

# Contemporary Issues in the Ethics of data analytics in Ride-hailing service

Victor Chang, Yujie Shi and Xuemin Li

International Business School Suzhou

Xi'an Jiaotong-Liverpool University, Suzhou, China

Email: victorchang.research@gmail.com

**Abstract:** Big data technology has brought spring to the establishment of transportation network companies(TNCs), such as Uber in the USA and Didi in China who provide the ride hailing service(RHS) which allows the individuals to act as independent contractors serving customers via smartphone app. The RHS system installed gigantic databases can store enormous data and its equipment can collect various kinds of data at any time. Additionally, data analysis technologies in TNCs such as behavioral analysis, heatmap optimization surge pricing and the automatic order assignment mechanism, can make the RHS more efficient. However, along with achievements big data technology offers, we are struggling for the ethical problems including privacy, inequity and unsafety.

**Key words:** ride-hailing services data analytics privacy ethical issues big data

## 1. Background and introduction

New big data technology, primarily is in the form of computing power and widespread access to the Internet through computers and mobile devices. Many services can

generate data and since data analysis has played important roles, analysis of patterns and correlations can be used to interpret meanings behind the scene (Mai, 2015). A suitable example is as follows. The traditional taxi service requires drivers to drive around the streets to find potential passengers according to their intuition and the adoption of beckon-to-stop model, an order-taking oriented approach. In this case, the application of electronic devices and high-end networking technology can turn the traditional taxi patterns into predictable movements, eventually this can be similar to forecasted automation (Zhou *et al.*, 2016). To make this happen, an online provider system for ride hailing service (RHS) can be used. It can play the role of matchmaker, broker or intermediary between the potential passengers and the “third party” private drivers (Ng, 2016). All services and systems should be integrated, so that it is convenient to use and more efficient to get things done. Similar to smart phones today, each smart phone can provide different functions beyond just calling and receiving phone calls. While the technology has advanced, this phenomenon of getting integrated product and services is unavoidable and is more common to happen.

In general, an online provider system for RHS involves three parties: drivers, a service provider (SP) and riders corresponding to driver mobile terminal, terminal processing server, and passenger mobile terminal, as shown in the Figure1 respectively (Pham *et al.*, 2017). Explanations for each component and how RHS can function is described as follows. Both drivers and passengers can download the smartphone apps, GPS and Network communication technology are used to collect and transmit data to support the service. There are databases, network communication module and trade processing module in terminal processing server (Ng, 2016). The gigantic database contains all the information of participants, and the information is delivered and received by the communication module in the drivers’ and passengers’ apps. Meanwhile, the trade processing module is responsible for

the trading connection between drivers and passengers (Chen *et al.*, 2017). The collaboration of these parts makes the complete system run efficiently. Functionality of this system is useful and convenient since it allows users to get a private car at anywhere and anytime with lower costs than taxis via their smart phone apps. The RHS can serve users the purpose of convenience to get private cars with lower costs. However, there are some ethical concerns for the RHS system. Millions of personal information can be obtained without any legal restrictions and with very low cost. This means that service providers and drivers know details of their customers. They can sell to any third parties under certain ways to cover their identities and responsibilities from legal bidding. Private drivers and companies can get more revenue and sources of income by providing these services. In some countries such as China, drivers need to have a separate application authorized by the local government in order to be on the list.

Ethical and privacy concerns are our focus of this paper, since the use of RHS system can raise a couple of questions – how to use it without invading privacy; how should we carefully handle confidential information; are users aware if sensitive information collected and analyzed; what are recommendations next? In this paper, we will present the all about the RHS and discussion about ethical issues with our recommendation. The remainder of this paper is organized as follows. Section 2 presents a review of the related literature. Section 3 presents several methods of data collection and application at different steps. Section 4 introduces three kinds of data analysis technologies and their functions. Section 5 points out three ethical issues: privacy, inequity and unsafety, followed by the solutions of scholars and us. Besides, there are some difficulties to solve ethical issues. Finally, we conclude the paper and provide directions for future study.

