demand for his students and efficient fast working pace. While some people might feel stressed with that, I cannot say how much I enjoy! I work very hard and play very hard! Not many supervisors take care of their students’ work-life balance, but Stephan does. I still remember he once said "Don’t forget to go for a run or swim". And yes, I came across him quite a few times when I was jogging around. By chance, Stephan noticed that I am an enthusiastic photographer, so he told me that he used to be a very passionate portrait photographer when he was in Europe. My eyes lit up when I heard about that!

My co-supervisor Dr Martin Tomko is such a funny and easy-going person. He always has creative and interesting ideas in both academic work and daily life. Something like...a hedgehog slice Wiki page! I realised he is an outdoor person, but didn’t know that he was an old OXO (i.e., a member of Melbourne University Mountaineering Club, MUMC for short) until a chance. He even showed me some photos when he was paddling on a river with MUMC. Wow! How exciting it is when you can chat about Midnight Ascent and the resume of rope sports with your PhD supervisor!

The other co-supervisor Dr Nicole Ronald is an interesting person as well. She sort of has a temperament of "so-what" and a strong character which I find very lovely. Her Master’s supervisor Dr Michael
Kirley from Computing and Information Systems is my committee chair, who is called a "question man" by Nicole because he always has lots of questions. Michael is very funny in many ways, including his profile picture on the website (unfortunately he changed it now...)!

There is a person who is not my co-supervisor, but has given so much support in my research. Dr Ronny Kutadinata was Stephan’s postdoc when I was in my second year. At the first glance I thought he was serious, but it turned out that he is easy to cooperate and communicate. Ronny is a great team mate focusing on details and the strictness of the model. The ACM SIGSPATIAL Best Paper Award paper is a great achievement where Ronny not only helped with beautiful mathematical functions, and supported me with optimisation knowledges through the whole work, but also stood by in the last week before the due answering questions and cheering me up. I am so grateful and blessed to have such team mate.

I am doing social network analysis, looking at small world networks, and my life experience once and once again provides evidences for this. Michael has a student called Yuan, whose wife is my best friend in Melbourne. This girl is Jingjing, a medical PhD student at St Vincent hospital and Peking University alumna. I got connected with her via one of my high school classmates. After we met each other, we felt immediately like twin sisters who had missed each other for ages. I took Jingjing as my portrait photography model. She introduced me to Peking University Alumni Association in Melbourne. We caught up with each other frequently, sometimes going for a run, sometimes grabbing a beer and sharing roasted chicken, sometimes having a Chinese tea party, sometimes a road trip until the road trip to Lake Tyrrell after which Yuan asked her to be his girlfriend. Then faster than I expected they built a family. I haven’t yet had as frequent contacts with Jingjing as I used to, but whenever I got stuck into some situations or she felt so down for something, we met and talked, infusing each other with energy. That still keeps going on, and I guess will last forever.

When I just got to Melbourne, there were quite a few Chinese students in my office. One of them was Ashley, doing a research assistant in hydrology. I was not so familiar with her for quite a while until some day I forgot how. We started having fun together – lots of events with a group of crazy people in my office, including my “academic twin sister” Zahra (who enrolled exactly on the same day as me) and my best male friend in Melbourne Alex. Ashley is so nice, kind, and caring. I enjoy talking to her because she’s such a good listener. There are so many good talkers in the world, but not so for listeners. Jiaying was Ashley’s housemate. I actually met her for the first time in an academic meeting when she was doing her Master’s, but didn’t have very deep impression until I knew that she was in the same flat as Ashley. She is sometimes very quiet in public, but other times re-
ally outgoing when facing like-minded people. As I do, Jiaying has lots of local and international friends in Melbourne, loves reading, has a strong character, and is very independent with liberal thoughts. I feel like Jiaying and I are always in the process of exploring each other’s thoughts and figuring out more and more similarities, both career-wise and life-wise. We were born for liberty and toughness!

In my office, people are from all over the world. It’s a united nation — fantastic! As mentioned above, Zahra enrolled on the same day as me. At the beginning, I somehow felt like she was a bit too cool until an accident (just a misunderstanding) happened, which, rather than ruining our relationship, made us much better. We exchanged information and at times “enjoyed” complaining together about something. Michael was in his early third year when I joint the team. He looked serious and quiet to me at the beginning. I was trying to initiate some talks no more than ‘Hi’ or ‘Good morning’ for a while until one day someone told me that Michael is a photographer as well! We started chatting more since after. The talks extended to some really funny stuff by which I realised that Michael is so weird in a good way, something entirely different from what he appeared. At that time, Rahul and Haifeng were sitting next to me. These two guys spent lots of time sticking together and telling jokes. Rahul and Haifeng did some language exchange at times, learning mandarin and Hindi from each other. They somehow seasoned my study life with funny moments. Surabhi joint the group in my second year. She is very outgoing and independent with interesting ideas. We chat frequently in spare time, and sometimes discuss deep topics like philosophy and cosmology.

Alex was originally Stephan’s student, but moved back to his background – Computer Science – after one year. He is such a weird and funny person. Of course he involved a lot in our office crazy group for fun. Alex is a passionate traveller as well, actually more experienced than me in backpacking who has been to almost 40 countries all over the world. I sought many tips from him for my travel. Alex talks so many jokes and I’m kind of immune to that now. Don’t say I’m too mean! I’m just stopping dropping my chin into my chest.

In my second year of PhD, Maryam came to our office as a visiting PhD student since her supervisor moved to Melbourne University. She was absolutely the No.1 workaholic I’ve ever seen. We didn’t really talk to each other until one day. I started to notice that she likes running. So we went for many weekend long runs to a bakery or a restaurant, had our lunch, and hopped on a tram back. We chatted about many many things, from politics to life decisions to academics. Sometimes we kept quiet when we were really busy, but sometimes we laughed our tears out talking about funny experience or “harassing” each other.

There was a lunch group of a bunch of Chinese friends when I just got to Melbourne. I knew Fuqiang, Qiaowen, Sha, Jun, Zhiyang,
and Jian from that group. These people all have different personalities. Fuqiang is kind-hearted and easy-going, Qiaowen is mad at times while naive other times – not like a mum, Sha is my first guitar teacher, Jun is quiet but has a stronger inner world, Zhiyang looks a bit laidback, and Jian is a cynical person with very profound thoughts. I enjoyed sharing with them stories and experience.

The first club I joint at Melbourne University, less than a month after arriving in Melbourne was Melbourne University Photography Club, where I met Danzi, an elegant and attractive PhD girl through whom I knew my current housemate, Jonathan, an OCD and workaholic but super friendly Aussie lawyer with whom I worked together as event photographer (He motivated me a lot even though we don’t see each other frequently), Dawei, a Chinese Australian superman who has three jobs at the same time and from whom I learnt portrait photography skills, Gabriel, a civil engineer as well as fabulous street and travel photographer who appears indifferent but very funny and introspective inside, and Michael, a super tidy and well-organised Aussie who is so good at driving and photography and with whom I’ve had a few pleasant road trips. Aggie, who is my housemate, was introduced through Danzi and was also a Photography Club member when she was studying. We knew each other in a Chinese New Year party. We were not so familiar until by a random chance when she saw my post on social media that I’ve got some decent Chinese tea while she told me that she’d like to visit me with some Chinese desserts from my hometown in exchange for the tea. What a chance! We went for a few road trips in Australia and also a trip to Tibet together. The second club I joint was the Bhakti Yoga Club where I met Xuelan. Immediately we became friends just because like-minded people know each other quickly. She is such an active and happy girl, but has a sensitive side hidden in her little universe. We are so much alike in this regard.

It was actually through Xuelan’s introduction that I joint Melbourne University Mountaineering Club in my last year of PhD. This is such a bunch of weird, funny, tough, and nice buddies. I learnt so much from them not only outdoor skills but more personality and internal characteristics - resilience, perseverance, toughness, eagerness for freedom, and awareness of the environment. I went for so many hikes and climbs with them, aiming to challenge and conquer my own fear of height, to break though my limit, to do something crazy but fun, and to start seriously caring about the environment (picking up trash while hiking, reducing meat intake, and stopping using plastic bags and cutleries). Luke, a seemingly quiet but very friendly software engineer, was the trip leader of my first hike with MUMC. Caio, whom I knew also in that trip, is a very easy-going, optimistic, and kind-hearted computer scientist who’s lived in a few countries. Rowan, who was doing her Master’s in the same department as me, hooked
me onto Rogaine. She was so encouraging as she was still hiking and climbing when she was 4 or 5 months pregnant. Then I got to know Natalie, who was the previous MUMC president, looked a bit harsh at the first glance, but is actually very caring and sensitive. She somehow inspired me to pay more attention to environment by her own behaviours as she is a vegan who tries not to use one-off stuff. Gina is an easy-going and thoughtful girl who likes writing articles, knitting, and flowers. Andriana is a pretty and super nice girl majoring in physics, who has very stronger and tough character, and is talented in climbing. I still remember her words saying “Look! That’s problem-solving skill” when she helped me to start climbing. Channa, an Electrical Engineering student with whom I went for quite a few hikes, is a very dynamic and happy guy. Isabel was the conservation officer who is passionate about academics and excited about doing her PhD in Geography. By a random chance I realised that she used to be my student in the GIS lab. Ana, a passionate climber, looks like a quiet girl but is full of energy inside. I worked with her on some MUMC publications. Evie is the artist in MUMC, going hiking, chilling out on the cliff edge, and sitting there painting. Another artist is Aiden, the old OXO bushwalking officer, who appears serious but is actually funny and weird. Julius looks more mature than his real age, experienced in outdoor activities with very impressive outdoor skills. It was in his birthday gathering when I first tried climbing and couldn’t stop since ever.

There are always awesome and sweet buddies everywhere. Throughout my solo travel in Europe in late 2017, I met a few great friends. While hiking in Italy, I met Jonas, a happy and passionate German entrepreneur who’s travelled to about 30 countries and lived in 4 countries. Living in China now, he introduced me to many new friends in Beijing who all together not only helped me re-recognised my hometown but also ignited my passion and dream. On the train from Firenze to Pisa, a beautiful girl, later I knew whose name is Andrea, sat next to me. We started talking and got to know each other quickly. She’s from Mexico with so profound mind beyond her age which I enjoyed so much. Roaming around in Rome, I came across an interesting restaurant where half of it is a library. I walked in and talked with the owner, Enzo, who is also an artist, for three hours straight on Buddhism, Yoga, and spirituality. He told me that he’s never talked with a Chinese for such a long time – if he ever talked to one. And in a hostel in Fuessen, I went for a dinner with my hostel roommate Anthony, a British guy from Wales who is very elegant and polite, with deep thoughts and shared interest in nature and classical music with me. So we never know when people come and go, never know who we will meet. I might never in this life be able to see these beautiful souls and faces, but I know they exist somewhere, driving me forward.
The only anchor I know that’s always ready for me there is home - mum, dad, granny, uncle, auntie, cousins. They are always backing me up for my dream, letting me go far to wherever I want, but are always ready to huddle me whenever I need one. That’s love!

Love (Yaoli)
February 2018
Melbourne, Australia
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>List of Figures</td>
<td>xviii</td>
</tr>
<tr>
<td></td>
<td>List of Tables</td>
<td>xx</td>
</tr>
<tr>
<td>I</td>
<td>WHY RIDESHARING FALLS SHORT</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
<td>3</td>
</tr>
<tr>
<td>1.1</td>
<td>Research background and motivation</td>
<td>4</td>
</tr>
<tr>
<td>1.2</td>
<td>Major contributions</td>
<td>8</td>
</tr>
<tr>
<td>1.3</td>
<td>Research questions and hypotheses</td>
<td>9</td>
</tr>
<tr>
<td>1.4</td>
<td>Major outcomes</td>
<td>11</td>
</tr>
<tr>
<td>1.5</td>
<td>Structure of the thesis</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>LITERATURE REVIEW</td>
<td>13</td>
</tr>
<tr>
<td>2.1</td>
<td>The concept of ridesharing and related terms</td>
<td>13</td>
</tr>
<tr>
<td>2.2</td>
<td>The concern of social networks in ridesharing</td>
<td>14</td>
</tr>
<tr>
<td>2.3</td>
<td>Incorporating platial semantics: An activity-based model for ridesharing</td>
<td>16</td>
</tr>
<tr>
<td>2.4</td>
<td>The lack of social networks and platial semantics in time geography</td>
<td>18</td>
</tr>
<tr>
<td>2.5</td>
<td>The state-of-art of ridesharing algorithms</td>
<td>20</td>
</tr>
<tr>
<td>II</td>
<td>FOUNDATIONS FOR COLLABORATIVE ACTIVITY-BASED RIDESHARING</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>THE THEORETICAL FOUNDATIONS</td>
<td>25</td>
</tr>
<tr>
<td>3.1</td>
<td>Time geography and user-centred time geography</td>
<td>25</td>
</tr>
<tr>
<td>3.2</td>
<td>Social network analysis</td>
<td>26</td>
</tr>
<tr>
<td>3.3</td>
<td>Collaborative activity space and new ridesharing prototypes</td>
<td>27</td>
</tr>
<tr>
<td>3.3.1</td>
<td>The elements in CAS</td>
<td>28</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Basic unit: the adjusted PPA</td>
<td>28</td>
</tr>
<tr>
<td>3.3.3</td>
<td>The flexibility of activities: Genres of places</td>
<td>29</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Other factors of platial semantics</td>
<td>31</td>
</tr>
<tr>
<td>3.3.5</td>
<td>The envision of a dynamic CAS: a conceptual model</td>
<td>31</td>
</tr>
<tr>
<td>4</td>
<td>EXPERIMENT MATERIALS</td>
<td>33</td>
</tr>
<tr>
<td>4.1</td>
<td>Agent-based simulation</td>
<td>33</td>
</tr>
<tr>
<td>4.2</td>
<td>Study area and dataset</td>
<td>33</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Victorian Integrated Survey of Travel and Activity</td>
<td>34</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Social network permutation</td>
<td>35</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Points of Interest from Yelp API</td>
<td>35</td>
</tr>
<tr>
<td>III</td>
<td>PROPOSED RIDESHARING MODELS</td>
<td>37</td>
</tr>
<tr>
<td>5</td>
<td>HOW MUCH IS TRUST</td>
<td>39</td>
</tr>
<tr>
<td>5.1</td>
<td>Major contributions</td>
<td>39</td>
</tr>
<tr>
<td>Chapter</td>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>5</td>
<td>5.2</td>
<td>The algorithm of ridesharing considering friendship</td>
</tr>
<tr>
<td></td>
<td>5.2.1</td>
<td>Conceptual experiment design and data structure</td>
</tr>
<tr>
<td></td>
<td>5.2.2</td>
<td>Detour cost and ridesharing feasibility functions</td>
</tr>
<tr>
<td></td>
<td>5.2.3</td>
<td>The Social Network based Ridesharing algorithm (SNeRs)</td>
</tr>
<tr>
<td>5</td>
<td>5.3</td>
<td>Simulation implementation</td>
</tr>
<tr>
<td></td>
<td>5.3.1</td>
<td>Theoretical implementation</td>
</tr>
<tr>
<td></td>
<td>5.3.2</td>
<td>Empirical study</td>
</tr>
<tr>
<td>5</td>
<td>5.4</td>
<td>Results</td>
</tr>
<tr>
<td></td>
<td>5.4.1</td>
<td>From theoretical analyses</td>
</tr>
<tr>
<td></td>
<td>5.4.2</td>
<td>From empirical studies</td>
</tr>
<tr>
<td>5</td>
<td>5.5</td>
<td>Discussions</td>
</tr>
<tr>
<td>5</td>
<td>5.6</td>
<td>Conclusions and future work</td>
</tr>
<tr>
<td>6</td>
<td>6.1</td>
<td>Major contributions</td>
</tr>
<tr>
<td>6</td>
<td>6.2</td>
<td>Methodology</td>
</tr>
<tr>
<td></td>
<td>6.2.1</td>
<td>Social network based taxi-sharing algorithm and optimisation</td>
</tr>
<tr>
<td></td>
<td>6.2.2</td>
<td>Construction of taxi-sharing-oriented social networks</td>
</tr>
<tr>
<td></td>
<td>6.2.3</td>
<td>The dynamic co-evolution of taxi-sharing and social network</td>
</tr>
<tr>
<td>6</td>
<td>6.3</td>
<td>Experiment design</td>
</tr>
<tr>
<td></td>
<td>6.3.1</td>
<td>Simulation runs and parameters</td>
</tr>
<tr>
<td></td>
<td>6.3.2</td>
<td>Generating recursive trips over 30 days from VISTA</td>
</tr>
<tr>
<td></td>
<td>6.3.3</td>
<td>Home distribution and distance decay function</td>
</tr>
<tr>
<td></td>
<td>6.3.4</td>
<td>Permutation of social networks</td>
</tr>
<tr>
<td>6</td>
<td>6.4</td>
<td>Results</td>
</tr>
<tr>
<td>6</td>
<td>6.5</td>
<td>Discussions</td>
</tr>
<tr>
<td>6</td>
<td>6.6</td>
<td>Conclusions and future work</td>
</tr>
<tr>
<td>7</td>
<td>7.1</td>
<td>Major contributions</td>
</tr>
<tr>
<td>7</td>
<td>7.2</td>
<td>Methodology</td>
</tr>
<tr>
<td></td>
<td>7.2.1</td>
<td>Model setting and assumptions</td>
</tr>
<tr>
<td></td>
<td>7.2.2</td>
<td>Trip chain construction</td>
</tr>
<tr>
<td></td>
<td>7.2.3</td>
<td>Alternative trip construction and matchup</td>
</tr>
<tr>
<td></td>
<td>7.2.4</td>
<td>Finding the optimal solution</td>
</tr>
<tr>
<td>7</td>
<td>7.3</td>
<td>Simulations of activity-based ridesharing in Yarra Ranges</td>
</tr>
<tr>
<td></td>
<td>7.3.1</td>
<td>Trip distribution and time horizon</td>
</tr>
<tr>
<td></td>
<td>7.3.2</td>
<td>Alternative destination choice set retrieval</td>
</tr>
<tr>
<td></td>
<td>7.3.3</td>
<td>Optimization with linear programming</td>
</tr>
<tr>
<td>7</td>
<td>7.4</td>
<td>Systematical tests and results</td>
</tr>
<tr>
<td>7</td>
<td>7.5</td>
<td>Discussions</td>
</tr>
<tr>
<td>7</td>
<td>7.6</td>
<td>Conclusions and Future work</td>
</tr>
</tbody>
</table>
8 COLLABORATIVE ACTIVITY-BASED RIDESHARING 95
  8.1 Major contributions ..................................... 95
  8.2 The Collaborative Activity-based Ridesharing (CAR) model 96
    8.2.1 The workflow of CAR ............................... 96
    8.2.2 Time geography and the Spatio-Social Index ........ 98
    8.2.3 The influence of spatial distribution ............... 100
  8.3 The comparison of different models and empirical studies 101
  8.4 Results .................................................. 104
  8.5 Discussions .............................................. 107
  8.6 Conclusions and future work ........................... 109

9 INTRODUCING RIDESHARING INTO PARKING PROBLEMS 111
  9.1 Major contributions ...................................... 111
  9.2 Background of Parking-and-ridesharing .................. 112
  9.3 Conceptual Design of Experiment ......................... 113
    9.3.1 Model context ....................................... 113
    9.3.2 Ridesharing strategy ............................... 117
    9.3.3 Decision-making on the user type .................. 117
  9.4 Model Implementation .................................... 119
    9.4.1 Value of time ....................................... 119
    9.4.2 Willingness for ridesharing ......................... 120
    9.4.3 Traffic flow rate .................................... 120
    9.4.4 Information of travellers on estimated waiting
          time ...................................................... 121
    9.4.5 Parking search in parking lot ....................... 121
    9.4.6 Duration of stay in city centre ..................... 122
    9.4.7 Modeling of public transit .......................... 122
    9.4.8 Choice of the rest parameters ...................... 123
  9.5 Results ................................................... 123
    9.5.1 Evaluation for default parameters .................. 124
    9.5.2 Evaluation for parameter variations ............... 125
    9.5.3 Evaluation of modified cost function for rideshar-
          ing passengers ....................................... 126
  9.6 Discussions .............................................. 128
  9.7 Conclusions and future work ........................... 129

IV THE MAJOR OUTCOMES ....................................... 131
10 DISCUSSIONS AND OUTLOOK 133
11 CONCLUSIONS ................................................ 137

BIBLIOGRAPHY ..................................................... 139
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The conceptual space-time prism</td>
<td>26</td>
</tr>
<tr>
<td>2</td>
<td>Three essential elements in Collaborative Activity Space</td>
<td>29</td>
</tr>
<tr>
<td>3</td>
<td>Potential path area in Collaborative Activity Space</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>Time constraint in the merged route</td>
<td>44</td>
</tr>
<tr>
<td>5</td>
<td>Flowchart of the Social Network based Ridesharing algorithm (SNeRs)</td>
<td>45</td>
</tr>
<tr>
<td>6</td>
<td>Spatial distribution of friendships</td>
<td>51</td>
</tr>
<tr>
<td>7</td>
<td>Simulation results with different scenarios</td>
<td>55</td>
</tr>
<tr>
<td>8</td>
<td>The dynamic interaction between taxi-sharing and social network</td>
<td>67</td>
</tr>
<tr>
<td>9</td>
<td>Output results from taxi-sharing over 30 days (coloured)</td>
<td>70</td>
</tr>
<tr>
<td>10</td>
<td>Distributions of social networks on day0, day24, and day29 (coloured)</td>
<td>72</td>
</tr>
<tr>
<td>11</td>
<td>Workflow of the activity-based ridesharing model</td>
<td>77</td>
</tr>
<tr>
<td>12</td>
<td>Activities of a day and trip chains</td>
<td>79</td>
</tr>
<tr>
<td>13</td>
<td>Potential path area and alternative destinations of a trip chain and of the trips on the chain</td>
<td>84</td>
</tr>
<tr>
<td>14</td>
<td>Combined route cost and visiting sequence</td>
<td>85</td>
</tr>
<tr>
<td>15</td>
<td>Distributions of number of matched pairs by population size</td>
<td>90</td>
</tr>
<tr>
<td>16</td>
<td>Workflow of the CAR model</td>
<td>97</td>
</tr>
<tr>
<td>17</td>
<td>The SSI and the 8-connected n-neighbourhood</td>
<td>98</td>
</tr>
<tr>
<td>18</td>
<td>The SSI overlapped area and PPA in time geography</td>
<td>99</td>
</tr>
<tr>
<td>19</td>
<td>The spatial distribution of the population and their friendships</td>
<td>105</td>
</tr>
<tr>
<td>20</td>
<td>The conceptual scenario of the P&amp;S system</td>
<td>114</td>
</tr>
<tr>
<td>21</td>
<td>Distribution of incoming vehicles over the day</td>
<td>120</td>
</tr>
<tr>
<td>22</td>
<td>Search time for parking depending on the parking lot’s occupancy rate.</td>
<td>122</td>
</tr>
<tr>
<td>23</td>
<td>Gamma distribution describing the parking duration at the central parking.</td>
<td>123</td>
</tr>
<tr>
<td>24</td>
<td>The comparison of the average total travel costs and the average travel time</td>
<td>124</td>
</tr>
<tr>
<td>25</td>
<td>Change in the ratio of different types of travellers and the number of free parking spaces at the central parking with the hour of day.</td>
<td>125</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>------</td>
</tr>
<tr>
<td>26</td>
<td>The changes in ratio of different types of travellers with the total number of travellers to and from the city centre.</td>
<td>126</td>
</tr>
<tr>
<td>27</td>
<td>The changes in ratio of types of travellers with the value of time.</td>
<td>126</td>
</tr>
<tr>
<td>28</td>
<td>The changes in ratio of types of travellers affected by different factors of public transit.</td>
<td>127</td>
</tr>
<tr>
<td>29</td>
<td>The changes in ratio of different types of travellers with the total number of travellers to and from city centre.</td>
<td>128</td>
</tr>
</tbody>
</table>
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>VISTA data structure</td>
<td>34</td>
</tr>
<tr>
<td>Table 2</td>
<td>Structures/Classes in matching algorithm</td>
<td>42</td>
</tr>
<tr>
<td>Table 3</td>
<td>The decision making matrix</td>
<td>48</td>
</tr>
<tr>
<td>Table 4</td>
<td>Statistics of generated social networks</td>
<td>50</td>
</tr>
<tr>
<td>Table 5</td>
<td>Setting of input parameters in ridesharing</td>
<td>52</td>
</tr>
<tr>
<td>Table 6</td>
<td>Friend ratio</td>
<td>57</td>
</tr>
<tr>
<td>Table 7</td>
<td>Results from empirical studies</td>
<td>58</td>
</tr>
<tr>
<td>Table 8</td>
<td>Variable catalogue</td>
<td>63</td>
</tr>
<tr>
<td>Table 9</td>
<td>Weights of terms in utility function</td>
<td>67</td>
</tr>
<tr>
<td>Table 10</td>
<td>Statistics of the tested samples</td>
<td>89</td>
</tr>
<tr>
<td>Table 11</td>
<td>Statistics of trips and matching results</td>
<td>92</td>
</tr>
<tr>
<td>Table 12</td>
<td>Catalogue of models and parameter settings</td>
<td>102</td>
</tr>
<tr>
<td>Table 13</td>
<td>Output variables to be compared</td>
<td>103</td>
</tr>
<tr>
<td>Table 14</td>
<td>Statistics of the direct friendship links in each social network (Link length in metres)</td>
<td>104</td>
</tr>
<tr>
<td>Table 15</td>
<td>Simulation outputs by each model</td>
<td>106</td>
</tr>
<tr>
<td>Table 16</td>
<td>Inputs and outputs of the model</td>
<td>116</td>
</tr>
</tbody>
</table>
Part I

WHY RIDESHARING FALLS SHORT

This part explains the motivation and background of my PhD thesis. Based on a review of current ridesharing markets and technologies, a discussion is developed on why ridesharing is not yet as popular as its expected potential by some academic studies. This part goes through the origin and the concept of ridesharing and other broadly related terms, the up-to-date status of the ridesharing market in both operation and technology backup, and the downsides of ridesharing that hinder its further promotion.
INTRODUCTION

The rapid increase of urbanisation and urban dwellers globally contributes to the exponential growth of travel demand, which becomes a hassle for city residents. Longer commuting time, heavier parking burden, higher fuel consumption, and worse pollution all add to people’s daily stress [International Transport Forum, 2013]. In many big cities, for example, Melbourne, car travel not only takes a significant portion of travel (80%), but also takes significantly longer time (e.g., 70% longer in Melbourne) during peak hours than in free flow¹. Beijing, as a global mega-city, is even worse with an average of 52 minutes one way commuting to work². What’s even worse, long time commuting is argued to cause harm to people’s brain slowly³. In this regard, ridesharing has a good potential to reduce urban traffic load, as well as fuel consumption, air pollution, and the stress with long commuting time⁴.

Ridesharing is an emerging transport mode that harnesses both private cars and taxis to combine two (groups) of travellers into the same vehicle in an ad-hoc manner, if all or part of the two groups’ travels are overlapped in space and time. The ad-hoc manner means rides are matched on real-time or short-notice demands instead of by pre-arrangement.

Despite the popularity expected by some academic studies, ridesharing turns out to have a much lower uptake rate in real life [Amey, 2010]. The contrasting phenomena, however, is the success of ridesourcing. Ridesharing is often confused with the terms of ridesourcing or ride-hailing. The latter two refer to the arrangement that helps people to find on-demand available vehicles, including private cars or taxis, no matter the vehicle is shared with other groups of travellers or not. Ridesharing, in this sense, is a subset of ridesourcing where trips are shared. The competition in ridesourcing leads to the top two “billion

¹ Melbourne’s traffic congestion is worse than Sydney, researcher finds. http://www.news.com.au/finance/work/at-work/melbournes-traffic-congestion-is-worse-than-sydney-researcher-finds/news-story/2a176e8e0b2adbe2238e9dd8ef3cebe1
⁴ Here’s how ride sharing can help save the planet, and your commute time. https://www.weforum.org/agenda/2017/01/heres-how-ride-sharing-can-help-save-the-planet-and-your-commute-time
dollar startups": Uber and Didi Chuxing. But unfortunately, the consequence of adopting ride-hailing, if not shared, might deteriorate the traffic congestion: For one thing, it attracts people off public transit; for another, drivers act for the passenger demands driving extra kilometres that might have been unnecessary. Seldom has attention been paid to how much share really goes to ridesharing in the whole ridesourcing market. In Uber’s business, for example, only UberPool is ridesharing.

In fact, coinciding with the socio-psychological factors discussed by Amey [2010], Chaube et al. [2010], Wessels [2009], and Gargiulo et al. [2015], a few articles point out that people subjectively dislike ridesharing, though they embrace ridesourcing. The main reasons are drawn to two factors: trust and flexibility. Trust is discussed in this context with regard to security, privacy, and comfort of sharing rides with others. Flexibility refers to not only the physical spatio-temporal constraints but also the subjective detour tolerance (i.e., the maximum detour time) for different people and travel purposes. The two essential factors are integrated into the ridesharing algorithm design by this thesis. With improved trust and flexibility, ridesharing should become a more attractive mode choice for city dwellers. Thus, the proposed new algorithms aim to increase ridesharing rates for such benefits as saved travel length. For example, it ideally can cut cumulative trip length by at least 40% in the case of New York City [Santi et al., 2014].

1.1 RESEARCH BACKGROUND AND MOTIVATION

Demand-responsive transport (DRT) was proposed calling for the coalition of the merits of private transport and public transit (e.g., Brake et al. [2004]; Ronald et al. [2017]). Its essential idea is to use vehicles of flexible schedules and routes to carry a group of people based on the estimated travel demand in a relatively short notice (e.g., a few hours or days). One type of DRT recently emerging is the semi-private transit mode of ridesharing or taxi sharing (i.e., ridesharing with a taxi).

---

7 The ride-hailing effect: More cars, more trips, more miles. https://www.citylab.com/transportation/2017/10/the-ride-hailing-effect-more-cars-more-trips-more-miles/542592/
8 Uber is a mobile application allowing travellers to call for a ride from taxis or private cars in real time, while UberPool is a particular service of Uber that allows people to call rides to be shared with other travellers.
Ridesharing originated as a general concept to carry more than two people in the car of one of the persons who share a portion or all of their time schedules and routes [Furuhata et al., 2013]. By the general definition, the pre-arranged daily routines such as commuting with colleagues or taking neighbours’ children together to school can be counted as ridesharing as well. However, the concept of ridesharing evolves nowadays as a real-time and ad-hoc arrangement of cars to combine people’s travels⁹. Real-time and ad-hoc refer to different aspects, respectively. Real-time means that the requests are gathered and matched in a relatively narrow time window, such as a few minutes, when new requests still come in during the process. However, real-time ridesharing algorithms are subject to the bias of the size of a time window. The longer the time window, the more approximate to the global optimal solution the result [Quadrifoglio et al., 2008]. For this reason, the experiments in the following chapters are not real-time ridesharing in a strict sense. Instead, the models were built for pre-planning of a whole day just to understand the best scenario of matched routes with less bias of the time window size. Ad-hoc means that the requests vary from time to time, for which a one-off planning is not applicable and matches should be customised to each specific scenario. The thesis is based on the concept of ad-hoc ridesharing with a time window as wide as a whole day. However, this does not conflict with ad-hoc ridesharing because planning would have been unnecessary if the ridesharing routes are repetitive everyday.

