

Citation: Xia J, Curtin KM, Li W, Zhao Y (2015) A New Model for a Carpool Matching Service. PLoS ONE 10(6): e0129257. doi:10.1371/journal. pone.0129257

Editor: Daqing Li, Beihang University, CHINA

Received: August 28, 2014

Accepted: May 6, 2015

Published: June 30, 2015

Copyright: © 2015 Xia et al. This is an open access article distributed under the terms of the <u>Creative</u> <u>Commons Attribution License</u>, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors have no support or funding to report.

Competing Interests: The authors have declared that no competing interests exist.

RESEARCH ARTICLE

A New Model for a Carpool Matching Service

Jizhe Xia¹, Kevin M. Curtin¹*, Weihong Li²*, Yonglong Zhao²

1 Department of Geography and GeoInformation Science, George Mason University, Fairfax, Virginia, United States of America, 2 School of Geography, South China Normal University Tianhe Park, Tianhe, Guangzhou, China

* curtin@gmu.edu (KMC); hongweili9981@163.com (WL)

Abstract

Carpooling is an effective means of reducing traffic. A carpool team shares a vehicle for their commute, which reduces the number of vehicles on the road during rush hour periods. Carpooling is officially sanctioned by most governments, and is supported by the construction of high-occupancy vehicle lanes. A number of carpooling services have been designed in order to match commuters into carpool teams, but it known that the determination of optimal carpool teams is a combinatorially complex problem, and therefore technological solutions are difficult to achieve. In this paper, a model for carpool matching services is proposed, and both optimal and heuristic approaches are tested to find solutions for that model. The results show that different solution approaches are preferred over different ranges of problem instances. Most importantly, it is demonstrated that a new formulation and associated solution procedures can permit the determination of optimal carpool teams and routes. An instantiation of the model is presented (using the street network of Guang-zhou city, China) to demonstrate how carpool teams can be determined.

Introduction

The work commute trip has long been known to be a critical element of transportation planning, and the most prominent contributor to traffic congestion. Since most commuters travel at regular times of day, this commonly leads to the morning and evening rush hours. The resulting congestion presents a cost to commuters in and of itself, but also influences pollution levels, energy consumption, and other related externalities. According to a report from the World Resources Institute (WRI), transportation accounts for nearly 14% of total greenhouse gas (GHG) emissions [1]. In addition to these economic and environmental concerns, energy efficiency—particularly associated with vehicle use—has become a high priority for governments around the world.

Numerous studies have been conducted to address inefficiencies in the traffic system. Models and solution procedures have been proposed to better manage the transportation system, including research into signal control optimization [2-3], and the integration of knowledge-based decision support systems [4]. Carpooling is another effective means of traffic management, through the direct reduction of the number of vehicles participating in the transportation system. Carpooling is loosely defined as the cooperation of two or more persons regarding

the use of a single vehicle to meet their mutual commuting needs. In addition to the societal benefits outlined above there are potential benefits to the individuals who participate in the carpool. These benefits could include reduced fuel costs, reduced toll costs, reduced time spent in the commute (if high-occupancy vehicle (HOV) lanes are available for use), and potentially reduced driving stress for the passengers in the vehicle.

Carpooling is officially sanctioned by most governments. Numerous HOV lanes have been constructed or designated for carpooling commuters. As of 2012, 345 HOV facilities (in operation or in some stage of construction or planning) have been identified across the metropolitan areas of the United States [5]. However, research has shown that carpooling has been decreasing precipitously [6]. Low fuel prices, high-quality transportation facilities (roads), strong incomes, and issues regarding quality of life are believed to have led to this decrease $[\underline{7-8}]$. Recently, with increasing fuel prices, rising environmental awareness, and an increasing number of policies directed toward carpooling activities, the awareness of the potential benefits of carpooling is believed to be increasing. A number of carpooling services have been built to provide carpool matching, such as Carpool World, Carpool Zone, and local government services. It has been found that a number of preference and policy variables can influence the success of these programs [9]. However, most carpooling services emphasize text-based information matching and integration, and seldom address the network analytic components of carpooling. Text-based carpool matching may result in a bad combination of carpooling team members, with high commute route costs. A network based carpool matching mechanism is needed to help commuters construct a carpooling team with minimum commute route costs.

In this paper, network-based carpool matching issues are addressed. In the literature review section, it is demonstrated that 3+ person carpooling is more effective in traffic management compared to 2-person carpool teams, but models, algorithms and services for 3+ person carpooling are not generally available due to the combinatorial complexity of the problems. This limitation motivates the research presented here, where a network-based carpool matching model for 3+ person carpooling is provided. In addition to the model, a range of solution procedures (both optimal and heuristic) are tested to determine their value in solving problems of this type. This paper is organized as follows: in Section 2 a review of the relevant literature regarding carpooling and carpool matching services is provided. Section 3 outlines a model for carpool matching and the variant solutions methods are applied to the carpool matching problem. Section 4 presents the experimental results across a large range of carpool matching problem instances, and Section 5 presents a practical application of the problem with a case study in Guangzhou city, China. Section 6 provides conclusions regarding the potential significance of these methodologies for policy and practice going forward, and suggestions for future research.

Literature Review

2.1 History and the current state of carpooling

Articles [6–7] provide extensive reviews of the history of carpooling activity in the United States. Since the world oil shortages in the mid-1970s, Federal and local governments have implemented a variety of policies and programs to encourage carpooling activity. After that period, carpooling became increasingly popular in United States, with nearly 20% of commute trips using a carpool in the 1970s. However, this decreased precipitously after 1980, and has dropped to approximately 10% in 2010. Decreasing oil prices, improving transportation facilities, and increasing incomes are potential reasons for this decrease in carpool trips.

To encourage carpooling, many transportation agencies have built extensive networks of HOV lanes. However, the efficiency of-and benefits from—HOV lanes remain topics of controversy [10]. [11] investigated the influence of carpool lanes on overall transportation network performance. That research found that, using actual trip data, the presence of a reserved carpool lane on a congested highway can increase commute time. Similarly, [12] have found that the overall efficiency of a highway decreases with the presence of an HOV lane. It has been observed that HOV lanes frequently carry fewer people than general-purpose lanes. [13] demonstrated that a high proportion of two-person family carpools, also known as "fampools", dramatically reduced the expected benefit of HOV lanes that were built with the expectation of carpools or vanpools carrying three persons or more per vehicle. This finding provides a primary motivation for the research presented here. If efficient means of generating larger carpools can be implemented, the expected benefits of HOV lanes can be more easily realized. Carpool matching services can lead to larger carpool sizes, and the ability to efficiently match multiple carpool users effectively multiplies the benefits of the carpool. This research presents methods for doing so.

2.2 Factors influencing carpooling activities

In the process of designing a carpooling matching service, it is critical to understand the factors involved in carpooling activities; including people's perceptions of the benefits and costs of carpooling, and their concerns regarding safety. [14] conducted two surveys to analyse people's views regarding carpooling activities. 40 undergraduate students (20 male, 20 female) and 48 staff members from the University of Iowa were selected, and asked to rate their concerns regarding carpooling. Responses were collected through both questionnaires and telephone survey techniques. The author concluded that the cost of time and convenience are the two decisive factors in the decision to carpool. People also expressed concerns regarding personal comfort and the gender mix among the carpool participants. [15] conducted a survey based on a significantly larger sample. The survey was conducted from May to July 2006 through the internet, targeting commuters in Dallas-Fort Worth and Houston. In total 4,634 responses were collected. Among those, 69.2% were solo drivers, 13.3% participated in a two-person carpool, 4.4% participated in a three-person carpool, and 2.3% rode in a vanpool. Results showed that access to HOV lanes and a reduction in driving stress were the two most important reasons for carpooling. Among the respondents who did not carpool, the most important reasons for driving alone were 1) the difficulty in finding someone with the similar location and schedule (55%), 2) the flexibility of solo driving (45%) and 3) needing a vehicle during the day (39%). The survey also showed that people are familiar with each other in most carpooling teams (75%). A survey (a total of 996 respondents) conducted in Lisbon, Portugal showed that the poor carpooling schedule and trust level between strangers are two major obstructions for carpool activities [16].

Location and schedule requirements seriously limit the convenience and flexibility of carpools. Although a large proportion of commuters likely share a similar commute route and schedule (thus leading to rush hour traffic), it is difficult for them to find each other and coordinate their travel. Therefore, an efficient carpool matching service which enables commuters to develop a potential carpool team can be a critical element in encouraging carpool use.

