Approaches to Successful Entry of the Ride-sharing Service for Startups

Sungkyu Park
Kanizsa Lab
Room #505, Building #9, 85,
Hoegi-ro, Dongdaemun-gu,
Seoul, South Korea
shaun@kanizsalab.com
shaun.park@kaist.ac.kr

Suyeon Cho
Kanizsa Lab
Room #505, Building #9, 85,
Hoegi-ro, Dongdaemun-gu,
Seoul, South Korea
suyeon.cho@kanizsalab.com

Jaram Park
Kanizsa Lab
Room #505, Building #9, 85,
Hoegi-ro, Dongdaemun-gu,
Seoul, South Korea
qkfka315@gmail.com

Jaeyeon Won
Kanizsa Lab
Room #505, Building #9, 85,
Hoegi-ro, Dongdaemun-gu,
Seoul, South Korea
ruunak@kanizsalab.com

CHI'17 Extended Abstracts, May 06-11, 2017, Denver, CO, USA
© 2017 ACM. ISBN 978-1-4503-4656-6/17/05 $15.00
DOI: http://dx.doi.org/10.1145/3027063.3053340

Abstract
Ride-sharing can be a key factor in reducing traffic volume and CO2 emission for public transportation especially within urban area. However, it is not easy for small ride-sharing startups, which typically has limited budget and manpower, to boost the service in the early stages. In this paper, we present our on-demand service whose main operation is based on a mobile application for taxi-sharing which utilizes the users' trust network. We also explain in detail the three phased approaches we have applied in an attempt to increase the matching rate among the users. A high matching rate is crucial for survival of a ride-sharing service, because once a service fails to match customers in real-time, the retention rate falls sharply. Next, we analyze and interpret the implication of committing to each approach. Additionally, we also introduce one last method, which employs an AI agent to guarantee match making among users at all times by overcoming the time-spatial constraints.

Author Keywords
Ride-sharing; on-demand services; trust networks; time-spatial constraints; traffic records; geo-social data.

ACM Classification Keywords
H.2.8 [Database Applications]: Spatial Databases and GIS; H.4.m [Information Systems Applications]: Miscellaneous; C.4 [Performance of Systems]: Measurement Techniques

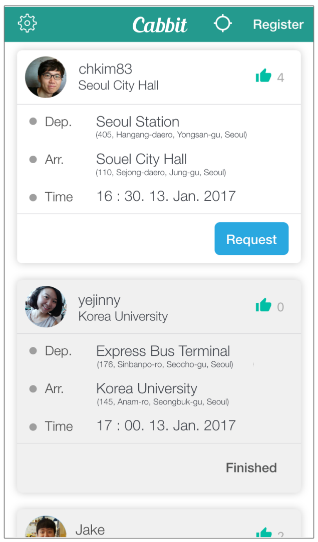

Figure 1: The home screen of Cabbit, showing registered trip info of users.

Introduction

Recently, people have started using the Web to connect offline and online entities on the real-time basis, and based on this movement, a new type of services called on-demand (meaning services that immediately provide its offline commodities when people request them online) has emerged [7]. The on-demand service model has been applied to various industries such as food delivery, home cleaning, car rental, as well as ride-sharing [1].

In this paper, we introduce our case, which is about practically running an on-demand service as a small startup, to share the insights we have gained with others who are interested in improving offline user experiences through online service, for example, by increasing user matching rate. We particularly pay attention to connecting people in terms of ride-sharing, one of the representative 'matching' market in the digital age [1], and to the development and operations of a mobile application named "Cabbit" (which is a portmanteau of a "cab" and a "rabbit") in the real world to allow people to easily find trustful companions in real-time.

Research on on-demand, ride-sharing, and HCI

The on-demand service is considered to be a better model than a traditional service having only one channel, since the on-demand can maximize the strength of online and offline by interactively combining two channels, as Internet (especially with the mobile capability) as well as e-commerce have risen [10]. However, the on-demand service also faces some difficulties as indicated by this study: one is that the on-demand service requires high level of localization, meaning it can mere provide services for people with a physically limited range. Another paper analyzed a survey and claimed that incorrect information about commodities, annoyance mainly caused by inordinate promotions, and etc. are the crucial factors hindering the continuance intention of users [3].

