Short Paper for hEART 2020

Cost Sensitive Estimation Methods of the RL-Activity-Scheduling Models under Disasters

Takuya Iizuka * Eiji Hato†

February 1, 2020

Keywords SP/RP data, activity scheduling model, evacuation, cost sensitive learning

1 Introduction

We aim to propose a method to estimate a reliable activity scheduling model under disasters using SP data, which is an indication of a respondent's behavioral options under virtual disaster situations in an area where a disaster has not yet occurred, and RP data obtained after an actual disaster in another area.

For evacuation planning, it is important to predict the activity scheduling under disaster situations. In order to obtain data for this purpose, SP surveys have been conducted in which a virtual disaster is assumed and subjects answer their actions. However, subjects cannot completely imagine the situation under disasters and often respond assuming that they will survive. Therefore, SP data contains biases (e.g. normalcy bias). We think that this problem can be solved by using RP data obtained by hearing etc. after the actual disaster for correction. However, another problem is the non-response bias that the data used for correction does not include the response of the victims.

The conventional estimation model maximizes the likelihood of prediction, but when using SP data including normalcy bias and RP data including non-response bias, people who cannot survive may be misclassified as survivable by activity scheduling model. In this case, when making an evacuation plan using models, it may cause an incorrect operation. Therefore, it is necessary to evaluate more sensitively the distribution of high-risk activities that do not survive (e.g. long-time activities in lowlands) than the distribution of safe activities included in SP/RP data. This is similar to the concept of cost sensitive learning by Elkan(2001)[3].

In this research, we propose introduce the concept of cost sensitive learning into estimation method using SP/RP data simultaneously for a more reliable activity scheduling model under disaster situations.

2 Methodology

We assume a Markov decision process for activity scheduling behavior under disaster situations, and model traveler's behavior as selecting actions sequentially on a time-space network.

As a sequential discrete choice model, Fosgerau et al.(2013)[4] proposes a Recursive Logit (RL) model. By expanding this to time-space networks, Oyama and Hato(2019)[6] removed the spatial cyclic, which is a problem in estimation, and made it possible to explicitly treat the behavior of staying at a certain point (Figure 1). Based on Oyama and Hato(2019), we describe an activity scheduling model in time-space networks.

Consider travelers in state s_t transitioning to the next state s_{t+1}. We assume that travelers chooses the next state to maximize the sum of the instantaneous utility $u(s_{t+1}|s_t)$ and the expected maximum utility

*Department of Civil Engineering, the University of Tokyo, Japan (e-mail: iiduka@bin.t.u-tokyo.ac.jp)
†Department of Civil Engineering, the University of Tokyo, Japan (e-mail: hato@bin.t.u-tokyo.ac.jp)

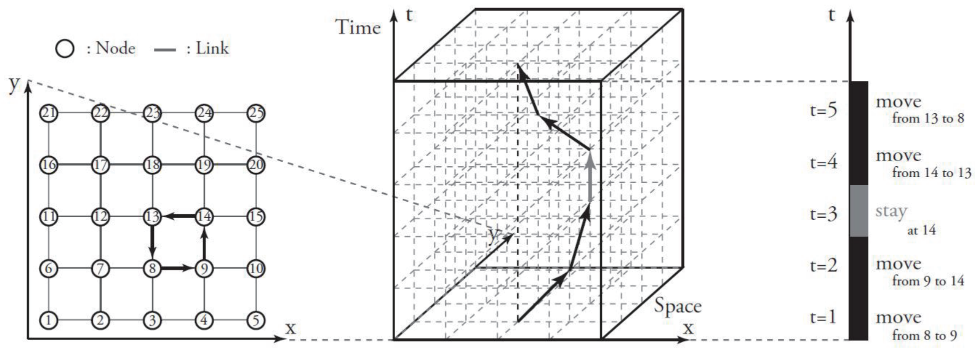

Figure 1: An example of activity scheduling path in time-space network

$V^{s_T}(s_{t+1})$ for the transition to the state s_{t+1}. Then, $V^{s_T}(s_t)$ can be formulated by Bellman equation[7] as follows:

$$V^{s_T}(s_t) = \max_{s_{t+1}} E\left[\sum_{t=\tau}^{T} \gamma^{t-\tau} u(s_{t+1}|s_t)\right] \quad (1)$$

$$= E\left[\max_{s_{t+1}\in C(s_t)} \{v(s_{t+1}|s_t) + \gamma V^{s_T}(s_{t+1}) + \mu\varepsilon(s_{t+1})\}\right] \quad (2)$$

where, ε follows the i.i.d. Gumbel distribution, μ is a scale parameter, and γ is a discount rate. $C(s_t)$ is a set of next states connected from state s_t, which means an alternative set in sequential route choice. Then, the choice probability of the states s_{t+1} from s_t is formulated as follows:

$$p(s_{t+1}|s_t) = \frac{\exp(\frac{1}{\mu}\{v(s_{t+1}|s_t) + \gamma V^{s_T}(s_{t+1})\})}{\sum_{s'_{t+1}\in C(s_t)} \exp(\frac{1}{\mu}\{v(s'_{t+1}|s_t) + \gamma V^{s_T}(s'_{t+1})\})} \quad (3)$$

