Modeling adoption and diffusion of new transportation system in station level

Cen ZHANG[1], Jan-Dirk SCHMÖCKER[2], Toshiyuki NAKAMURA[3], Masahiro KUWAHARA[4], Akira YOSHIOKA[5], Yukiko HOMMA[6]

[1] Doctor Candidate, Dept. of Urban Management., Kyoto University
(C1-2-436, Kyoto daigaku-Katsura, Nishikyo-ku, Kyoto 615-8540)
E-mail: zhang.cen.44v@st.kyoto-u.ac.jp
[2]Member of JSCE, Associate Professor, Dept. of Urban Management., Kyoto University
(C1-2-436, Kyoto daigaku-Katsura, Nishikyo-ku, Kyoto 615-8540)
E-mail: schmoecker@trans.kuciv.kyoto-u.ac.jp
[3] Member of JSCE, Designated Associate Professor, Institute of Innovation for Future Society., Nagoya University
(609 National Innovation Complex, Furo-cho, Chikusa-ku, Nagoya 464-8603)
E-mail: nakamura@trans.kuciv.kyoto-u.ac.jp
[4,5] Technical Research Division 2, TOYOTA InfoTechnology Center Co., Ltd.
(6-6-20 Akasaka, Minato-ku, Tokyo 107-0052, Japan)
E-mail:kuwahara@jp.toyota-itc.com, yoshioka@jp.toyota-itc.com
[6] ITS Planning Div., TOYOTA MOTOR Co., Ltd.
(1-4-18 Koraku, Bunkyo-ku, Tokyo 112-8701, Japan)

Understanding the adoption and diffusion of new transportation systems such as car sharing is remains challenging. We develop a methodological framework based on innovation and technology diffusion to describe the diffusion process focusing on spatial and social effects. Our objective is to be able to describe the expected new users over time for stations added to the network.

In order to do so, potential users are divided into two types: fast-adopters and followers. We formulate the fast-adopter model integrated with an information diffusion model to capture the initial changes. The follower model considers both the follower effect and a constant effect to present the effect that users enter the system over time. Further, to estimate the models not globally but at the station level, a redistribution model is established based on some reasonable assumptions. The least square method and gradient descent optimization algorithms are used to estimate the parameters.

The models were applied to the data from the Ha:mo RIDE car share system in Toyota city. The results illustrate that our model can fit the actual data well, both in terms of short-term sudden changes and long-term tendencies. Then, an initial spatial analysis of parameters was made. Finally, the reliability and problems of our method and further works are discussed.

Key Words : *Adoption , Diffusion, Car sharing, Social influences, Spatial effect, Demand forecasting*

1. INTRODUCTION

In recent years, with the emergence of new transport modes, such as the introduction of electric vehicles, high speed rail, car-sharing and bicycle-sharing, our transport systems are undergoing major changes, leading to profound impacts on travel behavior and lifestyles.

Both short-term and long-term demand dynamics forecasting remain important challenges for new transportation system, since currently we do not fully understand the mechanisms as to how a population adopts to new transportation systems.

Diffusion of innovation theory emerged in the first half of the 20th century [1)] and has received attention across multiple disciplines within economics and social sciences over the years [2)-4)]. The basic theory is simple and intuitive. Innovations diffuse into society following a logistic growth curve. Early demand for an innovation motivates additional future demand [5)].

Recently, disaggregate diffusion models, incorporates heterogeneity among decision-makers in the adoption process, in conjunction with discrete choice models to represent the utility maximization behaviour of decision-makers are becoming more and more popular, since the higher accuracy and more available personal data [6),7)]. Moreover, integrating stated preference (SP) with revealed preference (RP) data would be a better approach to forecast the adoption of new transportation services [8].

The aggregated models on system level can be implemented easily but could not account for more variables (e.g. spatial effects) and usually no good fitting for real data while disaggregated models in user level may be transferable across different conditions and have better prediction power, however, huge personal information collection is still the obstacle in practice.

The objective of this research is to develop a methodological framework to model the adoption and diffusion process by considering the spatial and social effect. Further, we aim to distinguish different adaptation behaviour for different types of stations and in different parts of the city leading to models at station level rather than on the system or user level. This allows to take the effect of facility extensions (new station establishment) into consideration which is the significant feature for transportation system diffusion.

The structure of this paper is as follows: After this introduction, in Section 2 we offer a description of the Bass model and our framework including decision making, information diffusion, station based adoption model and redistribution model.