In regard to ride-sharing, previous literature confirmed that there is a high potential for matching people to share taxi by analyzing the rides network based on the New York taxi open dataset, which contains information about over 12M records of taxi trips [8]. Another study focused on the benefits of ride-sharing and predicted the reduction in traffic volume when ride-sharing service would be introduced to a city, considering time-spatial and social conditions [4]. Santi et al. computed the optimal trade-off point between passenger discomforts and benefits for sharing taxi rides, also based on the New York taxi open dataset [6].

Several other studies focused on finding significant factors for increasing the matching rate, related to user behaviors on the ride-sharing service. Mitja et al. [9] emphasized the importance of a meeting point among multiple passengers and the driver, the passenger flexibility, and a high matching rate for a successful ride-share. They particularly stressed the importance of a high matching rate since only users who have positive matching experience will revisit the service and introduce the service to others. While Lee and Savelsbergh [5] emphasized that the large number of participants is of important in ride-sharing, Boyac et al. [2] argued that the efficient relocation of cars, such as being at the right place at the right time is crucial for the successful matching between the passengers and the vehicles.

Relying on the above works [2] [3] [9] [10], we believe that especially in the early stages of on-demand services, the imbalance between the number of providers and consumers and the time-spatial constraints are big obstacles to matching users with providers, and as a result, it makes them

[1]The World on Demand: 23 Startups That Deliver Everything From Sexy Man Servants to Marijuana, https://www.entrepreneur.com/slideshow/272805.

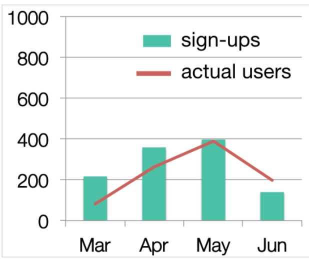

Figure 2: # of Sign-ups and actual daily users at the 1st approach stage (Mar-Jun).

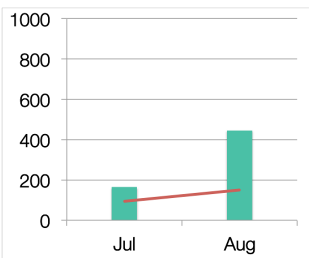

Figure 3: At the 2nd approach stage (Jul-Aug).

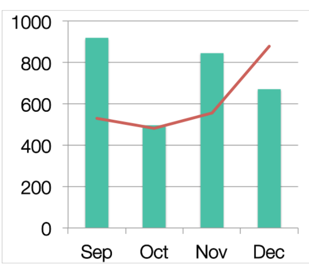

Figure 4: At the 3rd approach stage (Sep-Dec).

stop using the service. In that sense, we set up phased hypotheses toward the goal of a high matching rate upon passengers in an actual ride-sharing service and investigate to what extent our hypotheses fulfill the expectation.

Background: Mobile Taxi-sharing Service

Cabbit is a ride-sharing with using taxi (i.e., taxi-sharing) service that enables two passengers be matched on a pre-constructed trust network. Users are connected 1:1 with no participation from the taxi drivers, meaning that users should meet first and then find a taxi by themselves. Therefore, the departure point is the same for the both passengers, but their destinations does not necessarily need to be the same. Indeed, Cabbit has no provider side (i.e., one-sided market) so we need to make some consumers act a role of provider. This service model is different from Uber and Uber Pool, since many countries, including Korea, Uber and/or Uber Pool models are illegal [2]. Hence, we develop a more universal taxi-sharing service model that can be fit the needs of the people in various regions.

Figure 1 shows the home screen of Cabbit: the home screen consists of the list of registered trip information of users. Firstly, in order to find companions, users register trip information including points of origin-destination(OD), departure time, preferred gender (for safety reasons) into the application (See Figure 5 (a)), and then a list, made up of similar trip information from other users is recommended: the recommendation algorithm is based on the combination of time-spatial information of the trip, user demographics, and user social/behavioral context such as affiliation, recommender-recommendee relationship, matching history, and etc. After two users are matched, a chatting interface shows up and users can communicate about the detailed

[2]Uber in conflict with the law 71 places around the world, https://goo.gl/mofmW4.