This RL-activity-scheduling model in the time-space network can be estimated when the time series data of positions since the time of disaster occurrence is obtained (regardless of virtual or real). However, as described above, SP data which is an indication of a subject's behavior under a virtual disaster situation includes normalcy bias. In many cases, the subjects do not take into account activities other than evacuation, for example, delays in evacuation due to the collapse of furniture, the time to check the safety of family, and stop by the way to the shelter. Therefore, the time of activities other than evacuation is estimated short. We use RP data obtained after real disasters to correct it.

Ben-Akiva and Morikawa(1990)[2] proposed a methodology that uses the complementary strengths of both data by estimating a disaggregate behavior model by simultaneously using operable SP data and reliable RP data. Based on Ben-Akiva and Morikawa(1990), we describe an SP/RP simultaneous estimation framework. We consider RP model that generates RP data and SP model that generates SP data, and consider that the trade-off relationship between the main attributes is common between RP/SP model. Then, each instantaneous utility is formulated as follows:

RP model

$$\begin{aligned} u_{in}^{RP} &= \boldsymbol{\beta x_{in}^{RP}} + \varepsilon_{in}^{RP} \\ &= v_{in}^{RP} + \varepsilon_{in}^{RP} \quad (4) \\ \delta^{RP}(i,n) &= \begin{cases} 1 : \text{if individual n chose alternative i in the RP data} \\ 0 : \text{otherwise} \end{cases} \end{aligned}$$

SP model

$$\begin{aligned} u_{in}^{SP} &= \boldsymbol{\beta x_{in}^{SP}} + \boldsymbol{\alpha w_{in}^{SP}} + \varepsilon_{in}^{SP} \\ &= v_{in}^{SP} + \varepsilon_{in}^{SP} \\ \delta^{SP}(i,n) &= \begin{cases} 1 : \text{if individual n chose alternative i in the SP data} \\ 0 : \text{otherwise} \end{cases} \end{aligned} \tag{5}$$

where, $\boldsymbol{x}$ is x is an explanatory variable vector having a coefficient vector $\boldsymbol{\beta}$ common to the RP/SP model, and $\boldsymbol{w}$ is an explanatory variable vector having a coefficient vector $\boldsymbol{\alpha}$ specific to the SP model. In other words, $\boldsymbol{\alpha}$ represents the effect of the SP bias, which can remove the bias peculiar to the SP data. In order to share each coefficient in both models during estimation, a scale parameter is introduced. The scale parameter ν is expressed as the ratio of the variances of the error terms of each model as shown in Eq.(6). If the estimated value of this scale parameter is less than 1, the SP model has a more prominent error term than the RP model.

$$Var(\varepsilon_{in}^{RP}) = \nu^2 Var(\varepsilon_{in}^{SP}) \tag{6}$$

This method makes it possible to estimate parameters excluding effects of complementary bias between SP/RP data. However, when estimating the behavior model under disaster situations, the RP data obtained after disasters do not include the victim's response. This means RP data includes the non-response bias. It is conceivable that the proportion of victims who performed non-evacuation activities in dangerous areas before evacuation was higher than that of survivors. Therefore, both SP data and RP data have a bias that the duration of non-evacuation activities before evacuation is short. This is a non-complementary bias and is not corrected by traditional methodologies.

Then, we consider correcting this non-complementary bias by weighted estimation. Manski and Lerman (1977)[5] proposed to use a weighted estimator according to the bias of the sampled choice result to estimate the choice probability from data obtained by choice based sampling. However, in the activity scheduling model, the choice set changes according to the state because the scheduling path is described as a sequential choice sequence. Therefore, it is difficult to use a conventional sampled choice result based weights. Instead, we propose to formulate the risk of the state under disaster situations and use it as a weight. The weight is formulated as follows:

$$w(s_{t+1}) = 1 + \alpha \frac{d_{s_{t+1}}^{s_T}}{\tau_t},\ t \neq T \tag{7}$$

where, $d_{s_{t+1}}^{s_T}$ is spacial distance from state s_{t+1} to s_T and τ_t is the remaining time until evacuation should be completed (e.g. tunami arrival time). This weight becomes larger as the remaining time is shorter and the spatial distance to the evacuation completion state is longer. In other words, the greater the risk of a state, the greater the result of choicing that state in estimation by this weight (Fig 2). By Introducing this weights, it is possible to estimate more sensitively assesing the distribution of high-risk activities than the distribution of many safe actions included in SP/RP data.