In Sections 3, by using the least square estimation and gradient descent optimization algorithms, the parameters of stations can be estimated. In Section 4, as an example for adaptation to a car share system, data from Toyota city are used to illustrate how well new users to the system can be estimated. Finally, in Section 5 some initial conclusions and possible future research is discussed.

2. MODEL CONCEPTUALIZATION

In this section, at first, we will give a brief introduction of the Bass Model then offer a description of our framework to model the adoption in the station level.

2.1 The Bass Model

The Bass model is well-known in the marketing science. Adopters can be divided into two distinct groups: innovators and imitators. The technology diffusion literature stresses on the importance of the role of those two different types of adopters in shaping the market penetration rate of a new good or service. Innovators are individuals that “decide to adopt an innovation independently of the decisions of other individuals in a social system” while imitators are adopters that “are influenced in the timing of adoption by the pressures of the social system” [5]. Bass formulates the probability that a certain consumer will make an initial purchase at a given time t given that no purchase has been yet made by that specific consumer denoted as $P(t)$ in the equation below as a linear function of the number of previous buyers:

$$\frac{f(t)}{1-F(t)} = P(t) = p + q \cdot F(t) \tag{2-1}$$

with p as the coefficient of innovation; q as the coefficient of imitation; and $f(t)$ being the likelihood of purchase at t and $F(t)$ as the cumulative proportion of adopters by time t. p reflects the percentage of adopters that are innovators while q reflects the effect on imitators with an increase in the number of previous adopters.

In practice, we found that this model had the ability to predict the long-term tendency, but cannot explain the initial or short-term changes in the system.

2.2 Framework

2.2.1 The process of adoption

The process of adoption for individuals can be considered as three stages:

a. Information receive stage

Individuals receive some information about the transportation service at time t. For example, the information that a new station s was opened and was received by person j at time t.

b. Evaluation stage

Then person j will evaluate service and demand according to the attributes of service as well as station accessibility.

c. Decision making stage

Finally, person j will decide whether to adopt to the service or not at time t+1. In this stage, social effects, individual factors and the evaluation will both influence their decisions.

If one did not adopt to system, he or she will repeat the process *a.* to *c.* until the adoption.

2.2.2 Classification of individuals

In our models, depending on the willingness of adoption individuals can be classified into groups: Fast-adopters, Followers, and Non-adopters.

Fast- Adopters: Quickly adapt to new service after receive information of system.

Followers: Progressively adapt to the new service.

1.some of them may be affected by the other people around. If there are more people around using the service, then more people are likely to adopt to it. Here we call this the “follower effect”.

2.the other of them will decide to adopt the service independently. The probability to adopt would be a constant. Here we call this the “constant effect”.

Non-adopters: Have no potential demand and no interest in the new service, there is no way to encourage them to join.

2.2.3 Information diffusion model

We assume that the probability for individuals to obtain the information of stations is a constant.

$$g_s(t) = a_s \cdot \left(1 - G_s(t)\right) \quad (2\text{-}2)$$

$$G_s(t+1) = G_s(t) + g_s(t) \quad (2\text{-}3)$$

$$G_s(1) = I_s \quad (2\text{-}4)$$

$G_s(t)$: Cumulative the proportion of individuals received the information of station s by time t
$g_s(t)$: Proportion of individuals receive the information of station *s* at time *t*
a_s: Probability to obtain the information of station s (information diffusion speed)
I_s: Proportion of individuals receiving the information of station *s* when the station is opened.

2.2.4 Adoption model

As mentioned before, adopter can be divided into two groups, one is fast-adopters and other is the followers.

$$f_s(t) = f_s^1(t) + f_s^2(t) \quad (2\text{-}5)$$

$f_s(t)$: Ratio of adopters to the total market for station *s* at time *t*.
$f_s^1(t)$: Ratio of fast-adopters to the total market for station *s* at time *t*.
$f_s^2(t)$: Ratio of followers to the total market for station *s* at time *t*.

Fast-adopters are depended on the information diffusion and the percentage in the adopters.

$$f_s^1(t) = p_s \cdot g_s(t) \quad (2\text{-}6)$$

p_s: Proportion of fast-adopter for station s

There are more issues we should take into consideration in estimation of the followers, such as the follower effect and the constant effect mentioned before.

$$f_s^2(t) = (\mathrm{q}_s \cdot F_s(t) + \mathrm{c}_s) \cdot (G_s(t-1) - F_s(t)) \quad (2\text{-}7)$$

c_s: Coefficient of the constant effect
q_s: Coefficient of the follower effect
$F_s(t)$ Cumulative proportion of adopters at time *t* around station *s*

According to the estimated results at time *t*-1, updating the data at time *t*:

$$F_s(t) = F_s(t-1) + f_s(t-1) \quad (2\text{-}8)$$

The primary new users of stations can be calculated as:

$$N_s^p(t) = M_s \cdot f_s(t) \quad (2\text{-}9)$$

where M_s is the total number of potential users for station *s*.