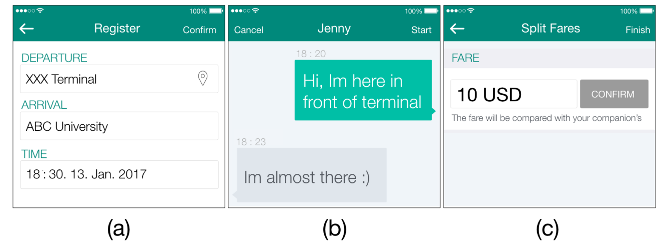

Figure 5: Three steps of using Cabbit: (a) registering trip info; (b) chatting with the matched companion; (c) splitting fares

meeting spot and time (See Figure 5 (b)). Lastly, users can split their fares for the shared trip with a one-stop fare distribution and payment system implemented into the application (See Figure 5 (c)).

Cabbit has been launched on March 16, 2016. Figure 2, Figure 3, and Figure 4 show the number of sign-ups and accumulated actual daily users in the stages of the first (March - June), the second (July - August), and the third approach (September - December), respectively. Each approach will be described in more detail later. If a user certifies their affiliation through affiliation mail account, we count them as a sign-up. The number of actual daily users is the number of users who register trip information on the application at least once a day. The total number of sign-ups is 3,793 and accumulated actual daily users is 3,621 as of December 31, 2016.

Throughout the first approach stage the number of actual daily users followed an increasing trend and so as the number of sign-ups. However both numbers sharply decreased after the summer vacation began in the middle of June in Korea (Figure 2). In case of the second approach, it didn't work well: Figure 3 shows that the number of actual daily users increased by a small margin compared to the number of sign-ups, and the absolute number is lower than all other stages as well. At the third approach stage the number of actual daily users steadily increased while the number of sign-ups are fluctuated over time (See Figure 4).

Fundamental Hypothesis: Building a Trust Network
Before start developing a taxi-sharing service, we wanted to know in advance the obstacles of ride-sharing. Although taxi-sharing has clear benefits such as splitting fares and saving time [3], this concept of service has not been much vitalized thus far in Korea. We conducted an online survey to understand people's attitudes towards taxi-sharing services. We arbitrarily sampled 200 university students and office workers in Seoul to ask questions and list up the main findings as follows:

- 56% of respondents were willing to use taxi-sharing services. They expected the following improvements from the service: 1) taxi should be available at certain time period (respondents mostly thirties to forties); 2) fares should be cheaper (mostly twenties).
- Among users who do not want to share a taxi, 90% of them answered that the reason was not wanting to take a taxi with someone they do not know.

Therefore, our fundamental hypothesis is that in order to create more demand and succeed a ride-sharing service it is essential to build a users' trust network, which let users feel safe and comfortable while sharing rides with strangers.

The First Approach: Providing Affiliation Information for Users Seeking Companions

- Period: March 16 - June 30.
- Hypothesis: Providing affiliation information of other users when choosing companions is enough for creating the initial user pool and making the users be matched within.
- Result: Not supported.

We set "affiliation" information as the baseline for connecting users. We set 'affiliation' as a mandatory information prior to signing up for Cabbit. Users' affiliations are shown with their trip information so that users can check other users' affiliations for their own safety (See Figure 1). As the user pool becomes larger and users start connecting with each other, our plan was to build a users' multi-layered trust network with various contexts such as connection histories, demographics, and time-spatial riding patterns.

Unfortunately, a crucial problem occurred soon after launching Cabbit with the initial setting: although we assumed that users from same universities would have similar moving patterns and be matched easily, users did not congregate in practice. Furthermore, once the users failed to meet companions they tend to leave the service, and we could not expect any viral effects. To overcome this obstacle, we planned the second approach, region-based clustering.

The Second Approach: Developing the Region-based Clustering

- Period: July 1 - August 31.
- Hypothesis: Clustering based on configured regions can expand the size of the user pool and make the users be matched within.
- Result: Not supported.

[3] In Seoul, Korea, it is hard to find a taxi due to imbalance between supply and demand in specific times such as late night on Friday.

Figure 6: Main taxi departure points (red) in Yeoksam 1-dong ($3.5km^2$): more than 60 companies located in the district.

Based on the open dataset of taxi trip logs [4] we clustered and selected one specific district, Yeoksam 1-dong Gangam-gu, from Seoul area as a starting point for taxi-sharing for the following reasons: there are 286,978 daily taxi passengers on average, and a total of 79% of the traffics are concentrated in the period from 7am to 10 am (i.e., commuting time to go to work). In addition, the points of origin are dense and close to subway stations (See Figure 6).

