Evaluation of collaborative consumption of food delivery services through web mining techniques

Juan C. Correa*, Wilmer Garzón, Phillip Brookerc, Gopal Sakarkar, Steven A. Carranza, Leidy Yunado, Alejandro Rincón

* Faculty of Psychology, Fundación Universitaria Konrad Lorenz, Bogotá, Colombia
b Escuela Colombiana de Ingeniería Julio Garavito, Bogotá, Colombia
c Department of Sociology, Social Policy, and Criminology at University of Liverpool, UK
b Escuela Colombiana de Ingeniería Julio Garavito, Bogotá, Colombia
b Department of Computer Applications, Raisoni College of Engineering, Nagpur, India

ARTICLE INFO

Keywords:
- Collaborative consumption
- Traffic conditions
- Google maps
- Online food ordering

ABSTRACT

Online food delivery services rely on urban transportation to alleviate customers’ burden of traveling in highly dense cities. As new business models, these services exploit user-generated contents to promote collaborative consumption among its members. This study aims to evaluate the impact of traffic conditions (through the use of Google Maps API) on key performance indicators of online food delivery services (through the use of web scraping techniques to retrieve customer’s ratings and the physical location of restaurants as provided by Facebook). From a collection of 19,934 possible routes between the physical location of 787 online providers and 4296 customers in Bogotá city, we found that traffic conditions exerted no practical effects on transactions volume and delivery time fulfillment, even though early deliveries showed a mild association with the number of comments provided by customers after receiving their orders at home.

1. Introduction

Collaborative consumption (CC) is a new form of consumer behavior with important implications for business research (Benoit et al., 2017). CC takes place when people coordinate the acquisition and distribution of a resource. This coordination is frequently done for a fee or other non-monetary compensation through trading, bartering, and swapping (Belk, 2014). It is “the peer-to-peer-based activity of obtaining, giving, or sharing the access to goods and services, coordinated through community-based online services” (Hamari et al., 2016, p. 2047). According to some scholars, it is not clear the future growth of CC and its impact on incumbent industries (Barnes and Mattsson, 2017). However, as an emerging phenomenon from the computer-mediated interaction between customers and providers, CC is present in a vast range of business, such as transportation (Uber, Zipcar), lodging (Airbnb), tourism (Couchsurfing), entertainment (Spotify) and online food delivery services (Just-Eat.com, Clickdelivery.com, UberEATS) (Pigatto et al., 2017).

Online food delivery services (OFD) offer opportunities for research as they are underrepresented in the literature of CC. According to Pigatto et al. (2017) these services can be characterized as business platforms that provide order services, payment and monitoring of the process but are not responsible for the preparation and order delivery operations. Although large fast-food chains like McDonald’s or Domino’s Pizza offer their delivery services, small or medium restaurants chains have seized the emergence of intermediaries that provide these sort of services (Yeo et al., 2017).

Despite the popularity of these platforms nowadays, their connection with CC remains neglected. The words “Online Food Delivery” are missing in most recent papers of CC (Pigatto et al., 2017; Hamari et al., 2016; Benoit et al., 2017; de Rivera et al., 2017), and the publications that aim the study of OFD (Hong et al., 2016; Gupta and Paul, 2016; Yeo et al., 2017) do not mention CC either. This lack of connection does not imply that they are independent and unrelated consumption phenomena. It only means that a relevant theoretical framework that illustrates the relationship between the two remains fragmented and ill-defined. Our aim here is to tackle this gap. Our approach posits the relevance of evaluating the impact of traffic conditions on the customer-provider relationship because OFD and CC are intrinsically related to urban transportation. This approach deviates from what we call...
the “standard methodological approach” that relies on surveys for data collection purposes. Although survey data is often intended to measure quality service through customers’ perception, this perception does not reflect delivery time accurately.

We do not claim that surveys are useless to studying the customer-provider relationship. In fact, they are still present along with other data collection techniques in the researcher’s toolbox (Garrett et al., 2017; Scaburro, 2015). Nonetheless, we grant fundamental importance to the sort of data that is available on websites or smartphone apps, as they reflect the spontaneous information of customers and providers who use a platform to enjoy the benefits of CC. The following example will illustrate this point. Imagine a client ordering a pizza, and the platform notes his pizza will be at home in less than 35 min, but it arrived an hour later due to a massive traffic jam. From a frequent customer’s view, this event might have none impact on the service perception. For a first-time customer, in contrast, this event might be perceived differently.