2.3 Carpool matching models and services

Numerous carpooling services have been developed with a wide range of approaches and functions. In this review, these services are loosely grouped into four categories: 1) services that list information but provide no explicit matching, 2) hardware- and communications-focused approaches, 3) services that use spatial information but not network-based information, and 4) services that use network-based spatial information to match users and provide good carpool routes. Each of these groups of services is reviewed in turn. **2.3.1 Carpool matching through lists of users.** The first category of carpool matching services contains those that provide information regarding users who wish to carpool, but provide no explicit carpool matching services. In this category, a number of carpooling services are based on textual information posting, searching, and integration. For example, Craigslist (www.craigslist.org) features free online classified advertisements and provides carpooling information for users in different metropolitan areas.

The main difficulty of such carpool services is that the system rarely suggests a match between drivers and riders. Rather, it simply provides a listing of those who may want to carpool, and the users need to search on their own to find other users with similar commute patterns. No actual carpool matching takes place through the system. If end users do successfully find a carpool team (often after significant trial and error using personal knowledge and perceptions of locations) they have no way of quantifying how efficient their carpool team is, nor can they view alternative carpool team options. These factors may combine to result in inefficient combinations of carpool teams.

2.3.2 Hardware- and communications-focused approaches. [17] designed an architecture to assist commuters in defining a group of people who usually have similar commute routes and schedules. This architecture is primarily hardware-based; using a carpool device to collect daily commuter routes and schedules. This idea can be extended for use with smartphones. PASS [18] is a parking-lot-based method for carpooling which differs from vehicular ad hoc networks. Wireless devices and accelerator sensors are used to collect commuters' information, and then establish a routing tree to deliver vehicle trajectory information to nearby parking lots. [19] investigated the network communications and the feasibility of using WiFi for carpooling activities in metropolitan areas. [20] employed a vehicle-to-passenger communications (V2P) approach to support communications between riders and drivers. These hardware- and sensor-based carpool matching approaches provide promising methods for the collection of commute information and they encourage communications between potential carpoolers. However, explicit matching algorithms are not their focus and commuting routes for suggested carpools may not be optimized.

2.3.3 Services that use spatial information but not network based matching. In contrast to those discussed above, examples of systems such as the Seattle Smart Traveller (SST) [21] system provide more structured carpool matching services, although not based on network information. SST not only collects spatial and temporal commute information, but also performs matching using SQL specifications. A temporal match is performed to define the overlap in departure and arrival times, followed by a spatial match that is based on the intersection of the buffers centred on commuter locations. Carpool World (www.carpoolworld.com) provides worldwide carpool matching functions including scheduling, consideration of special needs, travel behaviour recording and address matching. This matching algorithm is mainly based on the latitude and longitude of commuters rather than matching based on the commuter route itself. Zimride (http://www.zimride.com) is another schedule-based carpooling service. The system is focused on long-distance carpool matching based on the driver's schedule and the distance between driver and passenger.

2.3.4 Services that use network-based spatial information to match users and provide good carpool routes. WIGEOPOOL [22] is a GIS-based traveller information system that performs carpool matching based on network locations. It provides a number of carpool services such as driver and passenger searching, carpool team matching, and commute cost sharing. Carpool Zone (www.carpoolzone.smartcommute.ca) is a carpool matching service which serves commuters in the greater Toronto and Hamilton area, Canada. It includes functions such as intelligent route matching, interactive mapping, pinpoint geocoding, and security, privacy, and administrative functions for the contacting and matching process [9]. Carpool Zone

helps commuters to find other commuters who share a similar commuter route with specific needs (e.g. schedule, gender, language). The system gives the percentage of similarity between two commute routes through network analysis. "Let's Carpool" [23] is a similar carpool matching service for the Wellington region of New Zealand. It provides similar functions, but the similarity between two commute routes is not given to the end users. An evaluation study of "Let's Carpool" showed that the percentage of 1300 registered commuters who carpooled as their main commute mode increased significantly (from 12% to 27%) after the service was introduced.

"BlueNet" [24–26] is a carpool service in Taiwan with a mobile client and cloud-based carpool matching module. The system adopted a genetic algorithm approach in order to provide network-based carpool matching. [27] conducted a carpool matching simulation study based on 2008 travel demand data. Both matching strategies are designed to optimize the total system-wide vehicle miles travelled and the carpool matching rate, rather than individual preferences. Therefore, the interests of a single carpool user may not be guaranteed from a personal perspective. Taking the success of the above carpool studies, a number of mobile-based carpool applications have been developed, such as Lyft (https://www.lyft.com), Flinc (https://flinc.org), Carma (https://carmacarpool.com), and Tripda (www.tripda.com). These mobile applications provide basic real-time and network-based matching services, and improve the user experience by introducing driver-passenger communication tools, payment modules (e.g., fee calculator and payroll system), and security background check functions.

These carpool services use network-based algorithms to suggest a carpool team and commute route. The matching algorithms are mainly designed for a single driver–and a single passenger or system-wide carpool matching [24-27] optimization. In other words, the algorithms and carpool matching service may not provide an optimized solution when a commuter wants to carpool with more than one person. [12-13] have argued that two-person carpooling is so inefficient that it may even be considered a burden to the transportation system. Three-person carpooling and vanpooling are recommended. Models, algorithms and services for 3 + person carpooling are needed, but these are still missing in the current carpooling research and literature. This limitation motivates the research presented here, where a network-based carpool matching model for three or more person carpooling is provided. Moreover, in two recent reviews of issues surrounding ridesharing [28] and methods for optimally addressing ridesharing [29] systematic classifications were made in order to "foster the development of effective formal ridesharing mechanisms" [28]. This work presents an effort to develop one such formal mechanism and evaluate its potential.

Methodology

3.1 Conceptualization of Carpooling Activities on networks

Generally, there are two types of people who participate in carpooling activities: 1) "drivers" who have vehicles available for commuting and who want to share their vehicles with other commuters, and 2) "passengers" who are looking for a ride with drivers. Both the drivers and the passengers must participate in a two-sided matching in order to address all participants' needs. Further, a carpooling activity can be abstracted into two fundamental processes: 1) a carpool matching process that enables drivers or passengers to develop a carpooling team, and 2) a daily commute process where the driver needs to choose the order in which they will pick up and drop off all team members.

After a carpooling team is formed, the driver must pick up all team members from their residences in a pre-determined order, then drop off all team members at their workplaces in a pre-determined order, ending at the workplace of the driver. This process must repeat in roughly the reverse order at the end of the work period. In this process, the "residence" or "workplace" could also be a group pick-up area (e.g. a commute pooling lot, an apartment building), or an inter-modal transfer location (e.g. a bus stop or a train station) if such locations facilitate the carpooling process.

Common constraints on the carpool matching and daily commute processes are as follows:

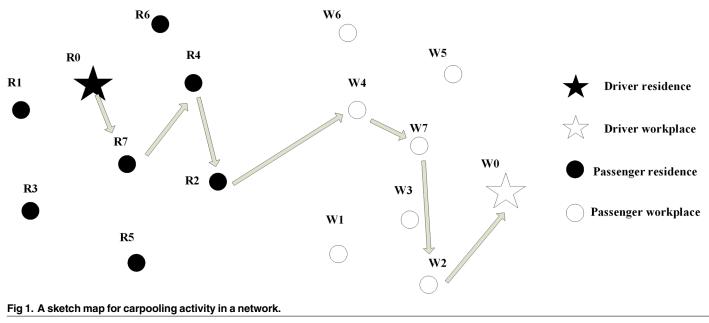
- 1. The total number of commuters in a carpooling team cannot exceed the maximum capacity of the commute vehicle.
- 2. The entire commute route must start at the residence of the driver and end at the workplace of the driver
- 3. Each team member must be picked up before they can be dropped off. While this constraint may seem obvious, it must be made explicit in the carpool matching model outlined below. However, this constraint does not require that all team members are picked up before some team members are dropped off. In other words, the driver could drop off a team member before picking up some other team member.

The objective of the carpool matching and daily commute processes is to minimize the cost (e.g. route length, commute time, or some combined cost) of the daily commute.

Fig 1 shows an example of this carpooling activity model in a network. Each person in the example is identified by a number, and the residence and workplace of that person share the same number. For example, Person 4 is termed P4, and that person lives at location Residence 4 (R4) and works at Workplace 4 (W4). In this example, Person 0 (P0) is a driver and P0 initiated a carpool matching process. P0 chose P7, P4 and P2 to be his or her carpooling team members. The order for picking up those members is P0->P7->P4->P2 and the order for dropping off is P4->P7->P2->P0. While this example shows only a single vehicle carpool matching process, an implementation of this system would permit many such carpools to be developed based on the needs of multiple drivers and many potential passengers.

