27th Bled eConference

eEcosystems

June 1 - 5, 2014; Bled, Slovenia

Spread like a virus. A model to assess the diffusion of dynamic ridesharing services

Riccardo Bonazzi

University of Applied Sciences of Western Switzerland

riccardo.bonazzi@hevs.ch

Fabio Daolio

University of Lausanne, Switzerland

fabio.daolio@unil.ch

Abstract

Dynamic ridesharing is a derivative of regular carpooling, which enables the formation of carpools on an as-needed basis, usually on very short notice the shared travel purpose also extends to a broad range of activities, beyond work or school. In this paper we propose a model to monitor the adoption of a dynamic ridesharing service, intended here as a mobile service that needs to achieve a critical mass to survive. Our theoretical model is inspired from the SIR model used in epidemiology to control the spread of an infectious virus. We test our model using real-data from two firms offering dynamic ridesharing services. Our model complements the view that innovative services evolve following an S-shaped curve, and it has practical relevance for managers and investors, who want to monitor and compare the evolution of competing firms in the field.

Keywords: dynamic ridesharing, mobile service, diffusion of innovation, SIR model, ouicar, blablacar

1 Introduction

This paper presents a research in progress and it is addressed to managers looking for a tool to assess the evolution of the user base of a dynamic ridesharing service.

Dynamic ridesharing (also known as real-time ridesharing) is a derivative of regular carpooling, which enables the formation of carpools on an as-needed basis, usually on very short notice and whose shared travel purpose also extends to a broad range of activities, beyond work or school (Siddiqi and Buliung 2013, p.480).

On the one hand, the particular nature of dynamic ride-sharing requires a sophisticated way to manage a large number of drivers available on short notice. Indeed, while classic carpooling allows users to define a meeting point some days in advance, dynamic ridesharing allows only few minutes to perform the matching between an interested rider and a potential driver, before the rider gives up and looks for another transportation option. On the other hand, dynamic ridesharing appears to be suitable for all those activities that do not require or allow planning in advance. Indeed, while classic car pooling is restricted to regular trips, such as home-work, or long-distance trips, such as going to a festival, dynamic ridesharing positions itself closer to taxi services, which are available at any time. Therefore, we consider dynamic ridesharing as a mobile service, which offers a new mobility solution and which might be cheaper and more ecological than taxi and private cars.

The website dynamicridesharing.org collects a large number of projects done by startup and companies in USA as well as around the world. Most of these projects have failed in the past, even though the recent examples of Lyft and Uber in USA, allows to be optimistic with respect to future development. Nonetheless, it is known that one of the greatest cause of failure is the lack of a critical mass achieved by the dynamic ridesharing service. Indeed, without a significant amount of drivers available, it is unlikely that a rider will find one driver close by on a short-notice. Consequently, if riders are not going to use the service, there will be no driver interested in declaring himself/herself available and the service will have no reason to exist.

Therefore, the monitoring of the user base appears to be a crucial element to monitor for managers of dynamic ridesharing services. We also believe that such key performance indicator has influenced the amount of money, which dynamic ridesharing startups have obtained in recent years by venture capitalists. Nonetheless, data about user base of dynamic ridesharing services is usually not fully shared with the public. Indeed, a statement like “Service X has 100’000 users” cannot be used for reliable assessments (for example, it would be useful to know when was this number collected and how it relates to the ones collected one month before and one month after).

Hence, it would be relevant to obtain useful metrics and a theoretical model to monitor the diffusion of a dynamic ridesharing service, to predict how many users will start -and how many will stop- using it. Therefore, our research question is: **how can we assess the diffusion of a dynamic ridesharing service among potential users within a target market?**

The rest of the paper proceeds as it follows. Section 2 briefly illustrates the existing literature, which address our research question. Section 3 illustrates our theoretical model and the methodology used to test it. Section 4 illustrates the results obtained. Section 5 summarizes the key elements of the paper and illustrates further directions of investigation.

2 Literature review

In order to address our research question we have defined a “literature search protocol”. We have translated our research question into keywords: *"dynamic ridesharing"* and *diffusion*. We have used the keywords to search on Google scholar any research article available online, written in English, which proposes a theoretical model to answer our research question.