Another important observation regarding these platforms is that they can be strategic tools for business competitiveness (Yeo et al., 2017; Jiang and Tian, 2016; Lindblom and Lindblom, 2017). For example, providers who track and adjust their delivery times according to traffic conditions might increase the number of customers, as their clients will be sure that they will receive their orders on time or a price discount for delivery delays. This example shows that the ever-changing nature of traffic dynamics is relevant for all business segments that rely on transportation to promote CC. The UberEATS application (https://www.uberEATS.com/), for example, represents a successful case that shows how a brand like Uber can also be compatible with a related service like food delivery.

The emphasis on collecting data from websites or smartphone apps demands the employment of recent computational techniques (Munzert et al., 2015; Landers et al., 2016; Lang, 2017), which we group under the umbrella name of web mining (Russell, 2014). We believe that the potential of this approach is its ease of implementation, ensuring the possibility of being adopted by business and social scientists for the following couple of reasons. First, as it relies on the records that both customers and providers do on the platform, it allows the analysis of real customers preferences rather than self-reports. Second, and not least significant, as these records tend to be massive, they could become an invaluable data source that facilitates the incorporation of Big Data techniques (Chen and Wojcik, 2016; Cheung and Jak, 2016; Sivarajah et al., 2017) in the CC literature. As the relationship between OFD and CC poses implications for different disciplines like economics, business, and consumer psychology, we deem necessary to provide an interdisciplinary approach to such connection.

2. A theoretical framework to the OFD-CC relationship

Botsman and Rogers (2010) published a book that summarizes the rise of CC and the “sharing economy”; a trend consisting of individuals’ interest in access to rather than owning products or services (Hamari et al., 2016). The popularity of these worldwide movements emerged because their benefits outperform those of buying and owning things, especially after the financial crisis in 2008. The cost-benefit ratio, widely employed by evolutionary biologists (Nowak, 2006), was an inspiring concept for the theory on the efficiency of commons-based society proposed by the economist Elinor Ostrom (1990). As put it by Botsman and Rogers (2010), Ostrom’s research “has demonstrated that even in capitalist societies, if simple rules are applied, a self-organized commons can work. Individuals will cooperate to act in the common good” (p. 19). These ideas are more recently explained from an interdisciplinary view addressing the question of how to make cooperation successful (Jaffe, 2017). Jaffe (2017) notions of temporal synergy explain why Uber has been incredibly prosperous. Roughly speaking, synergy occurs when the output of a system is not equal to the simple sum of the output of the interacting agents that compose the system (i.e., customers, technology-based platforms, and providers). Cases of synergy exist everywhere. In urban transportation, for example, synergy occurs as the augmentation of road capacity when motorcyclists ride in between the lanes of four-wheeled vehicles (Correa, 2017). Nonetheless, synergistic mechanisms that facilitate the emergence of CC are missing in recent theoretical frameworks (Benoit et al., 2017).

Identifying possible synergies between CC and OFD requires the understanding of the motivations and the meanings of using these platforms for its users. Barnes and Mattsson (2017) illustrate the utility of the theory of reasoned action (TRA) (Fishbein, 1979) to understand how volitional control and attitudes act as predictors of using a car-sharing platform. According to this framework, if a person has positive attitudes towards the safe use of CC platforms, then his intentions to use a platform will grow. Another relevant theoretical perspective is the technology acceptance model (TAM) (Davis, 1989). The TAM states the existence of two psychological factors that motivate the acceptance of technology innovations. These factors are “perceived usefulness” and “perceived ease of use”. The former refers to the perception of a user about the subjective probability that the use of technology will help increase his performance. The latter refers to the individual’s subjective appreciation that using a particular technology involves less efforts. The integration of TRA and TAM became a well-accepted proposition since Pavlou (2003) and Gefen et al. (2003) posited the role of trust and risk as the mechanisms that facilitate the adoption of technological innovations, such as OFD platforms. The primary constructs for capturing consumer acceptance of OFD platforms are the intention to transact and the on-line transaction behavior. These constructs, however, are related to trust and perceived risk given the uncertainties that exist in OFD platforms that do not provide feedback about orderings receipt and expected delivery times. Yeo et al. (2017) found that the behavioral intention towards using OFD was strongly associated with their perceived convenience and usefulness, and with the enjoyment resulting from using these platforms to receive meals at home. These authors emphasized that “consumer perceptions become positive when they are able to avoid dealing with the physical burden of traveling” (Yeo et al., 2017), p. 157. These ideas are compatible with the notion of “possessions as the extended self” proposed by Belk (1988) who pointed out that consumers often assign subjective meanings to the things they would like to use or have access to Belk (2014). Thus, CC not only represents an alternative way of consuming but a new business paradigm that suits the purpose of OFD platforms.