3.2 Basic Assumptions

The network-based carpooling model has the following assumptions:



doi:10.1371/journal.pone.0129257.g001

- 1. The carpooling model operates on a graph, or network. Residence and workplace locations can be abstracted as points along the network. Roads between residences and workplaces can be abstracted as arcs.
- All residences and workplaces are connected by roads. Therefore, the model assumes that all points can be connected by arcs. The arcs are undirected, meaning that travel is possible in either direction.
- 3. The impedance associated with any arc in the carpooling model can represent any cost function that is appropriate for the commuters. This can include the length of the road, the drive time (including congestion factors), road conditions, weather conditions, tolls, or any combination of cost factors that are pertinent to the commute. These impedances can be implemented by weighting the arcs or nodes in the network appropriately.

3.3 Optimal Solution Procedures

3.3.1 Optimal solution through enumeration. For small problems of this type it is well known that trying every possible arrangement of carpool routes is a solution method that is guaranteed to provide the optimal solution. Known as enumeration or brute-force search, algorithms of this type would test every possible carpooling team combination and carpooling route in order to find the best selection of team members and the associated best carpooling route. For this problem the enumeration process began with the identification of the pool of potential candidate passengers, and the number of passengers to be chosen from that pool. Based on those inputs the following steps were performed: 1) for each unique set of possible passengers out of the candidate pool identify all possible sequences of pick-ups from home locations and drop-offs at work locations; 2) eliminate any sequences compute the shortest path; 4) choose the minimum shortest path; 5) return to step 1 and repeat the process for each unique set of passengers from the pool.

The scale of the problem follows the following equation given the number of carpool members (s1) and potential candidates (*K*):

Problem Scale =
$$C_k^{s1} P_{s1*2}^{s1*2}$$
 (Equation1)

Assume there are 50 candidates and a driver wants to select three passengers for the carpooling team. There will be C_{50}^3 options for passenger selection, and P_6^6 permutations for picking up and dropping off those passengers. The magnitude of the set of possibilities will be $C_{50}^3 P_6^6 = 14, 112,000$, team/route combinations. If there are 100 candidates and 3 other carpool members, the magnitude of the set of possibilities will be $C_{100}^3 P_6^6 = 116,424,000$ carpool/ route combinations. If there are 50 candidates and 4 other carpool members, the magnitude of the set of possibilities will be $C_{50}^4 P_8^8 = 9,285,696,000$ team/route combinations.

As can be seen-and as expected with an enumerative approach to solving these problemsthe problem scale dramatically increases according to the increase in the number of candidates and the number of passengers in the carpool. In a real-world carpool matching process, there could be thousands of candidates. One would expect that any brute-force algorithm would require far too much computing time and resources for model execution on a problem instance of any significant size. Therefore a more efficient algorithm is needed for providing a realworld carpool matching service. Both a more efficient optimal solution procedure and heuristic approaches are outlined below. **3.3.2 Linear Programming solution procedure.** Integer linear programming provides a means of determining optimal solutions to problem instances that are beyond the reach of brute-force algorithms. In this section, a mathematical formulation of the carpooling route selection problem is presented and discussed.

Consider the following set of notation:

p is the index of passenger candidates and the driver; p = 1, 2, ..., N, where *N* is the total number of passenger candidates plus the driver;

k is the index of residences; k = 1, 2, ..., N;

l is the index of workplaces; $l = N+1, N+2, ..., N^*2$;

It is important to note that the indices of passengers, residences, and workplaces are all ordered. That is, passenger 1 (p = 1) lives at residence 1 (k = 1) and works at workplace N + 1 (l = N + 1).

Additionally, the residence of the driver is denoted as *s*, and the workplace of the driver is denoted as *t*, where t = s + N;

i and *j* are indices of nodes (residences or workplaces); *i*, *j* = 1, 2, . . ., N^* 2; following from above the node is a residence when *i*, *j* < = *N* and the node is a workplace when *i*, *j* > *N*;

 d_{ij} is the cost of the arc (e.g. length, impedance, or some other user defined cost) between node *i* and node *j*;

C is the number of passengers in the carpool team (defined by the driver based on the capacity of the vehicle);

r and *f* are indices of arcs comprising a route; r, f = 1, 2, ..., R;

R is the number of arcs in a carpool route, $R = C^{*}2 + 1$;

 x_{ijr} is a decision variable equal to 1 if the arc from *i* to *j* is chosen for step *r* in the route, and equal to 0 otherwise;

Using this notation, the general carpooling route selection formulation can be given as: Minimize

$$Z = \sum_{i=1}^{N+2} \sum_{j=1}^{N+2} \sum_{r=1}^{R} d_{ij} \mathbf{x}_{ijr}$$
(1)

Subject to:

$$\sum_{i=1}^{N+2} \sum_{k=1}^{N} \sum_{r=1}^{R} x_{ikr} = C$$
⁽²⁾

$$\sum_{i=1}^{N+2} \sum_{l=N+1}^{N+2} \sum_{r=1}^{R} x_{ilr} = C+1$$
(3)

$$\sum_{k=1}^{N} x_{sk1} = 1$$
 (4)

$$\sum_{l=N+1}^{N+2} x_{llR} = 1$$
 (5)

$$\sum_{i=1}^{N*2} x_{ikr} - \sum_{i=1}^{N*2} \sum_{f=r+1}^{R} x_{i(k+N)f} = 0, \text{ for } k = 1, 2, \dots, N; k \neq s; r = 1, 2, \dots, R-1$$
 (6)

$$\sum_{i=1}^{N*2} \sum_{r=1}^{R} x_{ijr} \le 1 \text{ for } j = 1, 2, \dots, N*2$$
(7)

$$\sum_{j=1}^{N*2} \sum_{r=1}^{R} x_{ijr} \le 1 \text{ for } i = 1, 2, \dots, N*2$$
(8)

$$\sum_{i=1}^{N*2} x_{ijr} - \sum_{i=1}^{N*2} x_{ji(r+1)} = 0 \text{ for } j = 1, 2, \dots, N*2; r = 1, 2, \dots, R-1$$
(9)

$$\sum_{i=1}^{N*2} \sum_{j=1}^{N*2} x_{ijr} = 1 \text{ for } r = 1, 2, \dots, R,$$
(10)

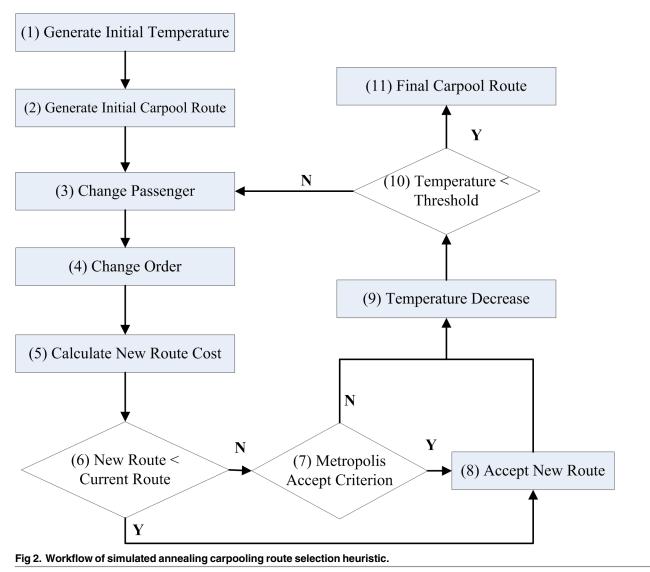
The objective function (1) seeks to minimize the cost of the entire carpool route. Constraint (2) requires the driver to pick up a given number (*C*) passengers at residences and (3) requires that the driver must drop off the given number (*C*) passengers and the driver at workplaces. Constraint (4) requires that the carpool route must start with the drivers' residence, and (5) requires that the carpool route must end with the drivers' workplace. Constraint (6) ensures that the driver will not drop off a passenger before picking her up. In order to ensure a complete, connected, and non-overlapping carpool route, constraint (7) requires that no more than one arc entering any node be selected, constraint (8) requires that no more than one arc exiting a node be selected; and constraint (9) ensures that if an arc enters a node on step *r* of the route, an arc exiting that node must be chosen for step r + 1 of the route. Constraint (10) ensures that there will be exactly one arc chosen for any step in the route.

This formulation can be used to generate optimal solutions for problem instances of hundreds of candidates for the carpool teams. However, even this method fails to find solutions for larger problems due to the lack of computing resources, or the time needed to determine optimality. Therefore, a practical application of carpool routing will almost certainly benefit from the implementation of heuristic methods.

