A Game Theoretic Approach for the Taxi Scheduling Problem with Street Hailing

Ruibin Bai, Jiawei Li, Jason A D Atkin, Graham Kendall

Abstract This paper describes a taxi scheduling system to improve the efficiency of taxi services which are mainly implemented by means of street hailing. This is of particular relevance to Chinese cities, where this is by far the most common way in which taxis are requested, since the majority of taxi drivers are independent rather than being affiliated with a specific company. The mobile phone and GPS-based taxi scheduling system which is described in this paper aims to provide a decision support system for taxi drivers and facilitates information exchange between taxi drivers and passengers, while allowing drivers to remain independent. The taxi scheduling problem is considered to be a non-cooperative game between taxi drivers and a description of this problem is given in this paper. We adopt an efficient algorithm to discover a Nash equilibrium, such that each taxi driver and passenger cannot benefit from changing their assigned partner. Two computational examples are given to illustrate the effectiveness of the approach.

1 Introduction

Taxis play an important role in modern public transportation networks, especially in countries such as China and India, where bus and underground train services are still under development and do not currently have the coverage required by passengers. Taking a taxi is a convenient way to get to your destination, although it is usually slightly more expensive than buses and subways. According to a 2006 survey, there are approximately one million taxis in service around mainland China and the two million people who are involved in the taxi industry turnover more than 29.5 billion US dollars annually [1].

We can broadly classify taxi services into three categories: street hailing, taxi-stands and prearranged bookings [21]. Street hailing is a common method for getting a taxi, especially in small cities where the taxi service is still at a relatively small scale. Street hailing is simple and has no requirement for expensive equipment or systems. However, as will be analyzed later in this section, street hailing causes several problems, including low taxi utilization, long passenger waiting times, road safety issues, and traffic congestion. Taxi-stands have fewer of these problems but can only serve a small group of passengers at fixed locations, for example, Rubin Bai
Division of Computer Science, University of Nottingham Ningbo, China, Ningbo, 315100, China.
E-mail: rubin.bai@nottingham.edu.cn

Jiawei Li
School of Computer Science, University of Nottingham, Nottingham, NG8 1BB, UK
E-mail: jwl@cs.nott.ac.uk

Jason A D Atkin
School of Computer Science, University of Nottingham, Nottingham, NG8 1BB, UK
E-mail: jaa@cs.nott.ac.uk

Graham Kendall
School of Computer Science, University of Nottingham, Nottingham, NG8 1BB, UK
E-mail: gxk@cs.nott.ac.uk
city centers, bus/train stations, airports, etc. Prearranged booking is widely considered to be a more advanced and efficient taxi dispatching method, but is impractical when there are many independent taxi drivers. The booking can be performed via traditional telephone calls or by utilizing other modern communication technologies (such as websites or text messages). A taxi dispatching center is often employed to collect all the passenger information and coordinate the taxi dispatching centrally. The dispatching systems which are used often utilize the latest GPS devices and wireless communication equipment. The advantage of such a pre-booking system is that the demands of passengers are clear and known in advance. One drawback is that the system is often expensive and, therefore, suitable only for large taxi companies. In addition, the system relies upon mutual trust between taxi drivers and passengers, that neither party will break the agreed appointment. Unfortunately, in reality, such breaking of appointments happens frequently in many countries.

Unlike in the EU and USA, prearranged booking is not a popular method for hailing taxis in China. Most taxi drivers are actually self-employed, although they may be registered as employees for some taxi companies. It is hard to build a prearranged booking mechanism where taxi drivers who are self-employed (or employed by many different taxi companies) would work cooperatively. An efficient communication network would also be needed to implement this type of taxi scheduling, especially in a city with a large population. In China, a very limited number of taxi companies provide a prearranged booking service, except in some large cities such as Beijing and Shanghai. Therefore, street hailing is still the most frequent method of getting a taxi in China and approximately 85% of taxi services are implemented in this way. However, as mentioned earlier, street hailing is inefficient and poses both safety and congestion problems. Taxi drivers are unable to predict exactly when and where to meet the next passenger and passengers cannot be sure how long they will have to wait for the arrival of a taxi, leading to high rate of wasted mileage for taxi drivers and longer waiting times for passengers. In addition, taxi drivers need to concentrate on both driving and finding the next passenger, increasing the risk of accidents. Frequently, too many taxis crowd the city centers and passenger hotspots, leading to congestion problems.

