

大規模言語モデルに基づく 都市擬似人流生成モデルの構築とその応用

張 鯤翼, 張 育榕, 龐 岩博, 関本 義秀

研究背景

人流シミュレーションは、日本の高齢化社会、温室効果ガス排出、インバウンド回復など、多様な交通課題において重要な役割を果たしている。ここ数年、人流のビッグデータと大規模言語モデル（LLM）を組み合わせたシミュレーション研究が行われてきた。しかし、プライバシーに関する制約により、豊富な個人情報と交通シミュレーションモデルの統合に注目する研究はほとんど行われて来なかった。この問題に対処するため、本研究では生成系AIを活用し、エージェントの行動パターンと個人情報の関連性を捉えられるMobility Generative Language Model（MobGLM）を提案し、より精密な人流の再現を目指した。

研究目的

- I. GPT2という大規模言語モデルに基づき、パーソントリップ調査におけるユーザーの行動を予測する生成型AIを開発する。
- II. 事前確率に基づくロジット調整法を利用し、ショッピングなど希少な非日常活動のロングテール分布に対処する。

研究方法

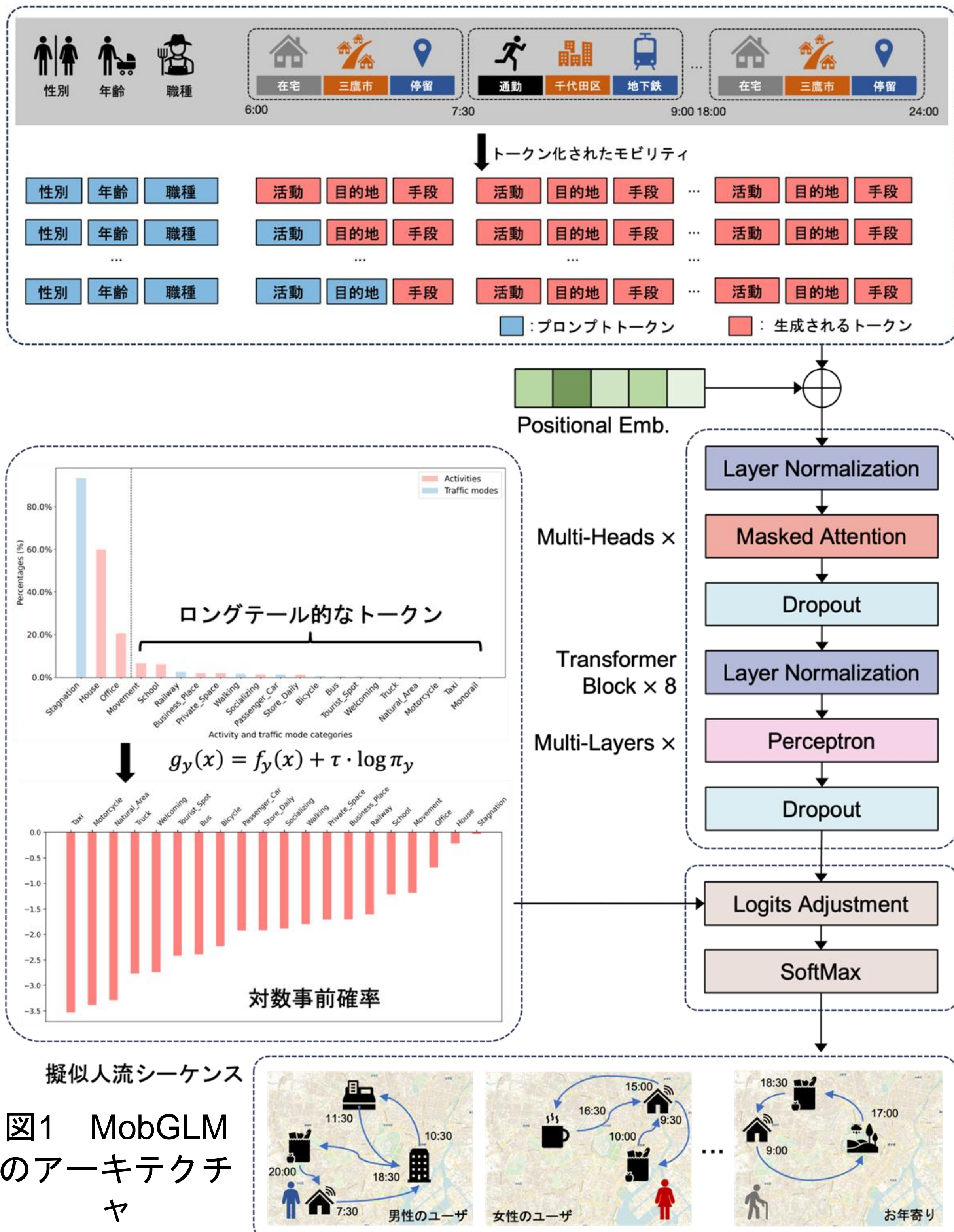


図1 MobGLMのアーキテクチャ

東京都における57万人を超えるユーザーの個人情報や位置情報をトークン化し、擬似的な移動シーケンスを生成した。主要な日常活動の生成を減らすため、活動と交通手段のロジット調整をパラメータ $\tau=5$ で行っている。