3.4 Heuristic Solution Procedures

It is well known that the demands on computer resources and solution times grow rapidly when searching for the optimal solutions to highly combinatorially complex problems. The search for suitable heuristic algorithms to resolve this issue has been ongoing for decades. While this is not the appropriate forum for a thorough review of heuristic techniques, those such as Tabu Search [30–31], ant colony optimization [32], genetic algorithms [33], and simulated annealing [34], are among the most prominent. For the carpool selection and routing problems addressed here, both a simulated annealing and a Tabu search heuristic have been implemented. In so doing, this research tests one method from the class of smart interchange heuristics (Tabu), and one method from the class of heuristics that are based on correlates to natural processes (simulated annealing), in order to explore the costs and benefits of each class.

3.4.1 A Simulated annealing heuristic solution procedure. Simulated annealing heuristics simulate the physical process of annealing. When metal is heated to a very high temperature, its molecules can move freely. When the temperature falls, the whole system returns to thermodynamic equilibrium and the molecules will stay in their altered state. By simulating this physical process, a simulated annealing heuristic can effectively avoid premature



doi:10.1371/journal.pone.0129257.g002

PLOS ONE

convergence to a local optimum and therefore possibly find a solution closer to (or ideally equal to) the optimal solution.

In the problem at hand, the length of the carpooling route corresponds to the energy of the metal, and the best route selection corresponds to the lowest energy state. We simulate the annealing process by setting the rule of temperature decrease and the rule for changing the carpooling route. Fig 2 shows the work flow of the simulated annealing carpooling route selection heuristic. The steps include: (1 and 2) generate initial temperature and initial carpooling route; (3 and 4) generate a new carpooling route by changing passengers and routes; (5) calculate the cost for the new route; (6, 7, and 8) if the new carpooling route is better than the current route, the current route will be replaced by the new route; otherwise, the new route will not be accepted if it also fails the Metropolis criterion test (Metropolis et al., 1953); (9) decrease the temperature; (10 and 11) when the temperature approaches a predetermined lower threshold, the current route will be chosen as the solution; otherwise, go back to step (3) to generate a new route. The pseudo-code of the simulated annealing carpooling route selection heuristic is shown in algorithm 1 to algorithm 4. The algorithm 2 is used for the Metropolis test in the

simulated annealing process. The algorithm 3 and algorithm 4 are designed to determine all possible permutations of feasible carpooling routes given the current passenger pool. This function adopts a classic recursive permutation algorithm. Algorithm 1. Carpool Matching Simulated Annealing Input: Initial Temperature, CoolingRate, Threshold Temperture, Passenger Number (c), Candidate List (P1, P2, ..., Pn), Distance Matrix Output: Optimized Carpool Route SA1: Passenger Pool = Randomly c number of candidate SA2: Current Route = a random route to pick and drop off passengers SA3: Best Route = Current Route SA4: Current Temperature = Initial Temperature SA5: While (Current Temperature > Threshold Temperture) SA6: Passenger Pool = Randomly replace a passenger with a candidate SA7: New Route = FindBestRoute (Passenger Pool) SA8: If (Cost (New Route) < Cost (Current Route)) SA9: Current Route = New Route SA10: Else SA11: Difference = Cost (Current Route) - Cost (New Route) SA12: Metropolis = MetropolisCriterion (Difference, Current Temperature) SA13: If (Metropolis is True) SA14: Current Route = New Route SA15: If (Cost (Current Route) < Cost (Best Route)) SA16: Best Route = Current Route SA17: Current Temperature = Current Temperature * CoolingRate SA18: Return Best Route Algorithm 2. MetropolisCriterion Input: Difference, Current Temperature Output: True/False MC1: Metropolis = Math. Exp (Difference / Current Temperature) MC2: If (Metropolis > Math.Random()) MC3: Return True MC4:Else MC5: Return False Algorithm 3. FindBestRoute Input: Passenger Pool Output: BestRoute with Current Passenger Pool FR1: Initialize route list with passenger _pool // e.g, { Pick1, Pick2, Pick3, Drop1, Drop2, Drop3} for the order of picking up and droping off passenger 1, 2, 3 FR2: BestRoute = Permutate routers (route list, 0) FR3: Return BestRoute Algorithm 4. Permutate routers Input: int[] route list, int start, Output: BestRoute PR1: If (start > = route list.length) PR2: If (route list is feasible) // not drop off before pick up the passenger PR3: currentDistance = Cost (route list) PR4: If (currentDistance < BestRoute)</pre> PR5: BestRoute = currentDistance PR6: For (j = start; j < = route list.length-1; j++)</pre> PR7: Swap the Step start and Step j in the route list PR8: Permutate routers (route list, start+1) PR9: Swap the Step start and Step j in the route list PR10: Return BestRoute

3.4.2 Tabu search solution procedure. Tabu search is a widely-used heuristic procedure that is designed to avoid local optimal solutions, and permit the search to find better–and

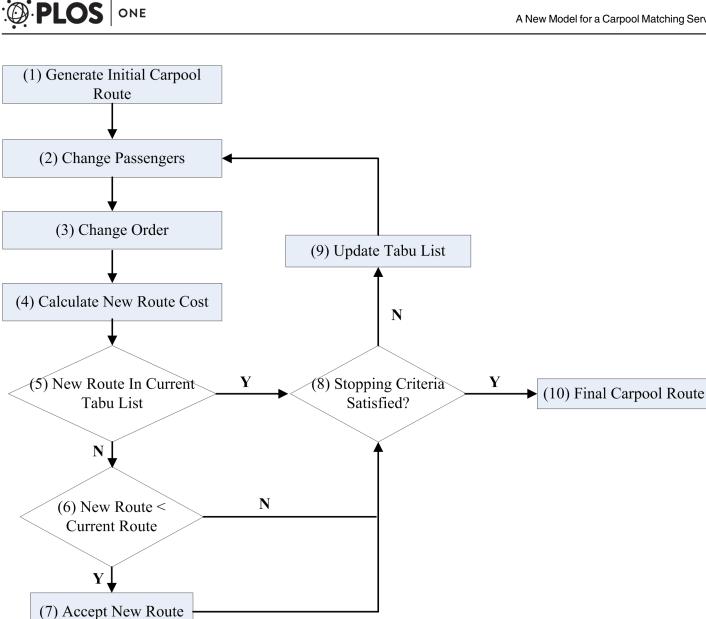


Fig 3. Workflow of Tabu search carpooling route selection heuristic.

doi:10.1371/journal.pone.0129257.g003

perhaps global optimal-solutions. While we do not provide a comprehensive review of the method of Tabu search, fundamentally, a tabu list is utilized to record the recent history of the search and direct the search away from recent solutions in order to more broadly explore the solution space. Parameters are set that prevent cycling back to previously visited solutions for a period of time (number of iterations).

Fig 3 shows the work flow of the Tabu search heuristic for the carpooling route selection problem. The steps include: (1) generate an initial carpooling route by randomly selecting passengers and routes; (2 and 3) generate neighbouring solutions (carpooling routes) by changing passengers and orders of picking up and dropping off; (4) calculate the cost for the new route; (5, 6, and 7) if the new carpooling route is not in the Tabu list and is better than the current route, the current route will be replaced by the new route; otherwise, go to step (8); (8 and 9) if the stopping criteria is not satisfied, update the Tabu list and go back to step (2) to generate neighbouring solutions; (10) Output the final carpool route once the stopping criteria is

satisfied. The pseudo-code of the Tabu search carpooling route selection heuristic is shown in algorithm 5 and algorithm 6.

Algorithm 5. Carpool Matching-Tabu Search Input: Tabulist MaxSize, Passenger Number (c), Candidate List (P1, P2, ..., Pn), Distance Matrix Output: Optimized Carpool Route TS1: Passenger Pool = Randomly c number of candidate TS2: Current Route = a random route to pick and drop off passengers TS3: Best Route = Current Route TS4: Tabulist = null TS5: While (! Stopping Criteria) TS6: Local Route = null TS7: While (Neighborhood Searching Criteria) TS8: Passenger Pool = Randomly replace a passenger with a candidate TS9: Current Route = FindBestRoute (Passenger Pool) // Algorithm3 TS10: If (Current Route not in the Tabulist) TS11: Local Route = Current Route TS12: If (Cost (Local Route) < Cost (Best Route)) TS13: Add Local Route to the Tabulist TS14: Best Route = Local Route TS15: While (Tabulist Size > Tabulist MaxSize) TS16: Expire Tabulist (Tabulist) TS17: Return Best Route Algorithm 6. Expire Tabulist Input: Tabulist Output: New Tabulist ET1: Index = 0; ET2: While (Index < Tabulist Size) ET3: Tabulist (Index) = Tabulist (Index + 1) ET4: Index = Index + 1 ET5: Return Tabulist

Implementation and experimental results

4.1 Experiment data and environment

The experiments were performed using both real-world and simulated data. For an underlying network representation the road network for Guangzhou city, China (S4 File) was employed. To determine the locations of carpool drivers and potential passengers, many sets of point locations were randomly generated within the bounding box of Guangzhou city. The sets contained an increasing number of potential passengers ranging from 50 to 1,000, with each set incremented by 50 (50, 100, 150, ..., 900, 950, 1,000) (S1–S4 File).