There have been several research projects and technological innovations for pre-arranged booking systems, often abstracted and formulated as the Dial-A-Ride (DAR) problem, which falls into the category of classic Vehicle Routing Problems (VRP), or the Traveling Repair Problem (TRP) [6]. Several on-line algorithms have been developed, using linear programming [7], heuristics [8-10] and meta-heuristics [11]. Global positioning systems (GPS) and Geographic information systems (GIS) have been used by vehicle dispatching systems in many cities, to improve the efficiency of the taxi service [2-5]. For example, a GPS-based Automatic Vehicle Location and Dispatching System (AVLDS) greatly improved the quality and efficiency of taxi services in Singapore. Once a request for a taxi is received, the control center will send signals to the taxis which are located within 10km of the passenger. If one of the drivers accepts the job, he/she can respond to the control center by merely pressing a button which is installed in his/her taxi. Once the job has been assigned, information about the taxi number and expected time of arrival will be immediately sent to the passenger [2, 3].

Research and innovation to improve the street hailing method has been very limited. The most relevant research that we are aware of includes models of the taxi market to analyze the demand and supply equilibrium [12, 13], the bilateral searching and meeting between passengers and taxis on a road network [14], and the taxi cruising policy, to improve the efficiency of the taxi service [15]. However, these models do not consider scheduling algorithms for street hailing. The lack of research in this field is probably due to the fact that most taxis that are involved in street hailing are self-employed and independent. Unlike in a centralized system, coordination and cooperation between these taxis is very difficult, or even impossible, without a good mechanism. The systems and dispatching methods which are used in the pre-arranged booking method are simply not suitable for street hailing. The independent nature of the taxi drivers using such a system cannot be ignored, and perhaps the most important consideration for such a system is ensuring that it allocates passengers to taxis so that it is always in the best interests of the taxi drivers to keep the passengers they have been
allocated. The focus of this paper is upon a novel system for helping with the street hailing problem, to complement existing taxi booking systems.

In the remaining sections, we first describe the street hailing problem that we want to address. Section 3 introduces a novel game theoretic model of street hailing, along with a solution method for the problem. We then adopt the proposed approach and provide two computational examples in Section 4. Finally, a brief summary of the conclusions of this study and the directions for future work are presented in Section 5.

2 Street Hailing Using a New System

We propose to use the following system to improve the taxi street hailing process. The system adopts the latest GPS, GIS, and 3G/GPRS wireless network and comprises a central server and independent client applications for the passengers and taxis (see Figure 1). The server and clients are connected via a 3G/GPRS network. In our implementation, the clients are smart phones with GPS and GIS support and automatically send location and status (occupied or vacant) information to the server via the 3G/GPRS network. The server periodically analyses the geographical distribution of clients on a digital map, generates a stable taxi-passenger scheduling and then broadcasts the allocations back to clients.

In practice, there may be several ways to perform the taxi scheduling under this new system. A straightforward approach would be providing passengers and taxis with a two-way communication channel to negotiate a manual schedule, although this may not be efficient. An improved design would provide automatic decision support for the taxis and passengers by taking full advantage of the geographical distribution information of the passengers and taxis on a digital map. This paper focuses on the latter implementation and proposes a scheduling procedure that is illustrated in Figure 2. As can be seen from the figure, the city map is divided into smaller zones to facilitate better taxi management and unnecessary complexity. A schedule is generated periodically (defined by the parameter $t_w$) rather than in real-time. To avoid confusion, the following assumptions are also made about the problem under consideration in this paper:

1. Taxis are independent and compete against each other in order to gain maximum personal return (i.e. there is no subset of taxis that plays a cooperative game).
2. A taxi’s primary goal is to find the next passenger within the shortest possible time.
3. Scheduling recommendations by the system are only supportive, not compulsory. Therefore, a taxi can decline a recommendation by the system in certain circumstances (e.g. if the recommended passenger is too far away, although the division of the city into zones should reduce these occurrences).
4. Locations of passengers and taxis are automatically retrieved from the embedded GPS chip and sent to the server by the software client.

5. A passenger’s destination information is not known when a schedule is generated. Therefore, this system cannot plan a taxi’s route beyond its very next service. This means the system is not a pre-booking taxi system but serves as a supplement system for the latter.

---

We note that some of these assumptions may be relaxed in practical implementations. The main purpose of their inclusion in this paper is to introduce a basic problem that can be studied initially to gain some insightful knowledge, and then be extended later. We acknowledge that the research reported in this paper is still in the early stages as far as fully addressing the practical problems and that several challenging but interesting practical problems have arisen. We believe that the research results reported in this paper, as well as the new problems raised...
in this paper, would be of interest for other researchers and could potentially lead to increased research attention for this important research area.

3 A Game Theoretic Model

In contrast to the usual algorithms for the DAR problem, we develop a game theoretic model that treats the taxi scheduling problem as a non-cooperative game between taxi drivers. The taxi drivers are assumed to be self-interested and to compete rather than cooperate with each other in serving passengers. Given the locations of passengers and other taxis, each taxi driver will have preferences for some passengers over others and each passenger may also have preferences for taxis (e.g. a preference to be picked up as quickly as possible). We aim to provide an algorithm to generate a matching of taxis and passengers that will be accepted by all taxis.

A matching of taxis and passengers is a Nash equilibrium if no taxi or passenger can find a better partner than their current one, so that nobody has any incentive to deviate from the suggested allocation. In other words, any alternative taxi a passenger may wish to switch to would prefer its currently allocated passenger and any alternative passenger which a taxi may wish to switch to would prefer its currently allocated passenger, so no unilateral decision by a taxi or passenger will result in a benefit for that taxi or passenger. Such an allocation would then be appropriate for utilization in a system to support street hailing.

Let \( T_i \) (\( i = 1, \ldots, n \)) denote the taxis, \( P_j \) (\( j = 1, \ldots, m \)) the passengers and \( t_{ij} \) time for taxi \( i \) to reach passenger \( j \). Let \( E_{ij} \) represent the preference of taxi driver \( i \) for passenger \( j \). We consider a computerized street-hailing system in this paper rather than a pre-booking system and we assume that the taxi drivers are not given information about the destination for any passenger. Taxi drivers can, therefore, only utilize the \( t_{ij} \) information for their passenger preferences and will prefer passengers which are closer to them, as previously discussed. We therefore assume that \( E_{ij} \) is entirely determined by \( t_{ij} \) and is negatively correlated with \( t_{ij} \). For the experiments for this paper, we assume that \( E_{ij} \) is defined as follows, in order to meet these requirements:

\[
E_{ij} = \frac{1}{t_{ij}}, \quad i = 1, \ldots, n; j = 1, \ldots, m \tag{1}
\]

The ability of taxi \( A \) to take a passenger off another taxi \( B \) to which it has been assigned relies upon both the passenger being willing and taxi \( A \) being able to reach the passenger before taxi \( B \). We assume that passengers have no preference for taxis other than that they wish to be picked up as early as possible, so these two factors can be considered to be identical in this case. A passenger preference, representing the ability of a taxi driver to take this passenger from another taxi driver, could then be determined by assuming that the preference of passenger \( P_j \) for taxi \( T_i \) is negatively correlated with \( t_{ij} \). Taxi \( A \) can then take passenger \( j \) away from another taxi \( B \) if and only if \( t_{ij} < t_{ij} \), i.e. if it can reach the passenger first.