According to Belk (2014), this paradigm is likely to shake established industries which, in turn, will show two types of knee-jerk reactions. On the one hand, they will show “flight reactions” consisting of diversifying out of the industry, and on the other hand, they will show “fight reactions” consisting of invoking the intellectual property rights to stave off the sharing economy. Beyond these reactions, industries can also show other types of adaptive responses. A third reaction assumes the creative destruction of old-fashioned business models to adopt new innovative ways of participating in the sharing economy. Once again, the transportation business shows these initiatives with the short-term car rental that BMW is conducting with the “Drive Now” mobility concept (https://www.drive-now.com/en/). A fourth reaction is to provide content for free and find other sources of revenue. Google services, for example, can be used at no cost, and the revenues come from customized advertising resulting from the content that matches user’s emails. Another reaction consists of buying up a leading company offering the platform, as it occurred with Zipcar which was bought by Avis, the American car rental company. In the case of OFD platforms, another reaction is the adoption of user-generated content (Huang and BenyouCEF, 2013; Goes et al., 2014) and urban mobility information services (Kahle and Wickham, 2013). These features allow customers the possibility of sharing their experiences with other users, who also welcome honest opinions about how traffic conditions affect delivery.

As recent business trends, both OFD and CC platforms are in their earliest stage. Nonetheless, the results of most recent studies show...
exciting implications for business research. For example, Gupta and Paul (2016) noted that OFD users in Eastern countries like China or Malaysia concentrate more on the possibilities these businesses offer regarding convenience and time-saving instead of their usage cost. In Western countries like Brazil or England, OFD are also overgrowing, providing added comfort to their customers when coupled with the relative ease of access afforded by the ubiquity of mobile Internet devices (Pigatto et al., 2017). Scholars characterized these platforms as an innovation of restaurants or food providers intended to increase their competitiveness (Yeo et al., 2017; Pigatto et al., 2017; Cavusoglu, 2012). This competitiveness might be evaluated by customers’ transactions actions volume and delivery time fulfillment, according to the typical traffic conditions that providers face. From a web mining perspective, it is imperative to point out how to collect these metrics. Up to date, Waze and Google maps are the most popular data sources for knowing local traffic conditions. Although they developed independently from each other, Google acquired Waze in June 2013. Both applications provide online visualizations of real traffic conditions for almost any city of the world. Waze has been used in Israel for detecting road safety events (Fire et al., 2012) and in Brazil (Silva et al., 2013) for characterizing traffic alerts at the city scale, and Google maps has been used for similar purposes (Kahle and Wickham, 2013). However, their potential use for OFD platforms remains unknown. Here we provide a procedure that illustrates their potential for new business models that rely on urban mobility to promote the use of OFD platforms.

3. Materials and method

We developed a procedure for retrieving key performance indicators of 1106 fast-food providers available at a Colombian OFD (https://domicilios.com/bogota). This platform allows providers to receive customer’s orders if they are within a radius of 6 kms. By using an advanced web scraper named “Agenty” (https://www.agenty.com/) we extracted the following indicators. First, we extracted the cost of the delivery, which reflects the amount of money charged for dispatching the food from the provider to the customer. Second, we obtained the expected delivery time, which is the providers’ declared times to deliver their orders to their customers. Third, we got the minimum ordering, that is, the minimum charge required for providers to deliver their orders to the customer. Fourth, we also collected the number of comments that customers have registered for each provider. The number of comments is the most relevant indicator of transactions volume. This number, in no way, equals the total number of customers who ordered a service. However, it shows the number of customers who ordered some service and left a positive or negative comment about it. As such, the number of comments is necessarily lower than the total of customers who made a transaction with the food provider, but it is sensibly informative about the customers who really care about allowing other customers know their experience with the food provider, which shows a conceptual match with the idea of collaborative consumption.