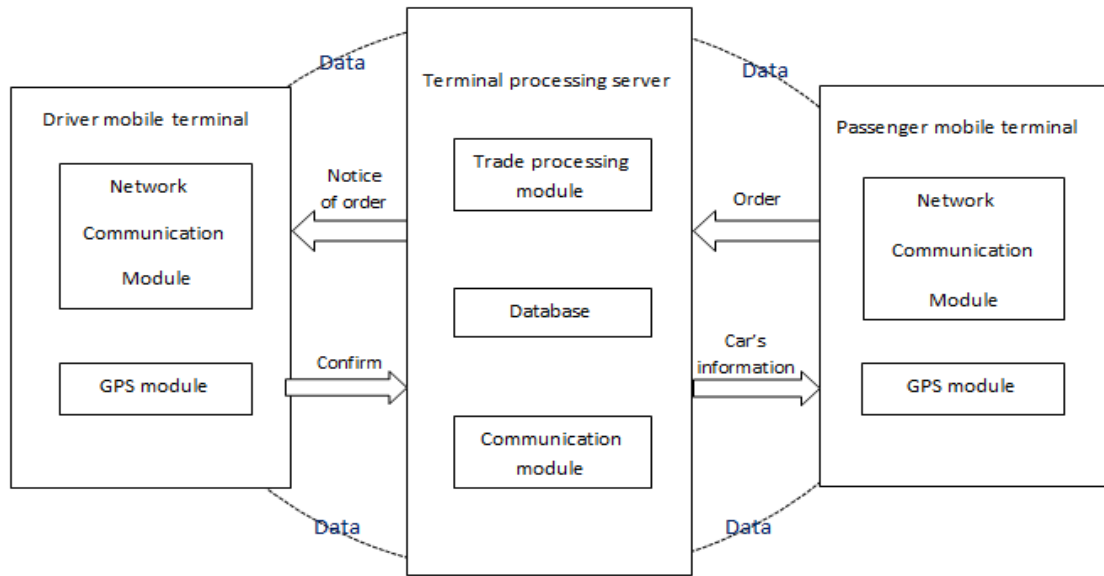


Figure1 Structure of ride-hailing service system

## 2. Literature review

Users can fill in many sensitive-personal information when they register in ride-hailing app (Li et al., 2011), GPS sensors can capture substantial raw data during the trip at any time (Zhou et al., 2016). Both these two kinds of data are used for matching passengers and drivers before the trip (Chen et al., 2017). Besides, GPS big data help the travel via an online anomalous trajectory detection method concluding two steps, which includes creating route recommendation algorithm to choose route, and real-time ongoing trajectory respectively (Zhou *et al.*, 2016).

The customers' behavioral analysis is a popular data analysis method in RHS to explore potential knowledge from analyzing the huge amount of source data (Khade, 2016). It helps the companies to imagine their users in a more fully dimensional way (Taylor, 2017), and thereby providing better service or for other commercial activities (Khade, 2016). For example, RHS can distinguish whether drivers are work

commuting drivers or essentially converted taxi vehicle drivers through their behavioral patterns (Yongqi Dong *et al.*, 2017).

In general, surge pricing mechanism in RHS is mainly determined by the real-time supply and demand(S&D) (Cachon *et al.*,2017). TNCs collect and analyze such internal data that reflect real-time (S&D) to calculate price elasticity (Cohen *et al.*, 2016). Iguazio (2017) points out that surge pricing is also related to external data sources such as weather and broadcast. Heatmap is an efficient tool used by TNCs to show the (S&D) to drivers, it can maximize drivers' profit as well as reducing waiting time of customers (iguazio, 2017).

Once the request of potential passenger is released, platform automatically matches the drivers and passengers according to matching algorithms (Dong *et al.*, 2018).

The dynamic ride matching problem is intractable, and many scholars have discussed different matching algorithms (Qian *et al.*, 2017). In general, the distance between customers and drivers, weather, timestamp, response of others, and the local supply and demand have impact on the order assignment of RHS (Chen *et al.*, 2017).

Privacy is a tough problem in RHS. Insecurity data storage of TNCs is the direct cause of the privacy threat (Pham *et al.*, 2016). Processing and analyzing data in service can obtain potential knowledge in different areas which have high value of utilization to commercial enterprises and state agencies (Zhou *et al.*, 2016). Mai introduces the privacy datafication model that the analysis of already available data enables others to generate new insights through (Mai, 2015; Pham *et al.*, 2016) propose a privacy-enhanced service called "PrivateRide" including four cryptographic tools, it can protect passengers' anonymity and location privacy, while maintaining most of the benefits of current service providers. (Yang *et al.*, 2018) take the big data technology and design a keyword match based on updating mechanism, which runs

the updating as well as ensuring the privacy security. (Banerjee *et al.*, 2017) highlight the government's responsibility in identification, monitoring and examination the data to solve privacy issue.

There are many reports about public worry that customers are unfairly discriminated because of dynamic pricing, and many people call these actions price gouging (Cachon *et al.*, 2017). Steven considers that the surge pricing is important to keep the effectiveness of the market, but the public cannot understand it. He proposes to change public opinion which requires government's help (Suranovic, 2016). Zhou (2016) also suggests lowering the price ceiling along with adding bonus points to offset customers' overpaying and thus alleviating inequity.

The car reservation model of RHS requires drivers to act as multitasking workers who need to operate phones frequently while driving (Tang *et al.*, 2017). It leads to drivers' distractions, which can increase the rate of accidents (B Weinstein, 2014). Besides, having a call during driving also brings negative performance of drivers (Fitch, 2013). Baptiste and David invent a new patent, which can inhibit the operation of electronic device when the speed of vehicle is higher than normal standard (Baptiste, 2011). There is also a front-facing smartphone camera can be used to detect distracted behavior of drivers while driving, and the application can take actions to stop the use of mobile upon detection of risky behavior (Wigton and Garner Iii, 2016).

### **3. The data collection and application in ride-hailing service**

#### **3.1 Match of passengers and drivers**

There are large amounts of data in various types for RHS (shown in Table1), and many ways can be taken to obtain these data. The simplest way is that when riders or drivers register their accounts in ride-hailing apps, they need to fill in a lot of

sensitive-personal information. For example, users are required to fill in their telephone number, name, gender, birth, occupations, identity card and drivers are also required to provide their driving licenses including their family address (Li *et al.*, 2011). All these information is stored in the gigantic database of terminal processing server shown in Figure1. Moreover, the GPS sensors installed in the smartphones can capture a substantial amount of GPS raw data during the trip at any time, which contains the data about the situation of the vehicles, such as geographical location information, the speed trend, timestamp, and the state of occupation (Zhou *et al.*, 2016). These two kinds of data are important for matching passengers and drivers. Explanations for the matching process is as follows. A passenger will use the GPS coordinate set in the smartphone app to find whether there is a spare capacity available and then sends request to the terminal processing server through communication module. After receiving the order, the server will make an algorithm according to the passenger's preference (willing to have car sharing), the distance between car and customers, whether there is a driver to take the order already, then filtrate the most suitable car for the passenger and deliver the notice of order to the driver. After the driver's confirmation, the server will retrieve the data of the car including license plate, name, location, unoccupied seats, telephone number from the database and then deliver it to the passenger (Chen *et al.*, 2017). There has been a connection between the two entities until the end of the journey.

