Modeling adaption to a new transport mode: Case study of Kyoto University's bicycle share system

Cen Zhang[1], Jan-Dirk SCHMÖCKER [2], Toshiyuki Nakamura [3],

[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] Assistant Professor, Dept. of Urban Management., Kyoto University
(C1-2-437, Kyoto daigaku-Katsura, Nishikyo-ku, Kyoto 615-8540)
E-mail: nakamura@trans.kuciv.kyoto-u.ac.jp

We formulate an approach based on stochastic state equations to describe the gradual change of behavior over time. At the start a user is likely to be in a low-usage state. A transition function then determines the likely change in behavior from one time period to another. We discuss time-homogeneity issues and possibilities to calibrate the transition function. The model is applied to panel data from Kyoto University's bicycle share system. Finally, the errors between actual and estimated value are analyzed to evaluate the model. The reliability and weaknesses of our method and further works are discussed.

Key Words : *Adaption , Panel data, Stochastic process, Markov Chain*

1. INTRODUCTION

Understanding long term demand dynamics remains an important challenge for transport. In most cases the future demand is uncertain, particular if there are changes in the supply system.

Considering situations in which new modes are being introduced or if major changes occur, "adaptation" is likely to take place, i.e. it requires time for travellers to adapt to the new available options. As these processes might even lead to different "final" equilibria, it is important that demand forecasting captures theses gradual processes.

Panel data are one way to observe these processes effective and accurately (Kitamura, 2003). Compared to time series and cross-sectional data which are the most frequently used ones for travel behaviour analysis, panel data have advantages in accurate inference of model parameters, greater capacity for capturing the complexity of travel behaviour and in identifying causalities (Yeun-Touh Li, 2016).

Therefore, to understand such adaptation we require panel data that track a person's behaviour over a longer period of time. Such data become increasingly available through smart card data as well as other "smart systems" that track the behaviour of users.

The objective of this research is to formulate an approach based on stochastic state equations that describes the change of behaviour over time by using panel data.

The process underlying the behaviour is assumed to be a Markov process in this study. In a Markov process, the behaviour at any given time point can be expressed by a set of discrete states and the process can be shown by transitions from state to state over time.

In previous studies, methods are developed to obtain the best fitting parameters of a Markov process based on minimizing the magnitude of the errors estimated for transitions between states of a system at different time points (Singer,1981; Kitamura,2003). Instead, in this study the maximum likelihood approach is utilised based on observations of all users' states at some or all discrete time epochs.

The structure of this paper is as follows: After this introduction, in Section 2 we offer a description of the stochastic process of discrete behaviours observed at discrete time epochs in a panel study. Then, how this process might be depicted by observations from a panel study is discussed. In Sections 3 and 4, by using the maximum likelihood estimation and the Newton iteration method, the parameters of the transition function can be estimated based on panel observations. In Section 5, an example of Kyoto University's bicycle share system is used to show the issue how accurately the parameters that characterize the process can be estimated and which method can be more effective. Finally, in Section 6 some initial conclusions and possible future research is discussed.

2. MODEL CONCEPTUALIZATION

Suppose the process to be measured in panel data can be represented by a set of discrete states. These discrete states can represent categories, frequency counts or measurements depend on the researches (Kitamura, 2003). Regardless of what it represents, it is assumed that for every person a transition from one state to another can take place between the adjacent time epochs and the process occupies one state between two successive transitions. It is also assumed that these transitions are probabilistic, and the transition function does not change over time (time-homogeneity).

Denote the discrete time periods by the letter *t* (with *t* = 0, 1, 2,...). Let $X(t)$ be the observed system state at time *t* with elements $x_{ij}(t) \subseteq (0,1)$ denoting whether a person *j* is in state *i* or not. *M* presents the set of all possible states and *m* is the number of states. t_n is the time period when the *n*th transition is made and $\mathrm{X}(t_n)$ presents the system state after the *n*th transition.