The experiments were performed on a desktop machine, running Intel Quad Cores i7-3770 @ 3.4 GHz, with 12 GB RAM, a hard disk with 7200 rpm storage seeking speed, and the Windows 7 operation system.

4.2 Optimal solution results

Fig 4 shows the execution times and objective function values for the linear programming procedure and the brute force enumeration procedure.

In Fig 4, the yellow highlighted cells represent problem instances where the linear programming procedure either solved the instance optimally where enumeration could not (within a 6 hour window) or the optimal solution was determined faster through linear programming. The results are as follows:

PL	.OS	ONE

	50	1.622	6.24	29.921	371.844	790.659	12291.530	20596.362	
	100	30.576	69.046	437.442	5746.375	21600.000	21600.000	21600.000	
	150	94.272	198.218	1553.877	21676.635	21600.000	21600.000	21600.000	
	200	215.842	270.343	2540.367	21600.000	21600.000	21600.000	21600.000	
	250	414.947	588.312	15940.463	21600.000	21600.000	21600.000	21600.000	
	300	637.920	755.112	19094.785	21600.000	21600.000	21600.000	21600.000	
tes	350	850.970	1115.213	21600.000	21600.000	21600.000	21600.000	21600.000	tes
lida	400	1010.886	1215.564	21600.000	21600.000	21600.000	21600.000	21600.000	lida
Number of candidates	450	1406.130	1561.216	21600.000	21600.000	21600.000	21600.000	21600.000	candidates
of c	500	1687.307	1924.254	21600.000	21600.000	21600.000	21600.000	21600.000	of
ber	550	1823.151	2150.212	21600.000	21600.000	21600.000	21600.000	21600.000	Number
Ium	600	2021.461	2316.910	21600.000	21600.000	21600.000	21600.000	21600.000	Ium
Z	650	2145.540	2593.161	21600.000	21600.000	21600.000	21600.000	21600.000	Z
	700	2348.654	2721.291	21600.000	21600.000	21600.000	21600.000	21600.000	
	750	2515.843	2810.552	21600.000	21600.000	21600.000	21600.000	21600.000	
	800	2695.231	2971.561	21600.000	21600.000	21600.000	21600.000	21600.000	
	850	2815.246	3055.267	21600.000	21600.000	21600.000	21600.000	21600.000	
	900	3012.002	3265.506	21600.000	21600.000	21600.000	21600.000	21600.000	
	950	3199.115	3353.610	21600.000	21600.000	21600.000	21600.000	21600.000	
	1000	3295.231	3539.545	21600.000	21600.000	21600.000	21600.000	21600.000	
		1	2	3	4	5	6	7	

Number of passengers	
a. Linear Programming Execute T	ime

50	0.071	0.182	1.943	1790.341	21600.000	21600.000	21600.000
100	0.027	0.04	17.596	21600.000	21600.000	21600.000	21600.000
150	0.093	0.096	59.133	21600.000	21600.000	21600.000	21600.000
200	0.124	0.171	646.238	21600.000	21600.000	21600.000	21600.000
250	0.177	0.312	1289.314	21600.000	21600.000	21600.000	21600.000
300	0.299	0.416	1703.984	21600.000	21600.000	21600.000	21600.000
350	0.426	0.639	2469.153	21600.000	21600.000	21600.000	21600.000
400	0.423	0.702	2994.675	21600.000	21600.000	21600.000	21600.000
450	0.517	0.884	4427.471	21600.000	21600.000	21600.000	21600.000
500	0.664	1.111	6477.418	21600.000	21600.000	21600.000	21600.000
550	0.787	1.395	9216.178	21600.000	21600.000	21600.000	21600.000
600	0.89	1.724	15824.271	21600.000	21600.000	21600.000	21600.000
650	1.022	1.922	21600.000	21600.000	21600.000	21600.000	21600.000
700	1.181	2.263	21600.000	21600.000	21600.000	21600.000	21600.000
750	1.36	2.646	21600.000	21600.000	21600.000	21600.000	21600.000
800	1.557	3.053	21600.000	21600.000	21600.000	21600.000	21600.000
850	1.78	3.509	21600.000	21600.000	21600.000	21600.000	21600.000
900	1.974	4	21600.000	21600.000	21600.000	21600.000	21600.000
950	2.275	4.5	21600.000	21600.000	21600.000	21600.000	21600.000
1000	2.486	5.074	21600.000	21600.000	21600.000	21600.000	21600.000
	1	2	3	4	5	6	7

Number of passengers

c. Enumeration Execute Time

Fig 4. Optimal solution results.

doi:10.1371/journal.pone.0129257.g004

50	31.567	40.914	45.878	58.691	75.975	83.091	117.312
100	17.321	35.93	48.732	57.398	N/P	N/P	N/P
150	21.530	33.179	43.74	57.898	N/P	N/P	N/P
200	41.385	45.558	50.707	N/P	N/P	N/P	N/P
250	20.495	30.352	35.349	N/P	N/P	N/P	N/P
300	28.219	31.773	36.552	N/P	N/P	N/P	N/P
350	24.021	28.687	N/P	N/P	N/P	N/P	N/P
400	33.682	34.368	N/P	N/P	N/P	N/P	N/P
450	32.445	35.721	N/P	N/P	N/P	N/P	N/P
500	33.847	34.737	N/P	N/P	N/P	N/P	N/P
550	12.775	18.316	N/P	N/P	N/P	N/P	N/P
600	26.333	28.841	N/P	N/P	N/P	N/P	N/P
650	13.673	15.801	N/P	N/P	N/P	N/P	N/P
700	14.421	20.944	N/P	N/P	N/P	N/P	N/P
750	11.448	17.964	N/P	N/P	N/P	N/P	N/P
800	16.699	18.685	N/P	N/P	N/P	N/P	N/P
850	9.628	15.634	N/P	N/P	N/P	N/P	N/P
900	15.613	18.868	N/P	N/P	N/P	N/P	N/P
950	14.75	18.884	N/P	N/P	N/P	N/P	N/P
1000	13.618	18.448	N/P	N/P	N/P	N/P	N/P
	1	2	3	4	5	6	7,

٠

Number of candidates

Number of passengers b. Linear Programming Objective Value

50	31.567	40.914	45.878	58.691	N/P	N/P	N/P
100	17.321	35.930	48.732	N/P	N/P	N/P	N/P
150	21.530	33.179	43.740	N/P	N/P	N/P	N/P
200	41.385	45.558	50.707	N/P	N/P	N/P	N/P
250	20.495	30.352	35.349	N/P	N/P	N/P	N/P
300	28.219	31.773	36.552	N/P	N/P	N/P	N/P
350	24.021	28.687	37.593	N/P	N/P	N/P	N/P
400	33.682	34.368	36.327	N/P	N/P	N/P	N/P
450	32.445	35.721	36.991	N/P	N/P	N/P	N/P
500	33.847	34.737	37.125	N/P	N/P	N/P	N/P
550	12.775	18.316	24.927	N/P	N/P	N/P	N/P
600	26.333	28.841	32.216	N/P	N/P	N/P	N/P
650	13.673	15.801	N/P	N/P	N/P	N/P	N/P
700	14.421	20.944	N/P	N/P	N/P	N/P	N/P
750	11.448	17.964	N/P	N/P	N/P	N/P	N/P
800	16.699	18.685	N/P	N/P	N/P	N/P	N/P
850	9.628	15.634	N/P	N/P	N/P	N/P	N/P
900	15.613	18.868	N/P	N/P	N/P	N/P	N/P
950	14.750	18.884	N/P	N/P	N/P	N/P	N/P
1000	13.618	18.448	N/P	N/P	N/P	N/P	N/P
	1	2	3	4	5	6	7,

Number of passengers

d. Enumeration Objective Value

Number of candidates

- With small-scale carpool route selection instances (smaller than four passengers), the enumeration procedure can provide the optimal solution with a lower execution time as compared to the linear programming procedure. With three passengers, the linear programming procedure can only provide optimal solutions within six hours when the number of commuters is less than 300. The enumeration procedure can solve 3-passenger instances up to 600 commuters.
- As the number of passengers increases (above 3), the linear programming procedure can provide additional optimal solutions that are beyond the reach of the enumeration procedure, when a six-hour time limit is imposed.
- For the particular instance of 4 passengers and 50 potential commuters, both methods find the optimal solution within six hours, although the linear programming procedure does so significantly faster.
- The solution time for the linear programming procedure increases faster with the increase in the number of potential commuters, while the solution time for the enumeration procedure increases faster with the increase in the number of passengers in the carpool team.