### **3.2 Online anomalous trajectory detection**

The GPS big data can help the travel via an online anomalous trajectory detection method consisting of two steps, and new data will also generate during the trip and after the trip. In the first step, when a passenger enters a car, the GPS module automatically locates the current position, and the destination has been sent by the passenger through the communication module at the beginning. A set of route candidates between these two points are calculated based on analysis of many

historical trajectories stored in the database system of GPS. In the second step, the new vehicle trajectory is detected in real time, and TNCs can also do behavioral analysis on the trip (such as estimating whether drivers have cheated on the passengers by comparing the driving routes with the GPS recommendation routes) to check whether this trip is abnormal to ensure the quality of the service (Zhou *et al.*, 2016). There are always new trajectory and behavioral data generating during the whole trip. Finally, when the trip is completed, passenger and drivers can rate each other to show their feelings and suggestions (Cetin and Deakin, 2017), which is valuable data to indicate the performance of users to TNCs, who can analyze these data to optimize the service. For example, if some passengers are dissatisfied to a driver and give him low ratings, the TNCs will dismiss the driver.

Personal information	Telephone number, name, gender, birth, occupation, identity card, driving license, address...
GPS raw data	Geographical location, speed trend, timestamp...
Trajectory detection	The current position, the destination, the spare capacities...
	Real-time ongoing car trajectory & online behavioral data
	Rating data, feelings & suggestions of users

Table1 Diverse kinds of data in RHS

## 4. The data analysis in ride-hailing service

### 4.1 Behavioral analysis behind the Uber's discovery

There is an interesting psychological phenomenon of passenger behavior. The remaining money and service fees left on their smartphones are the determinant for potential customers to ask for Uber's service or decide to wait for the fare to drop in



rush hours. Uber has adopted the use of behavioral analysis to forecast passengers' spending, routes and preference, which have triggered the public criticism that the data collection policies of RHS go far beyond the basic information (Fedde, 2016).

The customer behavioral analysis is a popular data analysis method in the big data era, where the volume of raw data from the digital world is rapidly increasing. It helps companies to explore the potential information from the huge amount of source data through observing the behaviors of users and analyzing their motivations (Khade, 2016). It can be concluded that behavioral analysis is an effective way for companies to understand their customers and then provide right service at right time. However, many companies have difficulty in behavioral analysis presently due to the shortage of data, while it is not a problem for TNCs. Because as mentioned above, the TNCs have many high-tech means to collect the massive volumes of data. Hence, they can use data analysis methods to turn raw data into useful information, such as the behavioral data of users which can help to imagine them in a more fully dimensional way (Taylor, 2017), and thereby providing better service or for other commercial activities such as improving sales, advertising, market optimization and many more applications (Khade, 2016). For example, from the viewpoint of individual behavioral data, TNCs can identify two kinds of internet-based ride-sharing drivers through analyzing users' raw event data. The data analysts analyze the data such as the service period, the length of service time, the trip location provided by drivers to distinguish whether they are part-time drivers (work commuting) or full-time drivers (essentially converted taxi drivers) (Yongqi Dong *et al.*, 2017). Based on this classification, TNCs can predict the supply of RHS in different time (the specific use will be discussed below).

## **4.2 Pricing mechanism based on data analysis**

The pricing mechanism of RHS is different from the traditional taxi industry, which can be realized by data analytics. Most TNCs take surge pricing while taxi's fare is fixed. In general, the determinant of pricing mechanism is the relationship between the supply and demand of RHS (Cachon et al., 2017). Hence, TNCs observe and analyze various kinds of internal data (e.g. geographic region, time and date, rider id, quantity of orders, waiting time) to reflect the supply and demand of RHS in various places at different time periods. To be specific, the number of rider id can reflect the supply and the quantity of orders can reflect the demand in those regions. The waiting time can be treated as the measurement of demand-and-supply, the longer the waiting time, less the supply. Besides, TNCs are not only concerned about every trip taking RHS, but also focus on customers who have not made the requests for RHS finally after searching in the apps. Because this can reflect that to which degree can consumers accept the surging prices, it is also a crucial part of pricing mechanism (Cohen *et al.*, 2016). Furthermore, surge pricing is also related to external data sources such as broadcast, the weather conditions and social network (Iguazio, 2017). Because these external data also have impact on the supply and demand of RHS.

Heatmap is a color-based graphical representation of big data (Alexandre Perrot et al., 2017) and it is always used by TNCs to show supply and demand of RHS. The intensity of the color represents the customers' demand, the darker color shows that more passengers are request for RHS and they must pay surging fares at that time (Cohen et al., 2016). Heatmap can improve the decision-making process of drivers, and then maximize their profit as well as reducing the waiting time of customers (iguazio, 2017). Because many RHS apps can display the local real-time heatmaps to drivers which tells them where have more potential passengers and can incentivize them moving from over-supply regions to over-demand regions, thus bringing more business to them (Ke et al., 2017).