2.2.5 Redistribution Model

We distinguish in our model so called “primary” and “secondary” new users. The difference between the two is that the latter are new users who in our model are (firstly) associated with other stations but who also start to use station *s*. In other words, users have a base station with which we associate them but users are obviously also using other stations. We refer to the new primary users as $N_s^p(t)$ and the secondary (or related) new users as $N_s^r(t)$.

Suppose that one new user is observed at a total of *A* stations in the initial time period when he joins the system, that is one primary, base station and A-1 secondary stations. Given this assumption we obtain following relationships where $N_s^o(t)$ is the estimated total number of new users at station *s*.

$$N_s^o(t) = N_s^p(t) + N_s^r(t) \quad (2\text{-}10)$$

$$\textstyle\sum_s N_s^r(t) = (A-1)\sum_s N_s^p(t) \quad (2\text{-}11)$$

Here we aim to redistribute the total number of related new users to different stations rightly. It is a reasonable assumption that for individuals, the probability to choose related stations depends on the general demand of the stations. Moreover, this general demand has a positive correlation with the current potential market ($G_s(t) \cdot M_s$). It can be explained as one station has more potential attracted users, usually has higher probability to be chosen as a related station by users attracted by other stations. $d_{s,t}$ denotes the portion of related new users for station *s* at time *t*.

$$N_s^r(t) = d_{s,t} \cdot (A-1) \cdot \textstyle\sum_s N_s^p(t) \quad (2\text{-}12)$$

$$d_{s,t} = {G_s(t) \cdot M_s}\Big/{\textstyle\sum_s G_s(t) \cdot M_s} \quad (2\text{-}13)$$

3. PARAMETERS ESTIMATION

In Section 2, we established some models to describe the process of adoption to the new transportation service at station level. Next, we aim to estimate parameters in the model by least square estimation. The objective is to adjust the parameters of the model function to best fit a data set. We formulate this as follows:

$$\mathrm{J}(\boldsymbol{\theta}) = \sum_{s=1}^{S}\sum_{t=1}^{T}(y_{s,t} - N_s^o(t)))^2 \qquad (3\text{-}1)$$

With $y_{s,t}$as the observed new users at time t for station s.

As shown in our models, denote θ_s as the parameters we want to estimate for station s. $\boldsymbol{\theta}$ the parameter set can present as following:

$$\theta_s = [\mathrm{a}_s \quad \mathrm{I}_s \quad \mathrm{p}_s \quad \mathrm{q}_s \quad c_s \quad \mathrm{M}_s] \qquad (3\text{-}2)$$

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_n \end{bmatrix} \qquad (3\text{-}3)$$

To search the optimize parameters of the function $\mathrm{J}(\theta)$, we make use of the Gradient Descent, which is a first-order iterative optimization algorithm. To find a local minimum of the function using gradient descent, one takes steps proportional to the negative of the gradient (or of the approximate gradient) of the function at the current point [9]. We start the process with some arbitrary initial value $\boldsymbol{\theta}_0$.Then we find the “better guess” $\boldsymbol{\theta}_1$.The process is repeated until a sufficiently accurate value is reached. γ is a small value to control the step length.

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \gamma \cdot \nabla \mathrm{J}(\boldsymbol{\theta}_k) \qquad (3\text{-}4)$$

4. CASE STUDY

4.1 Car sharing system (Ha:mo RIDE)

Ha:mo RIDE is a mobility sharing system providing better and more convenient access by small electric vehicles. It can be used for commuting, tourism, shopping or even transfer from public transportation smoothly and efficiently. Users can reserve vehicles by smartphone easily, and board at any stations and return the vehicle at whichever station is convenient.

This system was introduced in Sept 2013 in Toyota City as well as later in Tokyo, Grenoble and Okinawa. At beginning, Ha:mo RIDE in Toyota City had only few stations, but now has more than 4,000 registered members. By now, after several extensions there are 62 stations. In this case study, the time epoch for analysis was set as one month, and 40 months of individual rental records were obtained until Dec 2016.