Our query returned eight results. One article was written in French and it had to be dropped from the list (Bellet and Clavel, 2007), whereas a master thesis written by a French student in English (Arbouet, 2011) addressed mobility in general and did not enter too much into details with respect to dynamic ridesharing. One book (Handke and Jonuschat, 2012) offered interesting insights about how to build a real-time ridesharing service, but it did not offer any significant insight on how to monitor the evolution of the user base. The report of Schiavone (2006), describes a view of dynamic ridesharing as an emerging service in 2006, and that underlines the great improvements that this type of service has done in the last eight years. The recent article of Siddiqi and Buliung (2013) offers a great overview of the evolution of dynamic ridesharing in the recent years by means of cases studies, but it focuses on the complementary relationship between dynamic ridesharing and information and communication technology. It is worth noticing that the authors observe that “there appear to be few success stories, most of the projects were discontinued for reasons ranging from cost inefficiencies (high capital and operating costs), lack of use, poor service levels, usability, and technological limitations. One measure of success for these case studies is the number of rides matched and executed, that is, the number of trips completed after a match between driver and passenger was made« (Siddiqi and Buliung 2013, p 493). The rest of the articles addressed dynamic ridesharing as a technical issue (Knapp, 2005; Lasdon and Machemehl, 2005).

Therefore, we can conclude that there is a gap in the literature with respect to a set of useful metrics and a theoretical model to monitor the diffusion of a dynamic ridesharing service -intended here as a mobile service- to predict how many users will start and how many will stop using it.

Nonetheless, previous research allows us to address this problem in (at least) two possible ways: (1) by monitoring the diffusion of such innovative service and (2) by analysing the change in the structure of the users’ network. There is a large amount of literature that addresses diffusion of innovation and adoption of innovation. For sake of simplicity, we cite the seminal book of Everett Rogers (1962), which popularized the notions of critical mass and technology adoption. Most of literature in diffusion of innovation and technology adoption assumes an evolution following the famous “S-curve”, which has been used to describe some of most famous innovation of the last century (Moore and Simon, 2012). Nonetheless, scholars in dynamic analysis of networks have used different models to describe the change in the network. Since this study concerns the spread of a new approach to conceive transportation, we refer to the model of Susceptible-Infected-Recovered, which has been used in epidemiology for long time (Norman 1975) and that has recently used by scholars to model Facebook and Myspace users evolution (Cannella and Spechler, 2014).

3 Our theoretical model and chosen methodology

Our theoretical model uses six constructs. The susceptible contacts are named *potential candidates* and their number increase over time thanks to *new potential candidates*, which are acquired through marketing campaigns. The infected contacts are named *user* *base* and their number increase following the interaction among users and potential candidates. Indeed, if the probability that a potential candidate will become a user after having seen the service (that is, the *adoption rate* is very high) and if a potential candidate interact with a lot of users (that is, *the interaction rate* is very high) she will be more likely to be become a user. Table 1 shows the detailed formulas used in our model, whereas figure 1 shows the interaction among constructs.

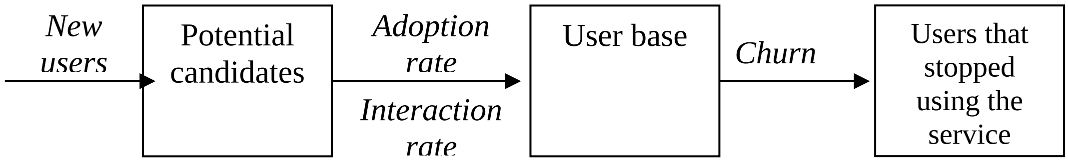

Figure 1: The constructs of our model

To operationalize our theoretical model, we need to have longitudinal data of different ridesharing firms, in order to test it over a certain span of time. Therefore, we have decided to use weekly scores obtained from Google Trends searching for “service x” and restricting the search on the country “Alpha”.

Construct	**Description**	**Formula**
User base	Weekly users for service X on the country Alpha	Change in users at time (t) = new users after the interaction among users and candidates - users that left = (Potential candidates at (t-1)*user base at (t-1)* Interaction rate) * Adoption rate - User base * Churn
Potential candidates	Potential candidates to be interested about service X in the country Alpha	Change in potential candidates at time (t) = new potential candidates - new users after the interaction among users and candidates = new potential candidates - Potential candidates at (t-1)*user base at (t-1)* Interaction rate) * Adoption rate
Adoption rate	Probability that a new user will start using the service, once it has discussed with an existing user	Fix parameter defined in the model: [0.0%-100%]
Interaction rate	Percentage of existing users that interact with each new user, with respect to the total amount of existing users.	Fix parameter defined in the model: [0.0%-100%]
Churn	Percentage of existing users that stop using the system, because they have found a better alternative.	Fix parameter defined in the model: [0.0%-100%]
New potential users	Percentage of new users that stop using the system, because they have found a better alternative.	Fix parameter defined in the model: [0.0%-100%]