We calculated a fifth indicator, called “delivery time fulfillment” (DTF). DTF is the difference between providers own declared delivery times (publicly available on the website) and the expected travel time provided by Google Maps API (Kahle and Wickham, 2013). Google estimations of travel times consider the distance between the provider physical location and the physical customer location. A positive DTF indicates that a provider can deliver the ordering before their own declared delivery time. A null DTF means that a provider delivers the ordering on time, while a negative DTF implies a delay. We provide a GitHub repository where we share the scripts that we used for data collection and data analysis. The access to this repository is allowed under request.

We queried the geographic location of these providers from Facebook, and we set its latitude and longitude coordinates with the “Batch Geocode Tool” (https://www.mapdevelopers.com/batch_ geocode_tool.php). We took the presence on Facebook as another behavioral indicator that shows that restaurant owners are interested in interacting with their clients through this social network. We discarded those providers who did not show their geographic locations on Facebook for further analyses. A total of 787 food providers with commercial operations in Bogotá City composed our second data set. Because customers’ information is private according to the website information policy, we generated a random sample of geographic points from Google Maps as a valid replacement of real customers’ addresses. To evaluate the relationship between traffic conditions and OFD operations, we classified our sample of food providers as follows. Firstly, we used Google Maps to identify the typical traffic around each restaurant during rush hours on Saturdays; preliminary analyses using Google Maps API revealed us those rush hours took place in mornings (between 8 and 10 am), noons (between 12 and 2 pm), and afternoons (between 6 and 8 pm). Then, following Google Maps visualizations, we empirically classified the typical traffic of rush hours in the following categories: free or green traffic (G), average or orange traffic (O), and heavy or red traffic (R). We generated letter triads that allowed us to characterize the typical daily traffic. Thus, for example, the sequence “R-O-G” means that the typical traffic changes from “red” in the morning to “orange” at noon and “green” in the afternoon, describing a place where traffic conditions improve as time passes.

4. Results

We firstly considered the relationship between OFD performance indicators and traffic conditions as captured by Google Maps API. Kolmogorov-Smirnov tests revealed that none of the OFD indicators show a normal symmetrical distribution.

Fig. 1. Statistical distributions of online OFD performance indicators.
Given the large statistical variance of both the number of comments and the minimum charge ordering, we applied a log-transformation to their raw values for analytical purposes. Fig. 1(A) shows the statistical distributions of the logarithm of the number of comments. Food providers with the highest number of comments were those with heavy traffic in the morning and at noon (R-R-G), while those with free traffic in the morning and in the afternoon (G-R-G) received the minimum number of comments. Fig. 1(B) shows the distributions of the expected delivery time in minutes. The average expected delivery times ranged from 38 min for food providers located in places with heavy traffic in the morning and at noon (R-R-O) to 60 min for food providers located in places with highly congested traffic (i.e., an R-R-R daily traffic). Finally, Fig. 1(C) shows that the average minimum charge ordering in Colombian currency ranged from COP 9440 (approximately 3 US$) to COP 19,400 (6.5 US$) for restaurants located in congested points of the city (e.g., those with an RRO or an ORR daily traffic). To further understand the relationship between traffic conditions and online food ordering, we focused on the number of comments. As the absolute number of comments showed a great statistical variance ranging from zero to 12,830, we analyzed the statistical distributions of their logarithm. Fig. 2 shows these distributions. The number of comments revealed statistical differences according to the typical traffic for mornings (F = 9.44; p < 0.01; \( \eta^2 = 0.004 \)), noons (F = 15.01, p < 0.01; \( \eta^2 = 0.006 \)), and afternoons (F = 15.52; p < 0.01; \( \eta^2 = 0.008 \)), though the size of these differences proved to be negligible.

