Herbert Wertheim
College of Engineering
UNIVERSITY of FLORIDA

Enhancing Travel Behavior Analysis with Generative AI: Integrating Survey and Location Data

May 30, 2025
Big Data, AI and Transportation Planning Applications at the University of Washington

Xilei Zhao
Assistant Professor of Civil and Coastal Engineering

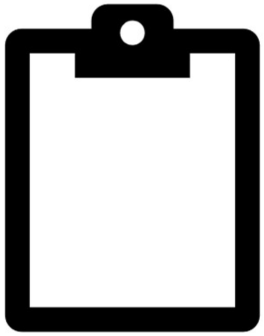
POWERING THE NEW ENGINEER TO TRANSFORM THE FUTURE

Introduction

- Both survey and location data have been widely adopted for travel behavior analysis and used to inform transportation planning.

- For example, conducted by FHWA, the NHTS is the authoritative source on the travel behavior of the American public, and recently, FHWA launched the NextGen NHTS OD Data Products constructed from raw mobile-device-location and GPS-device data obtained from multiple data providers (FHWA, 2023).

- However, survey and location data have their own limitations that hinder their full potential.



The Motivating Challenge: Survey v.s. Location Data

	Strengths	Limitations
Survey Data	• Detailed data on individuals' choices, trip purpose, sociodemographics, attitudes, perceptions, etc. • Ideal for individual-level causal analysis • Better integration with behavioral theories	• Small sample sizes • Limited spatiotemporal details on trips • High data collection cost • Sampling bias (esp. reporting bias) • Reliant on human memory
Location Data	• Large sample sizes • High spatiotemporal resolution • Relatively low cost to collect data • Relatively low sampling bias • Not reliant on human memory	• Lack of individual-level data on sociodemographics, attitudes, perceptions, etc. • Underrepresentation of people who have no/limited access to mobile devices • Uncertainty in results (e.g., activities are inferred by modelers) • Inability to conduct individual-level causal analysis (only associations)

Survey and location data seem to be highly complementary!

The Overarching Goal of Our Team

Developing new methodologies to integrate survey and location data for enhanced travel behavior analysis.

The Motivating Challenge: Survey v.s. Location Data

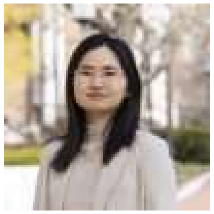
	Strengths	Limitations
Survey Data	• Detailed data on individuals' choices, trip purpose, sociodemographics, attitudes, perceptions, etc. • Ideal for individual-level causal analysis • Better integration with behavioral theories	• Small sample sizes • Limited spatiotemporal details on trips • High data collection cost • Sampling bias (esp. reporting bias) • Reliant on human memory
Location Data	• Large sample sizes • High spatiotemporal resolution • Relatively low cost to collect data • Relatively low sampling bias • Not reliant on human memory	• Lack of individual-level data on sociodemographics, attitudes, perceptions, etc. • Underrepresentation of people who have no/limited access to mobile devices • Uncertainty in results (e.g., activities are inferred by modelers) • Inability to conduct individual-level causal analysis (only associations)

Research Question

Can we accurately infer mobile device users' sociodemographics to unlock the full potential of location data?

Sun, Y., Xu, S., Wang, C., & Zhao, X. (2025). *Where you go is who you are: Behavioral theory-guided LLMs for inverse reinforcement learning*. arXiv. https://doi.org/10.48550/arXiv.2505.17249