### **4.3 Order assignment based on data analysis**

As mentioned above, the application of location-based service(LBS), communication technologies and GPS positioning efficiently solves the information asymmetry problem between drivers and passengers in RHS and makes it much easier for them to know the need of each other (Dailey *et al.*, 1999, Levofsky and Greenberg, 2001). All these technologies are work together to operate the order assignment mechanism in OHSPS, which acts as matchmaker between private drivers and potential passengers. As mentioned above, the general RHS operates as follows: people who plan to have trips release their trip information including starting location, destination, numbers of passengers and their preference (willing to share car with others) on the ride-hailing platform. The ride-hailing platform then automatically matches the drivers and passengers according to the matching algorithms and send the matched itinerary information to both drivers and passengers (Dong *et al.*, 2018). The large-scale dynamic ride matching problem is complex and intractable, and there are many scholars discuss different matching algorithms (Qian *et al.*, 2017). Although this problem is still understudied, we can be sure that the data analytics is indispensable. In general, when receive requests, TNCs will analyze requests according to a huge amount of data such as the distance between two entities (customers and drivers nearby), weather condition, timestamp, responses of other customers' requests and drivers, and the real-time supply and demand condition also has impact on the order assignment (Chen *et al.*, 2017). During this process, large-scale calculations are carried out, different methods and algorithms such as short-term passenger demand forecasting, surge pricing and path planning algorithms can be taken to assign orders between passengers and drivers in the dynamic ride-sharing (Qian *et al.*, 2017).

## **5. The ethical issues in ride hailing service**

RHS offers a new kind of sharing economy which has been favored by customers since its better performance, convenience and lower price comparing with the traditional taxi industry. However, there are some ethical problems about the RHS cannot be ignored. We will discuss these problems below.

## **5.1 Privacy**

### **5.1.1 Privacy issue caused by insecurity of data storage**

First, privacy concerns can make some people, organizations and countries nervous about who owns the data, how data will be used, and whether data will be sold, leaked, manipulated or abused for various “unthinkable” purposes. As mentioned in previous sections, there are many ways for TNCs collecting various kinds of data and the quantity of data can contain millions and billions of user information. Furthermore, some data can reveal significantly more about private and personal information. For instance, the data filled in the app when users registered can reveal their personal telephones, living places, genders, ages, jobs; and the trip data can reveal users’ social activities such as frequency, locations and expenses related to shopping, dining and watching movies (Pham *et al.*, 2017). However, it is very difficult to for TNCs to store such big data and is easy for them to suffer from data leakages, which pose a great threat to users’ privacy. There are several reports about the cases of privacy invasion caused by insecurity of data storage. For example, the employees in TNCs stealing sensitive data and violating individuals’ interest illegally for entertainment (Hill, Oct 3, 2014 ) or revenge (Guynn, 2014).

### **5.1.2 Privacy issue caused by data analysis**

In big data era, TNCs cannot only get the information that directly available, but also can access much deeper information through analyzing the raw data. The Uber’s discovery about the relation between phone batteries mentioned above is a typical case to violate passengers’ privacy through data analysis (Fedde, 2016).Uber

company discloses passengers' psychological privacy and make use of this for business purposes. As TNCs do much behavioral analysis every day, it is hard to imagine how much they understand Rideshare users. Besides, the gigantic database in OHSPS contains enormous data and other devices also capture data all the time. Hence, it is easy for TNCs to take many other methods of data analysis, which further deepens the possibility and seriousness of privacy invasion. Though data collection and processing seems to be compulsive without considering human rights (Bruce Weinstein, 2014), it can bring much benefits to TNCs and others. Specifically, except for having some interesting discoveries as Uber did, it is also possible for TNCs to get the knowledge in different areas such as the commercial behaviors, the traffic network dynamics and the rule of people's lifestyle, moreover, various value-added applications like commercial advertising, traffic management and personalized service (Zhou *et al.*, 2016). Hence, data is a valuable resource in the big data era ('The World's Most Valuable Resource,' 2017), and the results of data analysis can be treated as intangible assets to sell to the third parties for profit. This can be concluded by the privacy datafication model, which means that after analyzing and processing the already available data, others can obtain new knowledge or useful information. But these processes are not understood or permitted by the data owners (Mai, 2015)

### **5.1.3 The recommendations to privacy issue**

There are many technologies can be used to enhance privacy protection in Rideshare, such as using secure multiparty computation, which requires participants in calculations based on their respective secret input, it can set up an anonymous between customers and drivers (Ran, 1996). However, it is always impractical to apply them to the Rideshare for some reasons, such as their complexity or impracticality to implement; the difficult of charge a commission per ride for SP without enough information. Hence, Pham and other researchers introduce PrivateRide which

focuses on balancing the privacy of passengers, the convenience of services and the benefits of the SP, passengers and drivers are anonymously informed about the trip, only part of the ride information is revealed to the SP (Pham *et al.*, 2017). It combines four cryptographic tools for ride-hailing settings to properly protect privacy of riders and drivers. Besides, privacy issue can also be solved by big data technology. Considering the homogeneity of the privacy issue of data, learning from other industries is also advisable. A privacy-preserving system is proposed in health care industry, which combine Internet-of-things (IoT), big data and cloud storage. It also includes special encryption technique on IoT messages by using the IoT group key. The scholars design a keyword match based on updating mechanism, which can make the updating as well as ensuring the privacy security (Yang *et al.*, 2018). Moreover, the solution to the privacy issue is also dependent on government departments, who can create a “watchdog” unit to identify and monitor the different data collected. government also take the responsibility of examination the use of data (Banerjee *et al.*, 2017).