Convincing people to accept ridesharing is influenced by travellers’ psychology. For private car users, shifting to ridesharing (no matter whether as a driver or a passenger) means to compromise privacy and travel flexibility. For public transit users, security is more sensitive when it gets to such a small space as a car. The blog 7 Reasons Why I Hate Uberpool and Lyftline¹² listed seven reasons that repulse people from ridesharing, out of which five are drawn to insecurity or inflexibility. Amey [2010] has discussed that the factors contributing to or challenging ridesharing rates include a wide range from economic and behavioural to institutional and technological. Sarriera et al. [2017] compared the positive and negative influences of behavioural factors against the spatio-temporal and monetary cost. They revealed that the attraction of good social interaction still cannot compensate time cost, but the negative perception of unpleasant social interactions overtakes the benefit of saved money by ridesharing. However, the constraints that mutual trust and subjective space-time flexibility of the travel partners are not yet well understood in a systematic way. This thesis is focused on understanding the constraints that trust and space-time flexibility of travel partners impose on the viability of ridesharing.

---

¹¹ More related terms are elaborated in Chapter 2.
¹² https://therideshareguy.com/7-reasons-why-i-hate-uberpool-and-lyftline/
Though trajectory analysis has indicated good potential of ridesharing according to space-time concurrence [Santi et al., 2014], the participation rate of ridesharing is not yet satisfying [Amey, 2010]. The low willingness for ridesharing with strangers signifies that many of the existing ride opportunities are actually unacceptable for certain people, while the willingness to share rides with friends are significantly higher [Chaube et al., 2010]. Such observations call for the necessity of considering social networks in ridesharing [Wessels, 2009; Chaube et al., 2010]. Social network is not the only factor that contributes to the trust level between ridesharing travellers, but is a pragmatic measure of trust. Factors such as reputation accounting and privacy protection as proposed by Caballero-Gil et al. [2017] can be reduced to social networks. The definition of social network and friend in this thesis are based on the level of trust in ridesharing. A further discussion on these terms are given in Chapter 3.

**Trust**

The preference for travelling with friends seems to be a good reason to limit ridesharing to friends. Realising this argument, Facebook filed recently a patent called Event-based ridesharing [Richardson et al., 2016], which allows drivers to select users based on his/her social network connections for an online negotiation on carpooling to an event. However, a ridesharing system limited to friends has yet to address the impact of missing chances of getting a ride. Ridesharing exclusively with friends while declining offers from strangers may lead to fewer opportunities to get a ride within a given space-time budget (i.e., a feasibly allocated amount of time depending on when and where to go) and to higher detour costs. Chapter 5 aims to examine the theoretical costs and benefits of ridesharing with friends by systematic and comprehensive variation of parameters in a simulation beyond a particular context. Empirical tests are run as validations. As social networks are gaining attention in travel behaviour research [Arentze and Timmermans, 2008; Hackney and Marchal, 2011], this chapter helps gain insightful understanding into how spatial structures of social networks affect ridesharing results.

One step further, social networks are not static, but are intertwined with travel behaviours. Travel behaviours demonstrate heterogeneous patterns with different social engagements [Hackney and Marchal, 2011; Carrasco and Miller, 2009; Han et al., 2011]. Meanwhile, social networks will emerge from travel behaviours due to the co-presence of activities [Hackney and Marchal, 2011]. Recent emergence of taxi-sharing, as well as its future forms of shared mobility such as shared autonomous vehicles or mini-buses, reminds of the underlying significance of trust and safety [Chaube et al., 2010; Wessels, 2009; Fagnant and Kockelman, 2015]. For this reason, social networks evolve.
throughout time to preserve trusted people. Using taxi-sharing as an example, Chapter 6 uncovers how space-time networks are interwove into social networks by examining the mutual change of social networks and taxi-sharing outcomes. This work also yields important indications for policy makers why bringing social networks into taxi-sharing is crucial.

**Space-time flexibility**

The second fundamental factor of ridesharing investigated in this thesis is the space-time flexibility. People choose a transport mode because it is convenient. Since ridesharing requires coordination and bundled travel, any ridesharing algorithm should accommodate riders’ flexibility. Current ridesharing algorithms are *trip-based*, i.e., they respect *a priori* defined origins and destinations, but leave the spatial flexibility of riders un-exploited. The spatial flexibility arises from the fact that some activities can be participated at one of various destinations that are functionally similar [Kwan, 1998; Miller, 2005, 1991]. Accordingly *activity-based* travel planning broadly categorises activities as fixed location-wise and time-wise, e.g., work-place related activities, and flexible activities, e.g., shopping. By incorporating an activity-based planning approach into a ridesharing algorithm, it is likely that the matching chances rise as a result of the expanded destination choices. Chapter 7 takes the endeavour by proposing an *activity-based ridesharing* that employs time geography to build a choice set of alternative destinations and finds an optimal match fitting best into ridesharing partners’ schedules.

**Subjective space-time flexibility and collaborative activity-based ridesharing**

When thinking about social networks and space-time flexibility together, a social network’s limitation on ridesharing choices is a distortion on activity space [Wang and Winter, 2017]. On one hand, higher willingness of ridesharing and higher detour tolerance are granted to closer social acquaintances [Chaube et al., 2010; Wessels, 2009]. Relying on real-life social networks for ridesharing, on the other hand, might threaten the match rate by refusing offers from nearby strangers. Activity-based ridesharing [Wang et al., 2016b] hence offers a potential way to enhance the rides between friends: if a shared ride with a friend was originally not feasible, an alternative destination can reverse the situation. Even if there is still no feasible ride from a friend, alternative destinations can still potentially increase the overall ridesharing rates by matching more strangers. Accordingly, Chapter 8 proposes the *collaborative activity-based ridesharing*, a new solution for ridesharing to tackle trust and flexibility at the same time,
and discusses the influence of spatial densities and distributions of social network links and of travel demands.

_Last-mile ridesharing for urban centre parking_

Rather than an independent system, ridesharing contributes to other aspects in urban transportation. For instance, parking issue can leverage ridesharing for a better solution. Some strategies have already been established in various forms to reduce the number of vehicles looking for parking, for example, by shuttle bus transfers to a single destination at a particular time (e.g., an event), or by park-and-ride offerings (P&R) where public (mass) transport is co-located with satellite parking. Meanwhile, private ridesharing combined with managed parking has been tried but still remains a niche market\(^\text{13}\). One reason is likely that the popularity of ridesharing in reality [Dubernet et al., 2013b] is far less than its expected and proven potential [Tachet et al., 2017] due to the issues discussed above. Chapter 9 suggests an alternative called Park & RideSharing (P&S) that only leverages ridesharing from satellite parking spaces to urban centres instead of ridesharing from home front door. Since the shared portion of travel is shortened, P&S is expected to accommodate the unwillingness to ridesharing due to trust and flexibility, while it still reduces the traffic driving into urban centre.

1.2 **Major Contributions**

This thesis targets to build a ridesharing prototype called *Collaborative Activity-based Ridesharing* based on a revised time geography framework called *Collaborative Activity Space* to address the following two aspects in collaborative travels:

1. A collaborative travel model, e.g., a ridesharing algorithm, should embrace social network as it impacts both the participation willingness and the flexibility (detour tolerance) of the travel. The influence of social network on collaborative travel is not yet thoroughly understood. This thesis provides the first computational exploration of social network on the feasibility of ridesharing. The spatial distribution of social network, its topological structure, and its dynamic evolution over time based on historical travel experience are of particular interest.

2. A collaborative travel model, e.g., a ridesharing algorithm, should enhance travel flexibility by a reverse thinking of space-time budget: instead of calculating how much time to use for a pre-decided destination, people can allocate time budget first before deciding where to go. Time geography theories are adopted to transform from location-based ridesharing to activity-based

\(^{13}\) [http://liftango.com/](http://liftango.com/)
ridesharing. Inspired by activity-based travel demand analysis in transportation studies, the latter concept considers multiple location choices for the same activity and within a given space-time budget.

1.3 RESEARCH QUESTIONS AND HYPOTHESES

The overarching research objective is to propose an effective and innovative framework that encourages the public incentive of collaborative travel by addressing the socio-psychological issues from the perspectives of social networks and space-time flexibility. Ridesharing prototype algorithms are specifically investigated as an implementation of the collaborative travel framework: trust and space-time flexibility are harnessed to increase match rates of ridesharing. Nested under the overarching theme, this thesis comprises the following five research questions/hypotheses that are associated with five chapters:

i. **Would sharing rides with priority to friends hinder the chance of getting a ride?** To answer this, Chapter 5 conducts a systematic and theoretical overview of the influence of social network on social network based ridesharing (SNeRs), with the hypotheses:
   - Sharing ride with a friend does not necessarily increase the detour cost;
   - Prioritising friends in ridesharing can significantly increase the matching with friends, which relieves the concern of security and trust to some extent.

ii. **Are the social network and ridesharing behaviours co-evolving throughout time to get higher match rates?** Chapter 6 performs a systematic investigation into the evolution of social network topological structure and spatial distribution based on historical ridesharing experience, with the hypothesis:
   - The topological structure and spatial distribution of social network in ridesharing system evolves to reach significantly higher ridesharing rate and less detour cost.

iii. **Would ridesharing chances be increased if alternative destinations are considered?** Chapter 7 tries to take the advantage of time geography to increase ridesharing flexibility, with the hypothesis:
   - Activity-based ridesharing algorithm (ABRA) can significantly increase the successful rate of ridesharing compared with traditional trip-based method.

iv. **Are alternative destinations able to enhance the chance of matches between friends?** Chapter 8 does an integration of social network based ridesharing and activity-based ridesharing, called collaborative activity-based ridesharing (CAR), with the hypotheses:
v. **Is ridesharing an effective solution to relieve urban centre parking burden?** Chapter 9 proposes a parking model that leverages the benefit of ridesharing for higher parking efficiency, looking at the hypothesis:

• Introducing ridesharing into parking system is beneficial over both solo driving and even park-and-ride (with public transit), by saving the overall travel time.

The five research questions are interlocked and progressive. Trust and flexibility are intertwined in a complicated way. Trust in friends may boost the motivation for ridesharing but withholds ridesharing chances from strangers who are immediately nearby. Trust also augments the willingness to detour even more for a friend, leading to an overall detour cost increase. Both the willingness and the detour tolerance for a friend and for a stranger are different. To understand this relationship, Research question i builds up a static model involving social network on a basic and theoretical level. In a long run, the social network structure might change according to previous ridesharing experience with a person, which in turn affects the willingness and detour tolerance for the next round of ridesharing and in the future. Therefore, Research question ii extends question i to investigate the indicated social network evolution based on ridesharing experience.

From another perspective regardless of trust, Research question iii proposes ABRA that explores the potential to increase travel flexibility by expanding the destination choice set composed of locations that serve a similar travel purpose but are subscribed to space-time budget. In this way, even people whose original plans could not fit can be matched eventually. However, the expansion of destination choice set is aimless in terms of whom to share with. Bearing in mind the flexibility bounded by trust, it is unlikely to force a potential match between two strangers who apply tight tolerance for each other, even though the match would be fine if they were friends. Therefore, question iv incorporates social network and inspect their trade-off as the implementation of the collaborative activity-based ridesharing.

The Park and RideSharing (P&S) system in research question v is a pragmatic application of ridesharing to improve the operation of parking. It shreds light on at what scenarios ridesharing can overtake public transit systems and solo driving.
1.4 Major Outcomes

The feasibility of the overarching collaborative travel framework and the research hypotheses are backed up by the agent-based simulation results in this thesis. Here are the general highlights:

- Trust from exiting social networks or historical travel experience has a significant influence on collaborative behaviours in both short term and a long run. **Hint:** When designing collaborative travel systems, take the initial social network into account and consider the feedback from travellers.
- The alternative destinations for the same travel purpose subject to a given space-time budget can increase the chance of collaborative travels (i.e., ridesharing in this thesis). **Hint:** Travel purpose is more important than destination as it is the primary drive of travel. The space-time flexibility of a travel purpose should be exploited to suggest where to go.
- Leveraging space-time flexibility for alternative destinations contributes to more matches with preferred travellers. **Hint:** The flexibility indicates not only where to go but also whom to travel with, which can potentially increase the satisfaction rate of collaborative travels.
- Social network is a heuristic to refine space-time flexibility by adding subjective preferences to travel budgets. **Hint:** Subjective preferences in collaborative travels can eliminate unnecessary searches in space and time for potential matches.

1.5 Structure of the Thesis

This thesis is organised in the following structure:

Part i encapsulates Chapter 1 (the current chapter) and Chapter 2.
- Chapter 2 reviews the essential concepts related to ridesharing, the state-of-art of ridesharing research, the current algorithms designed for ridesharing, and more details of previous studies linked to the proposed new ridesharing algorithms.

Part ii is the general background, theories, and techniques of this thesis.
- Chapter 3 introduces the theoretical foundations of this thesis, including the existing theories from social network analysis and time geography, and more importantly, the proposed framework of collaborative activity space that incorporates the distortion of time geography by social networks.
- Chapter 4 elaborates the technical tools, study area, and datasets used in this thesis.
Part iii, including Chapter 5 to Chapter 9, specifies the specific models and experiments corresponding to each research question and hypothesis listed in Section 1.3.

- Chapter 5 introduces the *Social Network based Ridesharing*;
- Chapter 6 investigates the dynamic coevolution between the social network structure and its related taxi-sharing behaviours;
- Chapter 7 proposes the *Activity-based Ridesharing*;
- Chapter 8 elaborates how social network based and activity-based methods are blended as *Collaborative Activity-based Ridesharing*.
- Chapter 9, is an innovative model integrating ridesharing into parking systems.

Part iv is the final part including an overarching discussion on the major findings and outlook of the challenges (Chapter 10), and a conclusion of the works having been done along with the significant contributions (Chapter 11).
LITERATURE REVIEW

This chapter first does a review of the similar concepts to ridesharing in Section 2.1. From there on, Section 2.2 does a review of the trust issue in ridesharing which introduces the necessity and potential of social networks in ridesharing. Section 2.3 discusses the tradition of activity-based versus trip-based travel demand analysis in transportation studies, laying the foundation of activity-based ridesharing. Section 2.5 summarises the current situation of ridesharing algorithms on the market, based on which the contribution of this thesis is remarked.

2.1 THE CONCEPT OF RIDESHARING AND RELATED TERMS

Ridesharing is an emerging mode of transportation that individual travellers collaborate to share the same vehicle as well as the accompanied cost such as petrol for their own trips, which ideally has the benefits both of flexibility as private cars and of the reduced cost as public transit [Furuhata et al., 2013]. The key point is that the driver also has a travel purpose rather than taking passengers only. Furuhata et al. [2013] generally classified ridesharing into unorganised and organised ridesharing. The major difference is whether there is prearrangement to coordinate the joint trips by a central agency. According to the authors, the lack of central coordination in unorganised ridesharing means these trips are more likely to be built via existing relationships, while organised ridesharing can be promoted among strangers by prearrangement. In this sense, organised ridesharing is claimed easier to scale up due to its independence of previous history involvements [Dailey et al., 1999; Furuhata et al., 2013].

Under the category of organised ridesharing, the regular one arranged in advance is called carpool (e.g., Ferguson [1997]). Dial-a-ride is the service that matches and provides rides in short notice for any requested route within an area [Furuhata et al., 2013]. Semi-organised ridesharing is called flexible carpooling, where people pop up at a pre-determined location, such as a meeting point [Stiglic et al., 2015], to organise shared rides. Dynamic ridesharing is the on-demand real-time ridesharing that is centralised for prearrangement in short notice or even en-route [Agatz et al., 2012]. Dynamic ridesharing targets both commute trips (e.g., UberPool) and long-distance inter-city trips (such as Blablacar). The Internet technology seems plausible to encourage dynamic ridesharing, which, however, turns out not to be the truth for many other reasons [Furuhata et al., 2013; Amey, 2010],

13
including the two factors – trust and flexibility – to be addressed by this thesis.

Ridesourcing, on comparison, is the service to search for a private ride offered by the car owner with smartphone application [Rayle et al., 2016; Zha et al., 2016]. In ridesourcing, the driver serves as a private taxi driver. Taxi-sharing is also a type of ridesharing where the driver is a professional taxi driver and there are at least two (groups of) passengers, whereas in the private-car ridesharing the driver is both a driver and a passenger. Currently ridesharing is undergoing an expansion from the traditional driver-passenger private car ridesharing to taxi-sharing [Hosni et al., 2014; Ma et al., 2013; Martinez et al., 2015], and in the near future to autonomous vehicle (driverless) taxi-sharing [Krueger et al., 2016]. The latter two have in common that no rider is any longer bound to a certain vehicle, in contrast to ridesharing where the driver sets out first and arrives at her destination last. Taxi-sharing model hence can be transferred to autonomous vehicle model immediately.

This thesis is attempting to address on-demand (organised) ridesharing, but is not dynamic ridesharing in terms of modelling. The models perform a full-day prearrangement to get a full picture of the maximum potential when new factors are brought into ridesharing systems. Some chapters (Chapters 6, 7, 8) use taxi-sharing as the model scenario to release the bundle to a vehicle.

2.2 THE CONCERN OF SOCIAL NETWORKS IN RIDE SHARING

Based on trajectory analysis or matching model assumptions [Agatz et al., 2011; Amey, 2010; Dubernet et al., 2013a; He et al., 2012; Santi et al., 2014], a significant portion of inner-urban trips can be merged for ridesharing. For example, Amey (2010) reported a rate of 50% to 77% of the commuting trips to be merged; Dubernet et al. [2013a] informed by simulation that about 90% of the trips in Zurich metropolitan area can be matched as 2-person ridesharing. Such observations seem to prove ridesharing as a promising transportation mode with potential benefits such as lower costs to individuals, reduced traffic amount, and decreased emissions.

On the other hand, however, the participation rate in ridesharing is known to be a magnitude lower in reality. In the same MIT community, only 8.2% of the community population chooses to share rides [Amey, 2010]. Trust (for safety and comfort) is pointed as one of the most influential obstacles for ridesharing [Amey, 2010; Chaube et al., 2010; Wessels, 2009]. In their testing scenario for smartphone enabled collaborative travel, Dickinson et al. [2017] confirmed the importance of social network, particularly social ties and trust rather than age or sense of community. Brereton et al. [2009] also emphasised the necessity of involving social network into ridesharing based on their
technology review, but did not give a concrete solution. By microsimulation, Dubernet et al. [2013a] justified behavioural factor as the most limiting factor of ridesharing, which strongly substantiates the significance of the social network based ridesharing (Chapter 5).

People are significantly less willing to share a ride with strangers than with direct or indirect friends; and even if they do, they have much lower detour tolerance for strangers. Chaube et al. [2010] reported that 98% of the population of Virginia Tech university community would accept a ride from a friend, 69% accept from a friend of a friend, and only 7% from a stranger.

In addition and related to matching preferences, the ridesharing demand analysis on the University of Maryland campus points out the significance of service flexibility as a key factor to ridesharing [Erdoğan et al., 2014]. This is coped in this work as different detour tolerances with varied friends. Wessels [2009] reports the detour tolerance for ridesharing with a friend is about 25% of the shortest possible travel time, while it can be as low as about 6% for a trip with strangers. Thus, despite fewer opportunities for ride choices with friends, the higher tolerance might lead to a substantially higher uptake rate of ridesharing. The acceptance of shared autonomous vehicles is even more ambiguous considering the comfortability of sitting next to a stranger in a limited space, without the presence of a typically trusted taxi driver [Krueger et al., 2016]. Therefore, trust measure is a more prominent issue for autonomous taxi-sharing.

The mutual influence between social networks and travel behaviours has been highlighted before. Arentze and Timmermans [2008] pointed out that utility in social network and activity-travel pattern co-evolution is a function of dynamic social and information needs that drive the change of the social network structure, and of the similarity between the relevant characteristics of involved people. Farber et al. [2014] proposed the social interaction potential metric, a quantitative potential to participate in a face-to-face activity at a particular location. This metric is, however, about encounters at locations rather than bundled travels, the latter of which is addressed by Chapter 6.

Hillier [2007] coined reflexive level activity for graph-induced behaviour while non-reflexive level activity for graph generating activity. Social networks play a role on taxi-sharing in two reflexive ways: 1) Participation willingness and detour tolerance are dependent on social acquaintance [Chaube et al., 2010; Wessels, 2009]; 2) The geographic distributions of friendships and of the trips each day delineate the physical potential of taxi-sharing. Therefore, Chapter 6 investigates social networks from the perspectives not only of topological structure but also of spatial distribution. On the other hand, the non-reflexive level activity is the feedback from taxi-sharing experience on generating new social links.
Furthermore, the construction of potential social network links can be inferred by people’s spatial travel history. Crandall et al. [2010] built a probabilistic model and found co-location due to friendship is about 5,000 times by random chance. Cranshaw et al. [2010] confirmed a positive correlation between the number of social ties and the location entropy, a measure on the diversity of a place to which how many different people have visited. Correspondingly, Wang et al. [2015b] found significantly more movement overlaps between social network contacts versus random encounters. They also observed that small places with fewer people are more effective indications for friendship than more crowded places. Zheng et al. [2011] proposed a model to recommend friends based on the accordance of a sequence of visited locations with shared time slots, weighed by the popularity of locations. Pham et al. [2013, 2016] did not only infer the existence of social ties but also proposed methods to infer social tie strengths. However, these researches focus on the social network of joint activities (locations), which is orthogonal to a social network emerging from bundled travels (routes). Chapter 6 is interested in the emerging social network from taxi-sharing over time, and therefore only considers distance decay as a factor in social network evolution.

### 2.3 Incorporating Platial Semantics: An Activity-Based Model for Ridesharing

The understanding of locations and places in information science and location-based services has a history shifting from coordinated-based approaches to semantic-based approaches. Place, rather than being a synonym for a geographical location, is tightly linked with its service, i.e., its function, affordance and experience, while constraint by its spatial location [Tuan, 1974; Raubal, 2001; Gibson, 1977]. In accordance with this observation, travel demand analysis underwent an evolution from trip-based to activity-based, the latter regarding travel as a secondary or derived demand to satisfy the necessity of activity participation at the destination [Bhat and Koppelman, 1999; Pinjari and Bhat, 2011]. Ridesharing inherently is the bundle of activity-induced individual travels. However, ridesharing planning is predominantly trip-based (see Section 2.5), matching travels from one unique location to another, rather than from activity to activity.

Activity-based travel demand analysis covers a wide range of steps, including trip-chain generation, activity scheduling, and choice set selection. Trip chaining, yet by an ambiguous definition, is “a trip sequencing of activities”, as Thill and Thomas summarised based on their review of trip chaining studies [Thill and Thomas, 1987]. Chapter 7, however, focuses on the generation of destination choice sets in ridesharing, rather than deal with trip-chain generation and activity scheduling itself, which is another research domain (e.g., Arentze and
Timmermans [2009]; Charypar and Nagel [2005]). A full day schedule of activities, some fixed and some flexible, is assumed given in this thesis.

Timmermans et al. [2002] generalised four types of activity-travel modelling for destination choice set: constraints-based, utility-maximising, computational process, and microsimulation models. The first type, essentially time geography based, is the foundation of filtering potential alternative destinations in the proposed algorithm. It is, however, a restrictive modelling of the choice set that omits other criteria such as personal preferences [Märki et al., 2012]. A similar interest in time geography methods for ridesharing has been discussed by Rigby et al. [2013], although they looked at another issue of selecting potential pick-up areas based on space-time accessibility. Kwan and Hong [1998] initiated a network-based time geography method to formulate destination choice sets, which was expanded by Chen and Kwan [2012] into a computational model for multiple fixed and flexible activities’ choice set formation. Chen and Kwan’s method builds a space-time prism by using fixed activities as control points, which cuts down the number of potential combinations with flexible activity destinations.

There are two aspects to be considered in such a constraint-based choice set formation:

1. the identification and separation of fixed and flexible activities;
2. the set-up of the destination choice sets for flexible activities.

The boundary between fixed and flexible is vague as it depends on the perception of the users [Kwan and Hong, 1998; Raubal et al., 2004]. For example, some people only go to a certain restaurant for lunch, but others are more flexible. While Kwan and Hong [1998] defined flexibility only in regard to location, Raubal et al. [2004] also accounted for time elasticity, which is adopted by the algorithm presented in Chapter 7.

The affordance (or function) a place exudes for certain activities [Raubal, 2001; Gibson, 1977] can be collected from POI databases. However, the database might have an inconsistent classification system from the documented activity types used to retrieve flexible destinations. Such a problem can be fixed by the work of McKenzie et al. [2015] that mapped POI types into a present classification schema.

Another challenge is that people may carry out different activities at the same place, which makes the place semantically diverse to each person [Cano et al., 2011]. These problems are yet to be solved, but fall outside the focus of the current work.
2.4 THE LACK OF SOCIAL NETWORKS AND PLATIAL SEMANTICS IN TIME GEOGRAPHY

Traditionally, there is a gap between human decision modelling and spatial behaviour analysis. On one hand, current activity-based approaches are dominated by microeconomics that primarily concentrates on such decision-making aspects as people’s scheduling and travel demand generation [Axhausen and Gärling, 1992; Bhat and Koppelman, 1999; Pinjari and Bhat, 2011; Recker et al., 1986]. However, they do not explicitly deal with spatial behaviours (e.g., the detailed route planning). On the other hand, time geography [Hägerstrand, 1970], since 1970, has been a benchmark in studying more thorough disaggregated spatial movement. Nevertheless, as argued by Damm [1983], time geography is implicit in terms of human behaviours because of its lack of explicit modelling of human decision-making. Many factors in decision-making can lead to a distortion of time geography (e.g., Wang and Winter [2017]). However, because of the split between the modellings of decision-making and of travel behaviours, mostly the distortion is absent in time geography. More broadly speaking, time geography is more homogeneous than heterogeneous amongst agents, which is not the truth.

For several decades, studies of time geography have been heavily centred on the geometric properties and computational methods of space-time prism given pre-specified stations to visit and time constraint. Song and Miller [2015] have developed an analytical method to scrutinise details in space-time prism by cutting the whole prism into time slices (i.e., the projection of space at a time point). Their method enriches the old space-time prism analysis that only considers prism shape for accessibility, but does not address human decision-making, or the heterogeneity in decision-making such as varied preferences. In fact, human-spatial behaviours is a loop including human (decision) behaviours and their spatial (movement) behaviours, which is similar to the concept of Unwelt [von Uexküll and Kriszat, 1956] that considers perceiving and acting cyclic processes [Ortmann and Michels, 2011]. The heterogeneity in human decision-making, therefore, should be built into the time geography framework.

That said, there are some studies, such as the user-centred time geography by Raubal et al. [2004], that introduced human subjective factors into time geography. Raubal et al. [2004]’s work utilises the concept of “affordance” [Gibson, 1977] to address the subjective distortion specific to each individual. The proposed collaborative activity space in this thesis is a computational model derived from their conceptual framework. The decentralised time geography [Raubal et al., 2009] is another way to introduce heterogeneity into time geography by considering
local knowledge and communication range. It is, however, not improving the time geography model directly.

Ridesharing is a combined process both of decision-making and spatial movement in a collaborative context. Therefore, a more feasible computational time geography model tackling subjective heterogeneity should be proposed. Particular to this thesis, two factors are missing in the current computational time geography framework: social networks and platial semantics.

Social network makes a difference in two ways: on the shape of space-time prism and on the negotiation process. For the former, different tolerances are given to people of corresponding level of social proximity that distorts space-time trajectory to varied extents. The synchronisation of closer friends might yield more overlap between their individual space-time prism and travel bundle. Despite methods to quantify the overlapped of two prisms [Miller, 2005; Neutens et al., 2010; Song and Miller, 2015], social network and its strength have not yet been studied directly in time geography. Andris [2016] summarised some methods to integrate social network into GISystems, but it was done from an analytical perspective rather than modelling social networks into time geography. For the second aspect, higher acquaintance hypothetically contributes to higher probability of a successful cooperation [Chaube et al., 2010; Wessels, 2009]. Fang et al. [2011] posited a space-time prism model scheduling joint participation of activities with variable space and time preferences, given that preferences from all participants are known in advance. However, the prior knowledge of preferences cannot satisfy self-organised negotiation between friends and, broadly speaking, partners in ridesharing. For this reason, a semi-cooperative negotiation mechanism by Ma and Wolfson [2013] should be introduced into space-time prism to organise joint activity-travel scheduling.

Another problem is the lack of modelling of platial semantics in time geography. Place has a rich and comprehensive meaning in geography, including physical, cultural, functional, and conceptional dimensions [Relph, 1976; Cresswell, 2014]. Particularly to the context of this thesis, platial semantic refers to not only the functions and services of a place with regard to the travel activity at the destination, but also to the inherent space-time flexibility of that activity related to affordance. Even though activities are considered in travel behaviour [Raubal et al., 2004; Winter and Yin, 2011], there is no systematic or generic thorough understanding of platial semantics’ influence on user-centred space-time budget (e.g., subjective tolerance of the time budget), route planning, or mode choice. In general, the user-centred space-time budget can be adopted to accommodate the affordance due to social network preferences – travelling with someone looks more accessible than with someone else. The missing part is the un-
understanding of places by genres according to their space-time traits: fixed versus flexible location, and fixed versus flexible time.

2.5 THE STATE-OF-ART OF RIDESHARING ALGORITHMS

Many ridesharing algorithms have been developed so far. Furuhata et al. [2013] classifies these matching approaches into six categories: dynamic real-time ridesharing, carpooling, long-distance ride-match, one-shot ride-match, bulletin-board, and flexible carpooling. Dynamic real-time ridesharing that addresses short-term matching or even en-route matching (e.g., Agatz et al. [2011]; Amey [2010]; Deakin et al. [2010]; Ma and Wolfson [2013]) has a great advantage of flexibility of routes and time. Avoiding the bottleneck of a central planner for dynamic real-time matching, others have explored peer-to-peer solutions considering short-range radio communication between pedestrians’ and car drivers’ mobile phone applications [Wu et al., 2008].

However, not sufficient existing matching algorithms have dealt with passengers’ experience from the perspective of social psychology. Current social network and ridesharing studies make some progress but still fall short in several ways:

• Some studies recognise the importance of social relationships in ridesharing, but only in an indirect way. Håll et al. [2012] considered the cost of passenger discomfort in their simulation of dial-a-ride system by measuring that inexplicitly with excess of waiting and riding time. Kamar and Horvitz [2009] expected the impact of friendship would make a big difference on their simulation outputs, but only as future work.

• Existing algorithms integrating social networks into ridesharing do not investigate social networks in such a systematic and comprehensive way as this study does, nor do they focus on the spatial distribution of social networks. For instance, Gidófalvi et al. [2008] suggested a method to group users into ridesharing groups based on the social network structure, but did not discuss the influence of the spatial distribution of social network. An algorithm called Social-aware Ridesharing Group (SaRG) query [Li et al., 2015] implemented a swift strategy querying the matching by different combinations of groups of friends extracted from their social network, and found the best combinations to minimise detour cost by a branch-and-bound optimisation strategy. However, SaRG did not take into account potential matching chances with strangers or the competition between social contacts and strangers based on their spatial distributions.

• Bistaffa et al. [2015] defined a social ridesharing problem and provided a solution with a similar motivation as Li et al. [2015]. However, they neither systematically looked into the character-
istics of the geographic distribution of the social network, nor studied the dynamic evolution of the system.

- **Cici et al. [2014]** conducted a few empirical studies discussing the pros and cons of ridesharing with social network, which is a particular example of this present paper that systematically investigates the influence of a spectrum of social networks on ridesharing outcomes. More complicated but realistic factors, such as congestions, can be considered in ridesharing algorithms as well (e.g., Wang et al. [2015a]). They are, nevertheless, beyond the scope of studying the influence of social networks. Additionally, in terms of flexibility, none of the aforementioned algorithms allows the absence of unique destinations. This limitation (in choice) restricts the space-time budgets of individuals and thus the matching rates of ridesharing algorithms. But if the trips are treated as activity-based, the destination choice set can be expanded effectively.

