Estimation of Behavioral Asymmetry in Multi-Agent Interaction using Satellite Imagery

D. Ogawa[a], E. Hato[b,*]

[a] The University of Tokyo, Tokyo, Japan
dogawa@bin.t.u-tokyo.ac.jp
[b] The University of Tokyo, Tokyo, Japan
hato@bin.t.u-tokyo.ac.jp

Extended abstract submitted for presentation at the Conference in Emerging Technologies in Transportation Systems (TRC-30)
September 02-03, 2024, Crete, Greece

April 29, 2024

Keywords: Route choice, Social interaction, Equilibrium, Adversarial learning, Satellite image

1 INTRODUCTION

Satellite imagery has become increasingly crucial in analyzing the dynamic changes in urban spaces and their relationship with travel behavior. Particularly important is the analysis of asymmetric interactions among different modes, such as pedestrians and vehicles, within confined urban spaces, and accurately predicting mutual externalities. In this context, it is essential to appropriately address the endogeneity of interactions between modes within the model framework and learn parameters using satellite imagery data.

There are typically two methods to estimate the endogenous effect, structure estimation and mathematical programs with equilibrium constraints (MPEC)(Ferris *et al.*, 2005). However, interaction estimation by structure estimation is sometimes difficult because of the lack of distinguishability and heavy computation for the iterative calculation to obtain the equilibrium numerically(Aguirregabiria & Mira, 2007).

On the other hand, the MPEC-based methods are flexible and compatible with the machine learning framework. For example, multi-agent adversarial inverse reinforcement learning (AIRL)(Yu *et al.*, 2019) can reveal the user preference under a user equilibrium state, which can interpreted as an MPEC problem. However, there are still limitations in that the interaction value itself cannot be distinguished from the user reward function and the characteristics of the obtained equilibrium state are not revealed.

The main contribution of this study is as follows:

- The proposal of the stable interaction estimation method by introducing the Lipschitz normalization to the AIRL model.
- The analysis of the uneven multi-modal equilibrium between pedestrians and vehicles with the existence of the interaction effect using the multi-modal learning method.

2 METHODOLOGY

2.1 Notation

We consider the situation where the agents are in the state s and move to state s' by taking actions $\mathbf{a} = (a_1, \cdots, a_N)^T$. Context vector c is introduced to capture the trip destination(Zhao

& Liang, 2023). The context vector is the shortest path to the destination in this study.

2.2 Multi-agent adversarial inverse reinforcement learning with Lipschitz normalization

One of the two models to be trained in the AIRL-based route choice model is the generator model $\pi_{G_i}(a_i|s;c_i)$ called actor, and it calculates the probability of the agent i's action a_i given the state s. The other model is the discriminator model called critic, and it computes the probability that the pair of action and state belongs to the real data. The critic function is written as Eq. 1.

$$D_{\theta_i,\phi_i}(s, a_i, s'; a_{-i}, c) = \frac{\exp\{f_{\theta_i,\phi_i}(s, a_i, s'; a_{-i}, c_i)\}}{\exp\{f_{\theta,\phi}(s, a_i, s'; a_{-i}, c_i)\} + \pi_G(a_i|s; c_i)} \tag{1}$$

where θ_i, ϕ_i are the parameters of the critic function for agent i. The function $f_{\theta_i,\phi_i}(s, a_i, s'; a_{-i}, c_i)$ is the estimator of the advantage function and structured by g_{θ_i} and h_{ϕ_i}, which are the estimator of the utility function and the value function. The discount factor γ is introduced.

To train these two models, they share a part of the objective function, and the actor function tries to increase D_{θ_i,ϕ_i} and the critic function tries to decrease it. Under the estimation of multi-transportation utility, the user-equilibrium, the logistic stochastic best response equilibrium (LSBRE) (Yu *et al.*, 2019) as shown in Eq. 2, is satisfied.

$$\pi(a_i|a_{-i}, s; c) = \frac{\exp(\lambda Q_i^{\pi}(a_i, a_{-i}, s; c))}{\sum_{a_i' \in A_i} \exp(\lambda Q_i^{\pi}(a_i', a_{-i}, s; c))} \approx \mathrm{Softmax}(\lambda Q_i^{\pi_i}(\pi_{-i}, s)) \tag{2}$$

where a_i, a_{-i} are the action of the agent i and the action of the other agents, respectively. $Q_i^{\pi}(a_i, a_{-i}, s; c)$ is the Q function and λ is the parameter that shows the greediness of the agents.

