東京理科大学

Transport Demand Forecasting for Waterborne Using LLMs

LLMを用いた舟運導入に伴う交通需要予測

2026/09/09

Tokyo University of Science
Behavioral Computational Science Laboratory

Yuina Fujimori, Taketo Shimizu, Yura Yozawa, Haruhiko Murakami, Hayato Mizuno, Koya Sato, Yutaro Arita

東京理科大学

Background
背景

臨海部における既存公共交通の飽和

Saturation of existing public transportation in coastal area

- Residential development and population growth have caused chronic public transportation congestion in the Tokyo waterfront area.
- Adoption of water transportation as a new mode of transportation to enhance urban mobility.

Pilot Project for the Nihombashi–Toyosu Waterway

- A shuttle boat demonstration experiment was launched between Toyosu and Nihombashi in FY2021.(Currently discontinued.)
- Quantitative evaluation of demand and potential was needed for full-scale operation, due to challenges such as fares and operating schedules.

Quantitatively predict and evaluate the modal split resulting from the introduction of waterborne transport in coastal areas

臨海部における舟運の導入がもたらす交通分担率を定量的に予測・評価する

https://www.spt.metro.tokyo.lg.jp/tosei/hodohappyo/press/2023/10/06/21.html

Issues
問題

新たな交通手段の選択確率の推定

Estimating the probability of choosing an unknown mode of transportation

- SP調査をもとに新規交通手段の需要予測を行っている
- LLMが発達してきた現在，LLMを用いて交通需要予測ができないか
- PPデータから新規交通手段がある時のPPデータをLLMによって生成

Current Situation

Stated Preference (SP) Survey

Proposal

PP Data generated by LLMs

Advantages of Using LLMs

- Create a scenario in which a new transportation mode already exist.
- Eliminate bias in the design of explanations.

Data・Area
使用データ・対象地域

豊洲PPデータ

Toyosu Probe Person Data (PP Data)

- Mainly used trips involving **Toyosu and Nihombashi.**
- Also used to verify consistency by comparing with virtual PP data.
- PT surveys rely on respondents' memory, which can reduce data accuracy.
- PP data are limited in volume due to the high cost of data collection.

項目	観測値
調査期間	2019年7月9日 – 2021年11月30日
抽出トリップ	4515件
推定に使用した観測数	3686件

交通手段	件数
Train	2242件
Bus	260件
Car	420件
Walk	422件
Taxi	176件
Bike	166件

東京理科大学

Analysis flow
分析フロー

Bacic Analysis 基礎分析

Toyosu Probe Person Survey (2019–2021)
Extraction of trips made between Toyosu and Nihonbashi

MNL

explanatory variable : Move(Ride time), AE(Access and egress time), Fare(Cost), Trans(Transfer)

▼

Generation of PP Data Using LLMs

Generated synthetic PP data for the implementation of "Fune" using an LLM

▼

Model Comparison

Model of observed PP vs Model of generated PP

Generate and compare PP data across multiple scenarios with varying fares.

Model Specification

Trips directly connecting Toyosu and Nihonbashi (n=164)―3 alternative MNL

3 alternatives

- road (Car, Taxi)
- public (Train, Bus)
- active (Bike, Walk)

Choice probability and estimation

$$Pn(i) = \delta_{ni}\, e^{Vni} / \sum_j \delta_{nj}\, e^{Vnj}$$

- Standard MNL with the availability dummy δ carried into the denominator
- Maximum likelihood (Newton method)
- Standard errors: cluster-robust (sandwich) at the individual level, with a finite-sample correction

Utility function

$$V_j = ASCj + \beta MOVE\cdot(mj/10) + \beta AE\cdot(aj/10) + \beta FARE\cdot(cj/1000) + \beta TRANS\cdot trj$$

Variable definitions

m_j	Travel time [min] (in-vehicle for public)	all
a_j	Access + egress time [min]	public only
c_j	Cost [yen] (fuel, fare, taxi fare)	road, public
tr_j	Number of transfers	public only

東京理科大学

Estimation Results and Fit

Parameter	Estimate	Std. error	z	p	
ASC_public	−4.800	1.575	−3.05	0.002	***
ASC_active	−1.863	2.601	−0.72	0.474	
β_MOVE	−3.281	0.924	−3.55	<0.001	***
β_AE	−1.137	0.823	−1.38	0.167	
β_FARE	−10.348	2.353	−4.40	<0.001	***
β_TRANS	+0.391	1.779	0.22	0.826	

*** significant at the 1% level

$\rho^2 = 0.422$

Up from 0.294 for M0 (ASCs only)

≈ 1,900 yen/h

Value of time = 60·(β_MOVE/10)/(β_FARE/1000)

AIC 220.24

M0: 258.36. Log-likelihood improves by 23 with 4 extra df

Key findings

- β_MOVE: 10 more minutes of travel lowers utility by 3.28 — time sensitivity is clearly negative and significant
- β_FARE: 1,000 yen more cost lowers utility by 10.3. The cost coefficient, which flipped positive in other versions, is correctly identified here, making the VOT computable
- β_AE and β_TRANS fall short of significance with only 164 observations and are read as effectively zero; interpretation centres on β_MOVE, β_FARE and the VOT
- The standard error on β_FARE (±2.35) is large, so the 1,900 yen/h point estimate carries a wide band — report a profile-likelihood interval alongside it

東京理科大学

Generation of PP data

既存のPPデータからペルソナ文を生成し,エージェントを作成. LLMによりボートを加えたチョイスセットから移動手段を生成.