In our second approach, we expected that passengers would be able to share their rides in commuting times with others who work in the same districts. Hence, we mainly targeted the company employees within the selected business district as indicated in Figure 6. Although there were a lot of demands for taxis as predicted, our goal, which is increasing the number of matching, could not be accomplished and the second hypothesis was not supported, after the second approach had proceeded for two months. Regarding the case of 'time to go to work', passengers wanted to go to company without any delay, and saving of taxi fares was not the most important aspect for them, considering their economic level compared to the sacrifice of comfort. Regarding the case of 'time to go home', the points of origin are clustered within the district, yet the points of destination are not, which means users' homes are far apart and thus they cannot be matched. We concluded that while our initial targets (i.e., university members) were sensitive to taxi fares, company workers were relatively not. Moreover, most commuting passengers in Gangnam currently travel in short distance on a taxi, so taxi-sharing service was not so attractive for them.

Although we set the region-based clusters based on the public transportation logs, the configured extent of the clusters is still too large for the users to be congregated in terms of the time-spatial manner. Therefore, we developed another approach: clustering of the possible routes that can control time-spatial constraints.

[4] We analyzed a 100 million Korean taxi trip log in July, 2015, which contains OD info, time, weather, and so on so forth. The dataset was downloaded at http://data.seoul.go.kr (written in Korean).

The Third Approach: Developing the Route-based Clustering

- Period: September 1 - Present.
- Hypothesis: Clustering based on configured routes can expand the size of the user pool and make the users be matched within.
- Result: Supported.

The major reason for why the users could not be gathered has been analyzed, and we concluded that the configured size of clusters was too large. Hence, we analyzed the operation dataset (i.e., 19,736 points of searched and registered OD logs from March 16 to December 31) and experimented whether certain routes overlap within certain time-spatial range (i.e., 100 meter of the distance of origin and destination respectively and 15 minutes of the departure time) and could confirm there are some routes that meet the set condition.

Therefore, we contacted 42 student unions out of 110 candidate universities, and a total of seven student unions agreed to collaborate with us. Next, we started conducting online and offline promotions together with the unions. Consequently, the locations of the OD have become clustered across the country as we intended (see Figure 7).

User feedback on Cabbit
While preparing the third approach, we conducted a user survey with 10 sampled users, and the users gave us fa-

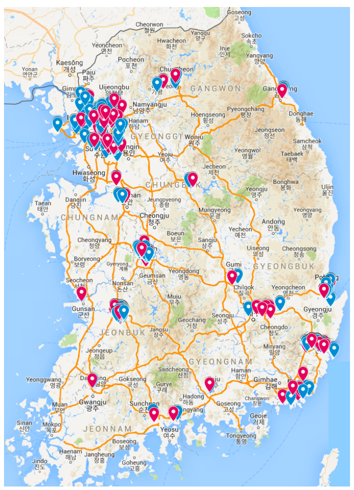

Figure 7: The location of registered trip info across the country from Mar to Dec (non-cumulative | origin: red | destination: blue)

vorable responses. Here are some quotes from the participants of the survey:
"I think the service is trustworthy since other users whom I can meet will be from the same univ."
"I like that you need to be certified through your affiliated email account. Due to this procedure, if some issues occurred during the process, it's easier to demand the responsibility from the users."
"I use taxi often especially at late nights when the buses stop running, and I like this service since I can save fares."

Yet, some participants adverted to the limitations of the service and suggested a few ways to improve, as follows:
"The matching should be enough for the service, and the fare payment system is not a must."
"I will use Cabbit only when I have plenty of time."
"I registered my trip info on Cabbit two to three times but I could not find other companions. After those experiences I have not used the application."

We found the biggest obstacle that prevents users from using Cabbit is that it cannot guarantee a 100% match between users. Users became reluctant to revisit the service after they had failed to benefit from the service. To compensate the limitation, we developed and introduced an artificial intelligence (AI) agent that can be connected with users who could not meet other companions with certain time-spatial condition. Through this agent, our service has become more reliable for users, meaning the users could always take a taxi with half-cut fare albeit the missing companions. Through locking in the users by guaranteeing the discounted taxi fares, we were able to let those consumers be our initial fans, so that they can perform as providers of the service and connect with other users followed by the initial fan group. We implemented this AI agent module for the third approach as well.

	Jul-Aug	**Nov-Dec**
# of Observed possible routes	8	12
# of Registered trip info within possible routes (A)	101	976
# of Total registered trip info (B)	620	2168
Percentage (A/B)	16.29%	45.02%

Table 1: Comparing # of registered trips for two month (Jul-Aug) from the 2nd approach and for two month (Nov-Dec) from the 3rd.