The dynamic route-planning and optimisation is frequently brought up as a challenge. Berbeglia et al. [2010] did a review of dynamic pickup and delivery problems, bringing up some interesting but unsolved questions, such as optimal waiting strategies, modifications of the objective function on a rolling basis, to name a few. Pillac et al. [2013] surveyed the broader class of dynamic vehicle routing problem. The present thesis, however, builds only static models to testify the usefulness of the new factors introduced into ridesharing systems. Dynamic ridesharing is beyond the scope of this thesis, but falls within the consideration of future work.
This part discusses previous theories of social networks and time geography as the foundations of my research. On top of that, the proposed framework by this thesis, called collaborative activity space, is elaborated as the general theory for the new ridesharing prototypes. The experiment materials and tools used by each prototype are explained as well.
3

THEORETICAL FOUNDATIONS:
INTEGRATING HETEROGENEITY OF
DECISION-MAKING INTO TIME GEOGRAPHY

The proposed theory of Collaborative Activity-based Space is based on the theory of time geography [Hägerstrand, 1970], but is a computational model for Raubal et al. [2004]'s user-centred time geography that infuses personal heterogeneity into time geography. As a solution to trust and flexibility issues in ridesharing, the two factors – social network and platial semantics – are internalised into the framework. This chapter first introduces the existing theories in time geography (Section 3.1) and social network analysis (Section 3.2), and then elaborates the Collaborative Activity-based Space framework (Section 3.3).

3.1 TIME GEOGRAPHY AND USER-CENTRED TIME GEOGRAPHY

Time geography is a set of theories on modelling accessible areas given space-time constraints [Hägerstrand, 1970]. This theory is useful to search for potential locations for a flexible activity in the sense that the activity is not bounded to a specific pair of geographic coordinates. Take the case in Song and Miller [2015] for example, a work-grocery-home journey is a trip chain with one flexible activity – grocery shopping. Time geography can be adopted to decide where to do the grocery shopping.

Some key concepts in time geography are listed below (Figure 1):

- **Space-time path**: This is one possible space-time trajectory linking the activities done by an individual.
- **Space-time prism (STP)**: This delineates the maximum accessible range and resources subject to spatial and temporal constraints. It encompasses all the feasible space-time paths.
- **Potential path area (PPA)**: This is the projection of an STP to the 2-dimensional planar space, covering the maximum accessible area during the entire time slot given the space-time budget.

The geometric properties of STPs and PPAs as well as their computational models in an Euclidean space are elaborated in Miller [2005]. Further extensions have been studied, including network-based time geography and probabilistic time geography [Song and Miller, 2015; Winter and Yin, 2010]. At every instant of time, there is a corresponding space projection (as the intersection of the STP and the planar surface at that time tick cutting through the STP). The union of the projections in space throughout the given time period is the PPA. The shape of an STP is dependent on the flexibilities of the origin and the
destination (see a further discussion in Section 3.3). Geometrically speaking, an STP is easy to build if there is only one flexible activity between two fixed (at-known-location) activities. This is, however, not always the case in reality. Chen and Kwan [2012] tried to address the multiple flexible activity issue in destination choice set. This inspires the design of the activity-based ridesharing algorithm in Chapter 7.

Raubal et al. [2004] proposed the user-centred time geography, where three types of affordances are introduced to model the heterogeneous space-time accessibility with regard to each individual. Physical affordance is a combined result of the invariant compounds and the physical constraints of an agent. For example, the distance between the current location and the destination given the accessible transport mode of a person. Social-institutional affordance is the perceived physical affordance given social-institutional constraints. For instance, a store closes at 9pm, so no one can visit the store when it is closed. Mental affordance is the subjective perception of the accessibility to a location given physical and social-institutional affordances. This is a comprehensive decision-making in spatial behaviours. Section 3.3 discusses based on the three affordances.

3.2 SOCIAL NETWORK ANALYSIS

Andris [2016] pointed out the separation between social network analysis and geographic studies, and suggested removal of the bipartition. Social network, on one hand, is an externally associated factor to travel behaviour, which is addressed by social network analysis; on the other, it can be internalised into time geography framework by modelling its distortion of STP, which will be elaborated by Section
3.3. This section will explain the foundations of the external association – social network analysis.

It is difficult to approximate a realistic or claim a representative social network structure. However, the small world model [Watts and Strogatz, 1998] is built for simulating social networks in the real world that satisfies small-world phenomena that, to put it simple, anyone in the world can be connected to any other person via only a few other persons [Milgram, 1967]. Small world networks have a higher clustering coefficient than a random network, with more triangles and still a few short-cuts between nodes. They do not have heavy-tail degree distribution as preferential attachment networks [Barabási and Albert, 1999]. Small world networks are thought better to approximate real-life social networks than preferential attachment because of the limited number of friends a person can hold in reality [Dunbar, 1992].

Particular to ridesharing, Chaube et al. [2010] suggested three levels (intensities) of social relationships: direct (1st-degree) friends, indirect (2nd-degree) friends, and strangers. Detour tolerance and willingness to share a ride varies according to different levels of friendship.

The spatial distribution of friendship is another factor of interest. It can be estimated using geo-located datasets, such as cell phone records, social network check-ins, to name a few [Cho et al., 2011; Liben-Nowell and Novak, 2005; Wang et al., 2015b]. However, the distributions from different datasets are highly dependent on the specific context, except sharing the common trait of spatial proximity. Liben-Nowell and Novak [2005]’s simulation demonstrated that one third of the friendships in the context of LiveJournal are independent of geography while the rest are spatially aggregated. Onnela et al. [2011] concluded that only small social groups are geographically very tight, while larger groups split into clustered smaller ones over space. Gao et al. [2013] detected spatially clustered community patterns in cell phone call networks. Shi et al. [2016], also with cell phone records, substantiated the existence of a distance decay in social networks. Wong et al. [2006] demonstrated that generating social network links based on a distance decay rule can capture the characteristics of small world networks. Ratti et al. [2010] detected clustered social communities within the landlines in the Great Britain.

Without losing generality, this thesis will initialise the social network as a small-world network, and will assign the spatial location of each node by random or by spatial proximity.

3.3 COLLABORATIVE ACTIVITY SPACE AND NEW RIDESHARING PROTOTYPES

The missing part in computational space-time path-prisms analysis [Miller, 2005; Song and Miller, 2015] is the human reasoning. As a
computational specification of Raubal et al. [2004]’s conceptual model of user-centred time geography, the Collaborative Activity Space (CAS) aims to fill in this gap in time geography with regard to collaborative travel.

3.3.1 The elements in CAS

The three essential elements in CAS (Fig. 2) are the spatio-temporal budget (ST), the social network (SN), and the platial semantic (PS). ST is the core of time geography, shaping the physical affordance. SN contributes to the mental affordance. The fact that a person subjectively chooses to enlarge the detour tolerance for a friend than a stranger means the mental affordance for a friend is higher. Social network is rarely internalised into spatial behaviours in this sense, though some studies (e.g., Wang et al. [2015b] and Andris [2016]) have analysed the influence of social network as an external factor on spatial behaviours. PS contributes to social-institutional and mental affordances. Travel is the secondary demand driven by the destination activity, which indicates that:

1. The activity for a person has a certain space-time flexibility (see Chapter 7). Some activities have to be conducted at a certain time and a certain location, while others are flexible location-wise or time-wise. The flexibility varies from person to person as well. This factor belongs to the social-institutional affordance but indirectly affects the physical affordance by changing the spatio-temporal budget.

2. The function/service provided by the destination defines the social-institutional affordance. A person aiming to buy some apples will not go to a hospital.

3. Personal preferences of which particular location to go contributes to the mental affordance. Some people, for example, only go to a certain café for breakfast.

The proposed ridesharing algorithms are rooted in this framework. Considering only ST and SN, social network based ridesharing is proposed (Chapters 5, 6). Activity-based ridesharing coalesces only PS and ST (Chapter 7). Incorporating all the three factors together leads to the collaborative activity-based ridesharing (Chapter 8).

3.3.2 Basic unit: the adjusted PPA

The basic unit in CAS is the adjusted PPA. Figure 3 shows the conceptual PPA with a pair of fixed origin (point O_i) and destination (point D_i). In classic time geography, the PPA is represented by the dash-dot ellipse only subject to physical affordance decided by ST. Given the influence of SN, two more ellipses – the bold and thin solid ones – are drawn that reflect the selectively shrunk affordances (i.e., detour
3.3 Collaborative Activity Space and New Ridesharing Prototypes

Figure 2: Three essential elements in Collaborative Activity Space.

The three major elements are: ST – spatio-temporal budget; PS – platial semantics; SN – social network.

tolerance) for 2\textsuperscript{nd}-degree friends and strangers (i.e., whoever beyond two degrees of separation). The adjusted PPA thus introduces the heterogeneity from social network into time geography framework when looking for accessible locations. Checking POI-1 to POI-4 in Figure 3 (POI stands for the Point of Interest, which represents a place/resource for a travel purpose), the varied detour tolerance has an obvious influence. While the four POIs are all feasible if travelling with 1\textsuperscript{st}-degree friends, POI-4 is out of reach for 2\textsuperscript{nd}-degree friends, and both POI-4 and POI-3 are out of choice detouring for strangers. The mechanism is further discussed in Chapters 5 and 6.

When the trip chain gets longer with multiple flexible activities between O\textsubscript{i} and D\textsubscript{i}, a simplified strategy is proposed in Chapter 7 that controls the total time budget between O\textsubscript{i} and D\textsubscript{i}, including detour tolerance. Each trip on that chain will correspondingly get a portion of the tolerance in proportion to its shortest travel time. The details are elaborated in Chapter 7.

3.3.3 The flexibility of activities: Genres of places

For a travel, the aim or activity is usually fixed while destinations can be diverse. Travel aim partially decides travel flexibility in both space and time. For instance, going to a party allows a late arrival but
going to an interview does not; dining out with a friend has multiple destinations to choose but going to school does not.

Based on the space-time flexibility of an activity happening at a place, four genres of activity-related places (AxP) are suggested, where A stands for an activity and P for a place:

1. Genre 1 (AxP1) has hard-time and hard-location (HTHL), such as going to schools or offices;
2. Genre 2 (AxP2) is characterised by hard-time but flexible-location (HTFL), e.g., meeting a friend at a coffee bar before the friend leaves to catch a train;
3. Genre 3 (AxP3) is marked by flexible-time but hard-location (FTHL), for example, visiting a certain museum but at any convenient time;
4. Genre 4 (AxP4) has both flexible-time and flexible-location (FTFL), say, going to the supermarket.

If an activity has the trait of a genre, the associated place is said to belong to that genre. Some places can belong to multiple genres depending on different situations of the activity and its participants. In a pragmatic case, Genre 4 is less likely to happen independently; people at least have some hard schedules at certain time points, so Genre 4 activities can be added to time slots between the relatively fixed events or based on people’s habit to do something at a specific time of day.
3.3.4 Other factors of platial semantics

In addition to the function and flexibility of a place, mental affordance is also affected by personal preference to a location. Some people only go to certain restaurants for dining, and some just go to that grocery store for fruits. The rigidity of choices is generally varied from person to person, which brings challenges to modelling. In practice, a possible way is to adopt rating system, e.g., Yelp and Zumato, to assign weights to locations. Weights are not yet modelled in the current work.

3.3.5 The envision of a dynamic CAS: a conceptual model

Potentially, the adjusted PPA can be applied to dynamic decision-making and travel behaviours. In the context of multiple flexible activity trip chain, the dynamic mechanism is to make decisions step by step (for each activity), which means the PPA is drawn for only a short period of time $\delta_t$. As shown in Figure 3, the perfect circle with the radius $\delta_t$ is a PPA$_t$, a short-term PPA dependent on time instant $t$ when a person makes a decision at point $O_i$. The potential locations for next step are points A and B in this case.

The accessibility to any point within the STP is dependent on the decision made at the previous step. When an agent makes a decision, for example, from point $O_i$ to point B, the remainder of the spatio-temporal budget changes accordingly. The affordances to other locations in the CAS also change. To guarantee that no event will be missed, the choice of next step should be within reach to the mostly likely destination at the step after next. As pointed out by Winter and Yin [2010], the probability to different locations within the STP is varied. Hence, the mostly likely visited points – even though the activities are by nature flexible – are leveraged as control points.

For the instance shown in Figure 3, this person is mostly likely to go via point POI-1 then finally to $D_i$. The task now is to decide a location for the first activity, A or B. In Figure 3, the double-dash-dot ellipse encompassing $O_i$, B, and POI-1 functions for this sake. B can be chosen if and only if POI-1 can be reached within the given time budget. A is left out because it is out of the ellipse. Note that POI-1 is not necessarily the location to be visited if there are better alternative choices. If POI-1 is substituted by a closer location, there can be more time to reach an alternative location to point B, say A. However, the safe guarantee mechanism has to assume POI-1 is to be visited to control the unforeseen risks.

There can be impromptu decisions popping up on the way. The dash-line ellipse anchored at B and POI-1, for example, allows the agent to change mind on the fly. At the point $M$ where the agent changes mind, a PPA$_t$ is created to filter feasible resources. The circle
centred at \( M \) with radius \( \delta_0 \) is for fast retrieval of resources, but the final selection should be subject to the constraints of the following stops as discussed above.

In general, the dynamic CAS is structured as a stack of PPA\(_t\)s. A PPA\(_t\) is generated at all decision points, i.e., most likely locations for the activities on the chain as well as the points where an agent changes mind. This process keeps going for each planned activity on the trip chain until the final destination \( D_1 \) is reached. Specific implementations might utilise different strategies.

The dynamic CAS is not modelled in this thesis since the proposed ridesharing models are static pre-planning for the whole day. The method proposed in Chapter 7 introduces a method to avoid the dependency in the step-wise decision-making. A dynamic ridesharing model is under consideration in the future work.
The proposed ridesharing models are all implemented with agent-based simulations and tested with data derived from a questionnaire-based travel survey dataset. This chapter explains the details of the shared computational methodology and dataset.

4.1 Agent-Based Simulation

As to implement the conceptual models, agent-based simulations with NetLogo and Java-based Repast Simphony are conducted. Agent-based simulations are selected particularly for the capability of modelling individual preferences and needs in travel behaviours in addition to trip demands [Ronald et al., 2015].

NetLogo is a multi-agent modelling environment based on a grid-lock context. It is easy and quick to start with and generally handy with simple conceptual/artificial models (e.g., Armendáriz et al. [2011]). Chapter 5 utilises NetLogo for the conceptual model in an absolute artificial environment.

When implementing the models with realistic dataset, NetLogo is not convenient as Repast Simphony in terms of its programming flexibility and computational speed. Therefore, Repast Simphony is employed to implement the empirical model in Chapter 5 and the models in Chapters 6, 7, and 8, all of which are based on realistic road network and travel survey dataset.

4.2 Study Area and Dataset

The models are simulated in the shire of Yarra Ranges, including several suburbs located east and northeast to Melbourne, Victoria. The travel demand is derived from Victorian Integrated Survey of Travel and Activity (Section 4.2.1). The population is synthesised based on this travel survey [Jain et al., 2017]. Social network is artificially generated following given mathematical rules for this study. Details are given in Section 4.2.2. One last dataset used is the POIs representing alternative destinations to model platial semantics. The dataset is introduced in Section 4.2.3.

The overall total travel time by car is reported to have an ideal value of about 15 minutes in Berkeley, CA, and the tolerable upper limits are about 36 minutes for car while 44 minutes for carpool [Milakis et

1 https://ccl.northwestern.edu/netlogo/
2 https://repast.github.io/repast_simphony.html
Considering the area of Berkeley is much smaller than the study area Melbourne, Australia, the upper limit can be higher than in Berkeley.

### 4.2.1 Victorian Integrated Survey of Travel and Activity

The Victorian Integrated Survey of Travel and Activity (VISTA) dataset [Victorian Department of Transport, 2011] is a government conducted questionnaire-based travel survey over the whole State of Victoria. The survey is a one-weekday survey of each person, but might be done on different days of different persons. The spatial unit of VISTA is the finest spatial level (SA1) of the Australian Bureau of Statistics. Trips in VISTA are recorded as origins and destinations at this zonal level for the sake of confidentiality. When implementing the models, the specific geographic coordinates are generated randomly in the corresponding zones. Travel duration is adjusted accordingly as well. Table 1 shows the data structure of VISTA data with only fields relevant to this thesis.

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Meaning</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORIGSA1</td>
<td>Trip origin zone</td>
<td>2128101</td>
</tr>
<tr>
<td>DESTSA1</td>
<td>Trip destination zone</td>
<td>2128101</td>
</tr>
<tr>
<td>PERSID</td>
<td>ID of the trip owner</td>
<td>Y09H061326P02</td>
</tr>
<tr>
<td>TRIPID</td>
<td>ID of this trip</td>
<td>Y09H061326P02-T01</td>
</tr>
<tr>
<td>START-TIME</td>
<td>Start time of trip (min)</td>
<td>512</td>
</tr>
<tr>
<td>ARR-TIME</td>
<td>End time of trip (min)</td>
<td>673</td>
</tr>
<tr>
<td>ORIG-PLACE2</td>
<td>Origin activity code</td>
<td>201</td>
</tr>
<tr>
<td>DEST-PLACE2</td>
<td>Destination activity code</td>
<td>301 (-2 is last trip of the day)</td>
</tr>
<tr>
<td>DURATION</td>
<td>Duration of the destination activity (mins)</td>
<td>131</td>
</tr>
</tbody>
</table>

Each entry in the table is a recorded trip of a surveyed person. It documents the starting and ending location, the corresponding starting and ending time of that trip, and the activities conducted or to be conducted at the origin and destination. It also records the activity duration at the destination. If a surveyed person has multiple trips
on that day, there will be multiple entries associated to that person’s ID.

VISTA surveys a representative sample of 1% of the households in Victoria. Despite the sample size being large enough, models in this thesis run with a synthetic population generated from further census data on the socio-economical information of each household [Jain et al., 2017], which adjusts the survey sample proportional to the population. Each synthetic person has a travel agenda for a day.

4.2.2 Social network permutation

The social network is artificially generated with the Small World module in Repast Simphony. The module, called WattsBetaSmallWorldGenerator, allows to set three parameters:

- **beta**: the probability to rewire a link randomly (the higher the value, the more shortcuts);
- **degree**: the average degree of connection (the higher the value, the more friends each agent has on average);
- **symmetrical**: boolean value to decide whether the generated edges will be symmetrical.

Particular parameter settings are given in each chapter.

The distribution of origin and destination points (O/D) of ridesharing partners crucially determines travel and detour cost. It has been found that people who have closer origins are more easily to share a ride than having similar destinations [Vanoutrive et al., 2012]. For this reason, only origin distribution of friendships is controlled.

The spatial distribution of friendships also matters when taking social relationship into account in ridesharing. As a specific implementation of the spatial aggregation of friendships discussed in Section 3.2, the models apply the distance decay law, that is, the geographical distance between a pair of friends over the sampled population generally follows a (exponentially truncated) power law distribution. The higher the distance decay coefficient, the more spatially aggregated the friendships. The chance of matching ridesharing vary with the spatial distribution of friendships.

4.2.3 Points of Interest from Yelp API

POIs in this thesis are retrieved to represent the corresponding locations that serve each type of travel activity. The POIs are extracted from Yelp Search API. Each search is indexed by the type of activity conducted at each trip’s destination, as per ORIGPLACE2 and DESTPLACE2 in Table 1. The weight of each POI is not modelled, but as discussed in Section 3.3.4, the heterogeneity of personal preference can be modelled by the rating on each location.

3 https://www.yelp.com/developers/documentation/v2/search_api
Part III

PROPOSED RIDESHARING MODELS

This part encapsulates four chapters elaborating the proposed ridesharing models from my PhD research. The four chapters are revised from four published or submitted publication manuscripts. Chapter 5 introduces social network based ridesharing, a model that considers heterogeneous ridesharing detour tolerance and willingness due to different social acquaintances. Chapter 6 is an extension of the work in Chapter 5, simulating the dynamic evolution of the ridesharing collaborative network affected by ridesharing histories. Chapter 7 proposes activity-based ridesharing, a model that integrates alternative destination locations for a similar travel purpose to encourage uptake of ridesharing. Chapter 8 develops the collaborative activity-based ridesharing based on the previous two models, leveraging alternative destinations to further enhance both general uptake rate and the special uptake rate between social contacts. Chapter 9 suggests a parking model that, on one hand makes use of ridesharing to reduce parking demand, and on the other restrains the length of the shared travel as short as possible for less inconvenience.
This chapter is based on the publication of the same title published in *Computers, Environment and Urban Systems* [Wang et al., 2017]. As the first author, I proposed the research idea, came up with the hypothesis, designed and implemented the experiment, and conducted the data analysis, discussions, and conclusions. My supervisor Prof Stephan Winter was actively involved in supervising me and discussing with me the research idea and design. My co-supervisor Dr Nicole Ronald was lending help in research idea brainstorm, dataset arrangement, network analysis, and technical support for the experiment implementation.

5.1 Major Contributions

Ridesharing with social contacts (i.e., “friends”) is substantially more accepted than with strangers. However, limiting ridesharing to friends while rejecting strangers also reduces ride choices and increases detour costs. This work studies, from a theoretical perspective, whether the additional detour costs of limiting shared rides to social network contacts would be prohibitive. It proposes a social network based ridesharing algorithm with heterogeneous detour tolerances for varied social contacts. The theoretical matching rates and detour costs are compared in a simulation for three levels of social connectivity: traveling with direct contacts only, with direct and indirect contacts, or with anyone. The simulation allows for a systematic and comprehensive testing of system behaviour when varying the parameters of social network structure, detour tolerance, and spatial distribution of friendship. Results show that for a clustered friendship – the expected spatial distribution of a social network growing with a ridesharing network – ridesharing with friends does not cause significantly higher costs. Furthermore, the algorithm prioritising friends can substantially increase the matching of friends. An empirical study justifies the findings.

This work is significant in at least two ways. First, none of the studies mentioned in Chapter 2 considered heterogeneous detour tolerances with different social contacts, though they introduced grouping strategies. Second, the simulation does a systematic check on the influence of spatial distribution of social network on ridesharing results.
To prove the benefit of ridesharing with friends, two null hypotheses against the objective are to be rejected. The first one is that sharing rides only with friends significantly increases detour cost. The second is that the matching rate is significantly lower with friends than with anyone. These hypotheses are not obvious since detour cost and matching rate are influenced by social similarity and spatial distribution of friends.

The implementation will use an agent-based transport simulation in order to be able to comprehensively vary all relevant parameters and get a theoretical insight into the complex system behaviours. The systematic scenario simulation is necessary for more transferable findings, because social networks and their spatial distributions vary widely from place to place. The theoretical simulation is set up using NetLogo, followed by a validation simulation based on realistic dataset with Repast Simphony for higher computation efficiency. The model systematically tests different parameter settings; the parameters include the social network structure (average degree of friendship), the spatial distribution of friends (spatially clustered vs. random), and the detour tolerance for different social connectivity levels. Three matching patterns are tested: 1) any driver and passenger must be direct friends, 2) any driver and passenger must be either direct or indirect friends, and 3) no one has to be friends. Two populations of 2,000 and 5,000 agents, respectively, with small-world social network structures are simulated. Core of the simulation is a proposed algorithm for social network based ridesharing (SNeRs). The detour costs and matching patterns will be collected and subjected to statistical analysis.

This chapter is organised with the following sections. Section 5.2 depicts the conceptual framework and Section 5.3 gives details of both the theoretical and empirical implementations, followed by Section 5.4 of results and Section 5.5 of discussions. The major conclusions and future work are given in Section 5.6.

5.2 THE ALGORITHM OF RIDESHARING CONSIDERING FRIENDSHIP

This section explains social network based ridesharing by going through the details of the proposed algorithm SNeRs. This section concentrates only on the algorithm, while details of a simulation deploying the algorithm will be given in the next section. The algorithm concentrates on the mobility demand of a population that has a choice between travelling with their own vehicle and travelling with somebody else by ridesharing. All other modes of transport, or other parts of a real-world population are irrelevant for the two hypotheses.
5.2 The Algorithm of Ridesharing Considering Friendship

5.2.1 Conceptual experiment design and data structure

The mobility behaviour of a population of a given size can be simulated with a static social structure and pre-specified travel plans. In this sense the algorithm studies the total (theoretical) potential of ridesharing in a population that has fully adopted ridesharing, in line with the hypotheses.

The algorithm computes the average detour cost of matched rides and the amount of matchings of all the posted rides. Travel cost is calculated in time. With a population allocated on a road network, the algorithm is trip-based. Each trip is made by one traveller. Since each traveller (modelled as Person in Table 2) is assumed to have a car, the person, if not matched, can travel alone. Each trip (called Ride in Table 2), consists of pre-defined origin, destination, and time stamps. Intermediate stops are not modelled, but form separate trips. Relationships between persons consist of direct friends (Friend), who are defined here as first-degree contacts in an appropriate social network, indirect friends (Ind-fri), who are defined as second-degree contacts in the social network, and strangers, i.e., those who are neither direct nor indirect friends.

The shortest travel cost (sTTC) is the algorithm’s computed time of individual travel with no detour. Detour tolerance (DetTol) is the maximal detour time a person can compromise. Willingness of ridesharing (Will) is the probability to join ridesharing. Initially, the role of a person is not fixed, being driver or passenger. The variable Role is only fixed after initial matching. A car can have multiple passengers if everyone’s tolerance is satisfied.

5.2.2 Detour cost and ridesharing feasibility functions

For each ride, the start point is associated with an earliest starting time (EST) and latest starting time (LST), while the other stops are only controlled by a latest ending time (LET). LST is determined by the LET at the next stop, with a given time for travelling. LET–EST = sTTC + DetTol. Note that EST, LET, sTTC, and DetTol are not independent of each other: only sTTC is externally given by the physical condition, while the others are adjusted for integrity as such: 1) EST or LET is decided due to whether the origin or destination (O/D) activity is time-fixed; 2) if only one end (O/D) is fixed, DetTol applies; 3) if both ends are fixed, always apply LET, and EST is adjusted as \( EST' = LET - DetTol \) only if \( EST' > EST \). A person \( p \)'s detour cost on a route \( r \) (\( Det_p | r \)) is the difference between the person’s travel cost \( TTC_p \) and his/her shortest travel time \( sTTC_p \) (Eq. 1). \( TTC_p \) is the time duration between getting out of the car (\( T_{p\text{out}} \)) and into the car (\( T_{p\text{in}} \)). Thus, for a driver \( TTC_p \) is the whole ride time; for a passenger in a single-passenger ride the detour cost is always 0; and for passengers
Table 2: Structures/Classes in matching algorithm

<table>
<thead>
<tr>
<th>Structure or Class</th>
<th>Meaning</th>
<th>Major variables</th>
</tr>
</thead>
</table>
| Person             | A traveller | - Role: “dr” (driver), “psg” (passenger), “” (not assigned)  
|                    |          | - sTTC: shortest total travel cost (in mins, original route)  
|                    |          | - TTC: total travel time after merging route with others  
|                    |          | - DetTol: detour tolerance (in mins, calculated by multiplying detour percentage by sTTC)  
|                    |          | - Will: willingness to share rides  
|                    |          | - drList: a list of potential persons as driver along with the detour cost caused by each. The structure is [[pID, cost],... ]  
|                    |          | - psgList: a list of potential passengers, with structure same as drList  
|                    |          | - Finalcost: the final total detour cost for this person, in percentage of sTTC  
|                    |          | - Finalist: the list of persons finally matched with this person, along with the detour cost for this person caused by each in the list. It has the same structure as drList. |
| Ride               | A trip of a Person | - SX, SY, EX, EY: starting and ending locations  
|                    |          | - EST, LST: earliest and latest starting time  
|                    |          | - LET: latest ending time |
| Friend             | The link between two Persons with direct friendship | `<PersonID1, PersonID2>` |
| Ind-fri            | The link between two Persons with a common friend but no direct friendship | `<PersonID1, PersonID2>` |
in multi-passenger rides there can be a detour cost caused by other passengers.

$$\text{Det}_{p|r} = \text{TTC}_p - s\text{TTC}_p = T^\text{out}_p - T^\text{in}_p - s\text{TTC}_p$$  \hspace{1cm} (1)

The optimisation objective (Eq. 2) of the matching algorithm is to maximise the number of matched persons subject to space-time constraint and detour tolerance (Eq. 3). Let $M$ be the set of individuals who have been matched, $P$ is the population set, and $\oplus$ represents the merged route of persons $p$ and $p'$ (whoever is driver or passenger). The merged route is $dr \oplus \text{psg}_j$ with single passenger, and $\bigcup_{\forall \text{psg}_j} dr \oplus \text{psg}_j$ with multiple passengers. Persons are assumed selfish in decision-making, so a global detour optimum might be sacrificed in searching for individual’s best choice.

$$\max |M| = \max |p \in P| \exists p' \in P, \text{ s.t. } p \oplus p' \neq \emptyset$$  \hspace{1cm} (2)

such that, for any passenger $\text{psg}_j$ with the driver $dr$,

$$s\text{TTC}_{dr} + s\text{TTC}_{\text{psg}_j} \geq \text{TTC}_{dr \oplus \text{psg}_j}$$  \hspace{1cm} (3a)

$$\text{Det}_{dr | \bigcup_{\forall \text{psg}_j} dr \oplus \text{psg}_j} \leq \text{DetTol}_{dr}$$  \hspace{1cm} (3b)

$$\text{Det}_{\text{psg}_j | \bigcup_{\forall \text{psg}_j} dr \oplus \text{psg}_j} \leq \text{DetTol}_{\text{psg}_j}$$  \hspace{1cm} (3c)

$$\text{EST}_{dr} + \text{TTC}_{n(0), n(1)} < \text{LET}_{n(i)} \leq \text{LET}_{dr}$$  \hspace{1cm} (3d)

Equation 3a shows the feasibility of a matching that the collaborative travel cost on the merged route, $\text{TTC}_{dr \oplus \text{psg}_j}$, should be lower than the sum of individual travel time by each person. Equations 3b and 3c ensure the detour costs on the merged route for the driver ($\text{Det}_{dr | \bigcup_{\forall \text{psg}_j} dr \oplus \text{psg}_j}$) or each passenger ($\text{Det}_{\text{psg}_j | \bigcup_{\forall \text{psg}_j} dr \oplus \text{psg}_j}$) satisfy their corresponding detour tolerances ($\text{DetTol}_{dr}$ or $\text{DetTol}_{\text{psg}_j}$) in the same car. Equation 3d is the time constraint of each node $n(i)$ along the merged route $\bigcup_{\forall \text{psg}_j} dr \oplus \text{psg}_j$ (Fig. 4). Assuming only drivers have an EST constraint, the earliest starting time of the driver ($\text{EST}_{dr}$) plus the travel time from the start point $n(0)$ to any node $n(i)$ on the merged route should be earlier than the required latest arrival time to $n(i)$ ($\text{LET}_{n(i)}$), and must be earlier than or equal to the driver’s ending time $\text{LET}_{dr}$ (ending point is $n(e)$). The nodes along the route are passengers’ start and end points, or the driver’s stop-by points, ranked by time sequence.
5.2.3 *The Social Network based Ridesharing algorithm (SNeRs)*

The matching process has two steps: finding candidates (Algorithm 1) and making decisions (Algorithm 2). Figure 5 is the overall flowchart: The left box shows Algorithm 1, and the right box Algorithm 2.

