

# GPT-2 を用いたテーマパーク内移動軌跡の時空間情報生成

○堀込泰三 水野貴之 (国立情報学研究所)

## Generation of spatiotemporal information on human trajectories in a theme park by GPT-2

\* T. Horikomi and T. Mizuno (National Institute of Informatics)

**概要**— Mizuno, Fujimoto, and Ishikawa (Front. Phys. 2022)は、緯度と経度で表された空間情報を相対的な位置を表すユニークなトークンに変換することで、自己回帰型言語モデル GPT-2 で、移動時間の解像度が一定の個人の移動軌跡が生成できることを示した。ところが、一般的なスマートフォンの GPS から得られる位置情報は、記録頻度が一定でないことが多い。そこで我々は、位置を表すトークンの前に出発地から目的地までの移動時間間隔を表すユニークなトークンを加えることで、移動軌跡の生成モデルの時間解像度を可変長に拡張する。

**キーワード:** 移動軌跡生成, 自然言語処理, GPT2, 時空間情報

### 1 はじめに

Mizuno, Fujimoto, and Ishikawa (Front. Phys. 2022)は、自然言語生成における Transformer モデルの 1 つである GPT-2 を用い、個人の移動軌跡の空間情報の生成モデルを構築した<sup>1)</sup>。彼らは、時間解像度を 30 分間と固定し、30 分毎の位置で表される移動軌跡を GPT-2 でモデル化している。しかし、世の中で扱われる多くの移動軌跡は、時間解像度が固定ではなく、出発地から目的地 1, 目的地 1 から目的地 2, … というように、Origin-Destination (OD) の各位置とその移動時間で表現される。従って、一般的な移動軌跡の情報と親和性の高いモデルは、位置と移動時間の両方を学習して、それらを生成する時間解像度が可変長な移動軌跡モデルである。我々は、位置を表すトークンの前に時間間隔を表すユニークなトークンを加えることで、個人の移動軌跡の空間情報と時間情報の両方を同時に生成するモデルを構築する。

### 2 位置情報データ

株式会社Agoopが提供する、2022年8月に千葉県浦安市を通過した計67万台のスマートフォン（1日あたり約2万2千台）の位置データ（1億2846万ログ）を利用した。位置情報には緯度、経度、日時、ユニークなスマートフォンIDが含まれる。位置情報の記録間隔の平均と中央値は、それぞれ、約7分間と約1分間である。

### 3 移動時間と位置を生成する GPT-2 モデル

Mizuno らは、緯度と経度で表された空間情報を相対的な位置を表すユニークなトークンに変換することで、自己回帰型言語モデル GPT-2 で、移動時間の解像度が一定の個人の移動軌跡が生成できることを示した<sup>1)</sup>。本稿では、時間モデルの移動時間を可変長に拡張するために、移動の時間間隔 $\Delta t$ を導入する。我々は時間間隔 $\Delta t$ を、 $\tau = \text{int}(\log_{1.5}(\Delta t)) + 1$ で離散化し、それぞれの離散化時間間隔 $\tau$ にユニークな文字を割り当て、位置を表す文字列 X の前に加える。我々は、東京ディズニーリゾート内の個人の 15.625m 解像度の移動軌跡を上記の文字列で表現して GPT-2 で学習する。東京ディズニーリゾートの面積は 201 万  $\text{m}^2$  であるため、15.625m 解像度、約 8233 グリットに分割される。2 節で紹介したデータ内における東京ディズニーリゾートに滞在した約 7 万の移動軌跡を、

4 : 1 で学習データとテストデータに分割する。これらのデータにおける東京ディズニーリゾートでの滞在時間は、平均値が 11 時間で、最頻値は 12 時間台である。

### 4 結果

Fig. 1 は、東京ディズニーランドにおける入場ゲートを初期座標として、学習した GPT-2 で移動軌跡を 3 本生成した例である。様々なアトラクションを、違和感なく巡っている様子が生成できていることがわかる。GPT-2 で生成される移動軌跡における移動時間間隔の確率密度分布は、実際の分布関数と統計的に一致する。

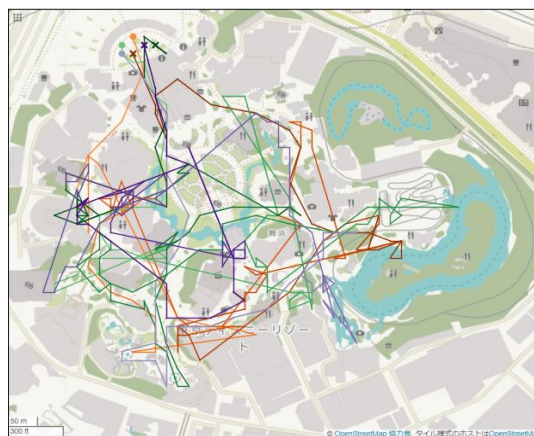


Fig. 1: 生成された移動軌跡の例。

### 5 まとめ

本稿では Mizuno ら<sup>1)</sup>が提唱した手法を拡張し、移動時間を可変長にした。また、文献 1 では都市、本稿ではテーマパークと、異なるスケールでの移動軌跡を、同様の手法を用いて生成することに成功した。今後、小売店舗内の動線といった、さらに異なるスケールでの移動軌跡生成モデルの構築に取り組みたい。

### 参考文献

- 1) Mizuno T, Fujimoto S and Ishikawa A: Generation of individual daily trajectories by GPT-2, Frontiers in Physics, 10:1021176 (2022)