Table 1: The constructs of our model and the formulas used to obtain them

The choice of Google Trends as a source of data might be challenged. Previous studies have already shown how google trends can be used to predict spread of influenza across the country: Hence, we are wondering if it could apply even to dynamic ridesharing services, which we intend to model as a contagion of a new approach to mobility. On the one hand, Google Trends data might extremely sensitive to marketing campaigns of firms, since they reflect the interest of potential users that have been reached. On the other hand, data obtained from ridesharing firms are rarely available and potentially biased, since they reflect the strategy of each company to communicate to investors and to not communicate to competitors. Accordingly, we state our null hypothesis, which would falsify our model:

H0: The weekly scores obtained from Google Trends searching for "service x" and restricting the search on the country "Alpha" are not a useful representation of the evolution of the user base.

In order to falsify our null hypothesis, we need look for evidences of strong correlation between these data trends and evolution in the success of ridesharing trends.

Assuming that H0 is falsified, we need to proceed and fit the parameters of our model with the data trends obtained. In order to avoid sample selection bias, we intend to split the collected data into two parts: (1) data used to obtain the parameters and (2) data used to test the model. Consequently, we can propose our hypothesis, which concerns the shape of the diffusion trend line. **H1: The diffusion of ridesharing services follows an S-shaped trend line.**

The implications of this hypothesis are significant for managers and investors. Indeed, if the diffusion of dynamic ridesharing does not follow an S-shaped curve, it might imply that continuous investments will be required to sustain the growth of the user base.

4 Test results and discussions

In this section we present a test performed on two dynamic ridesharing firms in France. We have preferred France to USA, since the dynamic ridesharing firms in France have started later in time and their data is less compromised by pivoting choices of the startup (that is, the firm decides to dramatically change the service that it is offering or to change name and brand). Nonetheless, we have compared the data from a new emerging startup (ouicar) with a startup that has existed for many years (blablacar, previously known as covoiturage.fr).

In order to falsify our null hypothesis (H0), we have observed the spread between the scores of the two companies on Google scholar in December 2013 (more or less 3:1) and we have compared it with the amount of money that the two companies have received from venture capitalists, as measured two months later: EUR 4.5 Million for Ouicar VS EUR 12.5 Million for Blablacar.

This result might be due to a coincidence, but if we assume that the final step of negotiation with venture capitalists often takes a couple of months, we can notice the same results in USA, when we compare the total funds obtained in August 2013 by Lyft (USD 83 Million) and Uber (USD 410 Million) and their ratio on Google trend in June 2013 (1:5). Therefore, we have some reasons to believe that score on Google trends might have some predicting power, since they might represent the change in interest of the potential candidates and the interactions with the user base. Hence, we consider that **the null hypothesis is rejected**.

Figure 2 shows the trends of Ouicar and Blablacar, together with the values predicted by our model. As one can see, there are two peaks in the trend of ouicar, which the model cannot predict. We believe that this is a consequence of the sensibility of the Google trends data to the effect of strong marketing campaigns.