**Algorithm 1** SNeRs: Finding candidates

1. \text{DetTol\_dr;} \quad \triangleright \text{Detour tolerance of a driver}
2. \text{DetTol\_psg;} \quad \triangleright \text{Detour tolerance of a passenger}
3. 
4. // Exhaustively matching each pair of persons
5. \textbf{for} each Person in Persons as Dr \textbf{do} \quad \triangleright \text{A driver}
6. \quad \textbf{for} each other Person in Persons as Psg \textbf{do} \quad \triangleright \text{A passenger}
7. \quad \textbf{Call} Dr::WorthDetour(); \quad \triangleright \text{Check whether worth detour}
8. \quad \textbf{Set} \text{DetTol\_dr;}
9. \quad \textbf{Set} \text{DetTol\_psg;} \quad \triangleright \text{Setting is based on social network type and shortest path length}
10. \quad \textbf{Calculate} detour costs for Dr and Psg respectively as Dr::Cost, Psg::Cost;
11. \quad \textbf{if} Dr::Cost < \text{DetTol\_dr and}
12. \quad \textbf{Psg::Cost < DetTol\_psg then} \quad \triangleright \text{Each person’s detour cost for the other is within their tolerances}
13. \quad \textbf{Add} Dr along with detour cost caused by Dr to Psg::drList; \quad \triangleright \text{Psg’s candidate list}
14. \quad \textbf{Add} Psg along with detour cost caused by Psg to Dr::psgList; \quad \triangleright \text{Dr’s candidate list}
15. \quad \textbf{end if}
16. \quad \textbf{end for}
17. \textbf{end for}

Algorithm 1 loops the population as driver, matching them with all other people as passenger (line 5-6 in pseudocode). Both driver and passenger candidates are recorded, but the role is not decided until Algorithm 2. Social network and related procedures are highlighted in grey background in Figure 5. Willingness is the probability
5.2 The Algorithm of Ridesharing Considering Friendship

Figure 5: Flowchart of the Social Network based Ridesharing algorithm (SNeRs).
Algorithm 2 SNeRs: Making decisions

1: **Declare** Stack;  \(\triangleright\) Data structure to check rides one by one
2: 
3: **Rank** Rides by time order;
4: **for** each Ride in Rides **do**  \(\triangleright\) Loop by time priority
5:  **Add** Ride to Stack;
6:  **while** Stack is not empty **do**
7:     **Pop** Ride from Stack;  \(\triangleright\) LIFO; Last one in Stack
8:  **Let** CurrP1 ← Ride::Person;  \(\triangleright\) Owner of the current ride
9: 
10:     **if** CurrP1 wants to be a driver **then**
11:         **Clean** CurrP1::psgList;  \(\triangleright\) Remove matched or failed
12:         **if** first time to take a passenger **then**  \(\triangleright\) 1st time, try to find best match
13:             **Find** best cost passenger of CurrP1 in CurrP1::psgList;
14:             **Match**;
15:             **Add** CurrP1 to Stack if car is not full;
16:         **else**  \(\triangleright\) Non-1st time, find one as long as can match
17:             **for** each person CurrP2 in CurrP1::psgList **do**
18:                 **Check** with all existing passengers;
19:                 **if** everyone in car is fine **then**
20:                     **Match**;
21:                 **else**
22:                     **Fail**;
23:                 **end if**
24:             **end for**
25:         **end if**
26:     **end if**
27: 
28:     **if** CurrP1 wants to be a passenger **then**
29:         **Let** CurrP2 ← first in CurrP1::drList  \(\triangleright\) CurrP1 wants CurrP2 to be the driver
30:         **if** CurrP2 wants to be a driver **then**
31:             **if** the first time CurrP2 takes passenger **then**
32:                 **if** CurrP1 is CurrP2’s best match **then**
33:                     **Match**;
34:                 **Add** the driver to Stack if the car is not full;
35:             **else**
36:                 **Fail**;
37:             **Add** CurrP1 to Stack;  \(\triangleright\) Check next
38:         **end if**
of joining ridesharing that varies with types of friends, and is set as a threshold on a random generator. Only if over the threshold will the matching be conducted (line 7). Detour tolerance is represented as the percentage of \( s_{TTC} \) a person would sacrifice for each type of friend, but is calculated in minutes by multiplying \( s_{TTC} \). The function \( \text{WorthDetour}() \) checks (line 14) matching feasibility (Eq. 3a) and physical accessibility (Eq. 3d) affected by socio-psychological constraints of detour tolerance (Eqs. 3b and 3c) and willingness.

Algorithm 2 decides the role of a person and with whom, according to the immediate best detour cost, they will travel. The greedy strategy is only one of many possible decision-making strategies. Alternative strategies such as an auction-based method [Asghari et al., 2016] can be applied. The major presumptions and suggested strategies are listed below:

1. The list of rides is ranked by ascending LST to match more urgent travellers first. In this order, each traveller negotiates with each possible candidate until matched or candidate list runs empty (then going alone). People who make decision earlier affect the choice set of others later. The candidate list and the role of people are updated dynamically.

2. People are egoistic to achieve their current optimal result instead of global optima. The decision on a person’s role is determined by the least detour cost with current available candidates. The decision matrix is shown in Table 3. Yellow colour in this table and in flowchart is the decision on the role of the person who initiated the negotiation, while the light red is for the person being asked. Match is achieved immediately only if driver and passenger are mutually each other’s best option. Otherwise,
Table 3: The decision making matrix of persons asking and being asked for a ride

<table>
<thead>
<tr>
<th>Chosen role of CurrP1 (asking)</th>
<th>Passenger</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Driver</strong></td>
<td></td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; time</td>
<td>Non-1&lt;sup&gt;st&lt;/sup&gt; time</td>
</tr>
<tr>
<td>Cannot happen</td>
<td>Algorithm 3a</td>
</tr>
<tr>
<td><strong>Passenger</strong></td>
<td></td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; time</td>
<td>Non-1&lt;sup&gt;st&lt;/sup&gt; time</td>
</tr>
<tr>
<td>Algorithm 3b</td>
<td>Algorithm 3c</td>
</tr>
<tr>
<td>Cannot happen</td>
<td>Algorithm 3d</td>
</tr>
</tbody>
</table>

the decision strategy (see 3) depends on matching status and who initiates the matching.

3. Drivers always try to get the least-cost passenger as the initial passenger. A driver who has not taken a passenger will reject the request from whoever is not the current best passenger, even though the driver may still lose the chance to match with the current best in later rounds. But if the driver has taken a passenger, the driver accepts any request as long as all the people in the car satisfy their detour tolerance. When a driver initiates a negotiation, the invited passenger accepts the offer as long as no detour cost is caused; otherwise, the person travels alone.

4. Avoid mutual request. It is possible that person A’s best driver is B, while B’s best driver is A. The one with less detour cost as driver will become the driver.

---

**Algorithm 3a** Passenger asking driver, 1<sup>st</sup> time

1. if CurrP1 is CurrP2’s best choice then
2. \hspace{1em} Match;
3. \hspace{1em} Check driver’s ride if having empty seats;
4. else
5. \hspace{1em} Fail;
6. \hspace{1em} Check CurrP2 again;
7. end if

**Algorithm 3b** Driver asking passenger, 1<sup>st</sup> time
5.3 Simulation Implementation

This section gives the implementation details of both the theoretical and empirical simulations using the algorithm from Section 5.2. The theoretical simulation is set in an artificial context with generic characteristics so that the results are context-free and the fundamental regularities are transferable. An empirical study using realistic travel demand data is done to validate the findings.

5.3.1 Theoretical implementation

The population is located onto an artificial road network of 100 by 100 grid lines with equal interval where travel speed is assumedly even and fixed. The number of friends per person and the spatial distribution of friends are two of the tested parameters. The simulation is implemented in NetLogo.

Parameter settings

The Watts-Strogatz Small World model in Gephi 0.8.2 (beta = 0.01) generates a group of travellers with static social connections of small-
Table 4: Statistics of generated social networks

<table>
<thead>
<tr>
<th>Population size</th>
<th>Avg. degree of friends</th>
<th>Avg. clustering coeff.</th>
<th>Avg. path length</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.704</td>
<td>4.448</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.712</td>
<td>3.415</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.721</td>
<td>2.763</td>
<td></td>
</tr>
</tbody>
</table>

world structure that arguably captures the characteristics of social networks reflecting the constraints imposed by the physical world. The size of population (#agents) is set to 2,000 with an average degree of direct friends (#dir_fri) equal to 30, 50, and 100. Table 4 shows the statistics of the generated social networks.

Social networks have been observed to follow distance decay [Liben-Nowell and Novak, 2005; Onnela et al., 2011]. Distance decay depicts how a variable decreases with the increase of geographical distance. Allowing for preference to friends and constraints on travel time, people with more nearby friends have a higher probability to get a satisfying ride. Therefore, the spatial distribution of friendships is assumed to affect ridesharing results. To test this, two types of spatial configurations of friendship are examined: randomly distributed and spatially clustered. Since Vanoutrive et al. [2012] concluded that having closer starting points plays a more significant role in successful ridesharing, this simulation only controls origins when locating the population P in space. This can be improved with more realistic strategies, e.g., controlling start and end points to follow certain distributions.

Despite not specific, distance decay is a general rule constraining social network in space. The decay coefficient is set as 1.5 and 2, showing the intensified trend of spatial aggregation within the range supported by literature [Liben-Nowell and Novak, 2005; Onnela et al., 2011]. To be precise, the distance (Dist(i, j)) between two agents i and j who are direct friends follows an inversion sampling function using an inverse power law, IPL (Eq. 4). This function generates distances following a power law distribution, with coefficient β taking values in the range between 1 and 2. In Equation 4, Y follows a uniform distribution, α is the generalisation parameter, c is a constant, and R is the minimum distance between friends (set to one unit of the grid, which is the finest resolution in this study). To accommodate randomness from trip generation, each combination of parameters is run 5 times. Figure 6 displays the random or clustered origin distributions (each
from one random sample) of the 2000 agents with an average of 30, 50, or 100 friends.

\[
\text{Dist}(i,j) = \text{IPL}(Y \mid Y \sim U(0, 1)) = \begin{cases} 
1 & \beta \neq 1 \\
\sqrt{\frac{\alpha Y - c + R}{c \cdot R \cdot e^{\frac{Y}{\alpha}}}} & \beta = 1
\end{cases}
\]

(4)

To substantiate the expected advantages of SNeRs giving higher priority to friends, the detour tolerance and ridesharing willingness parameters are set in different ways. To reflect the social network’s influence on ridesharing, the detour tolerance is assigned to 30% for travelling with direct friends, 25% for indirect friends, and 7% for travelling with strangers; the willingness to offer or accept a ride is set to 100% for travelling with direct friends, 80% for indirect friends, and 10% for travelling with strangers. Accordingly, to be agnostic of
Table 5: Setting of input parameters in ridesharing

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Meaning</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>#dir_fri</td>
<td>Avg. # of friends per person</td>
<td>30, 50, 100</td>
</tr>
<tr>
<td>spa_distr</td>
<td>Spatial distribution of friendship</td>
<td>Randomly distributed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Clustered, decay coefficient $\beta = 1.5$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Clustered, decay coefficient $\beta = 2$</td>
</tr>
</tbody>
</table>

### Tolerance and willingness setting

<table>
<thead>
<tr>
<th>Detour tolerance</th>
<th>Scenarios</th>
</tr>
</thead>
<tbody>
<tr>
<td>detTol_1st</td>
<td>For 1st degree friends</td>
</tr>
<tr>
<td></td>
<td>30%</td>
</tr>
<tr>
<td>detTol_2nd</td>
<td>For 2nd degree friends</td>
</tr>
<tr>
<td></td>
<td>25%</td>
</tr>
<tr>
<td>detTol_strg</td>
<td>For strangers</td>
</tr>
<tr>
<td></td>
<td>7%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Willingness</th>
<th>Scenarios</th>
</tr>
</thead>
<tbody>
<tr>
<td>Will_1st</td>
<td>For 1st degree friends</td>
</tr>
<tr>
<td></td>
<td>100%</td>
</tr>
<tr>
<td>Will_2nd</td>
<td>For 2nd degree friends</td>
</tr>
<tr>
<td></td>
<td>80%</td>
</tr>
<tr>
<td>Will_strg</td>
<td>For strangers</td>
</tr>
<tr>
<td></td>
<td>10%</td>
</tr>
</tbody>
</table>

Social networks, the detour tolerance is uniformly 30%, and willingness 100%. The adopted values are based on survey results [Chaube et al., 2010; Wessels, 2009]. A summary of these values is shown in Table 5.

There are three matching patterns, with direct friends only (MP1), with social contacts (direct and indirect friends (MP2)), and with anyone (MP3). Each combination of input parameters listed in Table 5 is run 5 times, of which each run yields three total detour costs corresponding to three matching patterns. In total, 270 runs are performed.

### Trip generation

Each person in the population is assigned one trip with origin, destination, time point, and travel length. Trips are generated with similar starting time to examine travel behaviours at a certain instant. The origins are generated as elaborated above, i.e., correlated to the population. The lengths of trips follow a distance decay law, and the directions are almost random, indicated by previous studies that hu-
man trajectories on an aggregate level basically follow a Lévy flight [Brockmann et al., 2006; González et al., 2008].

The destinations are assigned so that trip length distribution satisfies distance decay of exponential distribution:

$$TL(i) = IEXP(Y|Y \sim U(0, 1); \lambda) = -\lambda \cdot \ln Y$$  \hspace{1cm} (5)

Equation 5 is an inverse exponential function (IEXP), where TL is the trip length of agent \(i\), \(\lambda\) is the mean value (set 15 units when a proper spatial distribution of agents within the given space is achieved), and \(Y\) follows a uniform distribution. Considering the utility for driving, the cut-off minimum trip length is 5 units, or five blocks in the street grid. Aligning this distance decay is empirically fitted using Euclidean distance. However, the simulated travel lengths are calculated in Manhattan distance. The average trip length in Manhattan distance is about 40 units. Given an origin and a trip length on a grid in a 2-dimensional plane, the destination is uniquely defined by the direction from its origin. It is assumed that agents near the boundary tend to travel towards the centre of the study area, while agents located near the centre travel in any direction. No trip is generated to or from outside of the grid.

5.3.2 Empirical study

The dataset has been elaborated in Section 4.2. These simulations have been realised in Repast Simphony for its higher computation efficiency compared to NetLogo. A population of 714 agents is permuted with small world social networks of an average degree of 6. Three groups of experiments (“Group” in Table 7) are run considering different social network types in matching (of MP3): 1) SW: small-world social networks with random spatial distributions; 2) SWD: small-worlds with the distance decays; and 3) NSW: no social network considered (Two sub-cases: 3a assumes the detour tolerance is homogeneously 10% and willingness is 10%; 3b assumes the tolerance is uniformly 30% and willingness is 100%). The first two groups are run 5 times each for different social permutations.

5.4 Results

5.4.1 From theoretical analyses

The two major outputs are the successful matching rate (Fig. 7a) and the average detour cost (Fig. 7b). Each row of the plot matrix represents a spatial configuration of friendship, ordered by an ascending trend of spatial cluster from top to bottom. Each column corresponds to a certain number of direct friends. The horizontal axis of each plot
(a) Detour cost (% of shortest individual travel time)

Spatial, beta=2.0  Spatial, beta=1.5  Random

30 friends 50 friends 100 friends
Figure 7: Simulation results with different scenarios (population size = 2,000; MP1=direct friends; MP2=direct and indirect friends; MP3=anyone)
is a nominal variable denoting the three matching patterns MP1, MP2, and MP3, and the vertical axis is the result value. On each plot, the solid line is the result by prioritising friends while the dash line with homogeneous tolerances and willingness. Each point is the average value of 5 runs. The differences between each corresponding pairs of results, i.e., points on dash-line versus on solid line, are significant on a 0.05 significance level except the following: Figure 7a, row 1, column 3, MP3; Figure 7b, row 1, column 2, MP3; Figure 7b, row 2, column 1 and column 2’s MP2s. The friend ratios in MP3 are displayed in Table 6 as 1) fRate, the percentage of the total population who share rides with direct friends, and 2) fRate2, the percentage of the matched population who share rides with direct friends. Theoretically a driver can take multiple persons, but it turns out that drivers only take one passenger for feasibility in the current model. Passengers, hence, have no detour cost in such case. The statistics calculates the detour cost per car, i.e., per driver in one-passenger ride.

Figure 7 and Table 6 contribute three major findings. First, detour cost with friends is not always significantly higher than with strangers. Using the detour cost with homogeneous tolerance and willingness (i.e., dash lines in Fig. 7a) as the benchmark, it holds true that rejecting rides with strangers (MP1 vs. MP3) is not beneficial with randomly distributed friendships, while it actually reduces detour cost with aggregated spatial distribution of friendships. The trend (of MP1) from top to bottom of each column in Figure 7a tells the decrease of cost with friends as friendship gets more clustered. The method of SNeRs that assigns heterogeneous detour tolerance and willingness, on contrast, is seen to reduce the overall detour cost at most times except the last two plots in the third row (MP1 vs MP3 on solid lines in Fig. 7a).

Second, when including everyone in the matching choices (MP3), giving priority to friends drastically increases successful matching between friends (Table 6). With clustered friendships (Decay (1.5) and Decay (2.0)), preference to friends (Het) results in roughly 14% ~ 18% of the total population matched with direct friends, while otherwise (Hom) it is only about 2% ~ 7%. With random friendships, the two values are around 5% ~ 10%, and 0.4% ~ 1.2%, but the former is still significantly higher. The percentage of direct friends in the matched population rises even up to 50%, and is increasing as spatial clustering of friendships gets stronger.

Finally, the number of successful matchings (Fig. 7a) is not positively associated with the size of choice set: 1) comparing solid and dash lines, homogeneous tolerance and willingness is not guaranteed to yield higher matching rate than heterogeneous; 2) in homogeneous cases (comparing MPs 1, 2, and 3 on dash lines), considering anyone nor contributes to a higher matching rate than considering friends only. With randomly distributed friendships (1st row in Fig. 7a), a
larger choice set generally comes up with more matches except the outlier with 100 friends. However, when friends get aggregated with $\beta$ equal to 1.5 (2nd row in Fig. 7a), the homogeneous method even yields a lower number of matches than its heterogeneous counterpart at MP3 (i.e., dash line drops down while solid line slightly goes up or keeps level). Looking at the homogeneous result alone, there is no monotonic trend either from MP1 to MP3. As the aggregation gets stronger (last row), however, the trend turns to the opposite of $\beta$ equal to 1.5 at MP3. Opposite to the previous two spatial patterns, the heterogeneous method of equal to 2.0 results in a decreasing trend from MP1 to MP3.

Table 6: Friend ratio:
The percentage who are matched with a direct friend in MP3 (fRate: % of the whole population; fRate2: % of the population who get a ride Het: with heterogeneous detour tolerance and willingness; Hom: homogeneous tolerance and willingness)

<table>
<thead>
<tr>
<th>Spatial pattern</th>
<th>Avg. # friends per agent</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>fRate(%)</td>
<td>fRate2(%)</td>
<td>fRate(%)</td>
<td>fRate2(%)</td>
</tr>
<tr>
<td>Het</td>
<td>7.52</td>
<td>22.96</td>
<td>5.2</td>
<td>25.97</td>
</tr>
<tr>
<td>Hom</td>
<td>1.08</td>
<td>1.53</td>
<td>0.42</td>
<td>3.70</td>
</tr>
<tr>
<td>Decay 1.5</td>
<td>fRate(%)</td>
<td>fRate2(%)</td>
<td>fRate(%)</td>
<td>fRate2(%)</td>
</tr>
<tr>
<td>Het</td>
<td>14.48</td>
<td>41.84</td>
<td>14.88</td>
<td>42.41</td>
</tr>
<tr>
<td>Hom</td>
<td>1.84</td>
<td>7.10</td>
<td>2.62</td>
<td>10.18</td>
</tr>
<tr>
<td>Decay 2.0</td>
<td>fRate(%)</td>
<td>fRate2(%)</td>
<td>fRate(%)</td>
<td>fRate2(%)</td>
</tr>
<tr>
<td>Het</td>
<td>16.28</td>
<td>41.81</td>
<td>17.8</td>
<td>51.35</td>
</tr>
<tr>
<td>Hom</td>
<td>4.7</td>
<td>9.20</td>
<td>5.5</td>
<td>10.55</td>
</tr>
</tbody>
</table>

5.4.2 From empirical studies

The average distance between first degree friends in randomly distributed social networks is 20.1km, that of social networks with distance decay is 15.6km, while it is 4.2km between anyone within a 10km buffer and is 20.1km between anyone anywhere. From Table 7, the results of average detour costs of these two types of social networks, however, do not demonstrate significant difference ($p > 0.05$), but they both are significantly higher than not prioritising social networks. The detour cost with decay is slightly lower than randomly
Table 7: Results from empirical studies (average of 5 runs)

<table>
<thead>
<tr>
<th>Group</th>
<th>Detour cost (of)</th>
<th>#matches</th>
<th>fRate2 (as Table 6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Direct friends</td>
<td></td>
</tr>
<tr>
<td>1) SW</td>
<td>7.83%</td>
<td>11.10%</td>
<td>∼ 71</td>
</tr>
<tr>
<td>2) SWD</td>
<td>7.69%</td>
<td>9.48%</td>
<td>∼ 87</td>
</tr>
<tr>
<td>3) NSW*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3a</td>
<td>SW</td>
<td>3.89%</td>
<td>SWD</td>
</tr>
<tr>
<td>3b</td>
<td>NSW</td>
<td>3.37%</td>
<td>SWD</td>
</tr>
</tbody>
</table>

*In Group 3, SW and SWD mean the relevant values of direct friends assuming the underlying social network is small world or small world with distance decay, even though social network is not considered in matching.

distributed network, despite no significance (p = 0.21). On the other hand, the number of matches with distance decay are significantly higher than that from the randomly distributed networks (p < 0.05). The percentage point of matched persons who are friends when the social network is prioritised (SW, SWD) are at least 30 times when no social network is considered (NSW).

5.5 Discussions

The first theoretical finding that the detour cost with a friend is not always higher than with a stranger remarks the impact of spatial configuration of friendship and sharing preference on detour cost. While many cases observe the increased detour cost if choices are only limited to friends (MP1), the two outliers (row 3 last two columns) are the extreme cases when friendships are so highly clustered that strangers who are pushed further out contribute no benefit to detour cost (MP1 vs MP3). The more clustered the spatial configuration of friendships is, the more effectively SNeRs degrades the detour cost (compare MP3’s on solid and dash lines of the same column in Fig. 7a). This is also accepted by the empirical detour costs of small-world versus with distance decay in Table 7.

Regarding the friendship ratio (the second finding), SNeRs prioritising friends is advantageous over existing algorithms that treat everyone uniformly. The former raises successful matching of friends to up to about 18% of the total population. It is at least double, and in some cases even tens of times the latter. The empirical study supports this by its fRate2 (Table 7). Furthermore, the more clustered the social network is (SWD vs SW vs NSW), the higher percentage direct friends takes of the matched population. Such drastic increase of friends in matches potentially leads to lower negotiation and trust
cost, of which the benefit is not yet modelled here. Considering negotiation cost in future work may prove an even higher merit of the proposed algorithm.

A successful matching rate is not guaranteed even with a larger choice set (the third finding) probably due to disruptive chances that earlier matches have an impact on later ones. Two facts contribute to disruptive chances: that the order of rides to be matched is fixed by starting time sequence instead of least detour cost, and that a failed matched ride does not have a second round of matching. Given that each person chooses their current best match, a higher number of candidates means, on one hand more choices, but on the other, higher degree of competition with others. A choice set is introduced with disfavoured competition if not being enlarged with ideal candidates with whom a match can be completed immediately. Adding candidates to an already crowded neighbourhood (e.g., 100 friends with $\beta = 2.0$) or matching a candidate of lower rank (i.e., higher cost as in $\beta = 1.5$) forces a decision to be postponed for potentially better choices but eventually no luck. Therefore, low-cost pairs should be matched as early as possible before the chance is missed out. In the cases of $\beta = 2.0$ with homogeneous tolerance, however, the results of MP3 are better because strangers are introduced in mid-range neighbourhoods with fewer competitions and still within detour tolerance (which is not the case in $\beta = 1.5$, Fig. 6).

5.6 CONCLUSIONS AND FUTURE WORK

This work builds an agent-based simulation of ridesharing considering heterogeneous tolerance and willingness with different social contacts. A theoretical study varying variables provides insights into the influence of social network’s spatial distribution on ridesharing. Detour cost with friends is not always significantly higher than collaboration with strangers, depending on spatial distribution of friends. More importantly, when people have spatially clustered and proximate friends, an algorithm preferring friends advantageously and preferably yields significantly higher matching rate between friends.

This work points out new challenges to be addressed in future work. The first is to investigate more empirical studies in varies study areas, especially of different landscapes and population distributions for more thorough understanding. Additionally, a dynamic model is needed to accommodate requests lodged in real-time, which requires efficient computation to run large simulations. Current methods [Agatz et al., 2011; Rigby and Winter, 2015] choose to sacrifice resolution (to zone level) for large simulation or complex agent decisions. However, demand-responsive transport requires relatively high resolution plus agent-based processing, for which better scaling solutions are needed.
THE EVOLUTIONARY INTERACTION BETWEEN TAXI-SHARING BEHAVIOURS AND SOCIAL NETWORKS

This chapter is adapted from the paper manuscript of the same title as the chapter title, submitted to Transportation Research Part A: Policy and Practice for peer review. As the first author, I was the main researcher proposing the research idea, designing and implementing the experiment, manipulating the dataset, analysing the results, doing discussions, and writing the manuscript. My supervisor Prof Stephan Winter was regularly discussing with me on the research idea, design, and progress. My colleague Dr Ronny Kutadinata was contributing to mathematical modelling and optimisation formulation by his expertise.

6.1 MAJOR CONTRIBUTIONS

Taxi-sharing is a process to match (part of) passengers’ trips so that they travel in the same car, which is different from ridesharing that also considers drivers’ mobility demands. However, traditionally taxi-sharing matching and routing algorithms focus on computational, operational or service efficiency. Few solution for taxi-sharing, and ridesharing in general, was acknowledging social and behavioural issues until very recently, but only to a limited degree. Understanding the spatial characteristics of social networks is very important for taxi-sharing behaviours: the spatial distribution and the density of a social network directly affect the feasibility of socially induced matchings, while the feedback from collaborative travel experiences projects on future collaborations. This work, therefore, focuses on the spatial aspect of a taxi-sharing-oriented social network along with its spatial development. Understanding the co-evolution of taxi-sharing and its interdependent social network helps to address people’s incentive of commitment to new travel modes.

The research hypothesis is that, from a dynamic long-term view, a taxi-sharing method with social network preference is advantageous over those ignoring social ties, yielding significantly higher match rates and suggesting an efficient search heuristic.

This chapter introduces a social network based taxi-sharing method that enables to prioritise taxi-sharing with acquaintances over strangers while capping the detour cost to reasonable and varied limits. Furthermore, the social network structure evolves and is updated based on shared trips. An empirical simulation demonstrates the advantages
of social network based taxi-sharing, i.e., an increased match rate and a comparable satisfaction level to trip-based methods. The spatial aggregation of the emerging social network not only suggests a space-time searching heuristic for taxi-sharing, but also indicates how social factors conquer space while space constrains social interactions.

6.2 METHODOLOGY

This section first introduces a social network based taxi-sharing algorithm and optimisation (Section 6.2.1), followed by details for the social network construction (Section 6.2.2), and finally looks into the dynamic mechanism between taxi-sharing (out of mobility needs) and its emerging social network from travel experiences (Section 6.2.3). Table 8 lists the variables frequently referred in this section.

6.2.1 Social network based taxi-sharing algorithm and optimisation

A known influence of social network on taxi-sharing is that people allow for heterogeneous detour tolerances due to varied social affinities [Chaubé et al., 2010; Wessels, 2009]. A social network based taxi-sharing method is proposed that integrates heterogeneous detour tolerance and satisfaction quantified by a symmetrical utility based on historical taxi-sharing experiences between a pair of persons. On each day, the matching does a static pre-planning model alike to the algorithm in Wang et al. [2016a] except it does not consider alternative destinations but involves social networks. A match decides the sequence of (for simplicity) two trips’ pick-ups and drop-offs subject to both travellers’ space-time budgets, assuming no traveller walks and there are sufficient vehicles (not explicitly modelled) serving travel demands. If a travel fails to be matched with anyone’s, the traveller will travel alone.

Given a population \( P \) with a certain geographic distribution of their homes, each person \( p_i \in P \) has a set of full day trips \( \tau_{it}^p \) for day \( t \). A trip regardless of its owner is \( x_{t,a} \), where \( a \) is the index. All trips on day \( t \) form a full trip set \( X_t = \bigcup_i \tau_{it}^p \). Let \( O : X_t \rightarrow P \) be the function to find the owner of a trip. The matching involves four stages [Wang et al., 2016a]: 1) reading the whole day plan and initialising each person’s travel agenda; 2) assigning travel budget (including detour tolerance) for each trip; 3) matching every each pair of trips \( x_{t,a}, x_{t,b} \in X_t \), \( O(x_{t,a}) \neq O(x_{t,b}) \), and finding all potential (satisfying both travellers’ space-time budgets) pairs of matches \( u(x_{t,a}, x_{t,b}) \in U_t \) as the pre-computation matrix; 4) running binary integer programming weighed by utility to get the final solution \( M_t \).

Given the travel agenda \( X_t \) on day \( t \), each trip \( x_{t,a} \in X_t \) is associated with the earliest starting time \( \text{EST}(x_{t,a}) \), latest ending time \( \text{LET}(x_{t,a}) \), direct (shortest) travel time \( \text{dTime}(x_{t,a}) \), and maximum
### Table 8: Variable catalogue

<table>
<thead>
<tr>
<th>Variable</th>
<th>Meaning</th>
<th>Variable</th>
<th>Meaning</th>
<th>Variable</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P )</td>
<td>population</td>
<td>( p_i )</td>
<td>person i</td>
<td>( t_i )</td>
<td>person i’s trips on day t</td>
</tr>
<tr>
<td>( X_t )</td>
<td>trips on day t</td>
<td>( x_{t,a} )</td>
<td>a-th trip on day t</td>
<td>( O(x_{t,a}) )</td>
<td>owner of a trip</td>
</tr>
<tr>
<td>( U_t )</td>
<td>day t’s potential matches</td>
<td>( u(x_{t,a}, x_{t,b}) )</td>
<td>one potential match in ( U_t )</td>
<td>( M_t )</td>
<td>final matches on day t (from ( U_t ))</td>
</tr>
<tr>
<td>( M_{t}^{p_i,p_j} )</td>
<td>final matches between ( p_i, p_j )</td>
<td>( \delta_t(p_i, p_j) )</td>
<td>degree of separation on day t</td>
<td>( S_t )</td>
<td>social network on day t</td>
</tr>
<tr>
<td>( c_t(p_i, p_j) )</td>
<td>utility of ( p_i, p_j )</td>
<td>( G_t )</td>
<td>complete graph for utilities in ( P )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Detour budget \( \text{maxDetTrip}(x_{t,a}) \). The \( \text{maxDetTrip}(x_{t,a}) \) is calculated as \( \min(dTTime(x_{t,a}) \cdot \text{DetRate}(x_{t,a}, x_{t,b}), \text{LET}(x_{t,a}) - \text{EST}(x_{t,a}) - dTTime(x_{t,a})) \), which is the maximum travel time for detour satisfying both psychological willingness and physical travel budget. \( \text{DetRate}(x_{t,a}, x_{t,b}) \) (Eq. 6) is the percentage of a trip’s shortest travel time the person would like to sacrifice depending on the degree of separation between the two travellers, \( \delta_t(O(x_{t,a}), O(x_{t,b})) \).

\[
\text{DetRate}(x_{t,a}, x_{t,b}) = \begin{cases} 
30\%, & \delta_t(O(x_{t,a}), O(x_{t,b})) = 1, \\
25\%, & \delta_t(O(x_{t,a}), O(x_{t,b})) = 2, \\
10\%, & \delta_t(O(x_{t,a}), O(x_{t,b})) = \infty.
\end{cases}
\] (6)

Stage 4 maximises the number of matches. Each \( u(x_{t,a}, x_{t,b}) \in U_t \) is binary: 1 (matched) or 0 (not matched), subject to that one trip can be matched to up to one other trip. The matrix is weighed by \( c_t(O(x_{t,a}), O(x_{t,b})) \), the utility between two travellers. Multiple matches between the same pair of persons on day t are assigned the same weight. The final output from binary linear programming is \( M_t = \{ (x_{t,a}, x_{t,b}) | u(x_{t,a}, x_{t,b}) = 1 \} \). The matched trips between
\((p_i, p_j)\) is defined as \(M_t^{p_i, p_j} = \{(x_{t,a}, x_{t,b}) \mid (x_{t,a}, x_{t,b}) \in M_t; O(x_{t,a}), O(x_{t,b}) \in \{p_i, p_j\}\}.\) Hence, the total number of matches up to day \(t - 1\) is:

\[n_{t' < t}^{p_i, p_j} = \sum_{t'=1}^{t-1} |M_{t'}^{p_i, p_j}|.\]  

\((7)\)

The maximum value of \(n_{t' < t}^{p_i, p_j}\), used in Section 6.2.3 to normalise \(n_{t' < t}^{p_i, p_j}\), is approximated by:

\[N_{t' < t}^{p_i, p_j} = \sum_{t'=1}^{t-1} \min\{|\tau_{t'}^{p_i}|, |\tau_{t'}^{p_j}|\},\]  

\((8)\)

where \(|\tau_{t'}^{p_i}| = \{|x_{t',a} \mid O(x_{t',a}) = p_i\}\). Theoretically, the maximum number of matches cannot be larger than any of the travellers’ number of trips per day. Let the actual detour cost of a match \((x_{t,a}, x_{t,b}) \in M_t\) be \(\text{detCost}(x_{t,a}, x_{t,b})\), simply quantified by the average percentage of shortest travel time of \(x_{t,a}, x_{t,b}\) for detour, the total detour cost of \((p_i, p_j)\) until day \(t\) is:

\[\text{detCost}_{t' < t}^{p_i, p_j} = \sum_{\{(x_{t',a}, x_{t',b}) \in M_{t'}^{p_i, p_j} \mid 1 \leq t' < t\}} \text{detCost}(x_{t',a}, x_{t',b}).\]  

\((9)\)

6.2.2 Construction of taxi-sharing-oriented social networks

The taxi-sharing-oriented social network is a network representative of the trust level and collaboration experience of taxi-sharing among a population. On day \(t\), the social network \(S_t\) \((t \geq 1)\) encapsulates all pairs of persons who are directly connected, i.e., degree of separation \(\delta_t(p_i, p_j) = 1\). If two persons are only indirectly connected via a third person, \(\delta_t(p_i, p_j) = 2\); otherwise they are strangers \(\delta_t(p_i, p_j) = \infty\). The initialisation of a social network, \(S_0\), is to permute links with \(\delta_0(p_i, p_j) = 1\) in \(P\).