Yuran Sun, PhD Candidate
University of Florida

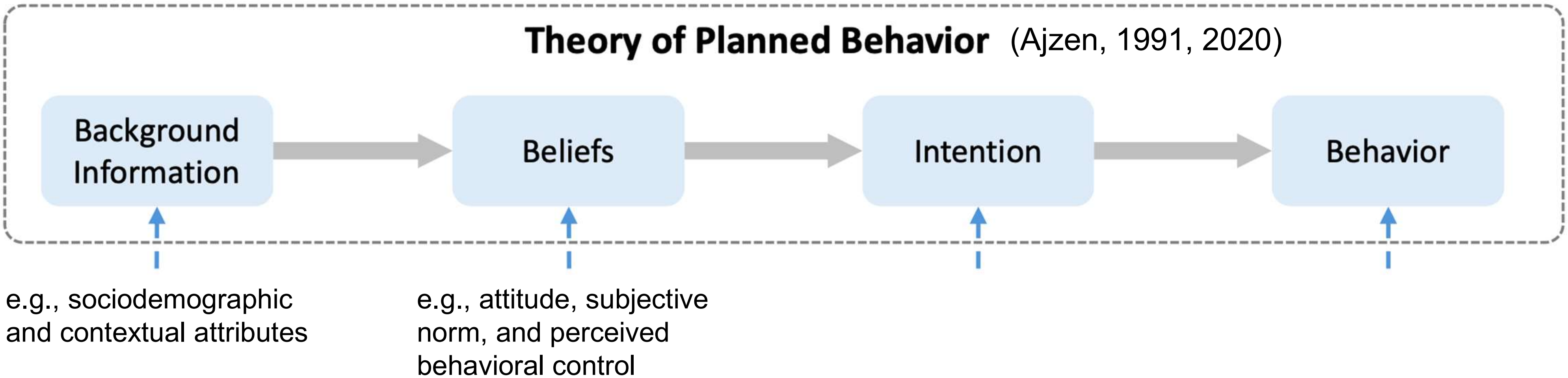


Susu Xu, Assistant Professor
Johns Hopkins University

Chenguang Wang, PhD Candidate
Stony Brook University

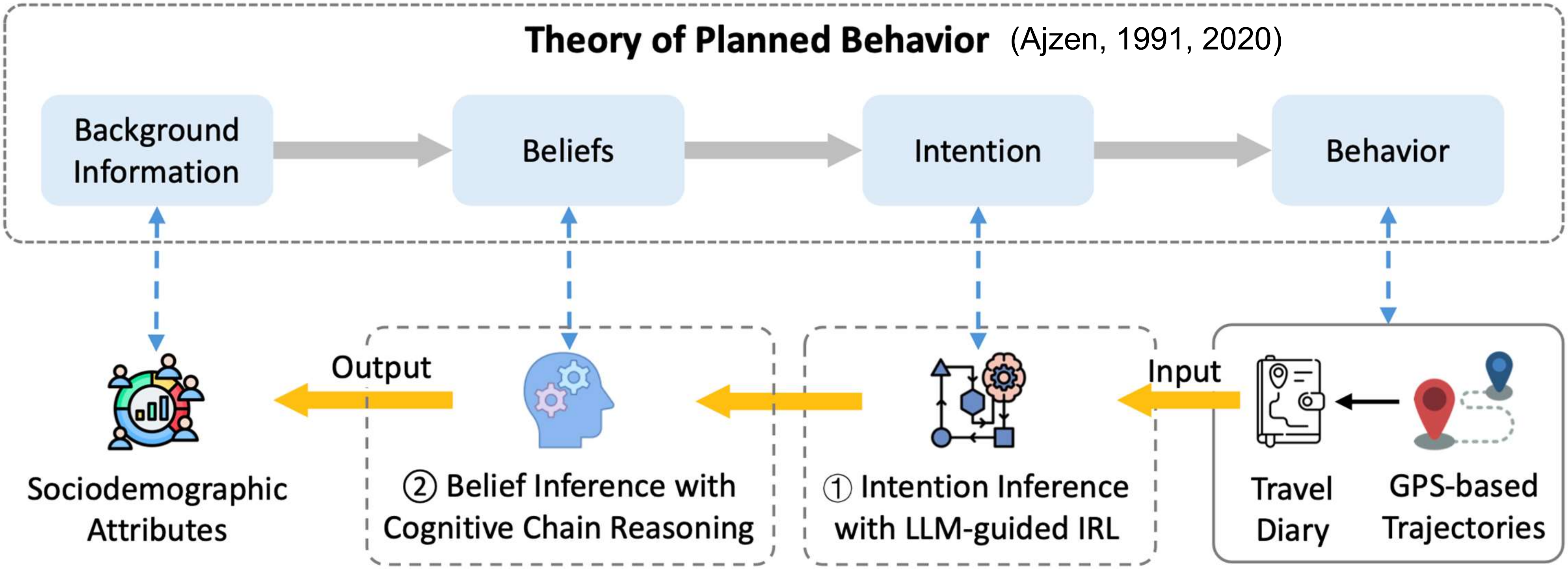
Xilei Zhao, Assistant Professor
University of Florida