Although data collection and analysis in RHS have violated the privacy, on the other hand, processing and analyzing the vast data can lead to a better society in some ways. Huge mobile data provide new perspectives and methods for measuring economic dynamics that are widening the areas of social science and economics (Lei Dong *et al.*, 2017). For example, analyzing the trajectory data of GPS module in RHS helps traffic control department better understand the road network dynamics so that officers can better manage transportation and take efficient project for roads design to reduce traffic congestions (Babar and Arif, 2017). Hence, it is advisable to utilize privacy when necessary instead of using various technologies to totally forbid it or just reveal only little part of it. What’s more, TNCs must ensure that these data sold to agencies are used in the proper channels and have contributions to the society. Meanwhile, the participants in RHS should understand every activity related to their

personal data. Besides, when riders and drivers register their RHS accounts, they also need to deeply understand and voluntarily sign an agreement about the transfer of their privacy right including the detailed purposes. Finally, the full process should be monitored by government agency.

## **5.2 Equity**

### **5.2.1 Inequity of surge pricing**

When we discuss about equity, as mentioned above, another distinct function of RHS is on surge pricing mechanism, which sets pricing according to the real-time local supply and demand situation and that means consumers will pay different fares for the same route but in different time. The based price (what we call the “no surge” or “1.0x” price) and surging multiplier are determined by TNCs (Cohen *et al.*, 2016). However, there are many reports about the public has worried that customers are unfairly discriminated because of dynamic pricing (Cachon *et al.*, 2017). In general, if the offered capacity exceeds demand, providers share the available demand equally, and if the offered capacity is less than demand, then demand is randomly assigned, consumers are also likely to receive the scarce service equally (Cachon *et al.*, 2017). But surge pricing mechanism breaks this equity because it allows drivers to only provide services to those who can pay more money, while the poor individuals will lose the right of receiving services. It makes RHS the same as auctions, where the highest bidder can get the treasures. However, transportation is the necessities of life rather than the luxury used to auction. Moreover, the TNCs can make more profit from the surge pricing. Hence, it fuels the public anger, and customers always accuse it of a price gouging which is conducted by the callous merchants for profiteering (Suranovic, 2016).

### **5.2.2 Inequity issue and market mechanism**

We are struggling for the inequity issue. Considering the large-scale of RHS currently, the government's intervention and management of surge pricing in the ride hailing market is the most effective solution the unfair issues. To be specific, government agency can control the pricing mechanism of RHS by forcing it to take the same pricing as taxi industry and issue price gouging price in this industry (Gallick and Sisk, 1987). These political works will assuage irritation of consumers, but at the costs of adding extra burdens on everyone. That is, participants will suffer avoidable losses in the pursuit of fairness (Suranovic, 2016). Because the rising pricing can attract more drivers to high-demand positions which is helpful to reduce the waiting time of passengers. Moreover, the price will go down with the increasing number of suppliers (Zha *et al.*, 2017). It is beneficial to equalize local supply and demand, and the RHS will thereby be allocated in a more efficient manner. From an economic point of view, surge pricing can bring economic efficiency which refers to the limited supply should be firstly allocated to the most valuable uses and then to lower valued users, thus maximizing the depletable resources. This is particularly remarkable when people realize the rule that their favorable outcomes generate from the pursuit of profit by enterprises and utility by consumers instead of controlled by any central direction (Cachon *et al.*, 2017). Hence, surge pricing is inevitable from the standpoint of market mechanisms. Furthermore, the price gouging laws are also not desirable because policy analysis requires more than a survey of public sentiment, and benefit of it must be weighed against the risk of it (Giberson, 2011).

Nevertheless, the willingness of market participants to accept the surge pricing mechanism also has impact on the effectiveness of it. Hence, it is necessary to change the public strongly negative sentiment on surge pricing and help them recognize the fact that the TNCs not only operate for profit but also provide a public service which is beneficial for social development. To perform this arduous task, it is advisable to have better education on the RHS users in economic courses and the



government has the obligation to persuade people to accept the rule that the free market in RHS is necessary (Cachon *et al.*, 2017). In addition, reducing the price ceiling and adding some bonus points can be helpful to offset passengers' overpaying and thus alleviating their negative emotion to surge pricing as well as reducing the degree of unfairness (ZhaoxiZhou, 2016). Furthermore, big data analytics technologies can be used to do an experiment about the effect of surge pricing which is more persuasive.

## **5.3 Safety**

### **5.3.1 Safety issue caused by order assignment of RHS**

Finally, we talk about the safety issue in RHS. The application of high technology (e.g. GPS, LBS) brings the online car reservation of RHS and the data analytics makes the matching algorithms more accurate, drivers just need to drive to the locations of passengers after receiving the notice of orders, which is much more convenient and efficient than the beckon-to-stop model of traditional taxi. However, we worry about its safety issue because the generation and transmission of orders information are totally rely on electronic equipment, and thus requiring drivers to act as multitasking workers who are notified immediately of passengers' riding request and then they need to operate their phones frequently while driving to deal with orders as soon as possible and thus providing high-quality service (Tang *et al.*, 2017). As a result, multiple driver distractions highly increase the rate of having traffic accidents (B Weinstein, 2014). Specifically, during the service process, drivers need to do visual-manual tasks, they need to local the phone, looking at the screen and touch the keys to response to passengers from time to time. While these behaviors transfer the drivers' attention from road condition to the ride-hailing business which is risky and is potentially deadly for ride-hailing participants, or even the passersby (Institute, 2013). Additionally, drivers usually need to call customers if there are

misunderstandings (such as the ambiguous location information because of the inaccuracy of GPS coordinate). Having conversations on a cell phone during driving can negatively affect driving performance and bring undesirable effects on the driving safety (Fitch, 2013). Though the phonetic functions enable drivers free their hands, they must deal with the receiving messages and responding to them which can still distract their attention to the traffic condition (B Weinstein, 2014).