A complete graph \(G_t = \langle P, E, C_t \rangle\) is built to represent utility between each pair of persons, where \(c_t(p_i, p_j) \in C_t\) is the utility for the pair \(e(p_i, p_j) \in E\). A direct social network link will be built if \(c_t(p_i, p_j)\) goes over a threshold \(\text{SocThres} = w^h + 0.6 \cdot w^{fr}\), i.e., \(\delta_t(p_i, p_j) = 1 \iff c_t(p_i, p_j) > \text{SocThres}\); if the utility drops below this threshold, the link will be dissolved. The threshold is built so that no link can be upgraded to 1st degree without any taxi-sharing history. Since 0.6 is the friendship utility of 2nd degree friends (to be specified in (Eq. 11b)), and \(t\) is the maximum utility of home location, \(\text{SocThres}\) guarantees that even if a 2nd degree link with the highest utility of home location cannot evolve into a 1st degree link unless taxi-sharing history adds up to it.
6.2.3 The dynamic co-evolution of taxi-sharing and social network

While Sections 6.2.1 and 6.2.2 clarify, respectively, how taxi-sharing happens and how a static social network is built on one day, this section elaborates the dynamic co-evolution of taxi-sharing and its induced social network. The co-evolution cycle goes in such a procedure (Figure 8): 1) update the utility to $c_t(p_i, p_j)$ for each $e(p_i, p_j) \in E$; 2) update the social network from $S_{t-1}$ to $S_t$ based on the latest utility; 3) run taxi-sharing simulation (Section 6.2.1) for $X_t$; 4) calculate $M_t$ and get the most up-to-date ridesharing history $H_t$ by appending $M_t$ to $H_{t-1}$.

Travellers’ preferences are based on the following assumptions:

- The proximity of origins counts more than that of destinations [Vanoutrive et al., 2012], which explains why home distance contributes to travel utility.
- Detour tolerance increases with social acquaintance [Chaube et al., 2010; Wessels, 2009], which reflects the trust issue in taxi-sharing.
- The collaboration network structure is affected by the initial social network structure that can be of any type as long as it gives a person higher trust to the connected persons.
- Trust and preference are gained or lost from taxi-sharing experiences, including:
  1) People prioritise lower detour cost;
  2) The more the times two persons share rides, the higher the trust;
  3) The above two rewards are subject to diminishing marginal utility [Kauder, 1965];
  4) Lacking sharing experience reduces travel utility and thus dissolves collaboration links.

Therefore, the travel utility encompasses three factors: the distance between riders’ homes that affects their possibility to share rides, current social network status, and taxi-sharing history (Figure 8). Taxi-sharing history up to date includes the average detour cost of past rides, the number of matches until yesterday, and the no-match penalty. As one feasible function that incorporates the aforementioned factors, the utility function for $e(p_i, p_j)$ is formulated as:

$$c_t(p_i, p_j) = w^h \cdot c^h(\cdot, \cdot) + w^{fr} \cdot c^t_{fr-1}(\cdot, \cdot) + w^{md} \cdot c^m_{\leq 1}(\cdot, \cdot) \cdot c^d_{\leq t}(\cdot, \cdot) + w^-n \cdot c^-n(\cdot, \cdot)$$ (10)

In Equation 10, $c^h(p_i, p_j)$ is the utility from home distance $d(p_i, p_j)$ of two persons, $c^t_{fr-1}(p_i, p_j)$ is the utility based on current friendship, $c^m_{\leq 1}(p_i, p_j)$ is the utility from the number of matches in history, $c^d_{\leq t}(p_i, p_j)$ is the utility from detour cost in history, and $c^-n(p_i, p_j)$, a negative value, is the penalty on friendship for no-match. $w^h, w^{fr}, w^{md},$ and $w^-n$ are the weights for each term. To show the overall ride ex-
experience throughout time rather than of one ride, the utilities from number of matches and detour cost have a joint function and weight.

The specific formulas for each term are defined below ($t \geq 1$):

\[
\begin{align*}
    c_h[p_i, p_j] &= f_h(d[p_i, p_j]) \\
    c_{t-1}^{fr}(p_i, p_j) &= \begin{cases} 
    0.9, & \delta_t(p_i, p_j) = 1 \\
    0.6, & \delta_t(p_i, p_j) = 2 \\
    0.1, & \delta_t(p_i, p_j) = \infty
    \end{cases} \quad (11a) \\
    c_{t-1}^{m}(p_i, p_j) &= \frac{\ln(n_{t'<t}^{p_i,p_j} + 1)}{\ln(N_{t'<t}^{p_i,p_j} + 1)} \quad (11b) \\
    c_{t-1}^{d}(p_i, p_j) &= \begin{cases} 
    1 - \frac{\text{detCost}_{t'<t}^{p_i,p_j}}{n_{t'<t}^{p_i,p_j}}, & t > 1 \\
    0, & t = 1
    \end{cases} \quad (11c) \\
    c_{t-1}^{n}(p_i, p_j) &= \frac{1 - e^{(t-t^*-1)}}{1 - e^{(D-1)}} \quad (11d)
\end{align*}
\]

In Equation 11a, the utility dependent on $d(p_i, p_j)$, the distance between the two persons’ homes, can be specified in particular contexts (see Section 6.3.3). Equation 11b indicates a higher utility (i.e., taxi-sharing priority) for closer friends. Equation 11c quantifies the influence of the total number of matches between a pair of persons in history $n_{t'<t}^{p_i,p_j}$ (Eq. 7). The logarithmic function indicates the diminishing marginal utility of each extra trip. The function is normalised with the utility value by plugging in the theoretically maximum number of matches $N_{t'<t}^{p_i,p_j}$ (Eq. 8). Equation 11d is the utility from the average detour cost of matches in history between $p_i, p_j$, where $\text{detCost}_{t'<t}^{p_i,p_j}$ is the total detour cost specified in Equation 9. Sub-utilities (Eqs. 11c and 11d) are positive feedback from taxi-sharing experience. Equation 11e, on contrast, is the negative feedback from taxi-sharing that punishes a friendship with no-match. $t^* \geq 0$ is the last day when the pair had a match, and $D$ is the total duration of the simulation. The exponential form reflects stronger retention of friends at the beginning, when people still memorise their links even though they have not met for some time. The penalty grows faster, however, when two persons do not have any contact for longer time.

6.3 EXPERIMENT DESIGN

The parameter settings are listed in Section 6.3.1 as a general overview. The simulation runs a period of 30 days ($D = 30$) with the processed datasets by Jain et al. [2017] as detailed in Section 4.2. The trips are assumed to be sampled for one weekday, since the surveys were conducted on a weekday. To tailor for this study, a trip-regeneration processing (Section 6.3.2) is used to permute trips for 30 days. Section
6.3.3 discusses the spatial distribution of the sampled trips along with its influence on the simulation. Section 6.3.4 details the implemented social network initialisation.

6.3.1 Simulation runs and parameters

In Eq. 10, the weight of each term in the utility function determines how crucial a role each term contributes. Higher ratios of $w^{md}$ to other weights from non-sharing factors (e.g., $w^h$, $w^{fr}$) mark higher influence of sharing experience on the consequential social network structure, while distance decay and current social network structure play smaller parts. The weights for each term in the simulation are designed as shown in Table 9. Setting#3 is the most impacted one by taxi-sharing. Each term in Equation 10 is normalised between 0 and 1 (inclusive), as formulated in Equation 11.

<table>
<thead>
<tr>
<th>Setting#</th>
<th>$w^h$</th>
<th>$w^{fr}$</th>
<th>$w^{md}$</th>
<th>$w^{-n}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.2</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

6.3.2 Generating recursive trips over 30 days from VISTA

Given that VISTA data only samples one-weekday travels, trips of the 30 days are assumed to have a recursive period of 5 days (excluding weekends). The first 5 days are permutated from the processed VISTA
data by introducing the randomness to flexible trips with either flexible travel time or location (defined in Wang et al. [2016a]) that may or may not occur on different days. With the permuted 5 days’ trips as the baseline, the rest days will repeat the trips on their corresponding day, e.g., day 5 repeats day 0. The only difference of recursive travels on different periods is the dynamically updated utility of taxi-sharing with a certain person, which consequentially leads to varied matching outputs.

### 6.3.3 Home distribution and distance decay function

As part of the utility function (Eq. 10), home distance contributes to the travel partner choice. The particular distance decay function is defined as Equation 12 because of the heterogeneity in terms of landscape and home distribution in Yarra Ranges. The average distance between two spatial clusters of population is about 10km. The step-function creates a heuristic that inclines to assign higher utility to people within the same spatial cluster ($\leq 10$ km) who have higher chances for taxi-sharing. Though home locations cannot fully capture the characteristics of human movement behaviours, they are one of the feasible representatives since people’s activity spaces are regularly bounded by such anchor points as home [González et al., 2008; Huang et al., 2010; Wang et al., 2015b].

$$f_h(d(p_i, p_j)) = \begin{cases} 0.1, & d(p_i, p_j) > 10 \\ \alpha \cdot (d(p_i, p_j) + 1)^{-\beta}, & d(p_i, p_j) \leq 10 \end{cases}$$

When plugging in the values that $d(p_i, p_j) = 0$, $f_h(d(p_i, p_j)) = 1$, and $d(p_i, p_j) = 10$, $f_h(d(p_i, p_j)) = 0.1$, we get $\alpha = 1, \beta = \log_{11} 10$.

### 6.3.4 Permutation of social networks

Two types of social network initialisations are designed to compare their corresponding taxi-sharing outputs: 1) a null model (from scratch) that assigns no 1st degree link at the beginning, and observes the emerging social network from taxi-sharing experience; and 2) a small world model that is topologically locally well-connected but with relatively short average distance between nodes, which is believed to capture the topological structure of real social networks [Milgram, 1967; Amaral et al., 2000; Watts and Strogatz, 1998]. The initialisation of small world social network is generated with the WattsBetaSmallWorldGenerator module in Repast. Despite of no universally recognised network density, the average direct contacts per person is set to 6 in the population of 714 people, yielding a low density of 0.84%, which is widely observed in real world networks [Faust, 2006]. The pair of persons with an initialised link is set with $\delta_t(p_i, p_j) = 1$. In
null model cases, no link is generated, and therefore, all $\delta_t(p_i, p_j) = \infty, (\forall i; j \neq i)$ at the beginning. Any $2^{\text{nd}}$ degree link is derived from the $1^{\text{st}}$ degree links.

### 6.4 Results

The dynamics of four variables are tracked over time (Fig. 9): 1) the average number (degree) of direct friends per person (Fig. 9a); 2) the average distance between homes of direct friends (Fig. 9b); 3) the number of matches on each day over the period (Fig. 9c); and 4) the fluctuation of the daily average utility of direct friends throughout time (Fig. 9d). In each figure’s legend, “1”, “2”, “3” represent the setting number in Table 9, and “sw”, “null” are for small world networks and null networks, respectively.

#### a. Degree distribution

In null network models, the average degree steadily grows to a stable plain (from 0 to 0.084, 0.232, and 1.188 for each setting). A similar trend (with a shift from the origin) appears at the early and middle stages in small world models until a sharp drop erupts on day 24 resulting from the broken links due to no-match penalty (Eq. 11e). While real social network links do not break down easily, the simulation shows only the fast-forward process of the collaboration links representing the feasibility of taxi-sharing partnership. Small world models eventually converge to a slightly higher degree (0.26, 0.478, and 1.384 for each setting) than their counterparts in null models at equilibrium.

#### b. Spatial aggregation

The average distance between direct friends’ homes in small world models is decreasing over time. On Figure 9b, an obvious nadir of the distance pops up on day 24 when unused links break down, followed by a bounce-back to the final equilibrium. Results from different weights of taxi-sharing experience are not alike. The more the experience weighs in (Setting#3 to Setting#1), the faster and lower the average home distance drops before the nadir. A similar order occurs at equilibrium as well. At the nadir, however, the rank is the opposite. On comparison, the average home distance from null models is smooth, slightly increasing to the equilibrium over time except a tiny hump at the beginning of each curve due to the scarcity of matches and randomness in space at the beginning. Nevertheless, Setting#1 from neither model yields distinguishable variation of average home distance owing to its low sensitivity to the reward of taxi-sharing.

#### c. Number of matches

The number of matches per day displays a recursive pattern as an inherent result from the repeated travel pattern of 5 days, but is not obviously related to the degree of friends, the average home distance, or the average utility (Figures 9a, 9b, 9d). Nevertheless, the daily number of matches from small world models
The evolutionary taxi-sharing behaviours

Figure 9: Output results from taxi-sharing over 30 days (coloured).

(c) Number of matches per day

(d) Avg. utility of direct friends

(a) Avg. degree of direct friends

(b) Avg. home distance of direct friends
is higher than their counterparts in null models. Higher weight of taxi-sharing also leads to higher match rate.

d. Average utility. The average utility between direct friends in small world models increases a little in the first few days and stays steady until day24 when a distinguishable rise appears. The sharp rise is consistent with the break-down of the unused links, a self-optimising process to maximise utility. Though the results from the null models demonstrate the opposite trend, decreasing steadily all the way, they converge to a statistically identical utility \((p < 0.05)\) to the small world models at the equilibrium.

6.5 discUsSions

Indicated from the simulation results, social-induced taxi-sharing and the social (collaboration) network mutually reinforce each other in favour of future taxi-sharing outcomes. This is a reflection of the deeper level phenomenon that society is driven both by socio-economic factors and by spatial laws (e.g., distance decay) [Hillier, 2007; Batty, 2008]. The major points are argued as below:

The equilibrium proves the existence of the stable status accommodating social factors and geographical constraints. Outcomes from simulations, no matter considering or ignoring social networks, all demonstrate an equilibrium in each case. Figure 9b shows that the consequential networks at the equilibrium from small-world social network initialisations are more spread over in space, which means higher detour cost sharing taxis with friends. However, the corresponding simulations with and without initial social networks converge to statistically identical average levels of satisfaction (i.e., utilities) with the rides (Figure 9d after day24). That is to say, higher detour cost with friends does not affect the eventual satisfaction level of travel. This is a consequence of balancing detour cost and travel partner preference. On one hand, for the cases with initial social networks, even if the matches prioritising friendships are not always of the lowest detour cost, they are traded off by trust. On the other, when no initial social networks exist, low trust is balanced off by lower cost shares with strangers nearby. The equilibrium is assumedly subject to the specific context of geography and population, which means more empirical studies of different contexts are needed when data is available.

Incorporating social networks into taxi-sharing can lead to a higher number of matches than ignoring social networks. In contrast to identical satisfaction levels and higher detour cost, the outputs with initial small world networks consistently yield significantly higher numbers of matches everyday than their corresponding results from null models (Figure 9c). The nonnull social network (i.e., small-world network) models embark the simulation with a higher level of trust,
Figure 1: Distributions of social networks on strata’s and diff/2 (coloured).
namely, higher detour tolerance for friends. Over the period of 30 days, the robustness of some initialised links is reinforced by collaborations. Discernible on Figure 10, the yellow links, which are the retained links in small world networks on day 29, exist between some relatively long-distance pairs. This is particularly obvious when social network weighs in more, i.e., in Setting#1 and Setting#2 (the first and second graphs in Figure 10). The null models that treat everyone as strangers, on contrast, are more conservative on detour tolerance from the beginning, yielding fewer matches (Fig. 9c). Taxi-sharing granting higher tolerance to friends beneficially enhances the number of matches, which indicates how social network conquers space. The initial social network does not have to be a consequent network particularly from taxi-sharing history, but can be adopted from any social network where trust has been built and hence can be reflected by taxi-sharing.

**Local strangers substitute remote friends to optimise the collaborative social network structure.** The existing spatially random social network (i.e., the small world network initialised without spatial distribution regulation) is observed growing spatially aggregated over time. Taxi-sharing is a type of spatially restricted social interactions that require co-presence. This is opposite to the observation by Hackney and Marchal [2011] that social networks are growing in space when people travel for spatially free social events. The extent of restriction is partially ascribed to how heterogeneous the landscape is. Melbourne built-up area is assumed to develop a spatially more spread social network from strangers than Yarra Ranges with many mountains and rivers. Spatially local taxi-sharing drives social link production by spatio-temporal co-presence (constraint), particularly, by overlapped travels. Therefore, a social network based algorithm not only reserves the existing social links sufficiently strong though of longer distance (while eliminating those long-distance unused links); but it also develops local short-distance links with those who were originally strangers but are proved efficient and trustworthy during taxi-sharing experience.

**Back to the hypothesis, the social network based taxi-sharing is advantageous.** Social network based taxi-sharing creates a heuristic filter that integrates the benefit of spatial proximity and the priority to friends. As aforesaid, instead of causing the loss of getting rides, preferring friends can even increase match rates. Moreover, feasible local links are constructed throughout the self-optimisation of the collaboration network. The heuristic, therefore, is that searching strangers for a match should be indexed by spatial proximity to reduce time consumption, while searching friends by social links satisfies travel preference.
6.6 CONCLUSIONS AND FUTURE WORK

This work investigates the interactive dynamics between people’s taxi-sharing behaviours and a taxi-sharing-oriented social network structure. The simulation is conducted with a suggested social network based taxi-sharing method that infuses varied utilities (rather than friend-only) for taxi-sharing with different partners into an existing taxi-sharing algorithm. The simulation is based on simplified assumptions only to lay the foundation for further studies into spatio-social systems. Getting back to the hypothesis, a social network based taxi-sharing method is superior to normal sharing methods because: 1) It increases the overall match rate while still keeping an identical satisfaction level of travel; 2) It utilises the attraction of social preference and trust to overcome the drawback of geographic barriers, which optimises the collaboration network structure; 3) It indicates a searching heuristic that feasible matches can be retrieved by both spatial and social indices.

As a generic category of algorithms, social network based taxi-sharing can be integrated into any other taxi-sharing or ridesharing algorithm. The social network also works as a heuristic for faster retrieval of potential partners, which may partially solve the calculation efficiency issues reported by Wang et al. [2016a]. Further studies can release the constraint of matching only two persons, and also explicitly model vehicles with a controlled fleet size. It is also worth examining a fully self-organised social-induced taxi-sharing system where no centralised taxi-sharing optimisation is conducted and the utility between a pair of persons is not mutual. Additionally, an empirical utility function, whenever the data is available, is urgently necessary to adjust the model.
This chapter is revised based on the paper of the same title published in ACM SIGSPATIAL Proceedings 2016 [Wang et al., 2016b], which received the Best Paper Award for ACM SIGSPATIAL 2016. I was conductive at every stage of this work, including research concept proposal, experiment design and implementation, data analysis, discussions, and paper writing. My colleague Dr Ronny Kutadinata was contributing significantly to research design, mathematical modelling, and mathematical formulation in optimisation. My supervisor Prof Stephan Winter was providing a sufficient amount of support in discussing research ideas, supervising the progress, and discussing the results.

7.1 MAJOR CONTRIBUTIONS

Ridesharing is an emerging travel mode that reduces the total amount of traffic on the road by combining people’s travels together. While present ridesharing algorithms are trip-based, this work aims to achieve significantly higher matching chances by a novel, activity-based algorithm. The algorithm expands the potential destination choice set by considering alternative destinations that are within given space-time budgets and would provide a similar activity function as the originals. In order to address the increased combinatorial complexity of trip chains, the paper introduces an efficient space-time filter on the foundations of time geography to search for accessible resources. Globally optimal matching is achieved by binary linear programming. The ridesharing algorithm is tested with a series of realistic scenarios of different population sizes. The encouraging results demonstrate that the matching rate by activity-based ridesharing is significantly increased from the baseline scenario of traditional trip-based ridesharing.

The research hypothesis is that, compared with the trip-based method, an activity-based ridesharing algorithm (ABRA) can efficiently increase matching rates by considering alternative destinations for flexible activities, while keeping detour costs comparable within tolerance. The algorithm makes a significant contribution to ridesharing by introducing time geography that inherently is capable of increasing travel flexibility and thus matching rate, and meanwhile to time geography from a computational perspective.

This research is significant in at least three ways:
By expanding potential destination choice sets, the algorithm should significantly increase the matching rates of ridesharing;

- By designing an efficient filter based on time geography, it limits the combinatorial explosion for trip chains on feasible activities;
- By linear programming, it finds optimal solutions for riders.

ABRA generally includes two steps: it initially builds up a pool of alternative destinations (and trips) for the targeted activities based on the trip’s space-time budgets; and then finds feasible matchings considering these alternatives in addition to the original ones. Matching is conducted as static preplanning with everyone’s daily schedules known. A daily schedule includes one or multiple trip chains, each of which includes multiple trips. Given its combinatorial computational complexity in deciding the destination choice set for a (especially long) chain of multiple flexible activities [Chen and Kwan, 2012; Arentze and Timmermans, 2004], ABRA proposes an efficient space-time filter for alternative destinations. The general principle is to impose a reasonable time window on these floating trips, while still allowing for detour tolerance for each trip. The assignment of time windows is generically dependent on the trip’s space and time flexibilities. ABRA has been set up to prove the hypothesis, which requires a non-heuristic solution to identify the theoretical potential of activity-based ridesharing. However, it is well recognized that in practical systems a dynamic, sub-optimal model is more applicable than the currently implemented static one. Details are elaborated below.

7.2 Methodology

The contribution of this work is an activity-based ridesharing algorithm (ABRA) that effectively expands trip destinations to choice sets for higher ridesharing matching rates. To prove the significance of the algorithm, it is compared to trip-based ridesharing planning as a baseline. The activity-based ridesharing algorithm computes centrally the global optimum of all feasible matches. Cheaper (i.e., heuristic, or decentralized) solutions may exist, but the global optimum is adopted here to reveal the theoretical capability of the activity-based approach. There are four modules (stages) in this model (Figure 11): person initialiser (M1), trip chain builder (M2), candidate matching (M3), and solution optimization (M4). M2 and M3 are the essential parts of activity-based ridesharing that make up ABRA.

7.2.1 Model setting and assumptions

The scenario is set with a certain population, each person of which has a complete list of full-day activities with predefined running order (i.e., the activity sequence and planning is out of the study scope). Assume that no two consecutive activities can be done at the same lo-
Figure 11: Workflow of the activity-based ridesharing model.
cation, and therefore must cause a trip, regardless of the activity’s space and time flexibility. \( M_1 \) initializes the population by assigning them these activities with originally planned locations and travel time. The output of \( M_1 \) is a population with initially planned trips.

The basic unit to be matched is a trip rather than a series of trips, regardless of the duration of the stops. Once a person gets to an activity destination, the ride is completed. Such design is beneficial in a way that nobody is forced to wait for another person conducting his/her activity. If one trip fails to be matched, that trip will be travelled alone. Transfer at a non-activity location is not allowed to avoid additional waiting time. Maximum amount of shared capacity is set to 2 passengers. This model also assumes self-driving vehicles operating as taxis, which makes ridesharing in this case a flexible form of taxi sharing, so that persons do not need to distinguish their roles as a driver or a passenger. This assumption releases the bonds caused by vehicle ownership: the person who gets into the car first (traditionally the driver) does not have to arrive at their destination at last. Consequently, there are no people (drivers) who, in addition to contributing their own vehicle, also have to make the most detours; instead, the algorithm can flexibly optimize the sequence of pickups and drop-offs solely based on transport demand.

\subsection*{7.2.2 Trip chain construction}

\( M_2 \): Trip chain builder is one of the essential parts of \textit{ABRA}. It mainly completes three tasks as shown by Figure 11: building trip chains, constructing space-time filters (STF), and retrieving feasible activity locations with these filters. The principle of \textit{ABRA} is to expand a person’s destination choice set for each location-flexible activity by finding places that are spatially diverse while functionally similar for that travel aim. For instance, if a person is looking for a supermarket, he/she can choose one from a few candidate locations depending on their time budget. There is a potential that different choices yield different chances to share a ride, which is to be proven by this work. Time-flexible but location-hard activities can also expand destination choice set, but indirectly. It allows for later arrival time, and thus is likely to provide more choices for its previous trip, if the previous destination is location-flexible. An example is when a person plans to shop for grocery on the way back home, he/she may stop at different grocery stores.

\textit{Building trip chains}

Accommodating a person’s space and time flexibility, \textit{fixed} activities are defined here as activities that can be conducted only at a fixed time and at a certain place. Release of either rigidity induces to a \textit{flexible} activity. Hence, the four types of activities are: hard-time-hard-
Figure 12: Activities of a day and trip chains.
location (HTHL), flexible-time-hard-location (FTHL), hard-time-flexible location (HTFL, which seldom happens and does not exist in the simulation), and flexible-time-flexible-location (FTFL).

Activities within a day are not independent; they are constrained by trips to other activities depending on the space and time flexibility of each activity. Say, a person has to start work at 9am in his/her office, before which he/she wants a coffee on the way from home. Then the activity of “grabbing a coffee” is limited by the next activity “working” in time, and thus in space. To determine where to buy the coffee, purely from a spatio-temporal perspective and omitting factors such as personal preference, an STF can be built. The construction of an STF requires two points with known location and time as control points: In this case, leaving home at a certain time and reaching work at 9am are the control points, and if this time window is sufficiently large to make stops or even small detours, the person has some flexibility to choose from a number of coffee places in between. Accordingly, a full day schedule must be split at control point activities, and be turned into a series of trip chains. Thus, the strong definition of a trip chain is a series of trips with two fixed activities at both ends and any number of flexible ones in between. Control point activities are hence named splitters (of trip chains) hereafter.

$M2$ is developed to find all splitters and to break a whole day schedule into trip chains at these splitters. In practice, a splitter is an activity isolating the trip chains on both sides (e.g., resistant to delay) and thus providing for space-time stability. In addition to HTHL activities (fixed activities), hard-location activities with a duration longer than a threshold (e.g., 1 hour) can also be used as splitters since they provide some capacity to absorb delays. Thus, a weaker version of a trip chain is that with at least a FTHL splitter.

Figure 12 shows the relations between activities, trip chains, and trips within a chain. A day contains at least one chain, and a chain must have at least one trip. Dash lines and dash-dot lines in the figure indicate some omitted details. The interior structure of a trip chain $i$ is expanded. Let $O_i$ be the origin of this chain, and $D_i$ the last stop. Let $N(i) \in \mathbb{Z}_{>0}$ denotes the number of trips within chain $i$, where $\mathbb{Z}$ is the set of integers. Hence, there are totally $N(i)+1$ activities on this chain, including those at the origin and the final stop. $\text{Act}_i^j$ denotes the $j^{th}$ activity on chain $i$. The corresponding locations for $\text{Act}_i^j$ form a set $\mathcal{S}_i^j = \{s_{i,k}^j\}$, where $s_{i,k}^j$ is the $k^{th}$ candidate location to conduct activity $\text{Act}_i^j$. The original location of an activity is indicated by $k=0$. Thus, $O_i = s_{i,0}^0$ and $D_i = s_{N(i),0}^i$. The following part will use this chain to explain the construction of the STF.

Construction of space-time filters

The construction of a trip chain depend on the shape of an STF. STFs are derived from the concept of space-time prisms in time geography
[Miller, 1991, 2005]. Related concepts are already discussed in Section 3.1. In Figure 13a, the black thick ellipse forms the PPA of the space-time prism spanned by the splitters $O_i$ and $D_i$. For a flexible activity in between, the geographic space offers multiple POIs, but only the ones within the ellipse are feasible.

An STF is a derived concept, but handles more complicated cases. Think about a situation that the length of a trip chain is longer than 2, which means there are more than one flexible activities between two splitters. In Figure 13a, the current chain contains $N(i)$ trips, i.e., $N(i) - 1$ flexible activities for which alternative locations should be searched. In Chen and Kwan [2012]'s work, a theoretical space-time model for multiple flexible activities are given and tested with a simple programme with only two flexible activities between splitters. From computation’s perspective, however, the combinatorial number of iterations will explode with the growth of trip chain length and the variety of location selections. The proposed STF solves this by imposing a hard time limit on each trip while still granting it flexibility. The solution is allowing for some detour for each trip, and fixing the time limits of earliest starting time and latest ending time of each trip. The advantage is that it reduces trips’ interdependency and thus leads to less computational complexity. However, such design requires a revision on its solution, because it lacks the time bounding by its following trip. The revision will be discussed in Section 7.2.3.

Let $t_{i,j,k,l}$, where $i,j \in Z_{>0}$ and $k,l \in Z_{\geq 0}$, denotes the trip within chain $i$ from $s_{i,j-1,k}$ to $s_{i,j,l}$. Thus, $t_{i,0,0}$ is the original 1st trip of this chain and $t_{i,N(i),l}$ is the original last trip to the last stop $D_i$. Also, define $EST(t_{i,j,k,l})$ and $LET(t_{i,j,k,l})$ as the earliest starting time and the latest ending time of $t_{i,j,k,l}$ respectively, which are used to delimit the the maximally tolerable time budget for that trip. Furthermore, the direct travel time of trip $t_{k,l}$ is denoted as $dTTime(t_{i,j,k,l})$. Moreover, start time (StartTime) and arrival time (ArrTime) are the actual travel timestamps of each person’s trip by their initial planning. $ActDur(\cdot)$ denotes the activity duration at a stop.