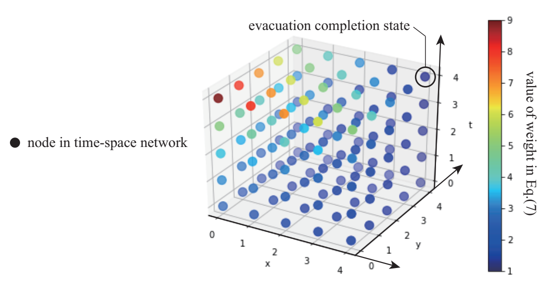

Figure 2: An example of risk weights of states in time-space network

Table 1: True parameter values used in simulation

	Explanatory variables			
	link cost	origin node stay dummy	family node stay dummy	shelter node stay dummy
RP model	-0.5	1.0	1.0	3.0
SP model	-0.5	0.5	0.5	3.0

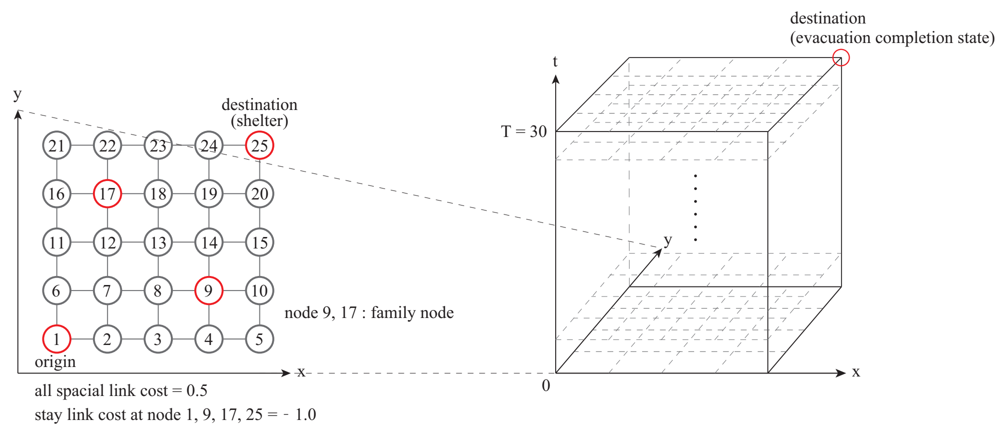

Figure 3: The network used for simulation

From the above, the log likelihood functions of RP model (Eq.(8)), SP model (Eq.(9)) and RP+SP model (Eq.(10)) are formulated as follows:

$$L^{RP}(\boldsymbol{\beta}) = \sum_{n=1}^{N^{RP}} \sum_{t=0}^{T^{RP}} \delta^{RP}(s_{t+1}, n) w(s_{t+1}) \log(p(s_{t+1}|s_t)) \tag{8}$$

$$L^{SP}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \nu) = \sum_{n=1}^{N^{SP}} \sum_{t=0}^{T^{SP}} \delta^{SP}(s_{t+1}, n) w(s_{t+1}) \log(p(s_{t+1}|s_t)) \tag{9}$$

$$L^{RP+SP}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \nu) = L^{RP}(\boldsymbol{\beta}) + L^{SP}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \nu) \tag{10}$$

The log likelihood function in Eq.(10) is maximized using the NXFP algorithm[1], the NPL algorithm[8], etc.

3 Numerical Analysis

Numerical experiments were performed to confirm the effectiveness of the proposed method. In a simple network, we compare estimation results of both conventional simultaneous estimation method and the proposed method. In the following, we set $\mu = 1$, $\gamma = 0.7$ in Eq.(2), Eq.(3), and $\alpha = 1$ in Eq.(7).

Firstly, we simulated 1000 and 200 scheduling paths for each of the RP and the SP model assuming true parameter values (Table 1) by performing Markov simulation of 30 time steps using Eq.(3) sequentially. The difference between the parameter values of origin node stay dummy and family node stay dummy in SP/RP model indicates the bias included in SP data. Figure 3 shows the network used for simulation. Of the samples generated from the RP model, 300 samples arriving at the evacuation node within 25 steps are randomly sampled and used for estimation. This implies a non-response bias for victims in RP data. All 300 samples generated from SP data are used for estimation. Our purpose is to estimate the true parameter value of RP model from SP data of 200 samples and RP data of 300 samples.