### **5.3.2 The recommendations to safety issues**

Though the distraction of drivers is quite serious issues, which pose threat to the safety of the drivers and passengers, it is impractical for implement of ride-hailing service without the use of mobile device at present. Consider the efficiency of this service, it is better to adopt measures to alleviate safety issue instead of totally prohibiting using mobile device. Considering the high driving speed is another reason for traffic accident, it is better to encourage drivers to slow down when they distracted from driving. Baptiste and David propose an invention, which has an anti-distraction system to detect vehicle speed and to inhibit the operation of any electronic device that can distract a driver. The operating principle of it is when detecting the speed of car is above the presetting safe speed, another system will work to inhibit the operation of electronic device (Baptiste, 2011). Then the messages from ride hailing platform or passengers will not be shielded until the speed of driver slows down. In addition, another patent is about detecting the distracted behavior by high-technology. It describes an invention which mainly using a front-facing smartphone camera to detect distracted behavior of drivers while driving. Upon detection of risky behavior, the application would take some actions such as locking the mobile screen or buttons to prevent drivers using electronic devices that distract them. Only when the car is stopped or pulled over for a time can the drivers reuse their cell phones (Wigton and Garner Iii, 2016).

Although the probability of traffic accidents caused by drivers' distraction is likely to be reduced by the solutions above, these inventions are complex and difficult to implement since it requires a large amount of research work and the big data analytics should secure patents. As a result, it is possible for TNCs to reject the application of these patents because of excessive cost. Moreover, it is undeniable that the application of these patents will bring inconvenience to drivers and passengers. Hence, government participation is necessary. Government should force TNCs to install systems that alleviate the negative effect to individuals caused by electronic device. In addition, government also have the responsibility to educate the drivers and passengers about the necessity and feasibility of these systems.

## **6. Conclusion**

In the big data era, human beings are beginning to realize the importance of data and the contributions of data analytics to improve efficiency. TNCs use electronic equipment and network technique to collect and store various kinds of valuable data. The GPS raw data help the system to match drivers and passengers; it also helps create a route recommendation algorithm that can calculate the most feasible route for the journey. The purpose of trajectory detection method is far beyond the monitoring purpose, since it offers an efficient way to obtain large amount of new data. Behavioral analysis turns raw data into useful information and thus helping companies deeper understand their users. Besides, surging pricing based on data analytics can equalize the demand and supply and thus making the service more efficient. Moreover, big data analytics generates the automatic order assignment mechanism, which can solve the information asymmetry between driver and passenger and makes the trips more convenient.

However, data manipulation policies in TNCs go far beyond basic data and cause ethical issues inevitably. Imperfect data protection mechanism causes the data

leakage, and the effectiveness of data analysis makes the privacy problem more complicated. Many technical solutions can solve privacy issue but hinder the social improvement. Therefore, analyzing private data is indispensable but it should allow ways to ask permission of relevant persons. Additionally, the surged pricing also causes the dilemma between equity and market mechanism. Since surged pricing is unavoidable, what we can do is taking actions to alleviate inequity. Moreover, the order assignment mechanism lead to the drivers' distraction and increase the rate of accidents. To resolve the unsafety issue, many scholars have invented some complex patents. However, they require government's power to implement and participants' understanding of its urgency on top of inconvenience. In conclusion, data analytics makes RHS more efficient, while the ethical problems of it are more complicated and inescapable in the big data era. Further study is exploring more ethics potential problems of data analytics in RHS and trying to relieve contradiction between development and ethical issues.

## 7. Bibliography

Alexandre Perrot , Romain Bourqui and Hanusse, N. (2017) 'HeatPipe: High Throughput, Low Latency Big Data Heatmap with Spark Streaming', 66. available:

<http://ieeexplore.ieee.org/document/8107949/>

Babar, M. and Arif, F. (2017) 'Smart urban planning using Big Data analytics to contend with the interoperability in Internet of Things', *Future Generation Computer Systems*, 65. available: <https://www.sciencedirect.com/science/article/pii/S0167739X17308993>

Banerjee, S., Hemphill, T. and Longstreet, P. (2017) 'Wearable devices and healthcare: Data sharing and privacy', *The Information Society*, 1-9. available:

<http://www.tandfonline.com/doi/pdf/10.1080/01972243.2017.1391912>

Baptiste, D. (2011) 'Method and apparatus for combating distracted driving'. U.S. Patent Application No. 12/590,582. available:

<https://patents.google.com/patent/US20110111724A1/en>

Cachon, G. P., Daniels, K. M. and Lobel, R. (2017) 'The Role of Surge Pricing on a Service Platform with Self-Scheduling Capacity', *Manufacturing & Service Operations Management*, 19(3), 368-384. available: <https://pubsonline.informs.org/doi/abs/10.1287/msom.2017.0618>