Results of The Third Approach

In order to assess to what extent our results could be improved with the third approach, we compare the results from the third Approach to those of the first and the second. By comparing Figure 2, Figure 3, and Figure 4, we can see that the performance metrics of the third approach (illustrated in Figure 4) outperformed those of the other two approaches. Particularly with Figure 4, the number of actual daily users has steadily increased. The number of sign-ups was the highest in September; however, the high number is mainly due to the seasonal effect: huge inbound of (i.e., from elsewhere to come into univ) movement came from the university students because of Chuseok (Korean thanksgiving day), the second biggest holiday in Korea.

Specifically to the impact of the route-based clustering, we observed whether the registered trip information was assembled by following the calculated routes, after we proceed with the third approach. As indicated in Table 1, we confirm that the third approach is properly working and users have congregated based on the designated routes.

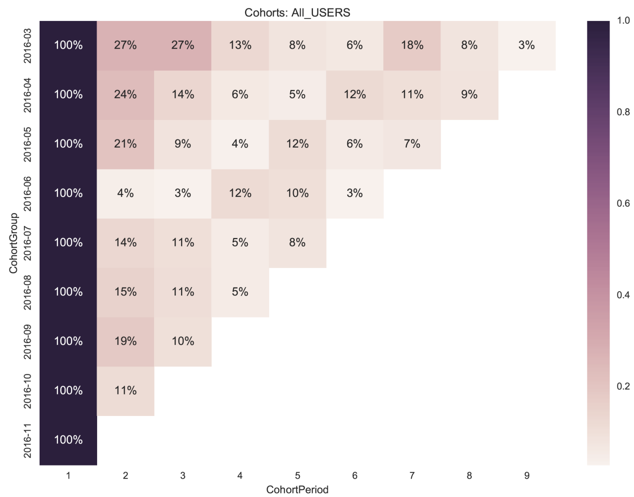

Figure 8: Cohort Group Analysis

Figure 8 shows the result of the cohort group analysis, which indicates how equivalent groups, split into months, after they firstly registered with their trip information, behave in the following months (i.e., showing how much percentage of users of one equivalent group registers trip information again in the following months, which is equal to the 'retention rate' for Cabbit). The retention rate is relatively high for the cohort group in March (i.e., the initial opening period): it can be considered as loyalty of the early adopters. Furthermore, the retention rate increases for every cohort group in August and September, because most of university members come back to universities after summer break in August as well as Chuseok in September. Thus, we can say that seasons noticeably affects user behaviors at Cabbit.

Yet, this result does not fully explain the effect of the AI agent, since some users could not be matched with the agent, even when they met every required condition for the third approach, due to the technical limitations in practice. Hence, to test the effect of the AI agent, we respectively set experimental group as a set of users who successfully connected with either other users or the AI agent at least once, and control group as a set of users who registered their trip information at least once but have not been connected at all. Then, within the experimental group we tracked the number of users who register trip information at least once more after the first matching, as for the control group, we observed the number of users who register trip information at least once more after the first registration of the trip information. The size of the experimental group is 1.60 times (point) larger than that of the control, as shown in Table 2. Based on the result, we can say that the users who have previous matching experience tend to reuse the service more than the users who do not have the experience.

User Group	Total #	# of Users who register trip info at least once more	Percentage
A	703	-	-
A-1	210	128	60.95%
A-2	493	188	38.13%

Table 2: Comparing # of revisiting users based on the user groups from Nov to Dec (Group A: Users who registered trip info at least once | Group A-1: Users who successfully connected with at least once | Group A-2: Users who registered trip info at least once but have not been connected with others at all).

Thus, it is essential to provide methods that can ensure the users to be matched in any circumstance. We are now running the AI agent for this purpose and, we are still seeking better methods that can be more efficient and cheaper than the current one.

Discussion and Future Directions

Our approach can serve as a reference to on-demand or ride-sharing startups eager to promptly penetrate the market. Here are our major findings from operating Cabbit:

- Users tend to continuously utilize the service only when they are convinced that they can use the service whenever and wherever they need: in that sense, the applied AI agent is one useful method.
- It is crucial to subdivide the possible routes as fine-grained as possible to maximize the matching rate.
- Startups should comprehend the detailed traits of the targets, since the actual demand of the ride-sharing services vary upon demographics, occupation, environment, and etc. In Korea, Cabbit model is more fit to meet the needs of the people in universities where the infrastructure for public transportation is weak.