We proceeded by analyzing food providers’ DTF. We found that delays ranged between 40 and 53 min, but the majority of online food providers showed a satisfactory DTF as they tended to dispatch the orders 23 min before their own declared travel times during Saturdays rush hours. We obtained these statistics as the average of the travel times between the physical location of Online food providers and the physical location of customers during rush hour. Statistical distributions of DTFs revealed significant differences according to the Google typical traffic for mornings (F = 8.96; p < 0.001; \( \eta^2 = 0.002 \)), noons (F = 4.88; p < 0.01; \( \eta^2 = 0.002 \)), and evenings (F = 7.71; p < 0.01; \( \eta^2 = 0.002 \)), though the size of these differences also proved to be negligible.
It could be argued that as OFD platforms might show their expected delivery times according to real traffic conditions provided by Google services, these expectations already incorporate congestion effects. To test if this is the case, we estimated the Spearman non-parametric correlation matrix between the declared delivery time (DTT), Google estimations of travel times (GoogleTT), the logarithm of customers comments (LogComments), and the delivery time fulfillment (DTF) (Fig. 3).

As can be seen in Fig. 4, providers’ declared times (DTT) are independent of Google estimations of travel times. And, while customers comments showed a significant association with DTT, this association is misleading when real travel times are taken into consideration. The correlation between DTF and the logarithm of customers comments also proved to be significant, revealing that the number of comments is sensitive to early deliveries.

5. Discussion

A scrutiny of the impact of traffic conditions on key performance indicators of online food delivery services was the aim of this paper. Since users of these services can share their experiences, by rating providers and posting their opinions about the quality of the received service, we regard that this evaluation is a necessary step for advancing our knowledge of collaborative consumption (Botsman and Rogers, 2010) and its relation to online consumers behavior (Roos and Hahn, 2017). As OFD are business models that rely on urban transportation to alleviate customers’ burden of traveling in highly dense cities (Yeo et al., 2017), we assumed that traffic conditions might impact key performance indicators of these business platforms during rush hours. The results suggested that this assumption finds partial support in the case of Bogotá City. Definitive conclusions will emerge by replicating our approach to other congested cities, and by designing controlled field experiments, in which, food delivery orders are carefully scattered across the city. We regard these considerations as the next step in our research agenda that aims to identify possible synergistic mechanisms by which OFD platforms emerge as successful new business models.

Compared with recent studies of OFD platforms (Jia, 2018), our work has shown the following contributions. Firstly, we have outlined an agenda that focuses on the role of urban mobility in consumer research from a web mining perspective (Correa and Forero, 2017; Correa, 2018). Indeed, our methodological approach has shown the advantages of web scraping (Munzert et al., 2015; Landers et al., 2016) to retrieve key performance indicators of food providers and evaluate their relationship with traffic conditions as provided by Google Maps API (Kahle and Wickham, 2013). We foresee that this topic will catch the attention of more scholars shortly as tracking traffic conditions might be relevant for other business models that rely on urban transportation, (e.g., Uber, Zipcar or Airbnb). Secondly, our study has shown the potential of developing and using APIs as relevant data sources that will change the way we collect data for the analysis of new business models. Given the fact that some of these APIs are already available for retrieving information related to the use of Uber in a city (Collier and Wu, 2017), we foresee that they become an essential tool for advancing our knowledge of platforms of collaborative consumption. In the case of the Uber API, for example, the possibility to track how dynamic fares change according to traffic conditions will provide us with valuable data to get better insights about patterns of consumers behavior that could not be observable before.

The managerial implications of our study are clear as well. Our approach is a first attempt that shows the potential of assessing the business value of the so-called “Big Data Analytics” (Côrte-Real et al., 2017) for collaborative consumption platforms that rely on urban transportation to support its commercial operations. Indeed, we concur with the idea that big data analytics “can be an effective aid to survival in competitive markets, particularly by supporting production and operations or product and service enhancement” (Côrte-Real et al., 2017), p. 387. The improvement of OFD services might be supported by analyzing customers comments with recent text mining techniques (Silge and Robinson, 2017) like the ones employed by Jia (2018). The potential for this approach also includes the possibility to collect data in different cities wherein global OFD platforms exist (e.g., UberEATS, Clickdelivery, etc.), which might put us in the situation to understand the existence of cultural differences in food preferences, and customers levels of satisfaction.

References

J.C. Correa et al.

Journal of Retailing and Consumer Services xxx (xxxx) xxx–xxx