Cetin, T. and Deakin, E. (2017) 'Regulation of taxis and the rise of ridesharing', *Transport Policy*.available:[https://s3.amazonaws.com/academia.edu.documents/48792801/Regulation\\_of\\_Taxis\\_and\\_the\\_Rise\\_of\\_Ridesharing.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1516285685&Signature=k7NBif90tOjclb7q4TNfU0nHvhA%3D&response-content-disposition=inline%3B%20filename%3DRegulation\\_of\\_Taxis\\_and\\_the\\_Rise\\_of\\_Ride.pdf](https://s3.amazonaws.com/academia.edu.documents/48792801/Regulation_of_Taxis_and_the_Rise_of_Ridesharing.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1516285685&Signature=k7NBif90tOjclb7q4TNfU0nHvhA%3D&response-content-disposition=inline%3B%20filename%3DRegulation_of_Taxis_and_the_Rise_of_Ride.pdf)

Chen, X., Zahiri, M. and Zhang, S. (2017) 'Understanding ridesplitting behavior of on-demand ride services: An ensemble learning approach', *Transportation Research Part C*, 76, 51-70. available: <http://www.sciencedirect.com/science/article/pii/S0968090X16302728>

Cohen, P., Hahn, R. W., Hall, J., Levitt, S. D. and Metcalfe, R. D. (2016) 'Using Big Data to Estimate Consumer Surplus: The Case of Uber', *Social Science Electronic Publishing*.available: <http://www.nber.org/papers/w22627>

Dailey, D. J., Loseff, D. and Meyers, D. (1999) 'Seattle smart traveler: dynamic ridematching on the World Wide Web', *Transportation Research Part C: Emerging Technologies*, 7(1), 17-32. available: <http://www.sciencedirect.com/science/article/pii/S0968090X99000078>

Dong, L., Chen, S., Cheng, Y., Wu, Z., Li, C. and Wu, H. (2017) 'Measuring economic activity in China with mobile big data', *EPJ Data Science*, 6(1), 1. available: <https://link.springer.com/article/10.1140/epids/s13688-017-0125-5>

Dong, Y., Wang, S., Li, L. and Zhang, Z. (2017) 'An empirical study on travel patterns of internet based ride-sharing', *Transportation Research Part C*, 86, 1-22. available: <https://www.sciencedirect.com/science/article/pii/S0968090X17302954>

Fedde, C. (2016) 'How Uber Knows Your Phone Battery Is Dying and More'.available: <https://www.questia.com/newspaper/1P2-39664588/how-uber-knows-your-phone-battery-is-dying-and-more>

Fitch, G. M. (2013) *The impact of hand-held and hands-free cell phone use on driving performance and safety-critical event risk : final report / Gregory M. Fitch [and nine others]* [Dictionaries Bibliographies Non-fiction].available:<https://trid.trb.org/view.aspx?id=1249880>

Gallick, E. C. and Sisk, D. E. (1987) 'Reconsideration of Taxi Regulation, A', *Plos One*, 6(7), : e22003. available: <http://heinonline.org/HOL/LandingPage?handle=hein.journals/jleo3&div=12&id=&page=>

Giberson, M. (2011) 'The Problem with Price Gouging Laws', *Regulation*.available: <http://heinonline.org/HOL/LandingPage?handle=hein.journals/rcatorbg34&div=12&id=&page=>

- Guynn, E. W. a. J. (2014) *Uber tracking raises privacy concerns* [online], available: <https://www.usatoday.com/story/tech/2014/11/19/uber-privacy-tracking/19285481/> [accessed].
- Hill, K. (Oct 3, 2014) '*God View': Uber Allegedly Stalked Users For Party-Goers' Viewing Pleasure* [online], available: <https://www.forbes.com/sites/kashmirhill/2014/10/03/god-view-uber-allegedly-stalked-users-for-party-goers-viewing-pleasure/2/#41a4f1b97fbe> [accessed].
- iguazio (2017) '*Grab, South East Asia's #1 Ride-Hailing Service, Selects iguazio's Unified DataPlatform*'.available: <https://www.jvpvc.com/press-releases/grab-south-east-asias-1-ride-hailing-service-selects-iguazios-unified-data-platform/>
- Institute, V. T. T. (2013) *New VTTI study results continue to highlight the dangers of distracted driving* [online], available: <https://www.vtti.vt.edu/featured/?p=193> [accessed].
- Ke, J., Zheng, H., Yang, H. and Chen, X. (2017) '*Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach*', *Transportation Research Part C*, 85, 591-608.available: <https://www.sciencedirect.com/science/article/pii/S0968090X17302899>
- Khade, A. A. (2016) '*Performing Customer Behavior Analysis using Big Data Analytics*', *Procedia Computer Science*, 79, 986-992.available: <https://www.sciencedirect.com/science/article/pii/S1877050916002568?via%3Dihub#>
- Levofsky, A. and Greenberg, A. (2001) '*Organized dynamic ride sharing: The potential environmental benefits and the opportunity for advancing the concept*', in *Transportation Research Board 2001 Annual Meeting*, 7-11.available: <http://ridesharechoices.scripts.mit.edu/home/wp-content/papers/GreenburgLevofsky-OrganizedDynamicRidesharing.pdf>
- Li, M., Cao, N., Yu, S. and Lou, W. (2011) '*FindU: Privacy-preserving personal profile matching in mobile social networks*', in *IEEE INFOCOM*, 2435-2443.available: <http://ieeexplore.ieee.org/abstract/document/5935065/>
- Mai, J. E. (2015) '*Big data privacy: The datafication of personal information*', *Information Society*, 32(3), 192-199.available: <http://www.tandfonline.com/doi/abs/10.1080/01972243.2016.1153010>
- Ng, S. (2016) '*Uber Technologies Inc. as a Disruptive Innovation*'.available: [https://s3.amazonaws.com/academia.edu.documents/52084016/Uber\\_as\\_Disruptive\\_Innovation.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1516287279&Signature=z7Zyh0gRxitTD4%2FPEY7jjiGLuqbE%3D&response-content-disposition=inline%3B%20filename%3DUber\\_Technologies\\_Inc.\\_as\\_a\\_Disruptive\\_I.pdf](https://s3.amazonaws.com/academia.edu.documents/52084016/Uber_as_Disruptive_Innovation.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1516287279&Signature=z7Zyh0gRxitTD4%2FPEY7jjiGLuqbE%3D&response-content-disposition=inline%3B%20filename%3DUber_Technologies_Inc._as_a_Disruptive_I.pdf)