Overview of Proposed Framework

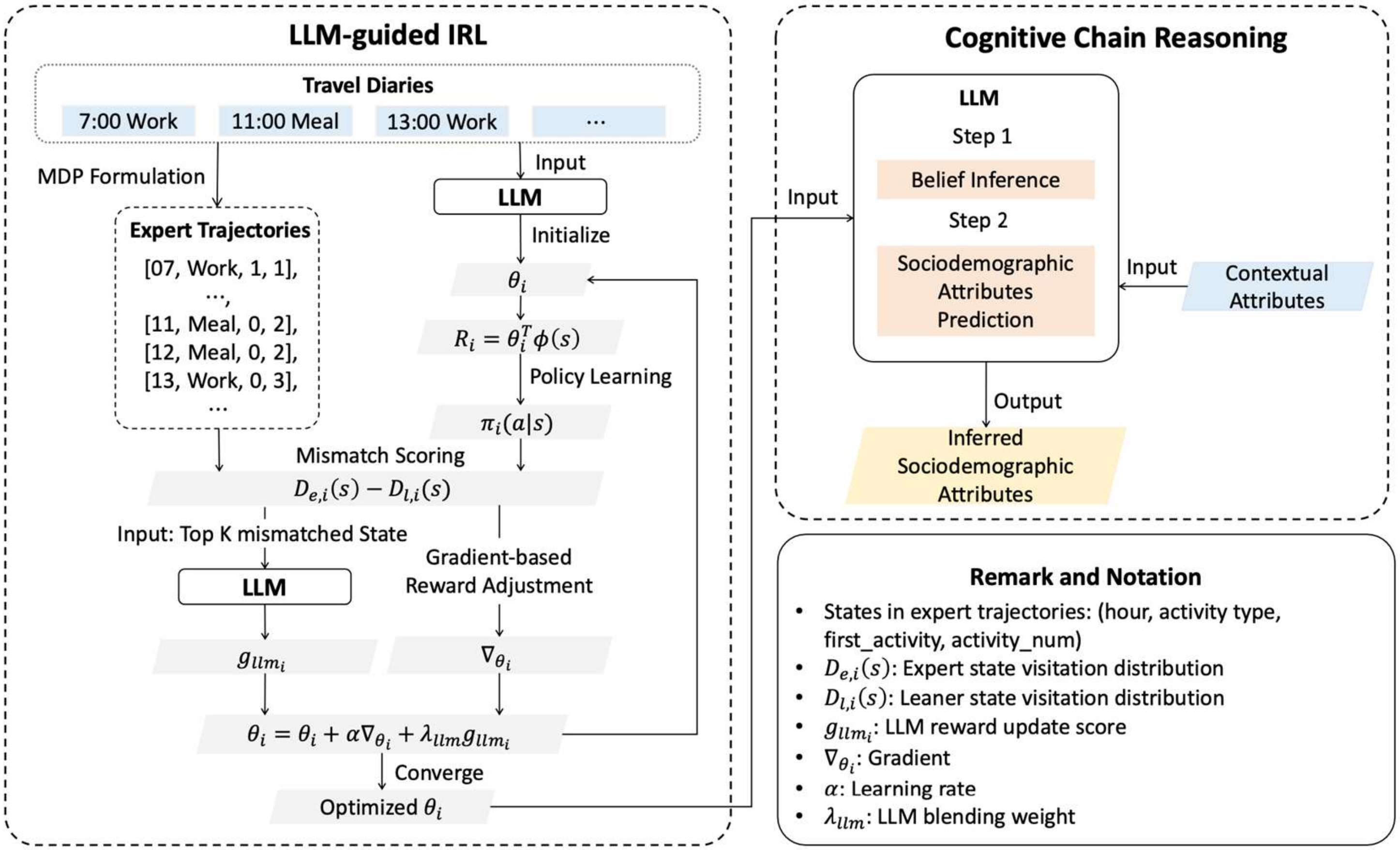


Can we inversely follow the cognitive pathway in the Theory of Planned Behavior (TPB) to recover individuals' sociodemographic attributes?