Each person must be in exactly one state at each time *t* so that $x_{ij}(t)$ takes binary values with

$$\sum_{i=1\ldots m} x_{ij}(t) = 1 \quad \forall j, t \tag{2-1}$$

We are interested in estimating the transition probabilities between subsequent time epochs. For this we define Q with vector $q_j(t)$ as the estimated probability mass function for person *j* which we refer to as the "state probability distribution". And define $q_{ij}(t)$ as the probability of person or person group *j* being in state *i* at time *t*.

$$\boldsymbol{q}_j(t) = \begin{bmatrix} q_{1j}(t) & q_{2j}(t) & \cdots & q_{mj}(t) \end{bmatrix} \tag{2-2}$$

Due to sampling limitations, including potential errors due to the discretization of the states, we presume that $q_{ij}(t)$ does not include the zero and one boundaries but instead takes following form:

$$0 < q_{ij}(\ ) < 1 \tag{2-3}$$

As we will discuss this assumption will be useful for our parameter estimation. Further, clearly each person must be in one of the *m* states at any time *t*.

$$\sum_{i=1\ldots m} q_{ij}(t) = 1 \quad \forall j, t \tag{2-4}$$

Our objective is to estimate a Markovian transition function ϕ that updates the estimated state probabilities. ϕ is a $m \times m$ matrix with elements p_{ki} denoting that the probability of state transitions from *k* to *i*.

$$\phi = \begin{bmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & \ddots & \vdots \\ p_{m1} & \cdots & p_{mm} \end{bmatrix} \tag{2-5}$$

The *n*-step transition function $\phi^{(n)}$ is also a $m \times m$ matrix with elements $p_{ki}^{(n)}$ denoting that the probability of transitions from the initial states *k* to state *i*.

$$\phi^{(n)} = \begin{bmatrix} p_{11}^n & \cdots & p_{1m}^n \\ \vdots & \ddots & \vdots \\ p_{m1}^n & \cdots & p_{mm}^n \end{bmatrix} \tag{2-6}$$

Due to the time-homogeneity assumption of the transition function ϕ, we can obtain following relationships:

$$\boldsymbol{q}_j(t) = \boldsymbol{q}_j(t-1)\phi \quad \forall t \tag{2-7}$$

$$(\phi)^n = \phi^{(n)} \tag{2-8}$$

$$q_j(t_n) = q_j(t_0)(\phi)^n \tag{2-9}$$

3. PARAMETERS ESTIMATION

In Section 2, we established the Markov Chain model to describe the stochastic process of discrete behavior with panel data. Next, we aim to estimate parameters in the model by Maximum likelihood estimation. The objective is to estimate the likelihood of correctly predicting the state of each person in the final time epoch t_n. We formulate this as follows:

$$L(\boldsymbol{q}|\boldsymbol{x}, \phi) = \prod_j \prod_i \left(q_{ij}(t_n)\right)^{x_{ij}(t_n)} \tag{3-1}$$

As we want to maximize likelihood we can consider the log likelihood function L:

$$\max L(\boldsymbol{q}|\boldsymbol{x}, \phi) = \max \ln L(\boldsymbol{q}|\boldsymbol{x}, \phi) = \sum_j \sum_i x_{ij}(t_n) \ln q_{ij}(t_n) \tag{3-2}$$

Since the parameters we want to estimate are the p_{ki} of the transition function ϕ, an expression of the log likelihood function L as functionality of these p_{ki} must be obtained.

First, let us define n_{ik} as the number of people who are at time t_0 in state *k* and in the target time n in state *i*,

$$n_{ik} = \sum_j x_{kj}(t_0) x_{ij}(t_n) \qquad (3\text{-}3)$$

Further, define $p_{ki}^{(n)}$ as the conditional probability that people who are at time t_0 in state *k* transfer into state *i* in the target time t_n with constraint

$$\sum_{i=1}^{m} p_{ki}^{(n)} = 1 \quad \forall k \qquad (3\text{-}4)$$

According to Eq. (2-9) we can obtain

$$q_{ij}(t_n) = \sum_k q_{kj}(t_0) p_{ki}^{(n)} \qquad (3\text{-}5)$$

Utilising Eq. 2-1 we first expand the log likelihood functions into

$$LL(\boldsymbol{q}|\boldsymbol{x},\phi) = \sum_j \sum_i \sum_k x_{kj}(t_0)\, x_{ij}(t_n) \ln q_{ij}(t_n) \qquad (3\text{-}6)$$

We then adjust the order of summation and introduce n_{ki} denoting the observed number of people who were initially in state *k* and are in time interval (*n*) in state *i*. The resulting function is shown in Eq.(3-7).