If we consider the taxi drivers’ preferences \( E_{ij} \) as the scoring of one partner and the aforementioned passenger preference as the scoring of the other partner then the taxi-passerger pairing problem can be considered to be a variant of the stable marriage problem. The stable marriage problem [17,18] involves \( n \) men and \( n \) women, each of whom has ranked all members of the opposite sex with a unique number between 1 and \( n \), in order of preference. The objective is to find a set of man-woman pairs such that there are no two people of the opposite sex who would both rather have each other than their current partners. The Gale-Shapley algorithm [16] was developed to solve this problem in polynomial time. This algorithm involves a number of ‘rounds’. In each round, each unengaged man proposes to the most preferred woman to whom he has not yet proposed. Each woman then considers all proposals to her and chooses the one she most prefers as her partner and becomes engaged to him, potentially ‘trading up’ from an existing engagement. This process continues until there is no unengaged man. The Gale-Shapley algorithm guarantees a stable matching.

The Gale-Shapley algorithm is still applicable when there are an unequal number of partners, i.e. \( n \neq m \), so can be applied to the taxi-passerger matching problem. The algorithm is
either man-oriented or woman-oriented. That is, each man (or woman) is paired with the best partner that they can have in any stable matching. In the case of \( n \neq m \), a stable matching is guaranteed, but some men (or women) will not be paired.

For the taxi scheduling problem considered here, given the definition of \( \{E_{ij}\} \) and the consequent correlation between the two rankings, the Gale-Shapley algorithm can be simplified. We have adopted an algorithm which first sorts all elements in the preference matrix \( \{E_{ij}\} \) into decreasing order, then considers each item in turn, from the first to the last. For each element \( E_{ij} \), if neither taxi \( i \) nor passenger \( j \) has been allocated then \( i \) is allocated to \( j \) and \((i, j)\) becomes part of the stable matching. The time complexity of the proposed algorithm is \( O(mn \times \log(mn)) \) (since the sorting is the slowest element). It is easy to verify that the stable matching is a Nash equilibrium, since the only passengers that a taxi driver could beneficially switch to will have been allocated to a taxi which is even closer to them.

There must be at least one stable matching (Nash equilibrium), given the preference matrix \( \{E_{ij}\} \), however, it could be difficult to find the Nash equilibria since the search space of potential combinations of rankings is so large. The Gale Shapley algorithm will find one of the equilibria, but it is probable that there is more than one stable matching and the others may be much harder to find. In game theory, the price of anarchy (PoA) is used to describe the ratio between the maximum social payoff and the payoff from an equilibrium [19]. We use it here to denote the difference between the total preference value of the stable matching which we have found and the preference value of the optimal stable matching. As an example, Figure 3 shows example locations for two taxis and two passengers, showing the selected and possible pairings. In Figure 2(a), both \( \{T_1, P_2\} \) \( \{T_2, P_1\} \) and \( \{T_1, P_1\} \) \( \{T_2, P_2\} \) are equilibria because \( T_1 \) is indifferent between choosing \( P_1 \) and \( P_2 \), but is closer to both passengers than \( T_2 \) is. However, \( \{T_1, P_2\} \) \( \{T_2, P_1\} \) has a better overall preference, thus a lower PoA. In Figure 2(b), both taxis \( T_1 \) and \( T_2 \) prefer \( P_2 \) to \( P_1 \) and there are two equilibria. The equilibrium \( \{T_1, P_2\} \) \( \{T_2, P_1\} \) has a higher total preference, so a lower PoA.

Figure 3 There are multiple equilibria because (a) taxi \( T_1 \) is indifferent between choosing \( P_1 \) and \( P_2 \) and (b) both \( T_1 \) and \( T_2 \) prefer choosing \( P_2 \) to \( P_1 \). The equilibrium with lower PoA is \( \{T_1, P_2\} \) \( \{T_2, P_1\} \) in both cases.