4.3 Heuristic Solution and results

<u>Fig 5</u> shows the execution times and objective function values for the simulated annealing and Tabu search heuristics.

The purple highlighted cells represent those instances where the respective heuristics are known to have identified the optimal solution. The red highlighted cells are those that are known to have identified sub-optimal solutions. The remaining values are the best solution found by the heuristic, although it is not known if the solution is optimal or not.

- In general, the Tabu search procedure requires more execution time than simulated annealing (but not by more than 3 seconds for any instance).
- When the number of passengers is smaller than 4, there is no significant difference between the Tabu search and simulated annealing results. However, Tabu search does find the optimal solution more often than simulated annealing (23 times v. 20 times out of 28 known optimal solutions).
- With larger problem size (number of passengers greater than 3), the Tabu search procedure generally provides a shorter carpooling route cost (about 5%), when compared with the simulated annealing procedure with similar execution time.

4.4 Comparison & Discussion

The results above have potentially significant consequences for the instantiation of a realworld carpool matching system. The consequences depend on the constraints under which the carpool matching must operate. For example, assume that a carpool is being established for long-term commuting purposes. A van holding 7 or more persons is to be employed for the commute, and there are 50 potential commuters. In this case, it appears that an optimal solution (and particularly the linear programming approach) is appropriate. The solution time limit of 6 hours would not be prohibitive (and could be increased, in fact), and the optimal solution would, of course, be guaranteed. The use of a heuristic solution procedure in this same case could result in a solution that is at least 61% above optimal (Tabu search returns 189 km while linear programming returns 117). Number of candidates

Number of candidates

<u> </u>							
50	0.175	0.230	9.121	11.534	12.300	12.300	12.301
100	0.025	0.168	9.062	11.544	12.300	12.300	12.300
150	0.022	0.167	9.113	11.566	12.300	12.300	12.300
200	0.021	0.167	9.114	11.584	12.300	12.300	12.300
250	0.021	0.168	9.188	11.669	12.300	12.301	12.300
300	0.021	0.170	9.284	11.619	12.303	12.300	12.300
350	0.022	0.172	9.266	11.720	12.300	12.301	12.301
400	0.021	0.172	9.232	11.650	12.300	12.301	12.300
450	0.021	0.181	9.220	11.657	12.300	12.300	12.300
500	0.021	0.187	9.301	11.668	12.301	12.300	12.300
550	0.021	0.192	9.312	11.680	12.400	12.400	12.401
600	0.021	0.195	9.266	11.658	12.401	12.401	12.401
650	0.022	0.195	9.319	11.627	12.301	12.301	12.301
700	0.022	0.203	9.323	11.648	12.401	12.401	12.401
750	0.022	0.203	9.340	11.658	12.401	12.401	12.401
800	0.023	0.203	9.326	11.635	12.302	12.302	12.301
850	0.023	0.205	9.304	11.712	12.401	12.402	12.401
900	0.022	0.204	9.327	11.682	12.301	12.301	12.302
950	0.022	0.208	9.317	11.712	12.302	12.302	12.302
1000	0.023	0.213	9.297	11.712	12.402	12.401	12.402
	1	2	3	4	5	6	Ţ.

	<u> </u>							
1	50	31.567	40.914	45.878	86.252	136.522	186.582	211.216
	100	17.321	35.930	48.732	90.811	126.192	177.201	166.222
	150	21.530	33.179	45.980	74.732	123.030	175.118	184.747
	200	41.385	45.558	51.932	64.297	140.619	166.648	210.137
	250	20.495	30.352	45.881	73.417	130.146	167.262	191.239
	300	28.219	31.773	43.078	69.139	122.970	153.563	216.019
SS	350	24.021	34.58	47.317	78.670	127.259	191.859	201.367
Number of candidates	400	33.682	34.368	37.437	84.493	124.304	183.368	196.695
andi	450	32.445	35.721	43.322	78.900	111.666	146.491	198.879
of cá	500	33.847	36.187	45.414	83.858	125.966	190.941	193.639
er o	550	12.775	19.836	33.107	81.323	115.584	141.614	196.818
hmb	600	26.333	28.841	42.371	77.514	123.381	182.799	164.802
ź	650	13.673	21.617	34.865	80.945	131.350	152.034	185.501
	700	14.421	20.944	38.624	75.365	108.211	143.944	182.485
	750	11.448	17.964	34.454	74.669	130.801	156.700	215.681
	800	16.699	18.685	36.290	81.069	139.352	173.785	184.933
	850	9.628	21.435	31.364	68.082	132.401	163.771	197.171
	900	23.848	25.892	44.815	77.443	136.116	164.536	218.236
	950	15.566	18.884	42.945	78.441	114.148	151.884	213.489
	1000	13.618	28.245	47.953	93.090	131.226	172.972	215.971
		1	2	3	4	5	6	7

Number of passengers

a. Simulated Annealing Execute Time

50	0.151	0.204	10.437	12.343	14.833	14.929	14.907
100	0.050	0.184	10.520	12.392	14.826	14.943	15.010
150	0.042	0.180	10.523	12.317	14.871	14.887	15.049
200	0.043	0.181	10.618	12.283	14.855	14.842	14.883
250	0.048	0.179	10.524	12.379	14.828	14.918	14.981
300	0.037	0.194	10.555	12.284	14.816	14.922	14.936
350	0.037	0.192	10.505	12.276	14.864	14.927	14.965
400	0.037	0.218	10.854	12.306	14.824	14.869	15.037
450	0.040	0.229	11.442	12.296	14.831	14.834	14.977
500	0.044	0.248	10.868	12.282	14.839	14.933	14.964
550	0.051	0.242	10.993	12.290	14.828	14.920	14.972
600	0.038	0.225	10.790	12.323	14.817	14.904	15.011
650	0.041	0.229	10.780	12.288	14.895	14.939	14.954
700	0.038	0.232	10.761	12.292	14.861	14.937	14.893
750	0.038	0.239	10.770	12.310	14.841	14.914	14.868
800	0.039	0.251	10.772	12.291	14.846	14.855	14.854
850	0.038	0.247	10.754	12.287	14.820	14.860	14.900
900	0.038	0.245	10.758	12.278	14.888	14.859	14.978
950	0.038	0.266	10.705	12.285	14.832	14.836	14.895
1000	0.038	0.261	11.009	12.309	14.898	14.929	14.981
	1	2	3	4	5	6	7

Number of candidates

Number of passengers

c. Tabu Search Execute Time

Fig 5. Heuristic Solution Results.

doi:10.1371/journal.pone.0129257.g005

Number of passengers

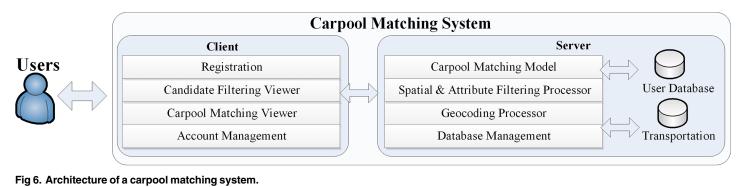
b. Simulated Annealing Objective Value

50	31.567	40.914	45.878	73.003	115.986	171.082	189.169
100	17.321	35.930	49.522	87.364	121.903	168.830	187.444
150	21.530	33.179	45.980	69.879	114.317	161.819	176.772
200	41.385	45.558	53.529	85.907	118.749	180.348	218.258
250	20.495	30.352	42.963	81.932	131.476	158.188	183.494
300	28.219	31.773	45.071	68.522	118.191	170.606	177.216
350	24.021	28.687	44.478	81.344	108.813	175.151	173.674
400	33.682	34.368	41.830	81.359	127.327	153.258	183.841
450	32.445	35.721	46.899	81.393	112.545	171.622	169.938
500	33.847	36.128	44.496	88.655	132.788	165.353	220.917
550	12.775	19.836	28.471	52.504	97.464	166.751	186.337
600	26.333	28.841	42.218	59.471	104.524	150.516	189.077
650	13.673	21.617	38.295	48.174	147.629	161.879	218.293
700	14.421	21.917	41.450	66.322	92.947	141.724	174.564
750	11.448	19.788	31.193	70.958	128.301	159.936	219.520
800	16.699	18.685	39.098	76.438	142.910	164.406	195.030
850	9.628	20.529	33.841	78.528	127.964	157.397	197.517
900	15.613	25.892	41.625	68.331	119.208	182.781	170.271
950	14.750	20.116	45.061	86.884	130.008	167.282	179.616
1000	13.618	18.448	35.427	82.007	110.130	154.980	187.145
	1	2	3	4	5	6	7

Number of passengers

d. Tabu Search Objective Value





doi:10.1371/journal.pone.0129257.g006

Conversely, if the carpools are to be determined on a daily/nightly basis, and only one or two passengers will be chosen, then the brute force approach is preferred. Finally, in the instances where larger carpools are desired, and the solution time is a concern, the heuristic approaches are the only option. Their rapid solution time allows for the near immediate determination of the carpool personnel and route, allowing for planning for the next day's commute.