$$\begin{aligned} LL(\phi^{(n)}) &= \sum_i \sum_k \sum_j x_{kj}(t_0)\, x_{ij}(t_n) \ln q_{ij}(t_n) \\ &= \sum_i \sum_k n_{ki} \ln(p_{ki}^{(n)}) \end{aligned} \qquad (3\text{-}7)$$

We continue by utilizing (2-8) and introduce functions $\boldsymbol{G}[g_{ij}]$ to present the relationship between $p_{ki}^{(n)}$ and the one step conditional transition probabilities p_{ki} as in Eq.(3-8)

$$\begin{cases} p_{11}^{(n)} = g_{11}(p_{11}, p_{12}, \cdots, p_{mm}) \\ \vdots \\ p_{ki}^{(n)} = g_{ki}(p_{11}, p_{12}, \cdots, p_{mm}) \\ \vdots \\ p_{mm}^{(n)} = g_{mm}(p_{11}, p_{12}, \cdots, p_{mm}) \end{cases} \qquad (3\text{-}8)$$

Finally, the resulting function to be estimated is:

$$L(\phi) = \sum_i \sum_k n_{ki} \ln(g_{ki}(p_{11}, p_{12}, \cdots, p_{mm})) \qquad (3\text{-}9)$$

4. EQUATIONS SOLVING

We aim to find the maximum value and best estimation for ϕ. There appear to be m^2 parameters in $L(\phi)$, however, due to constraints (3-4) there are only *m* (*m*-1) free variables p_{ki} to be estimated.

To search the maximize value of $L(\phi)$, the maximum likelihood estimation of p_{ki} usually satisfy the equation:

$$\frac{\partial L(\phi)}{\partial p_{ki}} = 0 \quad \forall k, i \qquad (4\text{-}1)$$

Since we can not solve this High Order Nonlinear Equations directly, we make use of Newton's iterative method for finding the roots of a differentiable function *f*. However, this methods is only suitable for non-binding equations. Since here we have constraints (4-2) and (4-3)

$$\sum_{i=1\ldots m} p_{ki} = 1 \ \forall k \qquad (4\text{-}2)$$

$$0 < p_{ki} < 1 \qquad (4\text{-}3)$$

new unconstrained variables z_{ki} are introduced as in (4-4)

$$p_{ki} = \frac{e^{z_{ki}}}{1+\sum_{i=1}^{m} e^{z_{ki}}} \quad (z_{ki} \in R) \qquad (4\text{-}4)$$

In this way the problem of solving equations with constraints variables is transformed into solving equations with unconstrained variables. $\boldsymbol{Z}_\mu[z_{ij}^{(n)}]$ presents the estimated values after the $\boldsymbol{\mu}$th iteration.

$$\boldsymbol{Z}_\mu = \begin{pmatrix} z_{11}^{(\mu)} \\ \vdots \\ z_{m(m-1)}^{(\mu)} \end{pmatrix} \qquad (4\text{-}5)$$

Denote finally

$$\frac{\partial L(p)}{\partial p_{\mathrm{ki}}} = f_{ki}(z_{11}, \cdots z_{m(m-1)}) \qquad (4\text{-}6)$$

$$f(Z) = \begin{bmatrix} f_{11}(z_{11}, \cdots z_{m(m-1)}) \\ \vdots \\ f_{m(m-1)}(z_{11}, \cdots z_{m(m-1)}) \end{bmatrix} \qquad (4\text{-}7)$$

$$f'(Z) = \begin{bmatrix} \frac{\partial f_{11}}{\partial z_{11}} & \cdots & \frac{\partial f_1}{\partial z_{m(m-1)}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{m(m-1)}}{\partial z_{11}} & \cdots & \frac{\partial f_{m(m-1)}}{\partial z_{m(m-1)}} \end{bmatrix} \qquad (4\text{-}8)$$

We start the process with some arbitrary initial value $\boldsymbol{Z}_0$.Then we find the "better guess" $\boldsymbol{Z}_1$.The process is repeated as

$$\boldsymbol{Z}_{\mu+1} = \boldsymbol{Z}_\mu - \frac{f(\boldsymbol{Z}_\mu)}{f'(\boldsymbol{Z}_\mu)} \tag{4-9}$$

until a sufficiently accurate value is reached.