4.2 Estimated New Users

The number of new users and estimated results over 40 months are shown in Fig.2. On an aggregate level one might distinguish following changes:

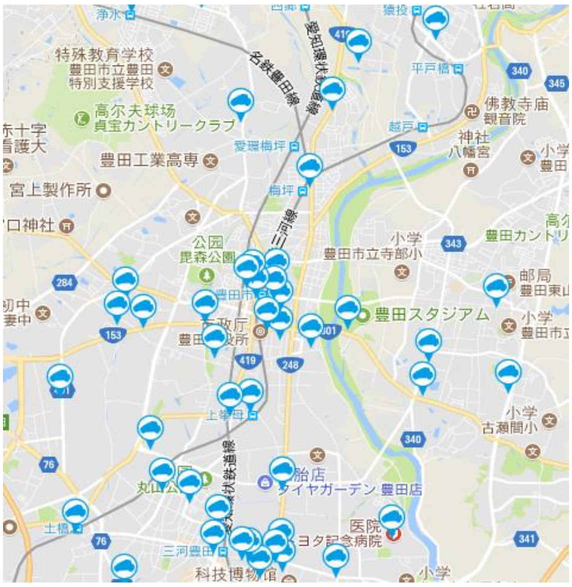

Fig. 1 Stations of Car Share System in Toyota (Ha:mo RIDE)

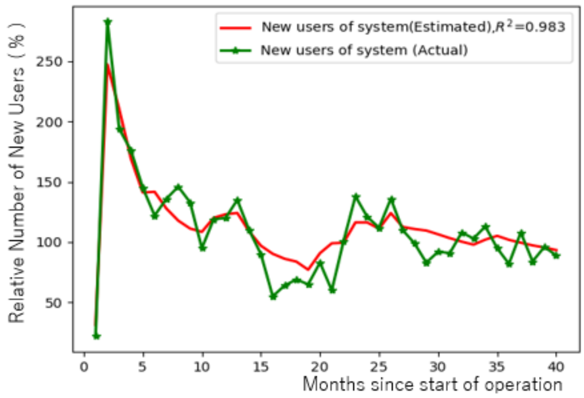

Fig. 2 Estimated Results
(set the new users in month 22 as the baseline = 100%)

In the initial two months we observed a sudden growth of new users. Then from month 2 to 15 (Oct 2013 - Mar 2015), the number of new users entering the system in a month quickly went down. But later the continuously growth from month 15 to 26 (Dec 2014-Nov 2015) was observed. Finally, the number of new users keeps roughly stable but a little drop during the remaining observed months.

The R-squared value, as a statistical measure that represents the percentage variation explained by the model, was calculated as 0.983. This illustrates the effectiveness of our model.

In general, our model can replicate the observed data quite well and reflect both the long-term tendency and short-term changes. However, we also notice that in the winter months 15-17 (Dec 2014 to Feb 2015), and 27-29 (Dec 2015 to Feb 2016), the estimated result is much lower than the actual. It may be caused by the low temperature in winter, which was not considered in our models.

4.3 Spatial analysis of parameters

4.3.1 Potential Market M_s

Fig. 3 shows the estimated potential market for the different Ha:mo RIDE stations. We observe that if there would not be Ha:mo RIDE stations at Toyota station or around Toyota Company a large number of Ha:mo RIDE users might not have started using the system.

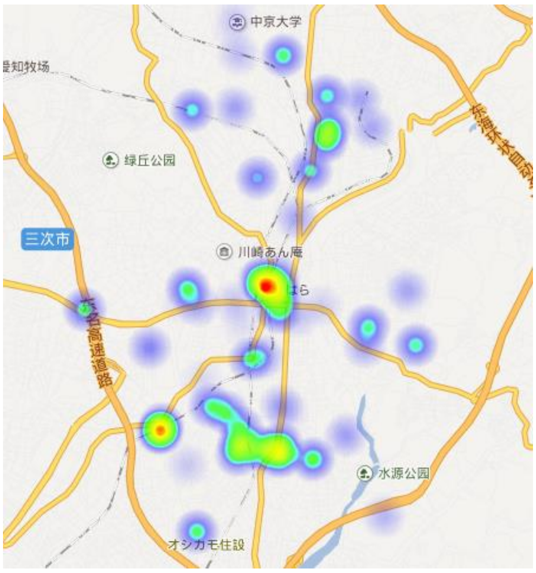

Fig. 3 Heat Map of Potential Market of Stations

4.3.2 Information Diffusion

Fig. 4 shows differences in information diffusion around different stations. Especially around railway stations our model estimates that the new user pattern can be best explained with high information diffusion. This seems reasonable as demand is dynamic and commuters are usually well-informed.