The Carpool Matching System

While the results above demonstrate the relative value of different approaches to solving the carpool matching problem, we are also concerned here with the practical implementation of those approaches. In this section the architecture of such a system is outlined, with an example case study from Guangzhou city, China.

The architecture of a carpool matching system is shown in Fig.6. The carpooling matching service contains two tiers (a client tier and a server tier). The **Client tier** provides a Graphical User Interface (GUI) and all interactive functions that are necessary to initiate a carpool matching activity, including registration, account management, candidate filtering viewer, and carpool matching viewer. The **Server tier** includes the carpool matching model (as discussed above), the model processor, a filtering processor, and a geocoding processor to support the carpool matching function.

5.1 Implementation

A carpool matching system has been developed using the Microsoft.Net Framework 4. Microsoft SQL Server is used for user information storage and management. The ArcGIS Engine is used for spatial data management and visualization. The city road network for Guangzhou city is shown in Fig 7.

5.2 Use case

In this section, we describe a use case illustrating how the carpool matching system can support carpooling activity among commuters. The process includes:

Registration. Users must input required or optional information and register with the system. Registration input includes spatial, temporal, and text-based information (Fig.8). A user must input the location of their residence and their workplace (spatial data), their work schedule (temporal data), and text-based information such as user name, gender, age, and any special requirements (e.g. female only). Address geocoding is employed to transfer text-based residence and workplace inputs into spatial locations. In this example, an alias name database is used to better support the geocoding process. For example, when a user inputs the address



Fig 7. Road network of Guangzhou city in 2009.

doi:10.1371/journal.pone.0129257.g007

PLOS ONE

User Nume	Hike	Time to Work	8:30	Be a Cand	lidate? Tes 💌
Telephone	703606322	Time off Work	5:30	Enuil	xiajizhe10gmail.co
Have Car?	Tes 💌	Special Need	fenale	Gender	Fenale
Resi dence	TianMe TianMe Town	Furry Hatch	Work Place Ba	nk nk of GuangDa	Furry Match
				No Matched	
	InPut Information	Change Info	ormation	InPut New	Address

Fig 8. Registration Interface Design (allow this image to be published CC BY 3.0 license).

doi:10.1371/journal.pone.0129257.g008

Target Layer workplace	•
CBD	Fuzzy Match
ZhuJiang Town	
Fig 9. Alias match function (allow this image to b	e published CC BY 3.0 license).

doi:10.1371/journal.pone.0129257.g009

"CBD", meaning the downtown area, the system will automatically match to "ZhuJiang Town" (Fig.9).

Attribute Filter. In this process, the user filters candidates according to text-based and temporal requirements. For example, a user may only want to carpool with female candidates whose schedule is 9:00a.m.-5:30p.m.

Spatial Filter. In this process, the user may filter candidates according to spatial requirements. For example, a user may only want to carpool with candidates who live within 2000 meters of the Zhongshan Road West ($\underline{Fig 10}$).

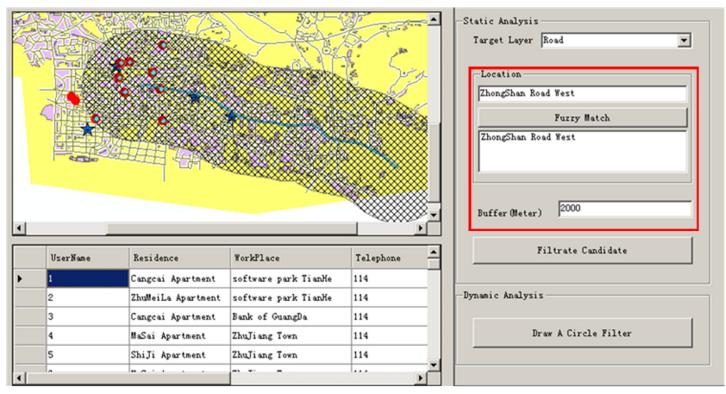


Fig 10. Spatial Filter (allow this image to be published CC BY 3.0 license).

doi:10.1371/journal.pone.0129257.g010



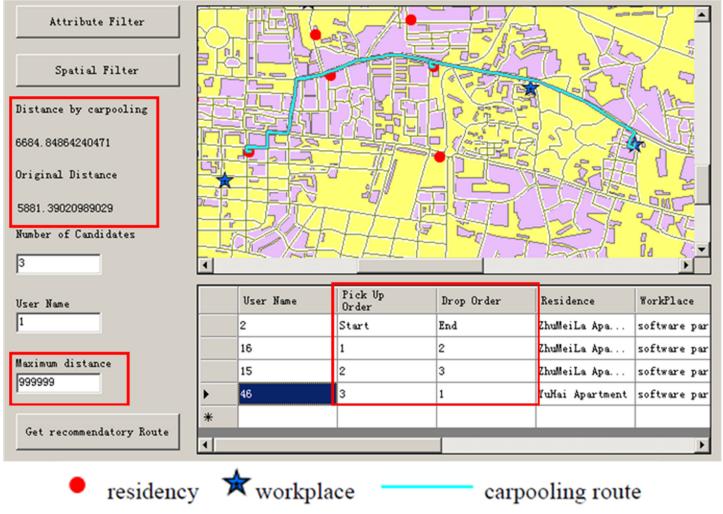


Fig 11. Carpool matching function (allow this image to be published CC BY 3.0 license).

doi:10.1371/journal.pone.0129257.g011

Determine carpooling passengers and daily commute route. After the attribute and spatial filters, the user must input the number of people in the carpooling team and the maximum commute distance. Then, the system executes the carpool matching model (based on the inputs above) and provides a recommended carpooling team. Information regarding recommended team members is shown in the system, such as user name, email, home address, workplace, etc. The daily carpooling commute route is shown on the map, specifying the order in which to pick up and drop off passengers. Moreover, the original commute distance (without carpooling), and the carpooling commute route distance are shown in the system so that the user can know the travel cost incurred through carpooling (Fig 11). If a passenger (e.g., P2) refuses the ride, the system will update the carpool route by choosing a back-up route that does not include passenger P2.

Conclusion and future work

One purpose of this research was to understand the issues regarding the provision of carpool matching services. From the literature we know that carpooling is an important element of the transportation structure of large cities, and we know that there are matching and solution

issues that must be addressed in any implementation. The research results do-in fact-expand our understanding in a manner that allows developers going forward to tune their carpool matching services to the kinds of problems that need to be solved (e.g. daily commute matching v. long term carpool determination). Specifically, this research provides an examination of a range of solution procedures that are appropriate under varying conditions. A carpool matching system prototype was constructed to demonstrate the efficacy of these methods. Perhaps most importantly, this research has shown that optimal 3+ person teams can be constructed with their associated routes for many common commute situations. This solution to a known technological problem represents a contribution to both the literature and practice of carpooling.

Several avenues for future research immediately present themselves. First, the finding that brute force search was nearly as effective as linear programming over a significant number of problem instances was surprising. This suggests that there is something idiosyncratic about the structure of this problem that lends itself to enumeration for a particular range of the parameters. If this can be generalized to other problems, that finding may change some approaches to practical problem-solving. Second, we know from the literature that ideas such as "comfort", culture, and level of trust are important in the success of carpooling activities. It may be that elements of social network analysis, qualitative research, and even psychology could be incorporated into a more sophisticated carpool matching model, in order to increase the carpooling population, improve carpool efficiency, and construct more lasting carpool teams. A related issue is the incorporation of more complex cost measures, including individual passenger costs, which may vary based on the pick-up and drop-off order. A related issue is that of directional variation in cost. It is well known that the costs to traverse a given arc can be very different based on the direction of travel, particularly during periods of heavy traffic (e.g. rush hour). If variable costs based on direction were implemented through the addition of arcs to the network, that could significantly increase the complexity of any given problem instance. Other means of integrating variable costs (directional, temporal, or otherwise) could similarly create greater difficulty in finding optimal solutions. Since these issues are so common they deserve immediate attention in future research. Finally, while the testing of methods, and even the construction of the prototype, can be accomplished in the research computer lab, any implementation must instead be exposed to the true population of a metropolitan area via the internet. That test will bring with it significant challenges of a different nature from what was undertaken in this research. It is hoped that the results presented here will encourage researchers to pursue these and other problems regarding carpool matching.