In order to minimize the PoA, a system may need to compute all Nash equilibria, to find the equilibrium with the highest total preference, but this would lead to an unacceptable amount of computation when the numbers of taxis and passengers are large. However, we have found that the differences between the total preference values for the equilibria are actually very limited in the taxi scheduling problem, as we will show in section 4.1.

4 Computational Examples

In this section, we simulate the taxi scheduling problem in a square area 50km x 50km. Taxi drivers and passengers are most likely to appear around hotspots such as business centers, airports, train stations and so on. The locations of passengers were approximately distributed around these hotspots according to the Poisson distribution [20, 21]. The distance between a
taxi and a passenger was determined by the Euclidean distance between their locations. Distances were considered to an accuracy of 0.15km, so that two distances were considered to be same if their difference was less than 0.15km.

4.1 Optimal matching

In this simulation, a taxi scheduling problem of 1,000 taxis and 1,000 passengers was solved using the simplified Gale-Shapley algorithm described earlier. The locations of taxis and passengers are shown in Figure 4. There were three hotspots.

![Figure 4 Distribution of 1,000 taxis and 1,000 passengers in a square area (taxis are denoted by x marks and passengers are denoted by dots).](image)

The optimal matching of 1,000 pairs of taxi and passenger was also computed and the distances between the taxis and the passengers are shown in Figure 5(a), ordered from smallest to largest. Figure 5(b) shows the number of taxis which fall into different distance ranges. It shows that most of the scheduled taxis travel only short distances to their passengers, while a small number of taxis need to travel long distances (more than 14km).
Figure 5 Optimal matching of 1,000 taxis and 1,000 passengers. (a) Distances to passenger of 1,000 taxis. (b) Statistic of number of taxis corresponding to travel distance.

We note that this scheduling system only provides decision support for taxi drivers. Taxi drivers will not necessarily choose to collect an assigned passenger or use the assigned route. In practice, a taxi should not be assigned a passenger who is very far away because doing so leads to a high, non-profitable mileage for the taxi driver and an intolerable waiting time for the passenger. Experiments were executed to compare the optimal solutions and the stable equilibria which the previously described algorithm obtained. We observed that, if we ignore the taxis that are assigned to distant passengers, there was little difference between the optimal matching and the stable equilibria. An example of such a comparison is given in Figure 6. The reason is that most taxis and passengers are located around the hot spots and the resulting short distances will be prioritized by the pairing method we use to achieve a stable matching. The amount of computation could be greatly reduced if we do not need to find the optimal matching and these results imply that this is practical.

Figure 6 A comparison between the optimal matching and a non-optimal stable matching.

Although taxis which are assigned distant passengers would not necessarily pick up the passengers, the system could still be of use to these taxi drivers because it provides information about the distribution of passengers and other taxis. It can, therefore, help the taxi driver to choose where to go next to find the next passenger.
For comparison, a FCFS strategy was developed where the passengers were allocated to taxis one at a time according to the time sequence of their requests, and each passenger was assigned to the nearest taxi which was available. FCFS is a ‘fair’ algorithm for passengers but it increases the average vacant mileage for taxi drivers compared with the optimal matching. We show the comparison of the first come first serve (FCFS) strategy against the optimal matching in Figure 6.

Figure 7 Comparison of the optimal matching with the result of FCFS strategy (the dashed line denotes the result of the FCFS strategy).