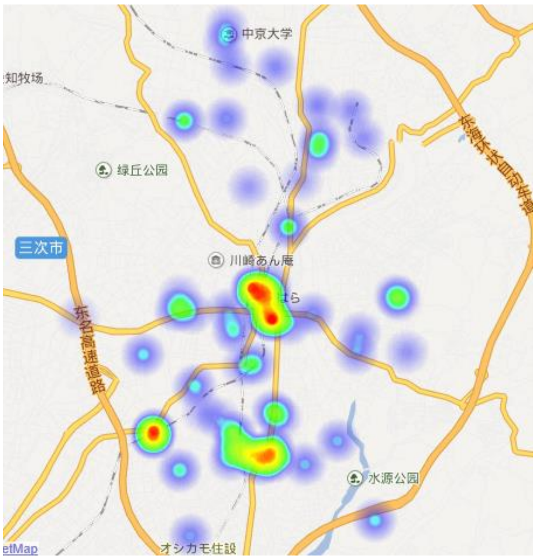

Fig. 4 Heat Map of Information Diffusion of Stations

5. DISCUSSION

The paper proposes a novel methodological framework based on innovation and technology diffusion model to describe the diffusion process by focusing the quantifying the spatial and social effects. We envisage this approach to be particularly useful for describing adoption to new transport services under the condition of the facility extension.

In this paper, we firstly discuss the aggregated model on system level, disaggregated model in user level and the comparison of these models. We therefore propose our framework including the process of adoption, classification of individuals, information diffusion model, adoption model and redistribution model.

Our models are applied to the data from the car share system in Toyota (Ha:mo RIDE). Over the 40 months of operation we could observe initial sudden changes and long-term tendency of new users. The results illustrate that our model can fit most of the adoption curves for stations over time with limited parameters very well. This curve can also be used to predict the future new users attracted to the system for existing stations. Therefore, our model can help targeting advertisement to specific locations and specific times to increase new users.

Beyond estimation, there are still some key issues that need to be discussed. Since we established our model in station level, there is an unavoidable problem that for the stations with few time periods and new users, the estimated parameter may be unreliable. Even they will not affect the results too much in system level analysis, we still need to deal with them carefully in the parameter analysis.

Finally, there are still a number of issues that need to be addressed in further work. Firstly, more analysis should be done on the relationship of station parameters. It will help to forecast the parameters for new stations. Secondly, the patterns of stations should be classified according to the estimated parameters. Besides, if the data are combined with population and land-use data our model might also be a basis for the demand prediction for new stations in existing schemes or even totally new markets.

REFERENCES

1) Rogers, E.M.: Diffusion of Innovations. Free Press of Glencoe, New York, 1962.
2) Sunding, D., Zilberman, D.: The agricultural innovation process: research and technology adoption in a changing agricultural sector. *Handbook Agric.Econom*, pp.207-261, 2001.
3) Delre, S.A., Jager, W., Bijmolt, T.H.A., Janssen, M.A.: Targeting and timing promotional activities: an agent-based model for the takeoff of new products. *J. Business Res*. 60, pp.826-835, 2007.
4) Dargay, J., Gately, D.: Income's effect on car and vehicle ownership, worldwide: 1960-2015. *Transport. Res. Part A*: *Policy Practice* 33, pp.101-138, 1999.
5) Bass, F.M.: A new product growth for model consumer durables. *Manage. Sci.* 15 (5), pp.215-227, 1969.

6) Higgins, Andrew, et al.: Combining choice modelling and multi-criteria analysis for technology diffusion: An application to the uptake of electric vehicles, *Technological Forecasting and Social Change* 79.8, pp.1399-1412, 2012.
7) El Zarwi, Feras, Akshay Vij, and Joan L. Walker.: A discrete choice framework for modeling and forecasting the adoption and diffusion of new transportation services, *Transportation Research Part C: Emerging Technologies* 79, pp. 207-223, 2017.
8) Jensen, A.F., Cherchi, E., Mabit, S., Ortúzar, J.de D.: Predicting the potential market for electric vehicles. *Transportation Science* Vol. 4, pp.427-440, 2016.
9) Snyman, Jan.: Practical mathematical optimization: an introduction to basic optimization theory and classical and new gradient-based algorithms, *Springer Science & Business Media* Vol. 97, 2005.

(Received July 31, 2017)