実験成果

表1 ベースラインモデルと提案されたMobGLMモデルのパフォーマンスの比較

モデル	TMR	LCSS	BLEU	ROUGE-L	DTW
3rd Order MM	0.791	0.545	0.700	0.676	10338.26
LSTM	0.855	0.449	0.835	0.772	9123.501
GRU	0.856	0.453	0.836	0.777	9071.662
Transformer XL	0.831	0.593	0.857	0.801	8728.593
MobGLM w/o LA	<u>0.885</u>	0.653	0.896	0.846	8410.800
MobGLM	0.884	<u>0.658</u>	<u>0.896</u>	<u>0.846</u>	<u>8410.76</u>

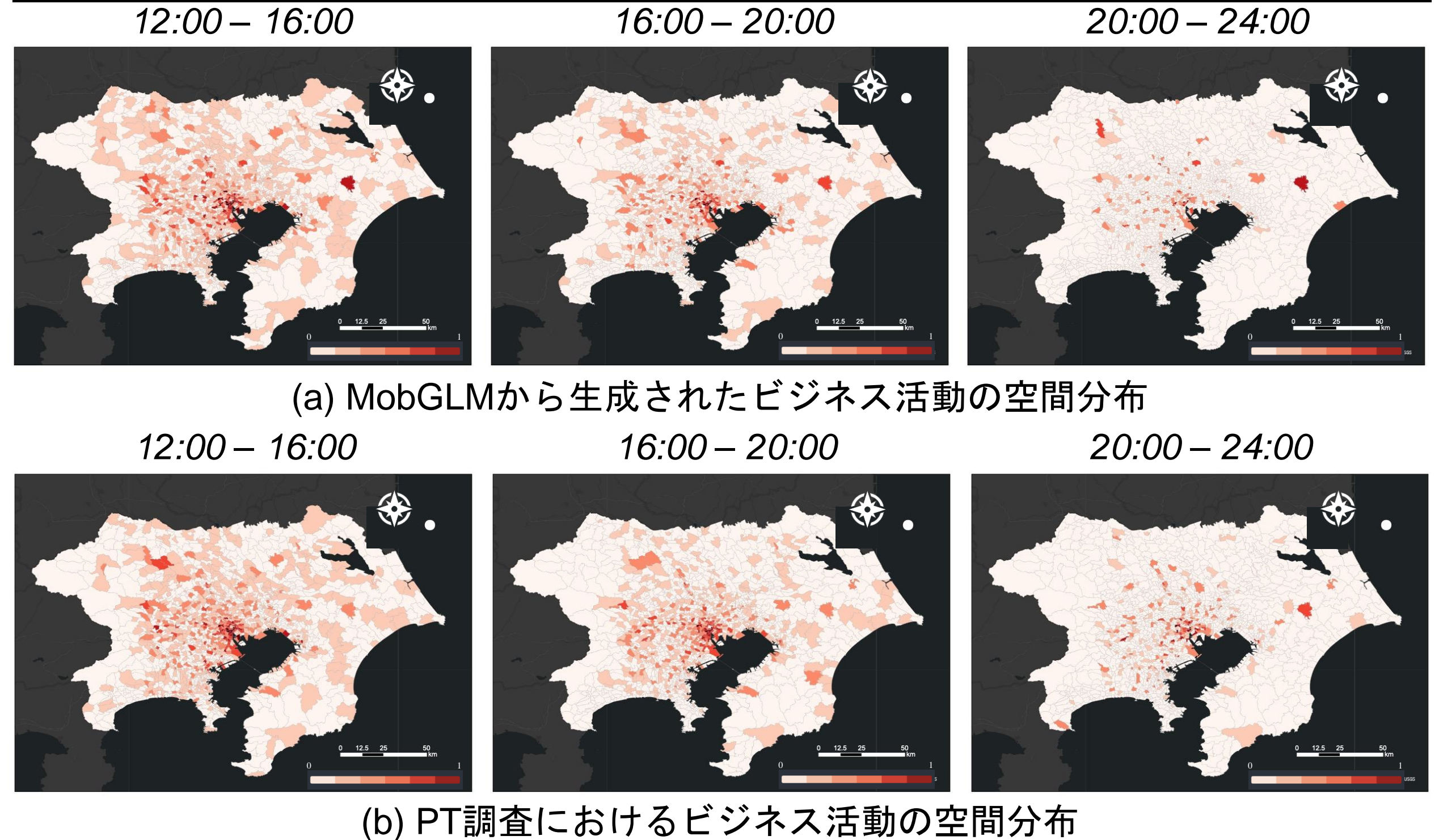


図2 生成された行動と実際のデータとの空間分布の比較

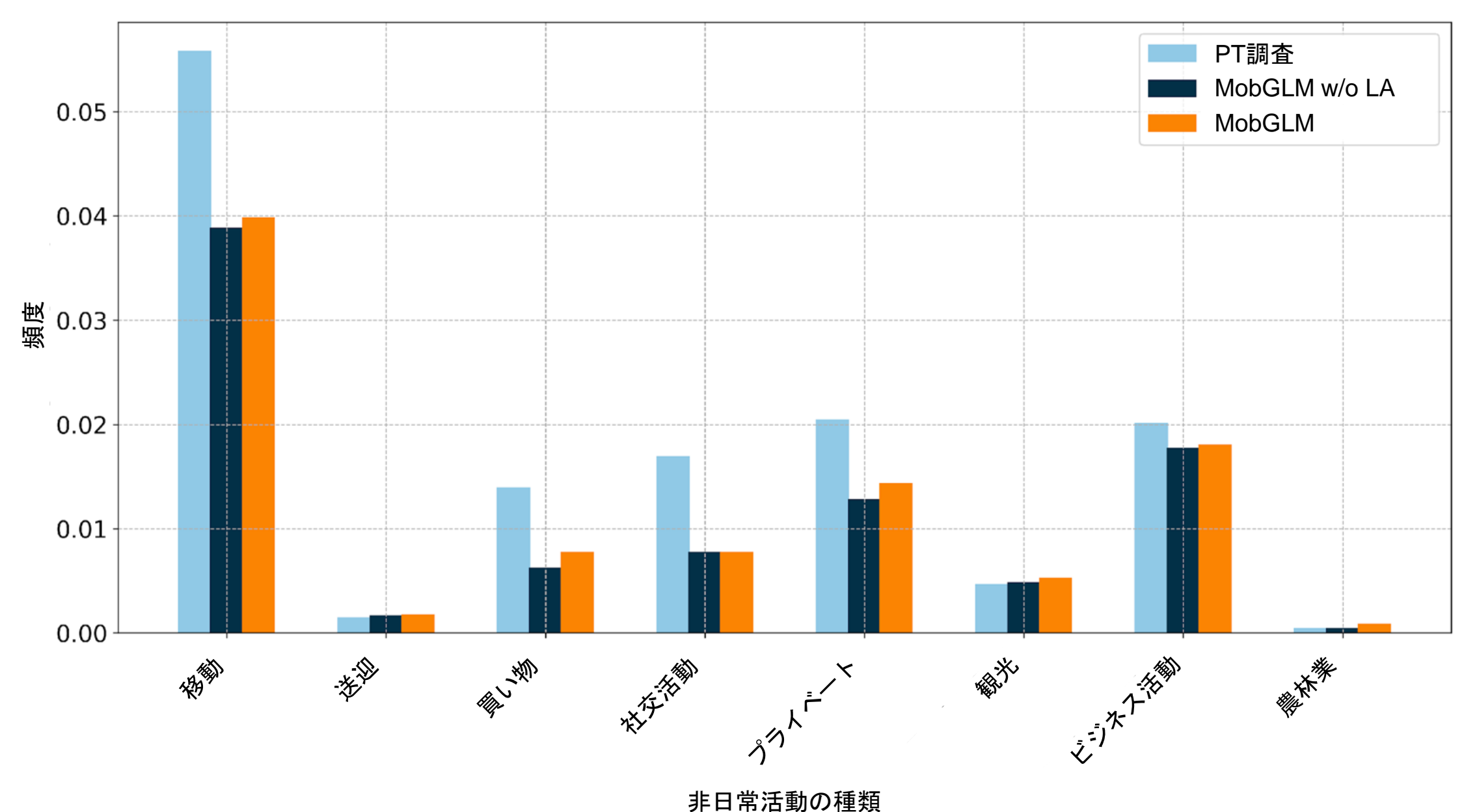


図3 非日常的な活動の頻度の比較

まとめ

人流シーケンスの予測のために最適化された大規模言語モデルに基づき、人間の行動をシミュレートする新しい生成的モデルを提案した。この研究により、擬似人流を生成し、人々のリアルな行動を再現する可能性が明らかになった。