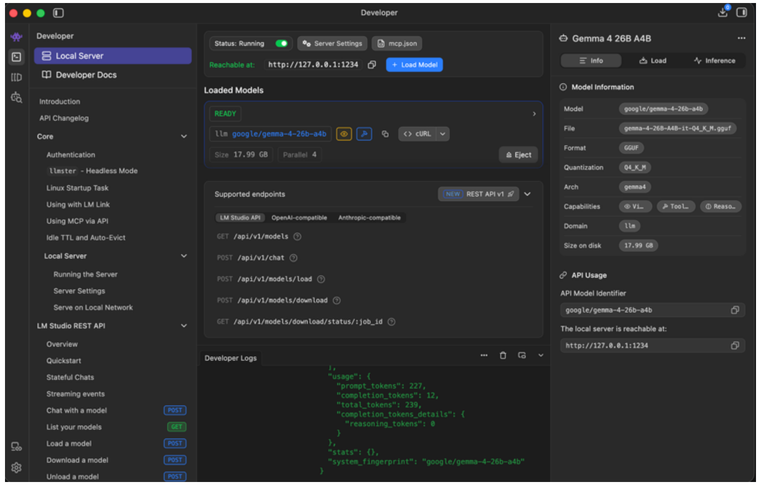
Generate persona profiles from existing PP data to create agents.

- Generated responses using a local LLM to handle detailed persona information.

  詳細なペルソナ情報を扱うため, ローカルLLMを用いて回答を生成した.

- Model used: **google/gemma-4-26b-a4b**(local)

  使用したモデルは, google/gemma-4-26b-a4b.



LM Studio

東京理科大学

Generation of PP data

LLMによって船を使うと回答されたトリップに対して船を使った感想を作成.

Boat-use impressions generated for trips where LLM selected the boat

▼

感想をペルソナに付与した状態で更にLLMによる回答を生成.(世代を複数回シミュレーション)

Further response generation via LLM, with impressions attached to persona (multi-round generation simulation)

- Example persona actually provided to the LLM

実際にLLMに付与したペルソナの例.

[You]Male, 60s or older, self-employed. Household income 15M JPY or more. Nearest station to home: Toyosu; to workplace: Ginza-itchome / Kyobashi.
Lives with spouse. Owns a car. Does not own a bicycle.
[Today's trip]
A weekday midday business trip from the Toyosu area to the Nihonbashi area.
Choose based on the conditions shown, not on your usual mode.
Select one of the following.
A Ferry: 35 min, 500 JPY, 0 transfers
B Rail: 23 min, 360 JPY, 1 transfer, incl. 12 min walking
Answer in JSON only.

{"choice": "A"}

- Example of impressions acquired and carried over by agents who chose the boat

船を選択したAgentが取得・引き継ぐ感想の例

- **Number of times boarded:** 1 time
- **Duration:** Felt longer than announced
- **Ride comfort:** Not bad
- **Wait time:** Around 22 minutes
- **Fare:** Felt expensive

東京理科大学

Generation of PP data

生成結果(5世代分)

Generation results (5 generations)

世代(generation)	船(boat)	鉄道(train)	バス(bus)	自動車(car)
g1	**42**	37	51	31
g2	18	56	51	33
g3	15	58	51	34
g4	13	60	51	34
g5	**12**	60	51	35

▶ Scenario obtained after 5-generation simulation
: **7.4% mode shift to boat**

5世代の計算後,7.4%を船が代替するシナリオの生成を得た.

Comparison: Observed vs Generated PP

Same 164 trips (Toyosu–Nihonbashi), 12 switched to boat by the LLM in the generated PP

FINDING 1

Time & cost sensitivity is preserved

	Observed	Generated
β_MOVE	−3.34 ***	−3.88 ***
β_FARE	−10.47 ***	−8.47 ***
VOT	≈1,900 /h	≈2,750 /h

Estimates agree within one standard error — the LLM's boat choices are consistent with the MNL utility structure.

FINDING 2

ASC_Boat becomes estimable

+6.45

z = 2.53 (significant)

Unidentifiable in the observed PP (0 boat choices). The generated boat choices reveal boat-specific preference directly.

✓

Takeaway

The LLM reproduces observed time/cost trade-offs, and its generated choices let us recover the boat ASC that the observed data alone cannot identify — a quantity otherwise only assumable.

Tokyo University of Science