Supporting Information

S1 File. DataAvailabilityStatement.doc. The Data Availability Statement (DOCX)

S2 File. ExperimentalData1.zip. The distance matrixes of randomly generated sets of potential passengers ranging from 50 to 700 passengers, incremented by 50 (ZIP)

S3 File. ExperimentalData2.zip. The distance matrixes of randomly generated sets of potential passengers ranging from 750 to 900 passengers, incremented by 50 (ZIP)

S4 File. ExperimentalData3.zip. The distance matrixes of randomly generated sets of potential passengers ranging from 900 to 1,000 passengers, incremented by 50, and the street network

for Guangzhou city, China ESRI Geodatabase format (ZIP)

Author Contributions

Conceived and designed the experiments: KMC JX WL YZ. Performed the experiments: JX KMC. Analyzed the data: JX KMC. Contributed reagents/materials/analysis tools: JX KMC. Wrote the paper: JX KMC. Experimental data: WL YZ.

References

- 1. Baumert K, Herzog T, Pershing J (2005). Navigating the Numbers: Greenhouse Gas Data and International Climate Policy. World Resources Institute.
- Abu-Lebdeh G, Benekohal RF (2003). Design and evaluation of dynamic traffic management strategies for congested conditions. Transportation Research Part A: Policy and Practice, 37(2), 109–127. doi: 10.1016/S0965-8564(02)00006-X
- 3. Lo HK, Chang E, Chan YC (2001). Dynamic network traffic control. Transportation Research Part A: Policy and Practice, 35(8), 721–744. doi: 10.1016/S0965-8564(00)00014-8
- 4. Ritchie SG (1990). A knowledge-based decision support architecture for advanced traffic management. Transportation Research Part A: General, 24(1), 27–37. doi: 10.1016/0191-2607(90)90068-H
- Federal Highway Administration (2013). High Occupancy Vehicle Facilities. U.S. Department of Transportation. Retrieved from http://ops.fhwa.dot.gov/freewaymgmt/fag.htm#fag7
- Chan ND, Shaheen SA (2012). Ridesharing in North America: Past, Present, and Future. Transport Reviews, 32(1), 93–112. doi: <u>10.1080/01441647.2011.621557</u>
- Ferguson E (1997a). The rise and fall of the American carpool: 1970–1990. Transportation, 24(4), 349–376. doi: 10.1023/A:1004928012320
- 8. Park H, Gebeloff R (2011). Car-Pooling Declines as Driving Becomes Cheaper. New York Times. New York.
- Buliung RN, Soltys K, Bui R, Habel C, Lanyon R (2010). Catching a ride on the information super-highway: toward an understanding of internet-based carpool formation and use. Transportation, 37(6), 849–873. doi: 10.1007/s11116-010-9266-0
- Shewmake S (2012). Can Carpooling Clear the Road and Clean the Air? Evidence from the Literature on the Impact of HOV Lanes on VMT and Air Pollution. Journal of Planning Literature, 27(4), 363–374. doi: 10.1177/0885412212451028
- Cassidy MJ, Jang K, Daganzo CF (2010). The smoothing effect of carpool lanes on freeway bottlenecks. Transportation Research Part A: Policy and Practice, 44(2), 65–75. doi: <u>10.1016/j.tra.2009.11.</u> 002
- Small KA, Winston C, Yan J, Baum-Snow N, Gómez-Ibáñez JA (2006). Differentiated Road Pricing, Express Lanes, and Carpools: Exploiting Heterogeneous Preferences in Policy Design [with Comments]. Brookings-Wharton Papers on Urban Affairs, 53–96. doi: 10.2307/25067428
- 13. Poole RW, Balaker T (2005). Virtual exclusive busways: improving urban transit while relieving congestion. Los Angeles, CA: Reason Foundation.
- Levin IP (1982). Measuring tradeoffs in carpool driving arrangement preferences. Transportation, 11 (1), 71–85. doi: <u>10.1007/BF00165595</u>
- Li J, Embry P, Mattingly SP, Sadabadi KF, Rasmidatta I, Burris MW (2007). Who Chooses to Carpool and Why?: Examination of Texas Carpoolers. Transportation Research Record, 2021(-1), 110–117. doi: 10.3141/2021-13
- Correia G, Viegas JM (2011). Carpooling and carpool clubs: Clarifying concepts and assessing value enhancement possibilities through a Stated Preference web survey in Lisbon, Portugal. Transportation Research Part A: Policy and Practice, 45(2), 81–90.
- Psannis KE, Hadjinicolaou M, Ishibashi Y (2006). Designing a Pervasive Architecture for Car Pooling Services. In Elleithy K., Sobh T., Mahmood A., Iskander M., Karim M. (Eds.), Advances in Computer, Information, and Systems Sciences, and Engineering (pp. 211–218). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/1-4020-5261-8 34
- Zhu J, Feng Y, Liu B (2013). PASS: Parking-Lot-Assisted Carpool over Vehicular Ad Hoc Networks. International Journal of Distributed Sensor Networks. doi: <u>10.1155/2013/491756</u>

- Chen R, Regan A (2009). Wireless Networks for Car- and Ridesharing Systems: Assessment of 802.11 Wi-Fi. Presented at the Transportation Research Board 88th Annual Meeting. Retrieved from <u>http://trid.</u> <u>trb.org/view.aspx?id=880820</u>
- Liu N, Liu M, Cao J, Chen G, Lou W (2010). When Transportation Meets Communication: V2P over VANETs. In 2010 IEEE 30th International Conference on Distributed Computing Systems (ICDCS) (pp. 567–576). doi: 10.1109/ICDCS.2010.83
- Dailey DJ, Loseff D, Meyers D (1999). Seattle smart traveler: dynamic ridematching on the World Wide Web. Transportation Research Part C: Emerging Technologies, 7(1), 17–32. doi: <u>10.1016/S0968-090X(99)00007-8</u>
- 22. Steger-Vonmetz C (2005). Improving modal choice and transport efficiency with the virtual ridesharing agency. New York: leee.
- Abrahamse W, Keall M (2012). Effectiveness of a web-based intervention to encourage carpooling to work: A case study of Wellington, New Zealand. Transport Policy, 21, 45–51. doi: <u>10.1016/j.tranpol.</u> 2012.01.005
- Jiau MK, Huang SC, Lin CH (2013). Optimizing the Carpool Service Problem with Genetic Algorithm in Service-Based Computing. In Services Computing (SCC), 2013 IEEE International Conference on (pp. 478–485). IEEE.
- Huang SC, Jiau MK, Lin CH (2014a). A Genetic-Algorithm-Based Approach to Solve Carpool Service Problems in Cloud Computing. IEEE Transactions on Intelligent Transpiration Systems, 16 (1), 352– 364.
- Huang SC, Jiau MK, Lin CH (2014b). Optimization of the Carpool Service Problem via a Fuzzy Controlled Genetic Algorithm. IEEE Transactions on Fuzzy Systems. 1(99), 1.
- Agatz NA, Erera AL, Savelsbergh MW, Wang X (2011). Dynamic ride-sharing: A simulation study in metro Atlanta. Transportation Research Part B: Methodological, 45(9), 1450–1464.
- Furuhata M, Dessouky M, Ordó00F1ez F, Brunet M-E, Wang X, Koenig S (2013). Ridesharing: The state-of-the-art and future directions. Transportation Research Part B: Methodological, 57, 28–46. doi: 10.1016/j.trb.2013.08.012
- Agatz N, Erera A, Savelsbergh M, Wang X (2012). Optimization for dynamic ride-sharing: A review. European Journal of Operational Research, 223(2), 295–303. doi: <u>10.1016/j.ejor.2012.05.028</u>
- **30.** Glover F (1989). Tabu Search—Part I. ORSA Journal on Computing, 1(3), 190–206. doi: <u>10.1287/ijoc.</u> <u>1.3.190</u>
- **31.** Glover F (1990). Tabu Search—Part II. ORSA Journal on Computing, 2(1), 4–32. doi: <u>10.1287/ijoc.2.1.</u> <u>4</u>
- Dorigo M, Gambardella LM (1997). Ant colonies for the travelling salesman problem. Biosystems, 43 (2), 73–81. doi: <u>10.1016/S0303-2647(97)01708-5</u> PMID: <u>9231906</u>
- Michalewicz Z (1994). Genetic Algorithms Plus Data Structures Equals Evolution Programs (2nd ed.). Secaucus, NJ, USA: Springer-Verlag New York, Inc.
- Brooks SP, Morgan BJT (1995). Optimization Using Simulated Annealing. Journal of the Royal Statistical Society. Series D (The Statistician), 44(2), 241–257. doi: <u>10.2307/2348448</u>