5. CASE STUDY

5.1 The COGOO Shared Bicycle Scheme

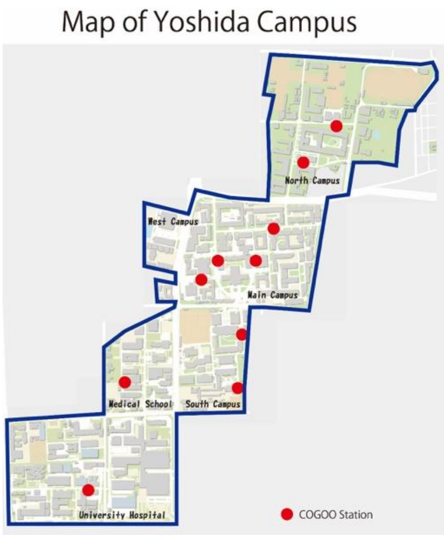

Fig.1 Kyoto University's Bicycle Share System

Bicycles are a laudable and sustainable form of transportation but they also create some problems on university campuses. Managing parked bicycles and specifically abandoned bicycles is a major challenge for Kyoto University and other Japanese universities. As part of efforts to encourage bicycle usage but at the same time reduce the numbers of abandoned bicycles, a bicycle sharing service called "COGOO" has been introduced. The COGOO service provides free bike rentals to Kyoto University students, faculty and staff. Users can register via mobile phone and pick up or drop off a bicycle from any of 10 COGOO parking lots. Since this system was introduced in March 2014, it has 1,600 registered members was used for more than 14,000 times accumulatively. We could obtain individual rental records for 13 months). (Unfortunately since April 2015 the service has been stopped due to bicycles being used inappropriately.)

COGOO is a good example to help us have a better understanding of how people will adapt to a new mode and predict the demand in the future.

In this case, we distinguish five different states for each person that are {never, few, sometimes, often, always} according to the use frequency (see Table 1). The time interval between subsequent time points is one month.

Table 1 State division

No	States	Frequency
1	never	0
2	few	1-3
3	sometimes	4-10
4	often	11-20
5	always	>20

5.2 DATA ANALYSIS

We focus on the adaptation of the new mode in the initial months from March to Jul 2014.

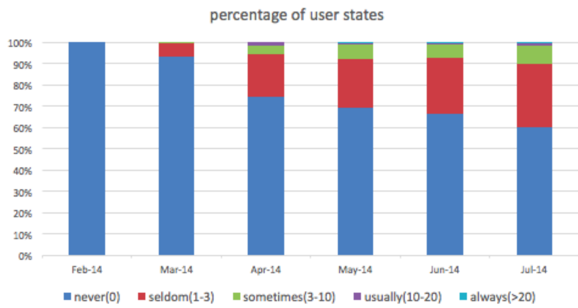

Fig.2 Distribution of user states

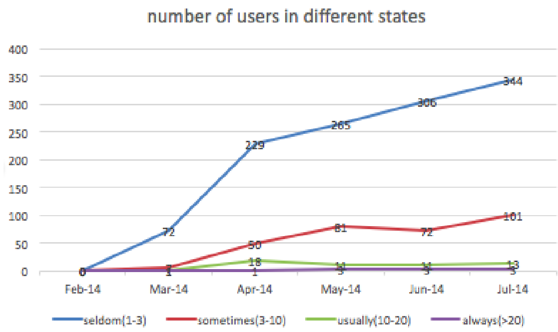

Fig.3 Number of users in different states

Percentage and number of users in all five states in different months are shown in Figs 2 and 3. It is clear to see that users of state 2 and 3 (low frequency user) at the beginning (the first month) increase very slowly then grow rapidly and later maintain steady growth. But high frequency users (states 4 and 5) after the first 1 or 2 month will keep quite stable though number of these users are few.

As shown in table 2, through the method we introduced before the estimated transition matrix can be calculated out. We can see that most of users in all the states maintain in the state which they were in the last period. Especially, all very high frequency users of state 5 keep their state. Besides, we find that the probability of jumps in more than one status level (e.g. state 1 to 3 or state 4 to 2) is very low. This suggests that people tend to change their habit gradually.