Here, we analyze the stability of the estimated equilibrium based on feedback from other agents when the slight perturbation affects the equilibrium state. The concept is shown in Figure 1. The time average of the derivative of the probability that the transportation i takes the action a_i for the transportation $-i$'s action function is written as Eq. 3.

$$E\left[\frac{\partial \pi_{i,a_i}}{\partial \pi_{-i}}\right] = \pi_{i,a_i} \sum_{a_i'} (\delta_{a_i,a_i'} - \pi_{i,a_i'}) E_{(s,a)} \left[\sum_t \gamma^t \frac{\partial r_i(s_t, a_{i,t}, \pi_{-i}; c)}{\partial \pi_{-i}}\right] \tag{3}$$

where r_i is the reward function of the agent i. The stability of the equilibrium can be analyzed by the maximum eigenvalue of the Jacobian of the action functions. The equilibrium is stable if the absolute value of the maximum eigenvalue of the Jacobian is less than 1.

$$\|\delta\pi_{-i}\|_2 \geq \|\frac{\partial \pi_{-i}}{\partial \pi_i} \frac{\partial \pi_i}{\partial \pi_{-i}} \delta\pi_{-i}\|_2 \tag{4}$$

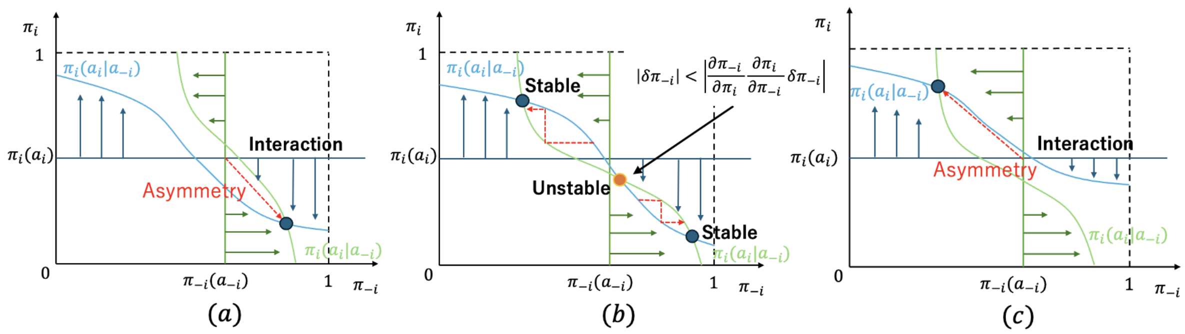

Figure 1 – *The equilibrium types based on asymmetry of choice probability and the stability analysis with the multi-agent interaction: (a) agent i is excluded, (b) two stable equilibriums and one unstable equilibrium, (c) agent i occupies.*

On the other hand, to reveal the effect of interactions quantitatively, the estimator of the advantage function for the agent i is rewritten as Eq. 5.

$$f_{\theta_i,\phi_i,\psi_i}(s,a,s',\pi;c) = g_{\theta_i}(s,a_i : c) + l_{\psi_i}(s,a_i,\pi_{-i}(a_{-i}|s);c) + \gamma h_{\phi_i}(s';c) - h_{\phi_i}(s;c) \tag{5}$$

where g_{θ_i} captures the direct utility of action under the state, and l_{ψ_i} captures the interaction term among multiple transportation modes. π_{-i} means the marginal distribution of other agents for the given destination distribution.

Now, we consider the estimation of the stable equilibrium using the multi-agent AIRL with an interaction term. The Jacobian of the reward function can be written only with l_ψ, and improving the smoothness of l_ψ can enhance the stability of the equilibrium. The concept of Lipschitz norm can be used as a measure of the smoothness of the neural network function, and the Lipschitz norm of neural networks can be efficiently regularized by the spectral normalization method (Miyato *et al.*, 2018).

2.3 Multi-modal learning for equilibrium analysis

From Figure 1, the effect of the interaction term tends to lead to the occupation or sparseness of the specific agent for each choice set, which means that each choice set can be classified by the occupation of the agents.

One candidate variable to explain the type of the choice sets is the spatial feature, such as the satellite images. The domain of the satellite image and that of route choice are completely different, but the concept of multimodal learning is expected to be useful in extracting the spatial feature from the satellite image. In this research, we trained the image feature encoder using the loss function of the actor function to compare the spatial feature of the mesh with different occupation tendencies. The overview of the satellite image integration process is shown in Figure 2.