4.2 Algorithms for dynamic scheduling

In practice, taxi scheduling is a dynamic process and the time window for the request of taxi service needs to be taken into consideration. We simulated a scheduling problem of 5,000 taxis and 50,000 passengers and present the results here. The locations of taxis, passengers, and their destinations were within a square area. Travel times were calculated based upon the distance between the taxi and customer positions. Of course, taxi drivers are able to reject passengers (for example if a passenger was too far away) and would normally do so straight away in that case. The travel times for taxis (and the waiting times for serviced passengers) would then be lower but there would be un-serviced passengers. We assume for these experiments that taxi drivers do not reject passengers and therefore some passengers have relatively large waiting times. In these experiments, the requests for taxi service from the passengers were randomly distributed around rush hour according to the Poisson distribution, in addition to the random distributions of their locations and destinations. Three strategies were studied:

a) First come first serve (FCFS). The nearest vacant taxi will be assigned immediately when the request for service is received.

b) Strategy 1: The scheduling is performed every $t_w$ minutes, where $t_w$ is a parameter specifying a service window size. For these experiments, $t_w$ was set to 5 minutes. The passengers whose requests for service are received within the 5 minute window are all scheduled together using the algorithm described in Section 3.

c) Strategy 2: A combination of the FCFS strategy and strategy 1. A passenger will be immediately served according to the FCFS strategy if there is a vacant taxi within 10km distance (the 10km being another adjustable algorithm parameter), otherwise, strategy 1 will be adopted and a taxi will be allocated at the end of the current service window.

The results of this simulation are shown in Table 1. With strategy 1, there is a delay between the passenger requesting service and the service being scheduled (of up to 5 minutes
in the worst case). Although the average vacant mileage is decreased for the taxis compared with FCFS, the saving in passenger waiting time is counteracted by the scheduling delay. Strategy 2 is better than the other strategies in terms of its low waiting time of passengers.

Table 1 Result of simulation

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Taxis</th>
<th>Passengers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average mileage</td>
<td>Average vacant mileage</td>
</tr>
<tr>
<td>FCFS</td>
<td>48.28</td>
<td>9.70</td>
</tr>
<tr>
<td>Strategy 1</td>
<td>52.01</td>
<td>6.45</td>
</tr>
<tr>
<td>Strategy 2</td>
<td>51.79</td>
<td>6.75</td>
</tr>
</tbody>
</table>

* Vacant mileage: travel distance for a taxi to reach the location of passengers.
** Waiting time: the time interval between receiving of the request of service and arrival of taxi.

The waiting times of passengers, mileages of taxis, and vacant mileages of taxis are shown in Figures 8-10 respectively.

Figure 8 Waiting times for passengers, (the passengers with waiting time more than 14 minutes are not shown).

Figure 9 Mileages for the 5,000 taxis.
The results of the simulation show that the proposed scheduling algorithm benefits both taxi drivers and passengers. It increases valid mileages and decreases vacant mileages for taxis. Low vacant mileages lead to low waiting times for passengers.

5 Conclusion and future research

Although this is only preliminary research into a scheduling algorithm for street hailing, we have described a low cost taxi scheduling system to improve the efficiency of taxi services. This would be particularly useful in Chinese cities, where the service is mainly implemented by means of street hailing. Based on mobile phones with GPS functionality, the system collects information about the locations of passengers and the locations and statuses of taxis and provides decision support for both taxi drivers and passengers. The simulation results in this paper show that our system could greatly improve the efficiency of street hailing.

The scheduling algorithm is based on a game theoretic model that considers the problem as a non-cooperative game between taxi drivers. An algorithm has been proposed which will discover a Nash equilibrium for this game, such that each taxi driver cannot find a better choice than their assigned passenger and route. The results of simulation show that this algorithm is efficient in computing the matching of taxis and passengers and it only needs small amount of computations. Thus, it is suitable for online taxi scheduling.

In addition to the time to pick up passengers, some other factors may influence a taxi driver’s preferences, such as individual experience about traffic conditions or favoritism for a specific area. It is worth taking these subjective factors into consideration, where possible, in implementing a practical system, although doing so will increase the complexity of the algorithm.

There is always uncertainty in the real world taxi scheduling problem. A passenger may change their plans or lose patience in waiting, or a taxi may be late because of unexpected traffic congestion. The influence of uncertainty on the proposed scheduling algorithm and its robustness in a real world environment in the presence of uncertainty will form part of our future research into this interesting and important problem.

References