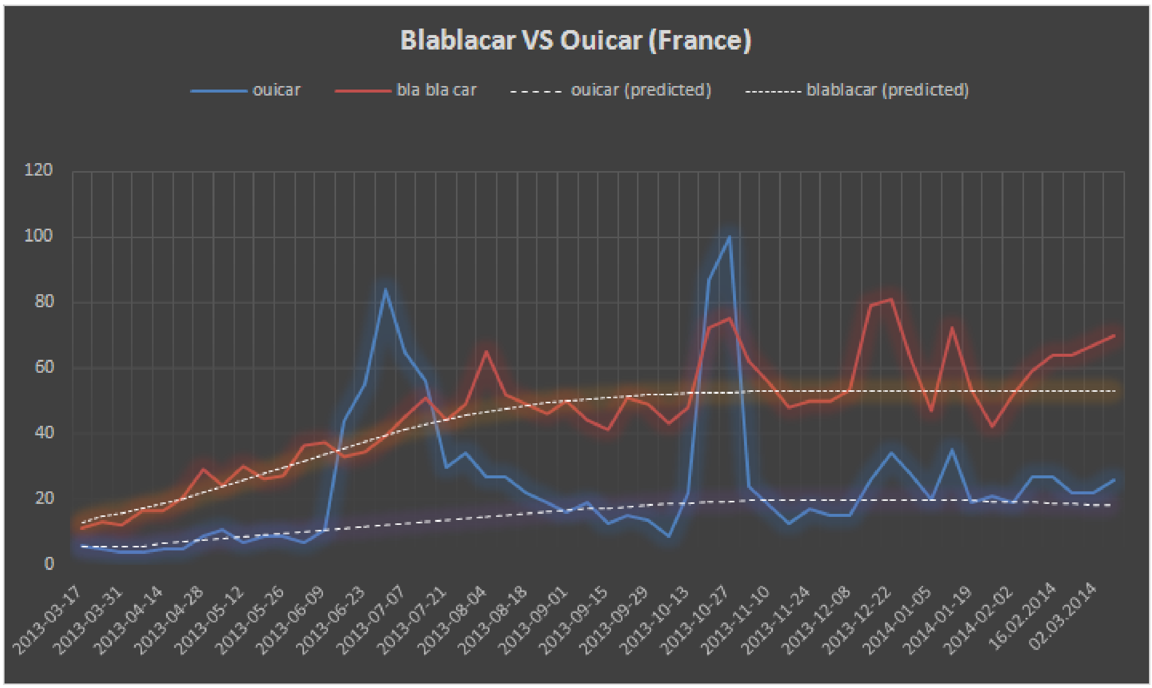

Figure 2: Google trends data for ouicar and blablacar

By fitting our model with the Google Trends data from March, 17th until November, the 24th, we induced the parameters presented in table 2. The detailed set of data is available in the annexes.

We have chosen to stop our training data set at the end of November 2013, since that was the model when the ratio between the two companies was consistent with their evaluations by investors. Nonetheless, as the figure 1 shows, our model appears to be able to represent the evolution of the Google Trends scores for the following three months.

	Potential candidates (Google Trend score)	**Infection Rate**	**Interaction rate**	**Churn**	**New users**
blablacar	64	3%	14%	4%	3%
ouicar	84	6%	3%	4%	50%

Table 2: Chosen parameters for the model

Assuming that the induced parameters shown in table 2 are reasonably close to the real values, some considerations arise. First of all, it appears that the potential target of ouicar is greater than blablacar. This might be due to the sensibility of google trends to marketing campaign and that might not imply that the real amount of potential

candidates is greater. Moreover, the difference between the two infection rates and interaction rate could be a sign of product differentiation, whereas the strong difference between the parameters related to new users is once again, related to the two peaks in the ouicar trend, and it might simply mean that the marketing and communication efforts of ouicar are greater than blablacar.
In the end, it appears that the evolution of the two services follows a S-shaped curve Therefore, **our hypothesis (H1) is partially verified** for the time being, even if the trend is quite volatile at the moment. This might be due to the continuous evolution of the service, which might require a more sophisticated model of diffusion to be represented.

5 Conclusions

In this study we intended to develop a set of metrics to monitor and predict the evolution of the user base to assess the diffusion of a dynamic ridesharing service among potential users within a target market.
By combining notions from the literature about diffusion of innovation and the literature of dynamic network analysis, we obtained a simple model, which treats innovation as a spreading contagion. We tested the model on data obtained from Google trends, which are freely available and that represent an alternative to data offered by firms, which are always torn between the need to rise interest from investors and the urge to not divulge too much to competitors.

The results, which we have obtained for the evolution of two French firms, show that Google trends score might be an alternative source of information to assess the evolution of the user base, and confirm that our model have predictive power to support strategic decisions within the firm. Nonetheless, the units used by Google to define its trends could not be obtained by the authors.
In the future, we intend to test the model with data from other firms and countries, and to address two major limitations of our model, namely (1) the intrinsic assumption that there is only one type of user and (2) the assumption that users will eventually stop using the service.

References

Arbouet, Marie (2011). “In Search of Mobility Services,”. Retrieved from http://www.diva-portal.org/smash/record.jsf?pid=diva2:436393.

Ballet, Jean-Christophe, and Robert Clavel (2007). “Le Covoiturage En France et En Europe: État Des Lieux et Perspectives,”. Retrieved from http://lara.inist.fr/handle/2332/1453.