Algorithm 4 outlines the process of calculating $EST(\cdot)$ and $LET(\cdot)$ of each trip. Note that the travel time for any alternative destination is always kept the same as its original’s, that is $\forall (k,l)$, $EST(t_{i,j,k,l}) = EST(t_{i,j,k,l})$ and $LET(t_{i,j,k,l}) = LET(t_{i,j,k,l})$. The algorithm works as follows.

1. Any activity $Act_i$ that is time-hard (TH) must be satisfied: The departure at a TH origin cannot be advanced, and arrival at a TH destination cannot be delayed.
2. Before finding the time bounds of each trip within a chain, the time bounds of the whole trip chain is first determined. If both splitters are TH, then the time bounds of the trip chain are simply the origin’s StartTime and the destination’s ArrTime. Otherwise, the following formula is used to determine the max-
The maximum allowable duration (without activity duration) of a trip chain:

\[
\text{maxDur}(i) = \min((1 + \text{DetRate}) \cdot \sum_j d\text{Time}(t_{0,0}^{i,j}), \text{GlobalDet}), \tag{13}
\]

where \(\text{DetRate}\) is the maximum detour time of a person in a trip (as a proportion to the direct travel time), and the resulting duration should never go over a global chain duration threshold \(\text{GlobalDet}\). As shown by Algorithm 4, if only one of the splitters is TH, then the corresponding time bound is imposed (either the origin’s StartTime or the destination’s ArrTime), while the other is adjusted according to the calculated maximum allowable duration in Equation 13. If both splitters are time-flexible (TF), the destination splitter is assumed as TH and the destination’s ArrTime is used as a time bound, just to avoid advancing departure.

3. Next, the time bounds for each trip are determined. By using time bounds of the trip chain defined above, the maximum detour time of a trip chain (i.e. the time budget) can then be calculated as follows,

\[
\text{maxDet}(i) = \text{LET}(t_{0,0}^{i,N(i)}) - \text{EST}(t_{0,1}^{i,j}) - \sum_{j \in \{1, ..., N(i) - 1\}} \text{ActDur}(\text{Act}_{ij}^i), \tag{14}
\]

According to the definition of \(\text{splitter}\), and given that no HTFL activities exist, it is easy to infer that any activity between the splitters on a trip chain are TF. Therefore, the maximum allowable duration of each trip is determined by distributing \(\text{maxDet}(i)\) in proportion to the direct travel time of each trip relative to the total direct travel time of the original/actual trip chain. Note that the maximum detour time of a trip \(\text{DetRate}\) is still imposed. Hence, the maximum detour time of a trip is calculated as follows:

\[
\text{maxDetTrip}(t_{0,0}^{i,j}) = \min(\text{maxDet}(i) \cdot \frac{d\text{Time}(t_{0,0}^{i,j})}{\sum_j d\text{Time}(t_{0,0}^{i,j})}, d\text{Time}(t_{0,0}^{i,j}) \cdot (1 + \text{DetRate})). \tag{15}
\]

Only need to loop the trips and assign their time bounds accordingly (Algorithm 4). Finally, recall that \(\forall(k, l), \text{EST}(t_{0,0}^{i,j}) = \text{EST}(t_{k,1}^{i,j})\) and \(\text{LET}(t_{0,0}^{i,j}) = \text{LET}(t_{k,1}^{i,j})\). Thus, the time bounds for all trips have been established.
Algorithm 4 Setting time budget for synthetic trips

1: for each trip chain $i$ do  \(\triangleright\) TIME BOUND FOR TRIP CHAIN
2: \hspace{1em} if Origin splitter is TH then
3: \hspace{2em} EST$(t_{0,0}^{L_1}) \leftarrow$ Origin’s StartTime;
4: \hspace{2em} if Destination splitter is TH then
5: \hspace{3em} LET$(t_{0,0}^{L_{N(i)}}) \leftarrow$ Destination’s ArrTime;
6: \hspace{2em} else
7: \hspace{3em} LET$(t_{0,0}^{L_{N(i)}}) \leftarrow$ EST$(t_{0,0}^{L_1})$
8: \hspace{2em} \hspace{1em}$+ \maxDur(i) + \sum_{j \in \{1,...,N(i)-1\}} \ActDur(\Act^j_i)$;
9: \hspace{2em} end if
10: \hspace{2em} else
11: \hspace{3em} EST$(t_{0,0}^{L_1}) \leftarrow$ LET$(t_{0,0}^{L_{N(i)}}) - \maxDur(i)$
12: \hspace{2em} \hspace{1em}$- \sum_{j \in \{1,...,N(i)-1\}} \ActDur(\Act^j_i)$;
13: \hspace{2em} end if
14: \hspace{1em} end if  \(\triangleright\) TIME BOUND FOR TRIPS
15: \hspace{1em} calculate maxDet$(i)$ as in (14);
16: \hspace{1em} for each trip $t_{0,0}^{L_{i,j}}$ of a chain $i$ do
17: \hspace{2em} \hspace{1em} calculate maxDetTrip$(t_{0,0}^{L_{i,j}})$ as in (15);
18: \hspace{1em} \hspace{2em} end for
19: \hspace{1em} for each trip $t_{0,0}^{L_{i,j}}$ (except last) of a $i$ do
20: \hspace{2em} \hspace{1em} LET$(t_{0,0}^{L_{i,j}}) \leftarrow$ EST$(t_{0,0}^{L_{i,j}}) + \maxDetTrip(t_{0,0}^{L_{i,j}})$;
21: \hspace{2em} \hspace{2em} EST$(t_{0,0}^{L_{i,j+1}}) \leftarrow$ LET$(t_{0,0}^{L_{i,j}}) + \ActDur(\Act^j_i)$;
22: \hspace{2em} \hspace{2em} end for
23: \hspace{1em} end for

Preliminary POI retrieval

A feasible POI set is built by a two-step retrieval: first by a rough screening with the time bounds of the whole trip chain, and then by fine-tuning with each trip’s time constraints. The hierarchical design reduces duplicated retrieval time on infeasible regions. The fine-tuning process is implemented by M3 elaborated in the next section.

In Figure 13a, the outer black ellipse draws the rough STF set by the whole chain’s time budget (Eq. 14). As long as a POI corresponding to one of the activity types of this chain falls inside this region, it will be added to the POI pool for fine-tuning. All the displayed POIs in Figure 13a are in this pool. It therefore builds the candidate destination choice set $S_j^i$. The double-line circle and thick circle POIs represent $S_j^i$, while those in grey shades are POIs for other flexible activities.
Figure 13: Potential path area and alternative destinations of a trip chain and of the trips on the chain.

7.2.3 Alternative trip construction and matchup

With the initially selected POI pool \( S^i_j \), \( M_3 \) uses the time horizon of each trip \( t_{k_i}^{i_j} \) to build up a finer STF that adjusts the choice set for \( \text{Act}_j^i \). Given the current start point \( s_{j-1,k}^i \) of the trip \( t_{k_i}^{i_j} \), the finer STF (illustrated by the red ellipse in Figure 13a) selects all the feasible POIs (thick circle POIs) from its candidate set \( S^i_j \) to construct a feasible destination choice set \( \mathcal{F}_{j,k}^i \):

\[
\mathcal{F}_{j,k}^i = \left\{ s_{j,k}^i \in S^i_j \mid \text{dTTime}(t_{k_i}^{i_j}) \leq \text{LET}(t_{k_i}^{i_j}) - \text{EST}(t_{k_i}^{i_j}) \right\}
\]  

The meaning of this feasible set is illustrated by Figure 13b, labelled “solution space”. Two points \( s_{j-1,k}^i \) and \( s_{j,k}^i \) are connected by a link.
Correspondingly, the set of feasible POIs for Act$^1_{ij}$ is the union of POIs with in-degree (defined as the count of feasible trips travelling towards that POI) greater than 0 in this solution space.

However, geographically this is not true. As shown by the node with two arrows pointing inwards but nothing outwards, the time bounds of its previous and following stops are not satisfied. This is illustrated by the POI located outside the red ellipse and linked by the grey dash-line in “geography space”. The problem is caused by the computation design lacking of the bounding by the following trip. Theoretically, the PPA delineated by two points is an ellipse, but delineated by one and the same point it degrades to a circle (see the dash-dot-dot circle on “geography space”). Let a node with out-degree (defined as the count of feasible trips starting from this node) equal to 0 be referred to as a dead-end node. Then there are two ways to handle dead-end nodes that are not final stops. After building this alternative trip network (i.e., the network in “solution space”), it is possible to either: 1) remove each dead-end node that are not a final node and all corresponding nodes that are linked to the removed nodes (by tracing back); and 2) use linear programming to remove such non-final dead-end nodes by imposing “flow conservation” constraints (explained in the next section). The latter is the approach taken in this work.

The next function of $M_3$ is to find all the potentially matched pairs of trips. The matching process attempts to match a pair of trips from two different trip chains, and builds a candidate set if the two trips satisfy the corresponding time budgets. Note that this implies that trips belonging to two trip chains of the same person will not be matched, since the time windows of these trip chains do not overlap. Thus, the feasibility of a match depends on the time budget of each person. A match is feasible if the travel time on the combined route between both person’s pick-up and drop-off points is shorter than each person’s total travel budget for that corresponding trip.

Figure 14 shows an example of the combined route of person $P_1$’s trip from $Act_{j_1}$ to $Act_{j_1+1}$ and person $P_2$’s trip from $Act_{j_2}$ to $Act_{j_2+1}$. 
The bold black arrows represent the feasible travel route and visiting sequence, among the four possible visiting sequences to the four locations. Note that a person’s drop-off location cannot happen before the other’s pick-up; otherwise, there will be no ridesharing. Judging the space-time feasibility of a visiting sequence with PPA, $P_1(j_1) \rightarrow P_2(j_2) \rightarrow P_1(j_1 + 1) \rightarrow P_2(j_2 + 1)$ is the only feasible sequence. An infeasible one, for example, $P_2(j_2) \rightarrow P_1(j_1) \rightarrow P_1(j_1 + 1) \rightarrow P_2(j_2 + 1)$, has its middle point $P_1(j_1)$ outside the ellipse with foci at $P_2(j_2)$ and $P_1(j_1 + 1)$.

### 7.2.4 Finding the optimal solution

The last module ($M_4$), optimization, adopts binary linear programming (BIP) to calculate the maximum number of feasible trip matching that can be achieved.

In order to do this, each trip $t_{k,l}^{i,j}$ is assigned a binary variable $x_{k,l}^{i,j}$, where $x_{k,l}^{i,j} = 1$ indicates that trip $t_{k,l}^{i,j}$ is to be taken, and $x_{k,l}^{i,j} = 0$ indicates otherwise. Furthermore, each feasible matching of two trips is assigned a binary variable $u(t_{k,l}^{i,j}, t_{k',l'}^{i',j'})$, where 1 indicates that trips $t_{k,l}^{i,j}$ and $t_{k',l'}^{i',j'}$ are to be matched and 0 indicates otherwise. The trip binary variables $x_{k,l}^{i,j}$ and the matching binary variables $u(\cdot, \cdot)$ are grouped in the vector $X$ and $U$ respectively. The objective of the BIP is then to get the maximum sum of $u(\cdot,\cdot)$, which can be formulated as follows.

$$\max_U \mathbf{1}^T \mathbf{U},$$

where $\mathbf{1}$ is a vector of ones, subject to:

\begin{align*}
\sum_{\{i,k,1\} \mid t_{k,l}^{i,j} \in \mathcal{F}_{i,j}} x_{k,l}^{i,j} &= 1, \quad (18a) \\
\sum_{\{k \mid t_{k,b}^{i,j} \in \mathcal{F}_{i,j}^b\}} x_{k,b}^{i,j} - \sum_{\{l \mid t_{l,1}^{i,j+1} \in \mathcal{F}_{i,j+1,1}\}} x_{l,1}^{i,j+1} &= 0, \quad (18b) \\
u(x_{k_1,l_1}, x_{k_2,l_2}^{i_1,j_1}) &\leq x_{k_1,l_1}^{i_1,j_1}, \quad (18c) \\
u(x_{k_1,l_1'}^{i_1,j_1'}, x_{k_2,l_2}^{i_2,j_2}) &\leq x_{k_2,l_2}^{i_2,j_2}, \quad (18d) \\
\sum_{\{(i',j'),k',l'\} \mid u(t_{k,l}^{i,j}, t_{k',l'}^{i',j'}) \in U} &\leq 1. \quad (18e)
\end{align*}

Equation 18a ensures that there is only one trip traversed between two consecutive activities within a chain. Equation 18b implies that if a particular location is visited within a chain, the person also has to leave the location (the aforementioned “flow conservation” constraints). Similarly, Equations 18c and 18d ensure that the matching only applies to trips that are actually carried out. Finally, Equation 18e implies that a trip can be matched to only one other trip.
7.3 Simulations of Activity-Based Ridesharing in Yarra Ranges

The algorithm is implemented and tested by simulating ridesharing behaviours with real world spatial and travel demand data. The study area is the Shire of Yarra Ranges, covering eastern and north-eastern suburbs of Greater Melbourne, Victoria, Australia. Assumedly, self-driving vehicles as taxis serve the population with ridesharing flexibility. Details of the travel demand dataset are documented in Section 4.2.1.

7.3.1 Trip distribution and time horizon

Time and distance are two key factors in ridesharing. With the synthetic population and their travel demand as input, randomness is introduced to generate diverse trips so that no trips of two synthetic persons inherited from the same surveyed person will be exactly the same. Randomness is implemented in the following ways:

1) Each synthetic person’s origin and destination (coordinates) are randomly generated, subject to the same SA1 zone (Section 4.2.1) as documented of its surveyed person. The direct shortest travel time ($dT\text{Time}$) between origin and destination is calculated in network travel time. Different time slots of travel are induced in such way.

2) The randomized assignment guarantees time integrity that a trip must be assigned enough time to be finished. Time assignment is shown by Algorithm 5. Randomness is introduced by $\delta$, a shift from the documented variables (e.g., time, activity duration ($\text{ActDur}$)). When adding $\delta$ to the recorded ending time ($\text{VS}_\text{ArrTime}$) for the randomized arrival time ($\text{ArrTime}_\text{rd}$), the result is ensured to be at least $\text{ArrTime}_\text{cal}$ since this is the physically minimal travel time, and move along the time of the following trips. After the assignment of trip time, the simulation follows Algorithm 4 to set time budgets. DetRate is set as 30%.

7.3.2 Alternative destination choice set retrieval

The simulation adopts Yelp API to retrieve spatial locations of each type of activity. All the POIs within the study area’s geographic boundary are drawn by querying with exactly the same words of the activity types documented in VISTA data, and saved in a local database. The simulation then simply searches this database, and uses the STF built from each trip’s space-time budget to construct its destination choice set. Required information by the simulation includes: Retrieval keyword (activity type), latitude, and longitude of POI. Alternative trips are then built based on these POIs.
Algorithm 5 Time integrity adjustment for synthetic trips

1: for each synthetic person \( p \) do
2:   for each trip \( t \) of person \( p \) do
3:     if \( t \) is the first trip of the day then
4:       \( \text{ArrTime} \leftarrow \text{VS\_StartTime} + \delta + \text{dTTime} \);
5:     else
6:       \( \text{ArrTime}_{rd} \leftarrow \text{VS\_ArrTime} + \delta; \)
7:       \( \text{ArrTime}_{cal} \leftarrow \text{previous t’s ArrTime} + \text{previous t’s ActDur}; \)
8:       \( \text{ArrTime} \leftarrow \max(\text{ArrTime}_{rd}, \text{ArrTime}_{cal}); \)
9:       if \( \text{ArrTime}_{rd} < \text{ArrTime}_{cal} \) then
10:          Shift all the following trips’ time by \( (\text{ArrTime}_{cal} - \text{ArrTime}_{rd}); \)
11:     end if
12:   end if
13:   \( \text{ActDur} \leftarrow \text{VS\_DUR} + \delta; \)
14: end for
15: end for

7.3.3 Optimization with linear programming

MatLab (R2016a)\(^1\) is used to solve the BIP problem and get the final solution for 1:1 matching. The objective function value \( (fval) \), which represents the number of matched pairs of trips, is of special interest to this work. A higher \( fval \) means a higher matching rate. The computational burden on this part is heavy. Since the growth of population size leads to the increase of matching between trips in a super-linear manner, the matrix will expand drastically. This is the reason why a series of small samples are tested. The computational complexity is discussed in Section 7.5.

7.4 Systematical Tests and Results

The experiment starts by initializing a large synthetic population and its trips, and is run with a series of subsampled populations to conquer the computational burden. The initial population has 714 agents that make up of 1% of the synthetic population generated by Jain et al. [2017]. Despite losing the demographic composition, using subsampling rather than the full population relieves the computational burden in searching for a global optimum.

Of the 52 simulated activity types, 28 are location-flexible activities that provide alternative trip chances. In total, 4,922 POIs are retrieved of all queried trips in the study area. The initial 714 population induces 2,185 original trips and 3,269 alternative trips, summing up

\(^1\) https://www.mathworks.com/products/new_products/release2016a.html
Table 10: Statistics of the tested samples

<table>
<thead>
<tr>
<th>Population size</th>
<th>20</th>
<th>50</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. no. of original trips</td>
<td>58.9</td>
<td>150.8</td>
<td>287.2</td>
</tr>
<tr>
<td>Avg. no. of total trips</td>
<td>272.0</td>
<td>1223.9</td>
<td>1580.2</td>
</tr>
<tr>
<td>Mean. no. of matched pairs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alts</td>
<td>2.45</td>
<td>11.5</td>
<td>28.3</td>
</tr>
<tr>
<td>No Alts</td>
<td>2.4</td>
<td>10.15</td>
<td>26.35</td>
</tr>
<tr>
<td>p-value</td>
<td>0.3299</td>
<td>0.0001</td>
<td>0.0000</td>
</tr>
<tr>
<td>Std.Dev. no. of matched pairs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alts</td>
<td>1.39</td>
<td>3.17</td>
<td>4.86</td>
</tr>
<tr>
<td>No Alts</td>
<td>1.47</td>
<td>3.12</td>
<td>5.19</td>
</tr>
</tbody>
</table>

to 5,454 trips. Regardless of location flexibility, home is the dominant destination. Considering only spatially flexible destinations, the population targets supermarket most, followed by petrol station, fast food, shopping center, food store, newsagency and bookstore, and restaurant or café.

Of the initial 714 people, the location-flexible activities only take a small percentage of the total activities. Each activity is associated with an original trip. The population yields 410 (18.8% of 5,454) original trips with flexible destinations. However, the less than one fifth original trips are dramatically enriched by the associated alternative trips that are about eight times their amount. Top-targeted activities after involving alternative trips become petrol station, fast food, shopping center, restaurant or café, supermarket, and hardware. These activities are of special interest of activity-based ridesharing.

With sizes of 20, 50, and 100, each population size is sampled randomly 20 times from the initial synthetic population pool of 714 agents. The matching runs twice on each random sample, one considering alternative trips and the other original trips only. The frequency distribution of the number of matched pairs for each population size is shown in Figure 15. From top to bottom, the population sizes are 20, 50, and 100 for each graph. Each curve is based on the randomly drawn 20 samples of that population size, with different approaches of matching: considering alternative trips (red solid curve) vs. not (black dash curve). Therefore, the statistical test aims to substantiate that, for each population size, the mean value of the red curve (marked by the vertical solid reference line) is significantly higher than its counterpart (dash reference line). Table 10 shows the statistics and statistical significance. Alts means the result by considering alternative trips, while No Alts refers to original trips only. Although Figure 15 does not demonstrate a visually apparent variation between the two curves, the test yields a significant result according to the dependent t-test. Choosing the dependent t-test allows for the inter-
Figure 15: Distributions of number of matched pairs by population size. From top to bottom, the population sizes are 20, 50, and 100. Red solid curves are the results of considering alternative trips, while black dash ones are considering only original trips.
dependence that the results from the two runs are out of the same random sample.

It is encouraging to see the 50 and 100 population size cases yield a significant increase of matches by activity-based ridesharing. The bold \( p \)-values in Table 10 indicate high significance for population sizes 50 and 100. Though the smaller, 20 people cases do not pass the statistical significance test, no activity-based ridesharing test yields a lower amount of matches than its counterpart, which is a consistency expected by the model. In a majority of times, alternative destinations contribute to an increase of successful matches.

7.5 DISCUSSIONS

The experiment highlights some interesting but also challenging issues.

**Representative sampling.** The optimization part is not scalable to large population sizes. Therefore, smaller random samples are drawn. As seen in Table 10, the significance of the method depends on the population size. With larger population, more opportunities emerge due to a denser spatial distribution of and thus higher overlap of trips. Population size of 20 is generally too sparse in space (and time) to show an effect, in contrast to the samples of 50 and 100. Besides sample size, as shown in Table 11, the total number of feasible trips to be matched (\(|X|\)) is not directly correlated to the extent to which ABRA can increase the ridesharing rate (\(\text{Gap}\)). \(\text{Gap}\) is calculated as the difference between the counts of matched pairs by considering alternative trips (\(\text{Alts}\)) and not (\(\text{No Alts}\)). Nor correlated with \(\text{Gap}\) is the number of potential matches (\(|U|\), the number of feasible matches before BIP). The irregularity might be caused by the space and time sporadicity of trips with the lack of space and time overlap, which is partially induced by the small sample sizes. The heterogeneous distribution of alternative trips can be another reason: if only one person has many alternative trips, the overall matching rate is not necessarily increased. The sample size of 100 is relatively representative with trips dense enough and widely spread in space and time. \(\text{Gap}\) is foreseen to increase until trips get saturated.

**Scalability and efficiency.** The current model is set as a static baseline model to investigate the benefit of activity-based ridesharing. It therefore searches for a global optimum to approximate the overall potential of activity-based ridesharing. However, questing for a global optimum makes it difficult to scale up. Let \(N, a, t, d\) be the population size, the average number of activities per person, and the average numbers of alternative trips and destinations per activity. Let \(\beta\) be a contingent constant such that the amount of potentially matched trips \(|U| = \beta t^a N\). The pre-computation complexity for matching candidates is \(O(|X|^2) = O((t^a N)^2)\). The optimization matrix has a size of
Table 11: Statistics of trips and matching results:
Samples of size 50 (X and U are explained in Section 3.4)

| #Matched pairs | ||X|| | ||U|| |
|----------------|------|------|------|
| Gap | Alts | NoAlts | Alts | NoAlts | Alts | NoAlts |
| 4   | 12   | 8     | 2008 | 169    | 3917 | 105    |
| 4   | 13   | 9     | 1076 | 161    | 2026 | 203    |
| 3   | 16   | 13    | 1172 | 179    | 1300 | 119    |
| 3   | 12   | 9     | 1022 | 155    | 2869 | 59     |
| 2   | 15   | 13    | 1619 | 144    | 11052| 21     |
| 2   | 10   | 8     | 1921 | 139    | 441  | 17     |
| 1   | 8    | 7     | 1016 | 141    | 1845 | 123    |
| 1   | 14   | 13    | 227  | 137    | 42   | 41     |
| 1   | 20   | 19    | 1023 | 164    | 1170 | 138    |
| 1   | 11   | 10    | 1015 | 154    | 1929 | 194    |
| 1   | 7    | 6     | 1106 | 157    | 3111 | 33     |
| 1   | 7    | 6     | 1824 | 131    | 1190 | 45     |
| 1   | 11   | 10    | 999  | 140    | 1923 | 132    |
| 1   | 9    | 8     | 1130 | 165    | 999  | 46     |
| 1   | 8    | 7     | 1084 | 116    | 1033 | 34     |
| 0   | 12   | 12    | 1021 | 147    | 1891 | 136    |
| 0   | 11   | 11    | 155  | 149    | 146  | 146    |
| 0   | 12   | 12    | 1961 | 151    | 318  | 16     |
| 0   | 10   | 10    | 1058 | 140    | 166  | 12     |
| 0   | 12   | 12    | 1059 | 163    | 127  | 26     |
O(t^2N) + O(daN), which takes too much time for an applicable system for BIP that strictly requires integer solutions solved by branch-and-bound. Consequently, only small samples could be drawn to address the computational burden. The constraint matrix for population sizes 50 and 100 can grow to tens of thousands rows by that many columns. With the initial full population size of 714, the constraint matrix jumps up to million by million.

**Returning to the research question.** Even with the limitations of scalability and sample sizes, the experiment substantiates the hypothesis that ABRA can significantly increase the successful matching rate compared with the traditional trip-based method. As aforesaid, the consistency is meaningful that ABRA is stably capable of increasing successful matches. With samples of population size 50 (Table 11), as many as four more pairs of trips can be matched. To the best case (the 1st entry), the matching rate is increased by 50% with ABRA. It therefore highlights the effectiveness of the proposed algorithm.

### 7.6 Conclusions and Future Work

This work proposes activity-based ridesharing as a novel method of ride-matching that aims to enlarge the chance of matching compared to trip-based methods. In activity-based ridesharing, people can lodge a request for a ride from an activity A to an activity B, rather than from location A to location B. The algorithm develops a space-time filter to construct the choice set of approachable destinations by extracting the POIs of the requested activity. This space-time filter is capable of handling multiple consecutive flexible activities, which is advantageous over simple space-time prisms. The experiments clearly prove the capability of activity-based ridesharing to increase successful matching rates. This outcome is trustworthy as the simulations are set in a real-world context. The implementation also demonstrates the correctness of the proposed (exact) solution of the global optimization problem.

However, it has also become clear that scalability is a serious challenge, for which a dynamic agent-based ridesharing model that employs real-time heuristics and accommodates human behavior heterogeneity is suggested as a future research direction: A dynamic system for ridesharing applications suits realistic scenarios better since people usually lodge a travel request on the fly. It can construct a space-time filter in a real-time manner, searching for nearby resources to quickly build a choice set. The candidate ride partner is consequently a local optimum approached by a decentralized decision process, which requires an agent-based model. Additionally, the agent-based model could accommodate heterogeneous human behaviors by developing heuristics, such as utilizing user ratings to filter out some POIs, or tailoring matches to the travel habits and visiting history of each person. Another interesting direction can involve the role
of social network as heuristics in activity-based ridesharing. Social network not only implicitly bundles people’s physical behaviors (e.g., Toole et al. [2015]; Wang et al. [2015b]), which affects the detour cost and chance of getting a ride, but also latently decides the preference to choose ride partners Chaube et al. [2010].

Another future work is the semantic accuracy. If the activity types of demand and supply data sets may not match exactly, the search of POIs by different activity types can actually be too narrow or too inclusive. In the experiment in this paper, activities documented in VISTA have not matched exactly with activities in the Yelp database. For example, fast food, food store, restaurants and supermarket are listed as different categories in VISTA, but are in one group in the Yelp database. Improving the matching quality will be an interesting topic for geographical semantics.
This paper is adapted from the manuscript of the same title submitted to the *Journal of Transport Geography* for peer review. I conducted all parts of the work, including research design and implementation, result analysis, discussions, and paper-writing. My supervisor Prof Stephan Winter was supervising the progress, contributing to research ideas, and actively discussing the results. My co-supervisor Dr Martin Tomko brought up critical thoughts and comments at the stage of experiment design, suggesting literatures to review, and discussing the results.

### 8.1 Major Contributions

A new ridesharing model called *collaborative activity-based ridesharing* (CAR) is proposed to enhance not only overall matching rates but also the matches between preferred ride partners. By coalescing the merits of two recently suggested innovative ridesharing models – the *social network-based ridesharing* [Wang and Winter, 2017] and the *activity-based ridesharing* [Wang et al., 2016b] – the new model leverages people’s preference to their social networks and the space-time flexibility of daily activities to improve the matching outcome. The capabilities and advantages of the proposed model are justified by a group of agent-based simulations in a realistic study area. The influence of geography on the match outcome is discussed in particular.

The hypothesis is that CAR can significantly increase the overall matching rate compared with social network-based ridesharing, and significantly increase the number of matches with friends compared with activity-based ridesharing.

Based on realistic travel demand data, an agent-based simulation for ridesharing pre-planning of a day is built to implement the CAR model. The simulation is run with the pre-generated social networks of small world topology embedded into space. Two spatial structures are investigated in the simulation: random distributions and distance-decays. Multiple methods were run for comparison: trip-based ridesharing, social network-based ridesharing, activity-based ridesharing, and CAR. With different geographic configurations of the underlying social network, results from each simulation are compared to investigate which algorithm comes out as the best in terms of detour cost, the numbers of overall matches and of matched friends. A special focus of the discussion is on the geographic characteristics
of the study area, of the population distribution, and of the social network distribution. The findings yield implications on how geography affects the performance of CAR.

8.2 THE COLLABORATIVE ACTIVITY-BASED RIDESHARING (CAR) MODEL

The workflow of CAR is elaborated in Section 8.2.1, of which the core part is a spatio-social index based on time geography theory (Section 8.2.2). The influence of spatial distributions of social networks and travel demands is discussed in Section 8.2.3.

8.2.1 The workflow of CAR

The workflow of CAR is presented in Figure 16, developed from Figure 1 in Wang et al. [2016b]. The major modification is the introduction of the social network file and the spatio-social index that affects the matching procedure in module 3 (M3, see the bullet point below). This workflow is applicable to a full day pre-planning model for ridesharing rather than a real-time one. The model is incorporating activity-based ridesharing [Wang et al., 2016b] and social network-based ridesharing [Wang et al., 2017]. The activity-based ridesharing model searches for alternative destinations of a similar travel purpose (activity at the destination) satisfying a person’s budget on each travel to enlarge trip choices. Alternative destinations are represented by Points of Interest (POIs) in the study area for each type of travel purpose. For instance, rather than simply going to a specific supermarket A, this algorithm might find feasible travels to any supermarket within 10 minutes’ drive from a current location and 20 minutes from the next fixed destination after the grocery shopping. It hence expands ridesharing choices. The activity-based approach is encoded in CAR with social network-based ridesharing, which renders higher detour tolerance to closer social contacts (e.g., a direct friend) while lower detour tolerance to a less familiar person (say, a stranger). CAR proceeds in the steps shown below:

- Module 1 (M1): At the beginning of each day, the model collects the whole population’s full day travel activity intentions and the social network structure permuted for the population.
- Module 2 (M2): The model divides each person’s full day schedule into trip chains, i.e., series of trips with short stop time at the destinations [Wang et al., 2016b; Thill and Thomas, 1987]. A maximum time budget is assigned to each trip, which is used to retrieve the POIs within that budget.
- Module 3 (M3): CAR fine-tunes the alternative destinations from a certain category of POIs, builds alternative trips between activities, and matches trips subject to the space-time budgets to
Figure 16: Workflow of the CAR model.
every partner in the same car. Different from the activity-based ridesharing that equivalently matches any pair, CAR renders priority to closer friends as the social network-based ridesharing does. The spatio-social index (Section 8.2.2) is introduced as a filter to efficiently decide: 1) the set of accessible POIs given a type of activity and a type of social network degree; 2) the feasible trips to be matched given the spatial and social constraints.

- Module 4 (M4): Input the potential matching pre-computation matrix into binary integer programming to decide the final matching. The optimisation maximises the number of matches subject to space-time budget and detour tolerance depending on the social network.

8.2.2 Time geography and the Spatio-Social Index

CAR introduces a spatio-social index (SSI) inspired by the computational entities in time geography. A space-time prism is a geometric entity delineating the boundary of an approachable area given the space-time budget (e.g., Miller [2005]). A potential path area (PPA) is the projection of the prism onto a 2-dimensional plane [Miller, 2005]. The prism itself encoding time is called potential path space (PPS). SSI can be regarded as a PPA at a coarser spatial resolution for fast retrieval. Taking advantage of the heuristic from social networks, SSI speeds up the search for: 1) feasible POIs to build alternative travels, and 2) potential trips to be matched satisfying social-network-dependent time budgets.