- Pham, A., Dacosta, I., Jacot-Guillarmod, B., Huguenin, K., Hajar, T., Tramèr, F., Gligor, V. and Hubaux, J.-P. (2017) 'PrivateRide: A Privacy-Enhanced Ride-Hailing Service', *Proceedings on Privacy Enhancing Technologies, Vol 2017, Iss 2, Pp 38-56 (2017)*, (2), 38. available: <https://www.degruyter.com/view/j/popets.2017.2017.issue-2/popets-2017-0015/popets-2017-0015.xml>
- Pham, T. V. A., Petrocelli, I. I. D., Jacot-Guillarmod, B., Huguenin, K., Hajar, T., Tramèr, F. and Hubaux, J. P. (2016) 'PrivateRide: A Privacy-Preserving and Secure Ride-Hailing Service'.available:[https://www.researchgate.net/profile/Kevin\\_Huguenin/publication/301204656\\_PrivateRide\\_A\\_Privacy-Preserving\\_and\\_Secure\\_Ride\\_Hailing\\_Service/links/571ea03508aefa648899a41a.pdf](https://www.researchgate.net/profile/Kevin_Huguenin/publication/301204656_PrivateRide_A_Privacy-Preserving_and_Secure_Ride_Hailing_Service/links/571ea03508aefa648899a41a.pdf)
- Qian, X., Zhang, W., Ukkusuri, S. V. and Yang, C. (2017) 'Optimal assignment and incentive design in the taxi group ride problem', *Transportation Research Part B*, 103, 208-226.available: <http://www.sciencedirect.com/science/article/pii/S0191261516306166>
- Ran, C. (1996) 'Studies in Secure Multiparty Computation and Applications'.available: <http://www.wisdom.weizmann.ac.il/~oded/PSX/ran-phd.pdf>
- Suranovic, S. (2016) 'Surge Pricing and Price Gouging: Public Misunderstanding as a Market Imperfection'.available: <http://www2.gwu.edu/~iiep/assets/docs/papers/2015WP/SuranovicIIEPWP2015-20.pdf>
- The World's Most Valuable Resource 2017, Makati City. available:<https://search-proquest-com.ez.xjtlu.edu.cn/docview/1898952789?accountid=153081>
- Tang, Z., Dang, S. and Zhen, W. (2017) 'The impact of the use of taxi-hailing apps on driving safety.', *Journal of transportation engineering and information.*, 15(1), 22-27.available: [http://www.wanfangdata.com.cn/details/detail.do?\\_type=perio&id=jtysgcyxxb201701004](http://www.wanfangdata.com.cn/details/detail.do?_type=perio&id=jtysgcyxxb201701004)
- Taylor, D. (2017) 'Buyer personas can help you understand customer perceptions', *Central Penn Business Journal*, 33(33), 12.available: <http://eds.b.ebscohost.com/eds/detail/detail?vid=1&sid=6d59195f-32d4-449e-9f9a-0444b66907f9%40sessionmgr101&bdata=JnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2I0ZQ%3d%3d#db=bth&AN=124921073>
- Weinstein, B. (2014) 'Opinion: Four other ways Uber is ethically challenged', *CNN Money*.available: <http://money.cnn.com/2014/11/21/technology/uber-ethics-oped/>
- Wigton, J. D. and Garner Iii, G. (2016) 'Method and use of smartphone camera to prevent distracted driving'.available: <https://www.google.com/patents/US9324149>
- Yang, Y., Zheng, X., Guo, W., Liu, X. and Chang, V. (2018) 'Privacy-preserving fusion of IoT and big data for e-health', *Future Generation Computer Systems*. available:<https://www.sciencedirect.com/science/article/pii/S0167739X17324391>

Zha, L., Yin, Y. and Du, Y. (2017) 'Surge Pricing and Labor Supply in the Ride-Sourcing Market', *Transportation Research Procedia*, 23, 2-21.available:

<http://www.sciencedirect.com/science/article/pii/S2352146517302703>

ZhaoxiZhou, Y. Y. (2016) 'Research and application of ride-hire software model', *Journal of Shanghai maritime university*, 37(3), 22-28.available:

[http://www.smujournal.cn/ch/reader/view\\_abstract.aspx?file\\_no=201506180166&flag=1](http://www.smujournal.cn/ch/reader/view_abstract.aspx?file_no=201506180166&flag=1)

Zhou, Z., Dou, W., Jia, G., Hu, C., Xu, X., Wu, X. and Pan, J. (2016) 'A method for real-time trajectory monitoring to improve taxi service using GPS big data', *Information & Management*, 53(8), 964-977.available:

<http://www.sciencedirect.com/science/article/pii/S0378720616300374>