Overview of Proposed Framework



SILIC: Sociodemographic Inference with LLM-guided Inverse Reinforcement Learning (IRL) and Cognitive Chain Reasoning (CCR)



Experiment

- **Baselines**: SVM, XGBoost, CatBoost, and GPT-4o.
- **Datasets**:
  - The 2017 Puget Sound Regional Council Household Travel Survey with the seven-day smartphone-based GPS diaries.
  - The contextual attributes are from the U.S. Census Bureau, the Washington State Geospatial Open Data Portal, and the EPA Smart Location Mapping dataset.

Results

Table 1: Model comparison for gender prediction. The proposed model outperforms all baselines in both overall and class-level performance.

Method	Class	Precision	Recall	F1-score	Accuracy	Weighted F1
SVM	Male	0.603	0.633	0.618	0.621	0.621
	Female	0.639	0.609	0.624		
XGBoost	Male	0.597	0.667	0.630	0.621	0.620
	Female	0.649	0.578	0.612		
CatBoost	Male	0.683	0.642	0.662	0.645	0.646
	Female	0.607	0.649	0.627		
GPT-4o	Male	0.627	0.700	0.661	0.653	0.653
	Female	0.684	0.609	0.645		
GPT-4o (Diaries)	Male	0.625	0.667	0.645	0.645	0.645
	Female	0.667	0.625	0.645		
SILIC	Male	0.920	0.767	0.836	**0.855**	**0.853**
	Female	0.811	0.938	0.870		

Table 2: Model comparison for age prediction. The proposed model outperforms all baselines in both overall and class-level performance, and effectively identifies age groups that baseline models struggle to classify.

Method	Class	Precision	Recall	F1-score	Accuracy	Weighted F1
SVM	18–44	0.752	0.989	0.854	0.742	0.644
	45–64	0.000	0.000	0.000		
	65+	0.333	0.200	0.250		
XGBoost	18–44	0.767	0.859	0.810	0.677	0.650
	45–64	0.250	0.148	0.186		
	65+	0.200	0.200	0.200		
CatBoost	18–44	0.792	0.870	0.829	0.718	0.700
	45–64	0.421	0.296	0.348		
	65+	0.250	0.200	0.222		
GPT-4o	18–44	0.870	0.870	0.870	0.710	0.732
	45–64	0.615	0.296	0.400		
	65+	0.000	0.000	0.000		
GPT-4o (Diaries)	18–44	0.863	0.891	0.877	0.734	0.740
	45–64	0.571	0.296	0.390		
	65+	0.067	0.200	0.100		
SILIC	18–44	0.887	0.989	0.935	**0.863**	**0.844**
	45–64	0.900	0.375	0.529		
	65+	0.500	0.800	0.615		

Results

Table 3: Ablation Study of LLM-guided IRL. It demonstrates the effectiveness of both LLM-based reward initialization and iterative update guidance.

	KL Divergence	L1 Distance
Random Initialization + Gradient Ascent	0.594	0.722
LLM-guided Initialization + Gradient Ascent	0.447	0.696
Random Initialization + LLM-guided updates	0.543	0.704
LLM-guided Initialization + updates	**0.419**	**0.654**

Table 4: Ablation study of CCR on gender and age prediction. CCR outperforms both pure inference and CoT.

Method	**Gender**		**Age**	
	Accuracy	**Weighted F1**	**Accuracy**	**Weighted F1**
IRL + Inference	0.718	0.708	0.815	0.797
IRL + CoT	0.766	0.766	0.823	0.812
IRL + CCR	**0.855**	**0.853**	**0.863**	**0.844**

Key Takeaways

- SILIC can accurately infer individuals' sociodemographics from observed mobility patterns by leveraging the TPB to guide the AI model development.
- LLMs are leveraged to provide heuristic support for IRL reward function initialization and updates, enabling the inference of individualized and well-posed reward solutions.
- The proposed CCR module can guide the LLM to make predictions in alignment with the TPB, by explicitly modeling the mediating cognitive constructs.
- SILIC offers a solution for enriching large-scale location data (real or synthetic) for various important applications in transportation and beyond.

Acknowledgement


UF
Herbert Wertheim
College of Engineering
UNIVERSITY *of* FLORIDA
POWERING THE NEW ENGINEER TO TRANSFORM THE FUTURE