Figure 17: The SSI and the 8-connected n-neighbourhood.

SSI is a gridlock index that partitions the whole area into grid cells (Figure 17), where each cell is a square with edge length $g$. The geographic location $(x, y)$ of a point $p$ is projected to its corresponding grid $g_i$ at row $r$ and column $c$, as long as $p$ falls inside that cell. The operation to find the index of a point is denoted as Equation 19. A
POI has a unique gi, while a trip has one each for its origin and destination. When querying a feasible potential trip, at least one of the origin or the destination should fall inside the approachable grids.

\[ I[p] = gi \quad (19) \]

**SINGLE-CONSTRAINT QUERY** Given a time budget (δ in geographic space and n in the gridlock space, where \( \lceil \delta \rceil = n \times g \)) as the total time for shortest travel and detour, SSI retrieves all the 8-connected n-neighbourhood cells of the centroid query point (Figure 17), if there is only one constraint (e.g., time and location constraints at the origin or the destination). These cells are utilised to query points fall within them. The operation of searching 8-connected δ-neighbourhood is defined by Equation 20.

\[ I_δ(p) = I(p) \bigoplus n \quad (20) \]

**DOUBLE-CONSTRAINT QUERY** SSI works in a derived way from the single-constraint query, searching the accessible range between two fixed points p₁ and p₂ (with fixed space-time constraints), for instance, leaving workplace no earlier than 5pm and getting back home no later than 6:30pm. SSI generates a δ-neighbourhood for both constraint points, and returns the cells in the overlapped area \( I_δ(p₁) \cap I_δ(p₂) \), ignoring the time dimension. In the case of the constraints by one trip’s origin and destination, the δ is the same for both ends given a particular trip. In Figure 18, the prisms ΔOO₁O₂ and ΔD₁D₂ are two identical PPS’s by the single constraints from the
origin and the destination, respectively. The PPA is the intersection of the projections of prisms \( \triangle OO_1O_2 \) and \( \triangle DD_1D_2 \), i.e., \( \text{PPA} = O_1O_2 \cap D_1D_2 = I_\delta(O) \cap I_\delta(D) \). The red rectangle, which is PPA times the time duration, is the bounding box of the accurate double-constraint PPS (the green prism \( OO_2DD_2 \)).

**Deciding the time budget** \( \delta \)  
SSI decides \( \delta \) depending on the degree of separation \( e \) in a social network so that a higher detour tolerance \( d = T(e) \) (in percentage of the shortest travel cost) is granted to topologically closer friends. \( T(e) \) is inversely correlated to \( e \), which is specified later in Equation 22. Particularly, direct friends (\( e = 1 \)) get the highest detour tolerance, followed by indirect friends (\( e = 2 \)) and finally strangers (\( e = \infty \)). The time budget is determined by the detour tolerance and the shortest travel time \( t_s \):

\[
\delta = (d + 1) \cdot t_s = (T(e) + 1) \cdot t_s, \quad e \in \{1, 2, \infty\}.
\]

Because \( \delta \) is the maximum accessible range from the current point, both single-constraint and double-constraint queries must return the grids covering the upper bound of travels. Such design guarantees that the grid index is unlikely to cut off possible selections. The returned grids do not compose an accurate PPA since the ridesharing matching will continue to refine the selections (POIs or trips) within those cells, but is sufficient to cut down the computational burden.

### 8.2.3 The influence of spatial distribution

Wang et al. [2017] discussed the varying results in social network-based ridesharing from different spatial settings of the social networks. The parameters describing social networks are the topological factor (the average degree of connections, denoted by \( \bar{n} \)) and the spatial factor (spatially aggregated or random distributions of nodes). In their observation, the likelihood to share rides is not monotonic with the density of rides, which means increasing the spatial density of rides may even result in a lower matching rate due to competition within a certain distance range. On the opposite, low density of friendships might put the efforts searching for friends in vain. The spatial density of friendships therefore is correlated with the success of ridesharing with social contacts.

The spatial configuration of social networks obviously affects the density of friends in space. Ridesharing social networks emerge from historical travel experience, which in reality may not yet be known. Thus, ridesharing social network has to be set for investigating the behaviour of these models. In this work, the classical small world network is applied as the topological structure of each social network. The spatial configuration is set on top of the small world structure in two ways, a randomly distributed network and a spatially aggregated
network (following the distance decay function Eq. 24). The clustering of any generated social network is likely to be in favour of matching friends. In addition, the topological structure contributes to spatial density of friends as well. Higher node degrees (more friends) on average lead to higher density of friendships in space, given the same study area. To compare the difference in ridesharing outcomes from varied spatial distributions of social networks, this model runs four combinations of friendship degrees and spatial configuration (Section 8.3).

8.3 THE COMPARISON OF DIFFERENT MODELS AND EMPIRICAL STUDIES

To testify the advantage of CAR over alternatives, a group of control experiments with different ridesharing models are run based on a realistic travel survey dataset in Yarra Ranges, Victoria. Composed of several suburbs located east and northeast to Melbourne, Yarra Ranges has a heterogeneous landscape with hills and rivers, and thus highly clustered spatial distribution of population. The area is selected due to its lack of convenient public transportation, especially for the last-mile problem.

The travel demand and activities are adopted from the processed datasets by Jain et al. [2017] based on the Victorian Integrated Survey of Travel and Activity (VISTA) 2009-2010 [Victorian Department of Transport, 2011]. Details are given in Section 4.2. The SSI grid is set with a unit of 1 minute’s drive for the simplicity of calculation.

There are four ridesharing models investigated and compared. All the models are implemented by Repast Simphony. Table 12 lists the characteristics each model considers. Social network-based ridesharing [Wang et al., 2017] and activity-based ridesharing [Wang et al., 2016b] are abbreviated by SNeRs and ABRA, respectively. Trip-based and ABRA models are run with the lowest and the highest detour tolerances ($d = 10\%, 30\%$) each time as baselines. SNeRs and CAR, in contrast, adopt varied detour tolerances and ridesharing willingness as a function of the degree of separation $e$ in a social network (Eq. 22). Previous studies [Wessels, 2009; Chaube et al., 2010] suggest significantly varied detour tolerances and travel willingness with different social contacts. In accordance with SNeRs, the detour tolerance is set per Equation 22 and willingness per Equation 23.

$$T(e) = \begin{cases} 
30\%, & e = 1 \\
25\%, & e = 2 \\
7\%, & e = \infty 
\end{cases} \quad (22)$$
Table 12: Catalogue of models and parameter settings

<table>
<thead>
<tr>
<th>Notation</th>
<th>Model</th>
<th>$\bar{n}$</th>
<th>Spa</th>
<th>Detour tolerance &amp; willingness</th>
</tr>
</thead>
<tbody>
<tr>
<td>t10</td>
<td>Trip-based</td>
<td>–</td>
<td>–</td>
<td>10%, 100%</td>
</tr>
<tr>
<td>t30</td>
<td>Trip-based</td>
<td>–</td>
<td>–</td>
<td>30%, 100%</td>
</tr>
<tr>
<td>ab10</td>
<td>ABRA</td>
<td>–</td>
<td>–</td>
<td>10%, 100%</td>
</tr>
<tr>
<td>ab30</td>
<td>ABRA</td>
<td>–</td>
<td>–</td>
<td>30%, 100%</td>
</tr>
<tr>
<td>sw6</td>
<td>SNeRs</td>
<td>6</td>
<td>random</td>
<td></td>
</tr>
<tr>
<td>sw6d</td>
<td>SNeRs</td>
<td>6</td>
<td>decay</td>
<td></td>
</tr>
<tr>
<td>sw12</td>
<td>SNeRs</td>
<td>12</td>
<td>random</td>
<td></td>
</tr>
<tr>
<td>sw12d</td>
<td>SNeRs</td>
<td>12</td>
<td>decay</td>
<td></td>
</tr>
<tr>
<td>car6</td>
<td>CAR</td>
<td>6</td>
<td>random</td>
<td></td>
</tr>
<tr>
<td>car6d</td>
<td>CAR</td>
<td>6</td>
<td>decay</td>
<td></td>
</tr>
<tr>
<td>car12</td>
<td>CAR</td>
<td>12</td>
<td>random</td>
<td></td>
</tr>
<tr>
<td>car12d</td>
<td>CAR</td>
<td>12</td>
<td>decay</td>
<td></td>
</tr>
</tbody>
</table>

$W(e) = \begin{cases} 
100\%, & e = 1 \\
80\%, & e = 2 \\
10\%, & e = \infty 
\end{cases}$ \hspace{1cm} (23)

These parameter values are representative of the social heterogeneity in ridesharing, but are not unique values and can vary from region to region.

Two parameters are considered in a social network: the average degree of friends $\bar{n}$ (note that this should be distinguished from the degree of separation $e$) and the spatial distribution (Spa) of the social network. The topological structures of the social networks in this study obey small world structures. The average degree $\bar{n}$ in a social network is set to be 6 and 12, which values are not universal but varying from case to case (e.g., Wang et al. [2015b]; Shi et al. [2016]; Schläpfer et al. [2014]). Whatever the particular value of $\bar{n}$, a social network generally yields a low density network [Faust, 2006]. These networks are then embedded in space with random spatial distribution or with distance decay in space. Let the home distance between each two persons be $l_{ij}$. The distance decay function, particular to the study area, is formulated by Equation 24. This is a piecewise function such that 1) the social ties of the middle-range distances between 500 metres and 10,000 metres follow a power law distribution $f(l_{ij}) = \alpha l_{ij}^{-\beta}$ (where $\beta = 2.0$ precisely for this case, and $\alpha$ obeys the normalisation specified later); while 2) the probabilities of very short (<500 metres) and very long (>10,000 metres) distance social ties are moderated. The whole range is divided into bins with an interval of 500
metres. The probability of a social tie of a distance is the cumulative probability of the corresponding bin. In Equation 24, \( F(l_{ij}) \) is the cumulative probability of \( f(l_{ij}) \) at each bin \([\lfloor \frac{l_{ij}}{500} \rfloor \cdot 500, (\lfloor \frac{l_{ij}}{500} \rfloor + 1) \cdot 500)\), where \( \alpha \) satisfies that \( \int_{500}^{10000} f(l_{ij}) \, dl_{ij} = 60\% \). The breakpoints 500 and 10,000 are decided due to the characteristics of the population distribution. Especially, the average distance between residential spatial clusters caused by natural barriers is about 10km.

\[
p(l_{ij}) = \begin{cases} 
20\%, & 0 \leq l_{ij} < 500 \\
F(l_{ij}) = \int_{\lfloor \frac{l_{ij}}{500} \rfloor \cdot 500}^{(\lfloor \frac{l_{ij}}{500} \rfloor + 1) \cdot 500} f(l_{ij}) \, dl_{ij}, & 500 \leq l_{ij} < 10000 \\
20\%, & l_{ij} \geq 10000 
\end{cases} \tag{24}
\]

The comparison between spatial and non-spatial networks contribute to understanding the interwoven spatio-social influence on ridesharing. Higher degree indicates higher social network density in both topological and spatial senses. Each CAR or SNeRs model for each different social network structure is randomly run 5 times with a randomly permuted social network of that structure, while trip-based and ABRA models with each detour tolerance are run only once since social networks are not considered in matching. However, the statistics in Table 15 for trip-based and ABRA models pertaining to social network structures (e.g., number of matched friends) are still calculated as the mean of 5 permutations for each social network structure.

Output variables to be investigated are shown in Table 13. The number of matches and the detour cost can be contradictory at times. However, given the detour tolerance, it matters more to investigate the number of matches.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>( |U| )</td>
<td>Size of the pre-computation matrix before optimisation</td>
</tr>
<tr>
<td>( n )</td>
<td>Number of matches after optimisation</td>
</tr>
<tr>
<td>( n_f )</td>
<td>Number of matches between direct or indirect friends</td>
</tr>
<tr>
<td>( n_{f1} )</td>
<td>Number of matches between direct friends</td>
</tr>
<tr>
<td>( s )</td>
<td>Detour cost, overall</td>
</tr>
<tr>
<td>( s_{f1} )</td>
<td>Detour cost, direct friends</td>
</tr>
<tr>
<td>( s_{f2} )</td>
<td>Detour cost, indirect friends</td>
</tr>
<tr>
<td>( s_{f\infty} )</td>
<td>Detour cost, strangers</td>
</tr>
</tbody>
</table>
Table 14: Statistics of the direct friendship links in each social network
(Link length in metres)

<table>
<thead>
<tr>
<th></th>
<th>SN6</th>
<th>SN6d</th>
<th>SN12</th>
<th>SN12d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td>2142</td>
<td>2212</td>
<td>4284</td>
<td>4271</td>
</tr>
<tr>
<td>Min. len.</td>
<td>46.31</td>
<td>28.43</td>
<td>30.12</td>
<td>23.68</td>
</tr>
<tr>
<td>Max. len.</td>
<td>41842.63</td>
<td>35160.04</td>
<td>37156.79</td>
<td>36243.97</td>
</tr>
<tr>
<td>Mean len.</td>
<td>15690.15</td>
<td>12018.69</td>
<td>15409.15</td>
<td>13270.13</td>
</tr>
</tbody>
</table>

8.4 RESULTS

As shown in Table 12, there are four social network structures under study, each being permuted five times. Table 14 displays the link statistics of the four social networks (average of five permutations): SN6 is the small world social network with the average degree of 6 and random distribution in space, and SN6d is its counterpart with distance decay in space (i.e., links are spatially clustered). The same holds for SN12 and SN12d of an average degree of 12. The link length is calculated as the Euclidean distance in metres between the home locations (the origins of each person’s first trips) of a pair. Given a certain population, higher friendship density reduces the average home distance between friends, denoting a higher potential of ridesharing. Distance decay by design is associated with shorter friendship distance than random distribution. Figure 19 displays the spatial distribution of the population and their friendship links. Though not visually discernible, the link density is higher in SN12 and SN12d (the bottom two sub-figures), and links are shorter in cases with distance decay (the right-hand sub-figures).

The calculation results are demonstrated in Table 15, where each value is the mean of simulation outputs from five permutations of each social network structure. The significance is tested with t-test on a 0.05 significance level. ∥U∥ is the size of the pre-computation matrix before optimisation, i.e., the number of all possible matches. n, n_{f}, and n_{f1} are the final number of matches, the number of matches between first and second degree friends, and that between direct friends only. The last four columns are the detour costs (as in percentage of each rider’s direct shortest travel time) of the matched rides, respectively, between anyone, direct friends, indirect friends, and strangers. Results from the same social network structure are highlighted in the same colour for comparison between different models.

For the number of matches, these are the major findings:

- Considering alternative destinations consistently yield higher number of matches (n), no matter whether social networks are taken into account or not (ABRA vs. trip-based, and CAR vs. SNeRs).
Figure 19: The spatial distribution of the population and their friendships.
Table 15: Simulation outputs by each model
(Bold numbers: significantly higher than its counterpart with random social network.)

<table>
<thead>
<tr>
<th>Model</th>
<th>$|U|$</th>
<th>n</th>
<th>$n_f$</th>
<th>$n_{f_1}$</th>
<th>s</th>
<th>$s_{f_1}$</th>
<th>$s_{f_2}$</th>
<th>$s_{f_{\infty}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>t10_6</td>
<td>227</td>
<td>119</td>
<td>0</td>
<td>0</td>
<td>0.0653</td>
<td>0</td>
<td>0</td>
<td>0.0653</td>
</tr>
<tr>
<td>t10_6d</td>
<td>1854</td>
<td>385</td>
<td>0</td>
<td>0</td>
<td>0.2089</td>
<td>0</td>
<td>0</td>
<td>0.2089</td>
</tr>
<tr>
<td>t10_12</td>
<td>1854</td>
<td>141</td>
<td>0</td>
<td>0</td>
<td>0.0666</td>
<td>0</td>
<td>0</td>
<td>0.0666</td>
</tr>
<tr>
<td>ab10_6</td>
<td>49452</td>
<td>409</td>
<td>0</td>
<td>0</td>
<td>0.2120</td>
<td>0</td>
<td>0</td>
<td>0.2120</td>
</tr>
<tr>
<td>ab10_6d</td>
<td>491.6</td>
<td>135.4</td>
<td>24.6</td>
<td>5.6</td>
<td>0.0946</td>
<td>0.1851</td>
<td>0.1943</td>
<td>0.0721</td>
</tr>
<tr>
<td>ab10_12</td>
<td>506.0</td>
<td>135.8</td>
<td>27.4</td>
<td>11.0</td>
<td>0.1000</td>
<td>0.2309</td>
<td>0.1946</td>
<td>0.0730</td>
</tr>
<tr>
<td>ab10_12d</td>
<td>634.2</td>
<td>164.0</td>
<td>75.2</td>
<td>10.0</td>
<td>0.1272</td>
<td>0.1702</td>
<td>0.1897</td>
<td>0.0761</td>
</tr>
<tr>
<td>ab30_6</td>
<td>656.2</td>
<td>165.4</td>
<td>84.8</td>
<td>18.0</td>
<td>0.1411</td>
<td>0.2347</td>
<td>0.1988</td>
<td>0.0721</td>
</tr>
<tr>
<td>ab30_6d</td>
<td>1429.2</td>
<td>198.0</td>
<td>36.8</td>
<td>9.2</td>
<td>0.0974</td>
<td>0.1984</td>
<td>0.2051</td>
<td>0.0730</td>
</tr>
<tr>
<td>ab30_12</td>
<td>13721.8</td>
<td>202.8</td>
<td>47.4</td>
<td>18.0</td>
<td>0.1051</td>
<td>0.2373</td>
<td>0.1926</td>
<td>0.0727</td>
</tr>
<tr>
<td>sw6</td>
<td>19471.0</td>
<td>233.6</td>
<td>108.6</td>
<td>12.4</td>
<td>0.1292</td>
<td>0.1869</td>
<td>0.1993</td>
<td>0.0690</td>
</tr>
<tr>
<td>sw6d</td>
<td>19003.4</td>
<td>239.4</td>
<td>122.2</td>
<td>22.4</td>
<td>0.1362</td>
<td>0.2154</td>
<td>0.1947</td>
<td>0.0721</td>
</tr>
<tr>
<td>sw12d</td>
<td>19003.4</td>
<td>239.4</td>
<td>122.2</td>
<td>22.4</td>
<td>0.1362</td>
<td>0.2154</td>
<td>0.1947</td>
<td>0.0721</td>
</tr>
<tr>
<td>sw12</td>
<td>19003.4</td>
<td>239.4</td>
<td>122.2</td>
<td>22.4</td>
<td>0.1362</td>
<td>0.2154</td>
<td>0.1947</td>
<td>0.0721</td>
</tr>
</tbody>
</table>
• Prioritising social networks in an algorithm (CAR vs. ABRA, and SNeRs vs. trip-based) can significantly increase the matches between friends (n_f and n_{f1}), given the same social network. As a baseline, the trip-based model and activity-based model yield no matches between friends by random chance.

• Considering alternative destinations further contributes to a significant surge of matches between friends compared with social network-based model (CAR vs. SNeRs).

• The number of matches between friends generally rises with the decrease of home distance between friends (from SN6 to SN12d). The bold numbers in Table 15 display the significant difference in the number of matched friends between corresponding spatially aggregated and random social networks, even though the difference in overall matching rate is not significant.

Regarding detour cost, the findings are mainly:

• Given the same social network structure, there is no significant difference in detour cost between the corresponding results of CAR and of SNeRs (rows in the same colour), since a detour is capped by its tolerance limit.

• Comparing the costs of direct and indirect friends (s_{f1} vs. s_{f2}), social networks with spatial aggregation constantly yield higher detour cost between direct friends while random social networks result in higher cost between indirect friends.

• The overall detour cost increases as social networks get denser from SN6 to SN12d, but all within detour tolerances.

The regularity of the pre-computation size ||U|| causes the high computational burden in the models considering alternative destinations (ABRA and CAR) compared with their counterparts (trip-based and SNeRs). In sacrifice of computation efficiency, however, the models encompassing alternative destinations lead to higher numbers not only of overall matched rides but also of rides between friends.

8.5 DISCUSSIONS

Based on the observations shown in Section 9.5, the proposed collaborative activity-based ridesharing demonstrates the following advantages and disadvantages that give hints on the feasibility of the model and its future improvements.

The hypothesis that CAR can significantly increase the overall matching rate compared with social network-based ridesharing is proved by the simulation results. Since the capability of activity-based ridesharing to increase matching rates (n) has been justified [Wang et al., 2016b], alternative destinations extend the destination choice set and thus the matching rate.

The most important advantage of CAR is its reinforcement of favoured matches, even compared with social network-based ridesharing (SNeRs).
Despite a higher overall matching rate, it is unsure to what extent activity-based ridesharing can match friends by random chance. In fact, the simulation results (ab10 and ab30) show no stochastic match between friends in any social network structure. What makes CAR remarkable is that it facilitates a significant increase in certain matches – the matches between friends – over random chances. The matching rate between friends by CAR is shown significantly higher than not only activity-based ridesharing, but also social network-based ridesharing, which is a stronger statement than the hypothesis. CAR therefore utilises alternative choices to match friends’ rides that would not be matched if the original routes were not substituted by the alternative ones.

Since detour cost is controlled by detour tolerance, there is no significant difference in detour cost between the outputs from CAR and social network-based ridesharing. The positive indication is that CAR effectively reinforces the increase of favoured matches at no cost of extra detour.

A characteristic of the study area (Fig. 19) is its heterogeneity in landscape (hills and rivers) and thus the population distribution. The highly clustered and uneven population, in contrast to the urban areas under study by Tachet et al. [2017] who assumed travel demands are evenly distributed, might lead to a significantly different result even with trip-based ridesharing. Embracing social network can lead to a more distinguishable difference in matching rate. With social network-based ridesharing, trials to share rides with a friend residing in another cluster might fail due to the infeasibility of space-time budgets. Despite CAR searches for alternative destinations subject to space-time budgets, the simulation results manifest the influence of spatially aggregated friendships, contributing to significantly more matches between friends (bold numbers in Table 15), though the overall matching rate is not significantly higher. At least the spatial aggregation of friendships denotes a higher satisfaction rate measured by the likelihood of ridesharing with friends. The regularity between spatial distributions of population and friendship and ridesharing outcome is still debatable, which calls for more empirical studies in diverse areas in terms of such factors as landscapes, population densities, friendship distributions, and urban configurations.

In spite of an increased matching rate, the downside of activity-based ridesharing is its heavy computational burden, as reported by Wang et al. [2016b]. The significantly expanded matrix size $\|U\|$ has already indicated the heavier computational burden in the searching and matching stage (M3 in Figure 16). SSI leverages the heuristic of social networks to effectively cut down the searching space of unfavoured potential matches from strangers. The sacrifice of the total number of matches by the design of the spatio-social index (SSI) is disputable, depending on the spatial distribution of social networks.
and trip density: When trips are dense in space, SSI is capable of narrowing down unnecessary searches by limiting the search range for strangers to a small area. But if trips are spatially sparse, especially when social networks are scarce as well, SSI may deteriorate ridesharing opportunities. For this reason, further investigation in the influence of space is needed with empirical data from diverse study areas. The design of a dynamic SSI is worthwhile for a real-time ridesharing application.

8.6 CONCLUSIONS AND FUTURE WORK

Coalescing two previously suggested innovative ridesharing algorithms – social network-based ridesharing and activity-based ridesharing – this study proposes a new ridesharing algorithm called collaborative activity-based ridesharing (CAR) to address the barriers of trust and flexibility in ridesharing. Agent-based simulation results based on empirical dataset substantiate the capacity of CAR to increase favourable rides without sacrificing more detour time, which potentially encourages public acceptance of ridesharing.

Since geographic studies are contingent, future work should involve more study areas to generalise the regularity between spatial configuration and ridesharing outcomes. Alternatively, the algorithms can be encapsulated into a tool to be applied in any locality. Harvesting coupled information of social networks and travel behaviours is necessary but difficult due to confidentiality. If possible to collect such coupled information, the feasibility of CAR could be measured more accurately and practically. Another direction of future work is the design of more efficient search indices for rides-matching considering spatial contexts. The upgrade of indices may also contribute to the dynamic modification of CAR. A dynamic CAR model would be more applicable in reality but is confronted with issues of computational efficiency and swift response. A possible solution based on search indices is a space-time partitioning of the problem for parallel computing.
INTRODUCING RIDE-SHARING INTO PARKING PROBLEMS

This chapter is revised from the paper of the same title submitted to AGILE’18 as an achievement from the Australia-Germany Joint Research Cooperation project with Leibniz Universität Hannover. My role in this paper includes initiating the topic, designing the conceptual model and the experiment, doing literature review, analysing the results, and drawing conclusions. My German colleagues were involved in designing the model; they implemented the simulation, and contributed to literature review, analysing and discussing the results. My supervisor Stephan Winter was helping coordinating the research process, consolidating the research idea, as well as result analyses and discussions.

9.1 MAJOR CONTRIBUTIONS

High demand for parking exists in many urban centres, contributing to traffic congestion and driver frustration. Various approaches to relieve parking pressure are implemented already in practice, which, despite the positive influence, has not yet eliminated the fundamental challenge. This paper presents a significant contribution – a park-and-ridesharing approach – that applies ridesharing only for the last mile in return for managed parking. Although generally ridesharing fails to gain popularity, in this regard it makes the most advantage of the comfort and flexibility of private cars.

Hypothetically, P&S (with private vehicles) has a number of advantages as good as or better than P&R (with scheduled public transport). P&S is self-regulating based on actual travel demand to the urban centre, which only takes place when the traffic gets crowded. P&S is temporally more flexible than P&R, benefiting from ad-hoc matching. P&S shares the economic benefits of P&R with regard to accessing cheaper parking spaces, reducing congestion charges, and reducing inner-city traffic. Finally, ridesharing comes with similar travel duration to driving individually.

P&S, as a two-sided matchmaking problem [Evans and Schmalensee, 2016], has to raise incentives for both parties: people who are willing to park at a satellite parking and give up their independence and convenience of their own car, and people who are willing to pick up travellers at a satellite parking for a lift to or from the city centre. The incentive for the former category, the passengers, can be a cheaper parking solution together with an acceptable predicted total travel
time. The incentive for the latter, the *drivers*, can include a guaranteed parking space in the centre, shared driving costs and/or lower parking fee in the centre. This study will exploit the incentives and investigate whether P&S provides any improvement to the parking issue at all, and in what scenarios.

To justify the advantage of ridesharing in parking problems over public transit and solo driving, this chapter looks at two conceptual scenarios for comparison:

1. The baseline scenario of solo-driving plus P&R: This baseline is the balance of people picking P&R and taking their cars. The balance exists where any additional traveller switching to P&R would not reduce their individual travel costs.

2. The alternative scenario of baseline plus P&S: If P&S is considered, the aforementioned flexibility compared to P&R kicks in, and the average travel time should go down further.

Results from the agent-based simulations show that the relevant ratios of travellers would use such a ridesharing system. In addition, park-and-ridesharing significantly reduces the waiting and searching time for parking in high travel demand scenarios, and thus potentially yields a high contribution to reduce parking burden in urban centres.

### 9.2 Background of Parking-and-Ridesharing

Parking in cities is a multidimensional problem requiring integrated approaches. Amongst the parking management strategies where many are collaboratively influential, mentioned by Litman [2006] as three strategies are remote parking, mobility management (e.g., change of transportation mode), and parking pricing that may contribute to 10% - 30% parking demand reduction. Introducing ridesharing into a parking system pertains to both remote parking at a satellite parking lot and mobility management in the last leg to the final destination. Some studies discussed the decision making between cruising for on-street parking and paying for off-street parking by saving time [Shoup, 2006; Van Ommeren et al., 2012]. Basically, the decision is a trade-off between time saving and money saving.

The uncertainty to predict cruising time was mentioned by Shoup [2006] as a factor adding to the complexity. However, new technologies are promising to provide a solution. A recent survey on smart parking [Lin et al., 2017] pointed out the importance of real-time information on parking space that facilitates people’s decision making. In addition to static sensors, crowdsensing can also be used to obtain real-time parking information [Bock and Di Martino, 2017]. Dynamic dissemination of parking information helps to reduce the competition for parking spaces [Chai et al., 2017]. Thus, this study considers
a smartphone application informing travellers about the parking situation in the present work.

Another aspect information technologies contribute to parking is for parking allocation and booking. Smart parking system, such as [Yan et al., 2008], allows drivers to reserve parking spot in real time. Smart parking is also integrated with pricing, for example, by associating the price of the same parking spot to different vehicles [Ayala et al., 2012]. The former idea is adopted in this model that ridesharing vehicles can lock in a parking spot in advance to save cruising time, while the latter is indirectly applied by assigning different prices to solo drivers’ vehicles and to ridesharing vehicles.

Ridesharing is understood here as a semi-private transport mode that combines the travels of a driver (in contrast to the commercial ridesourcing) with passengers in the same vehicle [Handke and Jonuschat, 2013]. Incentives for private car drivers can be various, from non-monetary (e.g., permission to use less congested high-occupancy lanes) to shared cost models (e.g., fuel or parking costs). There are many challenges on matching potential rides, including dynamic real-time matching and pricing, as well as identification of optimal meeting points [Furuhata et al., 2013; Agatz et al., 2012; Czioska et al., 2017]. To focus on the benefits of ridesharing in a parking system rather than a ridesharing method itself, a simplest ridesharing strategy is adopted in this study that drivers only pick up passengers who are queuing at a parking lot.

9.3 CONCEPTUAL DESIGN OF EXPERIMENT

The model is designed in a conceptual artificial world to test the hypothesis. This section explains the scenario and parameters of the model.

9.3.1 Model context

The simulation scenario consists of a city centre with a central parking lot and an inbound road with a satellite parking lot reasonably close to the city centre (see Fig. 20). People travel to the city centre, stay for a specific time, and leave the city back to the satellite parking. The satellite parking lot has a capacity sufficient to satisfy the demand. It allows people to park and look for a ride, either by a private driver passing by, or by public transit. The central parking lot with a capacity $N_c$ can be used by ridesharing drivers as well as solo drivers. However, if drivers need to wait for a free parking space in the central parking lot, ridesharing drivers are always preferred over solo drivers. Ridesharing drivers pay a reduced parking fee at the central parking. The satellite parking fee is either very cheap or free of charge.
Figure 20: The conceptual scenario of the P&S system, in contrast to solo-driving and public transit (referring to Table 16).
The traffic flow volume, the duration of stay in the city centre, and the searching time for an empty space in the central parking are given, which formulas are specified in Section 9.4. People at the satellite parking are informed of current estimated waiting time for a ride as well as for the average waiting and searching time for parking at the central parking lot to support their decision making. In practice, this information can be provided by an electronic sign or a smartphone application.

The model assumes also a public transit system between satellite and central parking lot as an alternative to driving to the central parking on one’s own. The public transit system guarantees that people have rides inbound and outbound in case ridesharing falls through. The efficiency of the public transit system is a parameter that affects travellers’ preference to ridesharing, which will be investigated by the model. The model also studies scenarios with no ridesharing available (which is the baseline in the hypothesis).

People who decide to park at the satellite parking lot have a choice between taking an offered ride or using scheduled public transit. In the model, ridesharing drivers do not wait for passengers; they only pick up passengers who are waiting, for which reason the matching between drivers and passengers is ignored.

At the satellite parking lot, each person chooses to be one of the four types of users:

1. solo driver with private car,
2. passenger taking public transit,
3. ridesharing driver with private car,
4. ridesharing passenger.

Each type of users is associated with a cost function \( C \) specified below. In the baseline scenario where ridesharing is not available, travellers can only choose between being a solo driver or using public transportation to get to the city centre.

Corresponding to the context specified above, the inputs and outputs of the model are documented in Table 16. Some of the parameters will be utilized as cost parameters to be elaborated in the following section.