Of course, there are some limitations as well. First, the user behavior logs tell that the seasonal effects cannot be ignored especially when focusing on the university groups: they are more sensitive to time change such as vacations and examination periods, compared to regular commuters.

Second, although we implemented the AI agent, it is difficult to retain this strategy, mainly due to the high maintenance cost. In this sense, it is inevitable to seek another method in order to steadily maintain the retention rate of the users.

Third, people feel uncomfortable when the ride-sharing service does not include a vehicle booking feature. Even after successfully meeting with other companions via Cabbit, users had to find a taxi by themselves. Corporate employees particularly complain about the situation, so if we aim to expand to them, we need to clear this hurdle beforehand.

In addition, although we construct and utilize a users' trust network, it is difficult to recommend possible companions based on the trust network, for now, due to the relatively small size of the user pool. However, it would become possible if more users flow into our service and the similar trip information becomes more available. In that case, we will be able to continue to enhance our companion recommendation algorithm with diverse models and settings.

Last, as we repeatedly heard from users that the fare payment process itself can be a burden to the newcomers, we plan to make the payment process to be optional, meaning users can choose whether they just want to meet up with companions or want to split the fare automatically. It would help more people join Cabbit and allow the service to have a richer user pool, which is always an essential part in providing an effective companion recommendation service.

We believe our approaches can be generalized to other avenues of services, because nowadays, every information is computationally tracked and can be seeds for the decision making process, such as deciding the characteristics of the audiences or the locations for promotions. Moreover, the gist of our research in the form of a case-study of taxi-sharing is showing how early startups (whether they are on-demand or not) can successfully enter the marketplace, with the following whole process: 1) establishing fundamental hypotheses by conducting surveys, prior to operating the service; 2) collecting actual user behavioral data by operating the service; 3) verifying the hypotheses by conducting cohort analysis and A/B testing, based on collected data; 4) resetting the hypotheses and reverifying them by continuously operating the service.

References

[1] Eduardo M Azevedo and E Glen Weyl. 2016. Matching markets in the digital age. *Science* 352, 6289 (2016), 1056–1057.

[2] Burak Boyacı, Konstantinos G Zografos, and Nikolas Geroliminis. 2017. An integrated optimization-simulation framework for vehicle and personnel relocations of electric carsharing systems with reservations. *Transportation Research Part B: Methodological* 95 (2017), 214–237.

[3] Yong-shou Chi, Moon-young Kang, Kyeong-seok Han, and Jeong-il Choi. 2016. A Study on the Discontinuance Intention on O2O Commerce: With a Focus on the Mediating Effects of Perceived Risk and User Resistance. *International Journal of u-and e-Service, Science and Technology* 9, 2 (2016), 207–218.

[4] Blerim Cici, Athina Markopoulou, Enrique Frias-Martinez, and Nikolaos Laoutaris. 2014. Assessing the potential of ride-sharing using mobile and social data: a tale of four cities. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 201–211.

[5] Alan Lee and Martin Savelsbergh. 2015. Dynamic ridesharing: Is there a role for dedicated drivers? *Transportation Research Part B: Methodological* 81 (2015), 483–497.

[6] Paolo Santi, Giovanni Resta, Michael Szell, Stanislav Sobolevsky, Steven H Strogatz, and Carlo Ratti. 2014. Quantifying the benefits of vehicle pooling with shareability networks. *Proceedings of the National Academy of Sciences* 111, 37 (2014), 13290–13294.

[7] John Y Sayah and Liang-Jie Zhang. 2005. On-demand business collaboration enablement with web services. *Decision Support Systems* 40, 1 (2005), 107–127.

[8] Erez Shmueli, Itzik Mazeh, Laura Radaelli, Alex Sandy Pentland, and Yaniv Altshuler. 2015. Ride sharing: a network perspective. In *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*. Springer, 434–439.

[9] Mitja Stiglic, Niels Agatz, Martin Savelsbergh, and Mirko Gradisar. 2015. The benefits of meeting points in ride-sharing systems. *Transportation Research Part B: Methodological* 82 (2015), 36–53.

[10] Xingang Weng, Liying Zhang, and others. 2015. Analysis of O2O model's development problems and trend. *iBusiness* 7, 01 (2015), 51.