Table 2 shows the estimation results for each method. RMSE is calculated between each estimated parameter and the true parameter values of RP model. Parameters estimated from SP data are close to the

Table 2: Estimation results

explanation variables	SP data parameter	SP data t-value	RP data parameter	RP data t-value	SP+RP data parameter	SP+RP data t-value	weighted SP+RP data parameter	weighted SP+RP data t-value
link cost	-0.384	-27.18	-0.193	-13.17	-0.293	-28.61	-0.347	-43.17
origin node stay dummy	0.544	35.17	0.895	63.69	0.853	65.41	0.876	93.08
family node stay dummy	0.527	26.86	0.985	60.38	0.939	60.80	0.962	76.40
shelter node stay dummy	2.932	66.38	3.232	103.53	3.136	113.19	3.049	117.45
origin node stay dummy(SP)	-	-	-	-	-0.262	-13.73	-0.262	-18.43
family node stay dummy(SP)	-	-	-	-	-0.358	-14.56	-0.386	-18.91
scale parameter ν	-	-	-	-	0.989	129.53	0.977	135.96
Num of Samples		200		300		500		500
initial likelihood		-210185.1		-239532.8		-820016.5		-1266126.0
final likelihood		-199700.6		-212290.4		-412245.5		-673614.1
likelihood ratio		0.050		0.114		0.497		0.468
RMSE	0.335		0.200		0.147		0.103	

true parameter values of SP model. Parameters estimated from RP data are higher in the shelter node stay dummy compared to the true values, and the others are estimated lower. It can be seen that they includes a bias caused by sampling 300 samples evacuated within 25 steps. RMSE between SP/RP common parameters estimated by simultaneous estimation and the true parameter values of RP model is smaller than that of only SP data or RP data. This indicates that SP data bias is corrected by the two parameters specific to SP model. RMSE calculated from SP/RP common parameters estimated by the proposed method is more smaller, then the proposed method is effective for improving the reliability of activity scheduling model estimation.

4 Conclusions

In this research, we proposed a method that combines the SP/RP simultaneous estimation and the method that uses the risk of the state on time-space networks as a weight in activity scheduling model estimation. Numerical analysis shows that the proposed method is effective in increasing the reliability of the estimation results.

Future work can be dedicated to applying this model to real data and confirm reproducibility. However, the behavior data of victims in disasters cannot be obtained, so it is difficult to evaluate the reliability using real data. Heuristic model evaluation techniques will be needed.In addition, future work can focus on formulating the risk of the state in the time-space network more appropriately. For example, Monte Carlo tree search can be introduced. It is conceivable to simulate whether the evacuation is successful or unsuccessful by playout and update the risk of the state by backpropagation. We believe that our proposed method can be applied to the correction of the estimation bias of choice-based sampling with or without participation in sharing, and its applicability is wide.

References

[1] B Y Victor Aguirregabiria and Pedro Mira. Swapping the Nested Fixed Point Algorithm : A Class of Estimators for Discrete Markov Decision Models Author (s): Victor Aguirregabiria and Pedro Mira Published by : The Econometric Society Stable URL : http://www.jstor.org/stable/3082006 Accessed : 04-. 70(4):1519–1543, 2002.

[2] Moshe Ben-Akiva and Takayuki Morikawa. Estimation of switching models from revealed preferences and stated intentions. *Transportation Research Part A: General*, 24(6):485–495, 1990.

[3] Charles Elkan. The foundations of cost-sensitive learning. *IJCAI International Joint Conference on Artificial Intelligence*, pages 973–978, 2001.

[4] Mogens Fosgerau, Emma Frejinger, and Anders Karlstrom. A link based network route choice model with unrestricted choice set. *Transportation Research Part B: Methodological*, 56:70–80, 2013.

[5] C. F. Manski and Steven R. Lerman. The Estimation of Choice Probabilities from Choice Based Samples Author (s): Charles F . Manski and Steven R . Lerman. *Econometrica*, 45(8):1977–1988, 1977.

[6] Yuki Oyama and Eiji Hato. Prism-based path set restriction for solving Markovian traffic assignment problem. *Transportation Research Part B: Methodological*, 122:528–546, 2019.

[7] Bellman Richard. Dynamic programming. *Princeton University Press*, 89:92, 1957.

[8] John Rust. Optimal Replacement of GMC Bus Engines : An Empirical Model of Harold Zurcher Author (s): John Rust Published by : The Econometric Society Stable URL : http://www.jstor.org/stable/1911259. *Society*, 55(5):999–1033, 1987.