The output of the system, total travel cost \( C_{T} \), includes three parts: the total time duration for travel converted into monetary unit, the monetary cost of the travel, and the intrinsic willingness cost to ridesharing converted into monetary unit. The total travel time is calculated from making the mode choice at the satellite parking lot to getting back to the car at the satellite parking lot, excluding the time for activities at the city centre. This time includes waiting time for rides or for public transit, travel time for the last leg, and search time for a parking space in the central parking lot (as a solo-driver or as a ridesharing driver). The monetary costs are the cost split for ridesharing, the cost for solo driving, and the fee for public transit. The willingness
Table 16: Inputs and outputs of the model

<table>
<thead>
<tr>
<th>Type</th>
<th>Parameter name</th>
<th>Parameter type</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>F(t)</td>
<td>Function</td>
<td>Traffic flow distribution with the time of the day (t)</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>Function</td>
<td>Distribution of the average parking duration</td>
</tr>
<tr>
<td></td>
<td>f_c, f_r, f_s</td>
<td>Constant</td>
<td>Parking fees for central parking; solo-driving and ridesharing parking, and parking fee for satellite parking</td>
</tr>
<tr>
<td></td>
<td>N_s, N_c</td>
<td>Constant</td>
<td>Maximum capacity at satellite and central parking lots</td>
</tr>
<tr>
<td></td>
<td>n_s, n_c</td>
<td>Variable</td>
<td>Current free capacity at satellite and central parking lots</td>
</tr>
<tr>
<td></td>
<td>f_pt</td>
<td>Constant</td>
<td>Public transit travel fee</td>
</tr>
<tr>
<td></td>
<td>f_d</td>
<td>Constant</td>
<td>Fees for driving either to the city centre or back (gas, car usage, congestion charge)</td>
</tr>
<tr>
<td></td>
<td>S_rs, S_pt</td>
<td>Constant</td>
<td>Seat capacity in ridesharing vehicle and public transit</td>
</tr>
<tr>
<td></td>
<td>w</td>
<td>Variable</td>
<td>Intrinsic willingness to do ridesharing (varied from person to person)</td>
</tr>
<tr>
<td></td>
<td>q_pt</td>
<td>Constant</td>
<td>Public transit frequency</td>
</tr>
<tr>
<td></td>
<td>d_pt</td>
<td>Constant</td>
<td>Public transit delay in comparison to driving</td>
</tr>
<tr>
<td></td>
<td>V(·)</td>
<td>Function</td>
<td>Value of time: individual per person; Necessary to add temporal and monetary costs</td>
</tr>
<tr>
<td>Output</td>
<td>c_c</td>
<td>-</td>
<td>Total travel cost including travel time, monetary cost, and the intrinsic willingness cost to ridesharing</td>
</tr>
<tr>
<td></td>
<td>n%</td>
<td>-</td>
<td>The ratio of each type of traveller</td>
</tr>
</tbody>
</table>
cost approximates the inconvenience of ridesharing converted to monetary cost, such as psychological uneasiness or time and money spent for cleaning, which only occurs to ridesharing participants.

9.3.2 Ridesharing strategy

Ridesharing in this simulation is organized in such a way that it introduces no extra waiting cost for ridesharing drivers at the satellite parking lot. Each driver arriving at the satellite parking decides whether (s)he will park (being a ridesharing or public transport passenger) or continue driving (being a ridesharing or solo driver). The decision is made with regard to the estimated costs of being one of the four types of users. The drivers who choose to share rides will join the ridesharing population at the satellite parking lot. They pick up ridesharing passengers, if there is a queue at the satellite parking lot, but do not wait for passengers. If the driver did not pick up any passenger, the driver recalculates the estimated costs for each travel mode and chooses the second best option next to being a ridesharing driver. A passenger (ridesharing or public transport) waiting for being picked up is constantly recalculating the cost for each travel mode and chooses the best option for him/her. This means (s)he may switch the mode and become another type of user. In this sense, the matching process of drivers and passengers does not apply in the present model.

9.3.3 Decision-making on the user type

Travellers’ decisions on which travel mode to take are based on the cost of each mode. The costs are calculated with the following functions assessing the (negative) utility of a particular decision takes monetary costs as well as monetized waiting time into account. An agent always chooses the travel mode with the minimal cost. The costs inbound and outbound are different.

Cost function for the way to the city centre

When deciding the travel mode to the city centre, travellers also consider the costs for the way back. Note that the driving time is omitted in all the functions because only the gap between different modes is considered. As a bonus for ridesharing, the parking fees for ridesharing vehicles $f_r$ at the central parking lot can be lower than the solo driver parking fee.

\[
\{ f_s \ll f_r \leq f_c \}
\]

The cost functions for being each type of user are:
Solo driver:
\[ C_S(\hat{t}_{cs}, f_c, f_d) = V(\hat{t}_{cs}) + f_c + 2 \cdot f_d \] (25)
\[ \hat{t}_{cs} \] is the estimated time waiting and searching for a parking space in central parking for solo drivers, measured based on the parking occupancy rate. Note that the estimates do not predict the time on arrival.

Since the cost function takes time as well as money into account, the value of time function \( V(\cdot) \) is needed to convert the time to monetary costs. \( f_c \) is the central parking fee for solo drivers. \( f_d \) is the one-way driving cost (including petrol, car usage, and congestion charge), which is multiplied by 2 to incorporate the driving cost outbound.

Public transit passenger:
\[ C_T(\hat{t}_{pt}, f_{pt}, f_s, R_p) = V(\hat{t}_{pt}) + f_{pt} + f_s + R_p \] (26)
\[ \hat{t}_{pt} \] is the estimated additional travel time caused by public transit on the way to the central parking, including the additional travel time and waiting time in comparison to traveling by car. \( f_{pt} \) is the public transit fee and \( f_s \) is the satellite parking fee. \( R_p = f_{pt} + V(\frac{q_{pt}}{2} + d_{pt}) \) is the estimated cost of returning from the city centre with public transit, where \( q_{pt} \) is the public transit interval and \( d_{pt} \) the delay of the public transit compared to the cars’ travel time. Note that public transit passengers might also use ridesharing on the way back, if this option has the lower costs.

Ridesharing driver:
\[ C_D(w, \hat{t}_{cr}, f_r, f_d, f_s) = w + V(\hat{t}_{cr}) + \frac{f_r + f_d + 2 \cdot f_s}{3} + \frac{f_d}{3} \] (27)
w is the intrinsic willingness cost for ridesharing. \( \hat{t}_{cr} \) is the estimated time waiting and searching for a parking space in central parking for ridesharing drivers, measured based on the time of other drivers. \( f_d \) and \( f_s \) are the same as above. In ridesharing, the monetary costs are split among the driver and passengers. Thus, the costs are divided by three, which is assumed to be the average number of persons in a ridesharing vehicle. The third term of the cost function refers to the monetary costs to the city centre and the fourth term to the monetary costs back to the satellite parking.

Ridesharing passenger:
\[ C_P(w, \hat{t}_{cr}, f_r, f_d, f_s, R_p) = w + V(\hat{t}_{cr}) + \frac{f_r + f_d + 2 \cdot f_s}{3} + R_p \] (28)
\[ \hat{t}_{cr} \] is the estimated time waiting for a ride to the central parking based on the time of other ridesharing passengers. \( w, f_r, f_d, f_s \), and \( R_p \) are the same as above. For the way back, there is a risk of no ride. Thus, the return cost is pessimistically modelled with the public transit return costs \( R_p \).
Cost function on the way back

On the way back, passengers share again the costs with the drivers. The ridesharing drivers have to pick up passengers, if they are waiting in the queue. Solo drivers remain to be solo drivers on their way back.

Ridesharing passenger:

\[ C_{PB}(w, \hat{t}_{sr}, f_d) = w + V(\hat{t}_{sr}) + \frac{f_d}{3} \]  

(29)

\( \hat{t}_{sr} \) is the estimated time waiting for a ride to the satellite parking. \( f_d \) is the same as above.

Public transit passenger:

\[ C_{TB}(\hat{t}_{pt}, f_{pt}) = V(\hat{t}_{pt}) + f_{pt} \]  

(30)

\( \hat{t}_{pt} \) is the estimated travel delay caused by public transit on the way to the satellite parking including the additional travel time and waiting time in comparison to travelling by car. \( f_{pt} \) is the public transit fee.

9.4 Model Implementation

To test and evaluate the benefits of P&S, the model described in Chapter 9.3 has been implemented\(^1\) using NetLogo. This section describes the parameters of the travellers’ value of time, their ridesharing willingness, and their visiting duration in the city centre. In addition, the model assumptions on factors such as traffic flow over the day, the estimation of waiting time, and the search duration in the parking lots are explained. For all parameters, the default settings are defined, based on which one parameter is varied at one alternative setting for subsequent sensitivity analysis.

9.4.1 Value of time

The value of time describes the individual monetary equivalent of a time unit for every agent. The model assumes that the value of time is normally distributed over all agents. The value of time is assumed non-negative, i.e., travel time is considered as a cost for all individuals. Thus, every negative value of the normal distribution is mapped to zero. The choice of the mean value of $25 per hour travel time is estimated based on Lam and Small [2001]. They found in a route choice

---

\(^1\) The implementation (by my German colleague Christian Koetsier) is publicly available under https://gitlab.com/Chr1ko/RidesharingSimulation
experiment in California in 1999 that the average value of travel time was $22.87 per hour.

9.4.2 Willingness for ridesharing

This parameter describes the monetary equivalent for which a traveller accepts to participate in ridesharing. The willingness is also specific for every person. It is assumed that this value is usually positive, meaning that the travellers only participate if there is a compensation. The distribution is modelled over the population as a normal distribution, but negative values are allowed in this case. This means that there are also travellers who accept additional costs (e.g., detours, additional car cleaning) when they offer a ride or become a ridesharing passenger. For example, they might enjoy the conversations during the ride, or drivers are just benevolent. The default value for the mean willingness is set to $5 with a standard deviation of $5.

9.4.3 Traffic flow rate

The simulation imitates a typical traffic flow coming into the city centre, with highest travel demand in the morning and the afternoon hours. The traffic inflow is modelled by the weighted sum of three Gaussian distributions. The parameters are chosen in such a way that the shape resembles the measurements by Hueper et al. [2009]: $w_1 = 0.6$, $\mu_1 = 8$, $\sigma_1 = 1$, $w_2 = 3$, $\mu_2 = 17$, $\sigma_2 = 3$, $w_3 = 4$, $\mu_3 = 12$, $\sigma_3 = 5$, where $w$ is a weighting, $\mu$ is the mean, and $\sigma$ is the standard deviation of the Gaussian distribution. The resulting distribution is visualized in Figure 21.
9.4.4 Information of travellers on estimated waiting time

When travellers arrive at the satellite parking lot, they are provided with information on estimated waiting time to facilitate the decision on which travel mode to take: ridesharing driver, ridesharing passenger, public transit passenger, or solo driver. The expected waiting time is calculated for ridesharing passengers waiting at the parking lots and for drivers waiting and searching for a parking space in the city centre. The estimates are based on a weighted average of the experienced waiting times $t_w$ of other travellers. Each time the waiting time is observed by a traveller, the mean waiting time $\bar{t}_w$ is updated first:

$$\bar{t}_{w,\text{new}} = (1 - w_u) \cdot \bar{t}_{w,\text{old}} + w_u \cdot t_w$$  

(31)

The parameter $w_u$ is the weight for a new measurement (0.5 in the simulation). If multiple ridesharing passengers enter a car, $t_w$ is the mean of their waiting time. When there is no recent waiting time measurement within a certain period (10 minutes), the mean waiting time is recalculated with an assumed measurement of $t_w = 0$ minutes. That way, the mean waiting time decreases for every time period without a waiting time measurement by a traveller to consider the case when nobody is waiting.

To avoid a decreasing waiting time estimation when all persons are stuck in the queue, the mean waiting time is compared to the individual (still increasing) waiting time $t_{w,i}$ of all persons $i$ waiting in the queue. The maximum of these values is used as the final waiting time estimation.

$$\hat{t}_w = \max(\bar{t}_w, t_{w,i})$$  

(32)

Note that there is always a time lapse between the decision made at the satellite parking space (which is based on the current waiting time) and the arrival at the central parking space. Hence the estimated waiting time may significantly change while driving from satellite to central parking.

9.4.5 Parking search in parking lot

With limited space in the parking lot in the city centre, the search time for an empty parking space in a parking lot depends on the occupancy level. This search time is modelled based on the formula by Axhausen et al. [1994], but capped by a maximal search time in this case.

$$t_s = \min\left(\frac{t_{s,\text{min}}}{1 - \frac{N_{\text{occ}}}{N_{\text{tot}}}}, t_{s,\text{max}}\right)$$  

(33)

In the simulation, the minimal search time $t_{s,\text{min}}$ is set to 1 minute, and the maximal search time $t_{s,\text{max}}$ is set to 5 minutes. $N_{\text{occ}}$ is the number of occupied parking spaces and $N_{\text{tot}}$ is the total number of
parking spaces in the parking lot. The function is visualized for these parameter settings in Figure 22.

9.4.6 Duration of stay in city centre

The parking duration can be modelled as a Gamma distribution\(^2\). The distribution parameters were chosen for a mean parking duration of 1.5 hours \((\alpha = 3, \beta = 2)\). For example, the average parking durations in some commercial areas in San Francisco are between 0.7 and 2.4 hours, if residential parking permits are not considered\(^3\). The distribution for the chosen parameters is visualized in Figure 23.

9.4.7 Modeling of public transit

The public transit is modelled as a direct connection, but with slower speed given the usual stops of public transit along the ride. Public transit is scheduled to operate at a regular interval (30 minutes in the default setting) throughout the day with a fix capacity (50 persons to simulate a bus). The travel time of public transit is set to be 15 minutes longer than for driving with the car to the city centre. Public transit is assumed to be always on time.

\[^3\] http://www.sfcta.org/sites/default/files/content/Planning/ParkingManagementStudy/pdfs/parking_study_final.pdf
Figure 23: Gamma distribution describing the parking duration at the central parking.

9.4.8 Choice of the rest parameters

The remaining parameters are set by default in the following way. The parking fees are assumed to be $10 for solo drivers and $5 for ridesharing drivers, respectively. Parking in the satellite parking is free. The fee for public transit is modelled to be $3, which is equivalent to the shared costs in ridesharing with two passengers in the default setting. The private cars are assumed to have a capacity for up to three passengers and one-way driving costs (petrol, car usage, and congestion charge) of $4. For the default setting, the capacity in the central parking lot is chosen in such a way that the waiting time in front of the parking lot is not longer than about 15 minutes even in the peak hours for the case without a ridesharing option (baseline scenario). The travel time between satellite parking and central parking is 30 minutes for cars.

9.5 Results

The advantages and disadvantages of introducing ridesharing into parking systems are evaluated by mainly two types of output variables under different scenarios: the total cost of travel and the ratio of each type of travellers. The total cost includes all monetary and time costs for the trips to the city centre and back. In addition, the intrinsic willingness for ridesharing is added to the costs for ridesharing participants. For 100 random simulation runs with the same parameter setting, the mean of the output variables are computed shown as one data point on each plot, and the error bars represent the standard deviation of these runs. The scenarios with and without ridesharing (the alternative and baseline scenarios in Section 9.1) are initially eval-
uated with the default parameter settings (Section 9.5.1). Then, the sensitivity of the parameters is investigated by varying one parameter at a time and evaluating the changes in Subsection 9.5.2. Finally, the ratio of traveller types is evaluated for a modified cost function of the ridesharing passengers on the way back to the satellite parking (Section 9.5.3).

9.5.1 Evaluation for default parameters

Figure 24 shows, given default parameter settings, the changes of average total costs and average travel time with the number of travellers in the simulation of the alternative and baseline scenarios, i.e., with and without ridesharing. The total travel costs and the travel time in the ridesharing scenarios are always significantly lower than in the scenarios without ridesharing (comparing the blue and red lines in solid line or dash line accordingly), regardless of the number of travellers. When the number of agents increases in the system, i.e., when the city gets more crowded and drivers start to wait in a queue in front of the central parking lot, the total cost of travel without ridesharing yields a significant surge from 5000 travellers, while the cost of ridesharing is relatively stable and even slightly decreases. The increase of the total costs results from the increasing travel time for large numbers of travellers.

Figure 24: The comparison of the average total travel costs (including the time costs and willingness for ridesharing participants) and the average travel time between the scenarios with versus without ridesharing.

Figure 25 displays the variation of the ratios of different types of travellers with the hour of day, with the empty parking capacity...
shown as the green dot line. In general, the ratios of the travel modes are quite stable over time with about two-thirds being solo drivers, about 30% ridesharing participants, and only a few public transit users. When the central parking lot gets more occupied (between 9 and 20), the ratio of solo drivers reduces and the ratio of ridesharing passengers increases by about 5%.

![Graph showing changes in parking spaces and traveler ratios](image)

Figure 25: Change in the ratio of different types of travellers and the number of free parking spaces at the central parking with the hour of day.

### 9.5.2 Evaluation for parameter variations

The participation rate for the different travel modes are illustrated in Figure 26 varying with different numbers of travellers for the scenarios with ridesharing. With increasing number of travellers, the ratio of solo drivers decreases and the ratios of ridesharing drivers and passengers increase. The ratio of transit passengers generally remains on the same level. Comparing the ways to and from city centre, it is observed that the ratio of ridesharing drivers is lower on the way back from city centre, while the ratio of solo drivers displays the opposite trend.

Besides the number of travellers, Figure 27 indicates that the higher the value of time, the higher the ratio of solo drivers, cumulated for both the way to the city centre and back. There are fewer ridesharing participants and public transit passengers as value of time increases.

The ratios of different types of travellers with regard to three factors of public transit demonstrate similar trends (Fig. 28). With the increase of public transit delay, of public transit fee, and of public transit frequency, the ratio of solo drivers increases, but the ratios
of ridesharing participants and public transit passengers drop down significantly. Thus, overall, lower public transit performance leads to higher ratio of solo drivers and lower ridesharing and public transit participation rates.

### 9.5.3 Evaluation of modified cost function for ridesharing passengers

It is a pessimistic thinking of ridesharing to assume that the worst scenario, i.e., taking public transit, is to happen on the way back to the satellite parking (see Section 9.3.3). Thus, an optimistic scenario is also evaluated for the return part of the cost function for ridesharing passengers. In the cost function, the public transit costs $R_p$ are replaced by the ridesharing passenger costs $C_{PB}$ with an estimated waiting time of 10 minutes (an estimation of the waiting time sev-
Figure 28: The changes in ratio of types of travellers affected by different factors of public transit. The default values are used for all other parameters.
eral hours ahead would be too error-prone). The result in Figure 29 shows that there are many more ridesharing participants compared to the default setting in Figure 26 (noticing the y-scale is different). The average waiting time turns out to be mostly lower than 10 minutes. Evaluations of the public transit parameters (delay, fee, interval) show that, in this optimistic scenario, the ratio of ridesharing participants is mostly independent of the public transit quality.

Figure 29: The changes in ratio of different types of travellers with the total number of travellers to and from city centre (assuming a waiting time of 10 minutes for a ride on the way back).

9.6 DISCUSSIONS

The performance of the park-and-ridesharing system relies on traveller’s calculation of the cost functions for each travel mode, considering time, money, and intrinsic willingness to ridesharing. Testing different scenarios, this study finds that P&S can significantly contribute to urban traffic mainly in the following three aspects.

Strongly supported by Figure 24, involving ridesharing into parking systems overall is beneficial to reduce both the average travel time and the average total travel cost, which includes monetary cost, travel time, and, in the case of ridesharing, the intrinsic willingness for ridesharing. Monetary costs are reduced, because ridesharing splits costs among people. However, travellers with a high value of time and/or a high cost value of ridesharing willingness incline to be solo drivers. As Figure 26 shows, the motivation for ridesharing increases with higher volume of travellers to avoid long waiting and searching time for the next available parking spot. This takes place when the central parking gets more crowded approaching the capacity limit of the central parking. Some of the ridesharing drivers become solo drivers on the way back since no passenger waits for pickup. In addition, the introduction of a park-and-ridesharing system leads to a better resilience of the traffic system. A well-running park-and-ridesharing system moderates the usage of the central parking: when the parking lot is relatively empty, more drivers travel alone; while
when it is busy, ridesharing saves the time waiting for an empty space and avoids a collapse of the complete system. It is, however, interesting that higher values of time lead to lower rate of ridesharing (Fig. 27). With increasing mean value of time, more travellers have such a high value of time that they avoid waiting for a ride from someone else or using public transit. That way, if there are fewer ridesharing passengers, there are also fewer chances to become a ridesharing driver and more people end up solo driving.

Another finding is the subtle competition-supplementation relationships between public transit and ridesharing for the default settings. Intuitively, public transit is thought as a competitor with ridesharing. This is, nevertheless, only partially true. Figure 28 shows that a more efficient (i.e., more frequent, less delayed, and cheaper) public transport system leads to a higher number of ridesharing participants. This is because public transport is a backup choice for ridesharing in case that ridesharing passengers fail to find a ride on the way back. For decision-making, each person assumes that, if being a ridesharing passenger, the worst scenario on the way back to the satellite parking is to take public transit, which increases his/her total cost assumption. The more efficient the public transport system, the less the increase of total cost. In this sense, public transit is supplementary to ridesharing and supports the attractiveness of ridesharing. However, when the cost function of the ridesharing passengers is modified to an optimistic scenario (Section 9.5.3), the competition-supplementation relationship between public transit and ridesharing is not observed. In this sense, the attitude and risk management towards ridesharing plays an important role.

Last but not least, designing an appropriate strategy to attract both ridesharing passengers and drivers is crucial. The design of lower fees and priority to ridesharing vehicles at the central parking is an important strategy to attract ridesharing drivers. In this situation, many people have the highest priority to ridesharing driver, because of not only low travel time but also lower costs. However, this is only possible if the volume of ridesharing passengers is sufficiently high. If being a ridesharing passenger is too infeasible (e.g. high value of time and long waiting time), the whole P&S system does not work.

9.7 Conclusions and Future Work

This paper looks at the benefits of a park-and-ridesharing system to reduce the parking demand in the city centre. The results show that a relevant ratio of travellers would choose the ridesharing system and that the system significantly reduces the travellers’ total costs (including travel time costs and costs for ridesharing willingness) averaged over all travellers. Most importantly, the P&S moderates the usage of the central parking, avoiding long waiting queues in front of the
parking lot in the case of high travel demands. In addition, the participation rate of ridesharing strongly depends on the assumption of the ridesharing passengers for the waiting time on the way back. If they are confident to get a ride within 10 minutes, the ratio of ridesharing participants strongly increases compared to the pessimistic assumption that they would need to take the public transit back.

The public transit system is rarely used as the experiment results demonstrate, but is still very important as a backup for ridesharing passengers on their way back to the satellite parking. Thus, for future investigation, it would be interesting to evaluate the potential of an on-demand taxi service bringing ridesharing passengers back to the satellite parking, when they wait too long for a ride. This solution could be a promising alternative to the public transit system. It is necessary to evaluate whether the costs for such a service would be lower than for the public transit system. Moreover, based on the current scenario, a more realistic setup should be evaluated for the benefits of the park-and-ridesharing system.
This part concludes the thesis. It includes the chapter of high-level discussions based on the experimental results, and projects into the future the potential trend of ridesharing and technical challenges. A general conclusion of the thesis along with the major contributions is drawn at the end.
DISCUSSIONS AND OUTLOOK

Looking into the future of the Collaborative Activity-based Ridesharing initiated in this thesis, there are still many problems unsolved. This chapter discusses and indicates the potential challenges and future directions for research.

Ground truths

The first challenge is the ground truth validation of the models. Though VISTA data was deployed as the foundation of the simulations, the sample size of the population that is applicable in the thesis, i.e., the surveys from car drivers, is not yet large enough, which is why synthetic population was generated [Jain et al., 2017]. Besides, this is only a one-day survey on a zonal level. Compared with smartphone-based data collection, the sample is quite limited. Realistic data on social networks falls short more. The accessibility is very low to social networks connected in physical life (that are approximate enough for collaborative travel) rather than online virtual life. It is even more difficult to harvest social network associated movement behaviour data due to the challenges in survey design [Axhausen, 2008] and its high sensitivity to privacy.

Additionally, since spatial configuration is a crucial factor deciding the outcome of ridesharing, a comparison across different study areas will be meaningful. In this thesis, the experiment results are proved in a heterogeneous landscape where population is clustered in space. Different patterns, however, might emerge in a busy city centre where population density is high everywhere. The urban form, road network topology, and many other spatial properties can contribute to the difference in ridesharing results. The regulation can only be detected with sufficient ground truths.

Autonomous vehicles

The proposed models are transferable into autonomous vehicle (AV) scenarios. As demonstrated in Chapters 6, 7, and 8, the taxi-sharing scenarios are mathematically identical to AV scenarios. That said, an emerging issue is the necessity of trust in an AV sharing system due to the absence of a third party (used to be the driver). The gender issue has already been a problem for a single-rider-single-driver scenario¹, even though taxi drivers are normally trusted by people be-

cause of their licence and the security screening. In an AV scenario, the concern is likely to be higher without the in-person witness of a trusted third party (as the driver used to be). Even though CCTV can observe the whole ride, psychologically passengers might not feel as comfortable as when there is a human observer. Additionally, a CCTV camera will be observing passengers all the time, which makes people feel awkward, while a human observer can use other sensations such as hearing to avoid looking at the passengers all the way. The issue of the absence of a third party witness as a human also highlights the value of Chapter 5 which leverages social network for higher level of trust in ridesharing.

Another interesting research question is to model the dispatch locations of those AVs, and the demand for the fleet. The service area covered by each vehicle is omitted in this thesis, but is worth studying in real logistic planning.

Additionally, whether Vehicle Kilometre Travelled (VKT) is increased or decreased still stays uncertain. While some argue that theoretically parking is no longer necessary in the era of AV, the practical operation of AV fleet has to take into account aspects related to VKT such as energy saving and financial benefit. Prediction of travel demand and simulation based scenario planning hence become crucial topics.

**Dynamic ridesharing and the algorithms**

Another issue yet to be addressed is to build a dynamic ridesharing system. As aforementioned, this thesis mainly focuses on proposing innovative factors for ridesharing systems and testing the full potential of the system. A dynamic system requires time-dependent matching strategies, which might lead to bias depending on the choice of time window. To get a full picture of the system free from such impact, only static ridesharing models are built. From the baseline, further dynamic models can be proposed.

Looking forward, one challenge in dynamic ridesharing is the computational efficiency. When scale is growing, strategies such as parallel computing (e.g., Attanasio et al. [2004]), short-range communication (e.g., Winter and Nittel [2006]), and spatio-temporal index/partitioning (e.g., Ma et al. [2013]; Nourinejad and Roorda [2016]) should be utilised. The problem exists at how to make a reasonable spatio-temporal partitioning strategy to decompose the problem depending on the density of demands, which is also brought up by previous studies [Winter and Nittel, 2006; Agatz et al., 2012]. Some areas are less sensitive to the partition of geographic space because the demand is dense everywhere, while it is likely that low demand areas can be hindered by an unreasonable partition. The shape of the partition/index might also play a role in addition to its size. This is drawn to the question to what scale / level ridesharing matching should be decentralised, if
not finding a global optima, so that the algorithm can optimise and scale up at the same time. Perhaps there should be dynamic partitioning heuristics that decide how to partition the geographic space based on the real-time observed traffic load.

Integration of social and cognitive factors into mobility spaces

The last and maybe most challenging aspect is to come up with more systematic models that build social and cognitive factors into mobility behaviours. Mobility is understood as the moving of people, ideas, and things with regard to humanities [Cresswell, 2011], which differs from transport geography that emphasises the physical infrastructure and constraints. Although many studies have applied the term mobility, social and cognitive factors that drive the behaviours are in fact rarely modelled to the extent such that the influence can be clearly understood except as a correlated factor [Andris, 2016; Sarkar et al., 2016].

There are many directions further studies should look into. For example, the cognitive distortion of mobility space caused by social networks is presented in a specific case by the social network based ridesharing (Chapter 5). However, there might be a more universal model to quantify the distortion [Wang and Winter, 2017]. For another instance, the attraction of travelling depends on the type of activities performed at the destination [Axhausen, 2008]. But the attraction is ambiguous to be modelled or quantified. One step further, the current artificial intelligence (AI), and particularly, GeoAI, is more to apply machine learning or deep learning to explore pattern. In terms of decision-making, AI still has a very long way to go to understand human decision-making. The improvement of GeoAI cannot be accomplished unless the social and cognitive factors can be modelled in mobility spaces.
This thesis proposes three innovative ridesharing algorithms that aim to accommodate the issues of trust and flexibility in people’s ridesharing behaviours. Ridesharing is a particular example in which multiple persons need to collaborate in their decision-making on spatial behaviours. To address the collaborative decision-making, a revised framework called Collaborative Activity Space, based on the theory of time geography, is proposed (Chapter 3) that integrates the subjective preferences (in comparison with objective physical constraints in time geography) and space-time flexibility with regard to travel purposes (i.e., the semantics of a place).

Nested under the theoretical framework, the highlights of this thesis are three ridesharing algorithms and a ridesharing application listed below.

1. The Social Network based Ridesharing (Chapter 5 and Chapter 6) satisfies people’s preference to familiar acquaintances when choosing a ride partner by prioritising the matches between social network connections. The special point is to involve the subjective heterogeneity in setting the space-time budget for a shared trip with people of different acquaintances.

2. The Activity-based Ridesharing (Chapter 7) slacks the constraints from space-time budget on people’s travel options by exploring alternative destinations that fit better into the schedule. This method highlights the importance of considering the activity for a travel purpose (i.e., the motivation for an urban travel) and the semantic of a place rather than only the location.

3. The Collaborative Activity-based Ridesharing (Chapter 8) coalesces the advantages of the Social Network based Ridesharing and the Activity-based Ridesharing, which harnesses alternative destination to increase matches with social network connections.

4. The Park and Ridesharing (P&S) model (Chapter 9) relieves the urban parking burden by making use of ridesharing. The new thing is this model takes into account people’s unwillingness to do ridesharing, so it only matches rides for the last leg of the travel rather than from the front door.

The agent-based simulation results confirm the effectiveness of the proposed models with increased number of matches, increased number of matched friends, or increased number of successful parking. The essential improvements and innovations of ridesharing lie beneath a better insight into how people set their space-time budgets with regard to with whom and for what. The research outlooks elab-
ated in Chapter 10 point out a few directions where research can further help exploit the market niche of ridesharing. Maybe someday, the ownership of cars will be revolutionised.


Andrew M Amey. Real-time ridesharing: exploring the opportunities and challenges of designing a technology-based rideshare trial for the MIT community, 2010.


B. Hillier. What do we have to add to a social network to get a society? Answer: something like what we have to add to a spatial network to get a city. In Stephan Winter and Garry Robins, editors, *International Workshop on Social Space and Geographic Space, in conjunction with COSIT’07*, pages 35–50, 2007.


Grant McKenzie, Krzysztof Janowicz, Song Gao, Jiue-An Yang, and Yingjie Hu. POI pulse: A multi-granular, semantic signature-based information observatory for the interactive visualization of big


DECLARATION

I declare that this thesis has been composed solely by myself Yaoli Wang. Some chapters are revised based on published or submitted manuscripts of which I am the first author and my contributions have been explained individually. Except where states by reference or acknowledgement, the thesis is entirely of my own. The thesis in whole has not been submitted in any previous application for a degree.

Melbourne, Australia, February 2018

Yaoli Wang
Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Wang, Yaoli

Title:
Collaborative activity-based ridesharing: linking human movement, social network, and platial semantics for future transportation

Date:
2018

Persistent Link:
http://hdl.handle.net/11343/211580

File Description:
Collaborative activity-based ridesharing: linking human movement, social network, and platial semantics for future transportation

Terms and Conditions:
Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.