Table 2 Estimated Monthly Transition Matrix

states	1	2	3	4	5
1	0.8568	0.1313	0.0097	0.0022	0.0000
2	0.2152	0.6987	0.0861	0.0000	0.0000
3	0.0000	0.2553	0.7009	0.0438	0.0000
4	0.0000	0.0000	0.3660	0.5917	0.0422
5	0.0000	0.0000	0.0000	0.0000	1.0000

It is important to identify the difference of our estimated states and actual states to evaluate our model.

$\boldsymbol{E}_j(t)$ and $A_j(t)$ represent the estimated number of users and actual number of users at time t respectively. $\Delta_i(t)$ is the difference between actual value and estimated value of state i at time t. MAE (Mean Squared Error) MAPE (Mean Absolute Percent Error) and MSE(Mean Square Error) can be calculated by the following equations:

Table 3 Error analysis results

Time epoch	Actual value	Estimated value	RMSE	MAE	MAPE
March	(1081,72,7,1,0)	-	-	-	-
April	(863,229,50,18,1)	(942,193,21,3,0)	2.70	158	13.63%
May	(801,265,81,11,3)	(848,264,43,5,1)	1.83	95	8.20%
June	(769,306,72,11,3)	(784,307,63,7,1)	0.53	31	2.67%
July	(700,344,101,13,3)	(738,334,80,9,2)	1.28	73	6.30%

$$\Delta_i(t) = E_i(t) - A_i(t) \quad (5\text{-}1)$$

$$\text{MAE} = \sum_{i=1}^{m} |\Delta_i(t)| \quad (5\text{-}2)$$

$$\text{MAPE} = \left(\sum_{i=1}^{m} |\Delta_i(t)| \Big/ \sum_{i=1}^{m} A_i(t) \right) \times 100\% \quad (5\text{-}3)$$

$$\text{RMSE} = \sqrt{\sum_{i=1}^{m} \Delta_i^2(t) \Big/ \sum_{i=1}^{m} A_i(t)} \quad (5\text{-}4)$$

From the analysis results in Table 3, it is clear that estimation of the time epochs which are close to our final time epoch (June and July) have higher accuracy. The reason may be that our methodology only use the initial time epoch and final time epoch, the estimated values of intermediate time epochs would have larger errors.

Generally, when MAPE <10% it can be considered as high accuracy prediction. This encourages us to further investigate the use of Markov Chain models to estimate the adaption to new transport modes also for other applications.

6. DISCUSSION

In this paper we focused on formulating an approach based on stochastic state equations to describe the gradual change of behavior over time. We also discuss time-homogeneity issues and possibilities to calibrate the transition function. The model is applied to panel data from Kyoto University's bicycle share system. The results of this case study offer indications that stochastic, discrete behavioral processes can be estimated in a panel data study. There are still a number of issues that need to be addressed in further work.

Firstly, our methodology only use the initial time epoch and final time epoch in the process to estimate the parameters, but this ignores the “path" users take in their adaptation process. Therefore, as a next step, we should also take the intermediate time steps into consideration.

Besides, even through the transit function in different time epochs are similar, there still are differences. In fact, time does appear to affect the transition function as shown in the figures which violates our time-homogeneity assumptions. We therefore suggest to give different weights for the prediction values based on intermediate time epochs.

Connected to this, we found that in our case study the summer break influences the usage frequency. It would lead to significant errors in the prediction. So the recurring impacts should be take in consideration in order to make more accurate prediction. Establishing a suitable model to describe the process of adaption accurately will be one of the critical tasks.

REFERENCES

Cinlar, E., 1975. Introduction to Stochastic Processes. Prentice-Hall, Englewood Cliffs, NJ.

Kitamura,R., 2003.The effectiveness of panels in detecting changes in discrete travel behavior. Transportation Research Part B,37,191–206.

Li, Y.-T., Schmöcker, J.-D., 2015. Demand adaptation towards

new transport modes: the case of high-speed rail in Taiwan. Transportmetrica B: Transport Dynamics, 3(1), 27-43.
Li, Y.-T, Schmöcker, J.-D., 2016. Adaptation patterns to high speed rail usage in Taiwan and China. Transportation, pp 1-24.
Schmöcker, J.-D., Hatori, T. and Watling, D. 2014.Dynamic process model of mass effects on travel demand. Transportation, 41, 279-304
Singer, B., 1981. Estimation of nonstationary Markov chains from panel data. Sociological Methodology, 319–337.