Cannarella, John, and Joshua A. Spechler (2014). “Epidemiological Modeling of Online Social Network Dynamics.” *arXiv:1401.4208*. Retrieved from http://arxiv.org/abs/1401.4208.

Handke, Volker, and Helga Jonuschat (2012). *Flexible Ridesharing: New Opportunities and Service Concepts for Sustainable Mobility*. Springer, 2012.

Hwang, Mimi, James Kemp, Eva Lerner-Lam, Nancy Neuerburg, Paula Okunieff, and Schiavone, John (2006). *Advanced Public Transportation Systems: The State of the Art Update 2006*. US Department of Transportation. Retrieved from http://www.fta.dot.gov/documents/APTS_State_of_the_Art.doc.

Knapp, Geoff (2005). “Integrating Traveller Services: The Ride Points System,”.

Lasdon, Leon S., and Randy B. Machemehl (2005) “Improving ITS Planning with Multicriteria Decision Analysis,”. Retrieved from http://www.lib.utexas.edu/etd/d/2005/wangz36672/wangz36672.pdf
Moore, Stephen, and Julian L. Simon (2012). “The Greatest Century That Ever Was: 25 Miraculous Trends of the Past 100 Years,”. Retrieved from http://object.cato.org/publications/policy-analysis/greatest-century-ever-was-25-miraculous-trends-past-100-years.
Mote, Jonathon E., and Yuko Whitestone (2001). “The Social Context of Informal Commuting: Slugs, Strangers and Structuration.” *Transportation Research Part A: Policy and Practice* 45 (4), pp 258–68.
Siddiqi, Zarar, and Ron Buliung (2013). “Dynamic Ridesharing and Information and Communications Technology: Past, Present and Future Prospects.” *Transportation Planning and Technology* 36(6), pp. 479–98.

6 Annexes: Weekly scores from Google trends

Training data

Week	ouicar	bla bla car	ouicar (predicted)	blablacar (predicted)
2013-03-17	6	11	*6*	*13*
2013-03-24	5	13	*6*	*14*
2013-03-31	4	12	*6*	*16*
2013-04-07	4	16	*7*	*17*
2013-04-14	5	16	*7*	*18*
2013-04-21	5	20	*7*	*20*
2013-04-28	9	29	*8*	*22*
2013-05-05	11	24	*8*	*24*
2013-05-12	7	30	*9*	*25*
2013-05-19	9	26	*9*	*27*
2013-05-26	9	27	*10*	*29*
2013-06-02	7	36	*10*	*31*
2013-06-09	11	37	*11*	*33*
2013-06-16	44	33	*11*	*35*
2013-06-23	55	34	*12*	*37*
2013-06-30	84	39	*12*	*39*
2013-07-07	65	45	*13*	*41*
2013-07-14	56	51	*13*	*43*
2013-07-21	30	44	*14*	*44*
2013-07-28	34	49	*14*	*45*
2013-08-04	27	65	*15*	*47*

2013-08-11	27	52	*15*	*48*
2013-08-18	22	49	*16*	*49*
2013-08-25	19	46	*16*	*49*
2013-09-01	16	50	*17*	*50*
2013-09-08	19	44	*17*	*51*
2013-09-15	13	41	*18*	*51*
2013-09-22	15	51	*18*	*51*
2013-09-29	14	49	*18*	*52*
2013-10-06	9	43	*19*	*52*
2013-10-13	22	48	*19*	*52*
2013-10-20	87	72	*19*	*52*
2013-10-27	100	75	*19*	*53*
2013-11-03	24	62	*20*	*53*
2013-11-10	18	55	*20*	*53*
2013-11-17	13	48	*20*	*53*
2013-11-24	17	50	*20*	*53*

Test dataset

2013-12-01	15	50	*20*	*53*
2013-12-08	15	53	*20*	*53*
2013-12-15	26	79	*20*	*53*
2013-12-22	34	81	*20*	*53*
2013-12-29	28	63	*20*	*53*
2014-01-05	20	47	*20*	*53*
2014-01-12	35	72	*20*	*53*
2014-01-19	19	53	*20*	*53*
2014-01-26	21	42	*19*	*53*
2014-02-02	19	52	*19*	*53*
2014-02-09	27	59	19	53
2014-02-16	27	64	19	53
2014-02-23	22	64	19	53
2014-03-02	22	67	18	53
2014-03-09	26	70	18	53