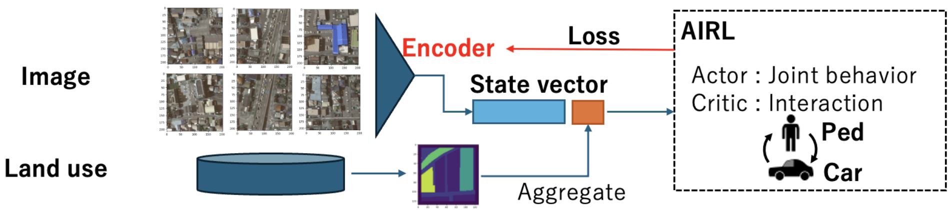

Figure 2 – *Overview of the multi-modal learning process combining satellite images and route choice data.*

3 RESULTS

3.1 Experiment settings

The proposed method is applied in Matsuyama City, Ehime, Japan. The route choice data of pedestrians and cars are obtained from GPS data from the survey in 2007. The mesh network is defined with 100m square grids, and each GPS point is allocated to the mesh network. Training (test) data contains 1675 (1689) route choice behaviors for pedestrians and 1759 (1334) for vehicles, respectively. The 2-layer convolutional neural networks with kernel size is 3×3 are used for the actor functions and critic functions.

The aviation image data with 0.5m resolution are used instead of satellite image data and obtained from the Geospatial Information Authority of Japan. The image data is compressed into 10-dimensional vectors using ResNet37(Bello, n.d.). The extracted feature is concatenated with the original mesh variables which are obtained from the basic survey of urban planning in Matsuyama City in 2018.

3.2 Evaluation metrics

The log-likelihood functions with and without the Lipschitz normalization and image feature are shown in Table 1. The initial log-likelihood is -2811.40 for pedestrians and -2025.38 for vehicles. In addition, some meshes with strong negative interaction effects are sampled and shown in Figure 3. The SHAP values for each model are also shown.

Table 1 – *Log-likelihood of AIRL-based route choice models*

Model	Pedestrian	Vehicle
No image, No normalization	-2707.01	-1776.24
No image, Normalization	-2703.66	-1766.62
Image, No normalization	-2700.47	-1773.86
Image, Normalization	-2694.94	-1760.09

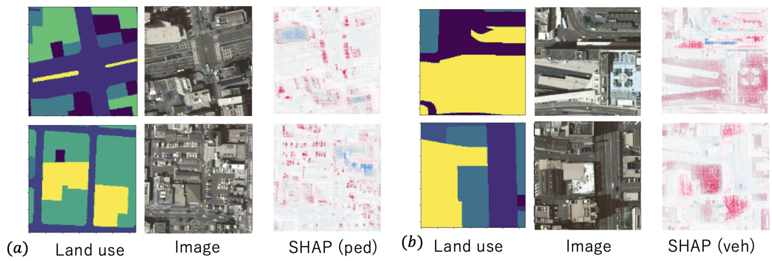

Figure 3 – *Image, land use, and SHAP values for the encoder for the large interaction meshes: (a) negative interaction for pedestrians, (b)negative interaction for vehicles.*

4 DISCUSSION

From Table 1, the Lipschitz normalization on the interaction term and the image feature improves the log-likelihood of the model. The revealed interaction effect tends to be different between pedestrians and vehicles. From Figure 3, for example, the negative effect is observed at the intersection and parking area for pedestrians and the area with high buildings such as the central area or the area near the station for vehicles.

In future work, the incorporation of more detailed spatial features such as street-view images is expected to be useful for the analysis of the multi-modal equilibrium.

References

Aguirregabiria, Victor, & Mira, Pedro. 2007. Sequential Estimation of Dynamic Discrete Games. *Econometrica*, **75**(1), 1–53.

Bello, Irwan. *LambdaNetworks: Modeling Long-Range Interactions Without Attention.*

Ferris, Michael C., Dirkse, Steven P., & Meeraus, Alexander. 2005. Mathematical Programs with Equilibrium Constraints: Automatic Reformulation and Solution via Constrained Optimization. *Pages 67–94 of:* Kehoe, Timothy J., Srinivasan, T. N., & Whalley, John (eds), *Frontiers in Applied General Equilibrium Modeling*, 1 edn. Cambridge University Press.

Miyato, Takeru, Kataoka, Toshiki, Koyama, Masanori, & Yoshida, Yuichi. 2018 (Feb.). *Spectral Normalization for Generative Adversarial Networks.* arXiv:1802.05957 [cs, stat].

Yu, Lantao, Song, Jiaming, & Ermon, Stefano. 2019. Multi-agent adversarial inverse reinforcement learning. *Pages 7194–7201 of: International Conference on Machine Learning.* PMLR.

Zhao, Zhan, & Liang, Yuebing. 2023. A deep inverse reinforcement learning approach to route choice modeling with context-dependent rewards. *Transportation Research Part C: Emerging Technologies*, **149**, 104079.