

GPT-2 を用いた位置情報ビッグデータに基づく 交通需要推定技術

小林 亮博¹・上坂 大輔¹・武田 直人¹・南川 敦宣²・森本 章倫³

¹ 非会員 KDDI 総合研究所 (〒356-8502 埼玉県ふじみ野市大原二丁目 1 番 15 号)
E-mail: [ak-kobayashi, da-kamisaka, no-takeda]@kddi-research.jp

² 非会員 KDDI 株式会社 (〒102-0072 東京都千代田区飯田橋 3 丁目 10 番 10 号)
E-mail: at-minamikawa@kddi.com

³ 正会員 早稲田大学教授 理工学術院 (〒169-8555 東京都新宿区大久保三丁目 4 番 1 号)
E-mail : akinori@waseda.jp

近年発達が目覚ましい深層学習技術を用いて、スマートフォン上で取得した大規模な位置情報（ビッグデータ）をモデル化し、交通需要を予測する試みが始まっている。その多くは RNN ベースであり、ユーザ毎の長期間のコンテキストを活用して行動を予測することが困難であった。本研究では自然言語の生成に用いられている言語モデル GPT-2 を人流のモデル化に適用するため、訓練タスクの設定方法とコンテキストに基づく行動生成技術を提案する。実際のスマートフォン位置情報から宇都宮市の人流をモデル化し、生成した OD 表と観測値を比較したところ、提案手法は従来法に比べて最大で 40% 高い精度で都市の人流を再現できた。

Key Words: Location Big Data, Traffic Demand Forecast, ABM, GPT-2

1. はじめに

近年、人中心のモビリティネットワークを構築することが、これまで以上に都市交通施策の検討で重要となってきており、人々の活動や暮らしが表現できる Activity Based Model (ABM) を用いた交通行動のシミュレーションに注目が集まっている¹⁾。ABM は、乗り継ぎ施策など多様な施策が評価できるとともに、評価にあたっては交通量の指標だけでなく、個人の活動の変化（外出率や活動時間、アクティビティパターンの変化等）や滞留人口等の多様な切り口で影響を把握することが可能である。

一方で、交通行動のシミュレーションを利用データの面から考えた場合、ある 1 日の交通行動のみを調査するアンケートに代わり、継続的に観測可能なスマートフォンの位置情報の重要性が高まってきている²⁾。特に、深層学習技術の発展に伴い、大量の系列データを訓練することで後に続く系列を予測可能な Recurrent Neural Network (RNN) を用いて ABM を構築する試みに着目が集まっている³⁴⁾。Yin ら³⁾は、RNN の一種である LSTM を用い、スマートフォンで観測された朝のトリップ系列からその日一日の交通行動を予測し、都市全体で集計することで夕方の渋滞を予測している。しかし、RNN は取り扱える系列長に限界があるため、ごく短い系列あるいは初期

位置のみから 1 日の行動を予測しており、ユーザ間の違いを十分に扱えない。

例えば、図 1 に示すユーザの行動を予測する状況を考える。A→B→C→X という系列を訓練した場合、RNN に A→B→C という系列を入力すると、次にくる系列として X を出力する。しかし、訓練データ中に含まれるユーザの大半が A→B→C→D という行動を取っている場合、図 1 のユーザがたとえ毎日 A→B→C→X という行動を取っていたとしても、RNN が正しく行動を予測することは困難である。

これに対し本研究では、スマートフォン位置情報の継続性に着目し、長期間の行動系列そのものをユーザの特性を示すコンテキストとして Neural Network (NN) に入力し、次の 1 日の行動を予測する手法を提案する。図 2 は

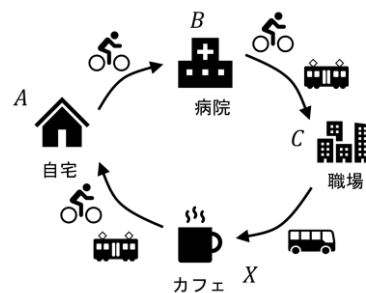


図 1 ユーザの 1 日の行動

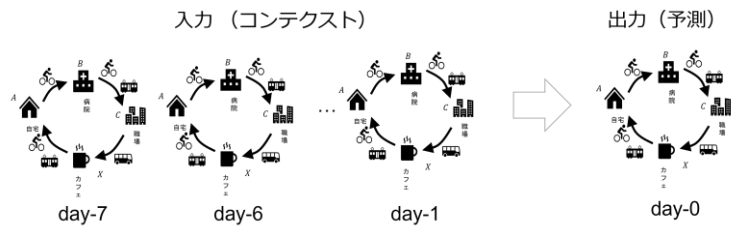


図2 コンテキストを用いた行動予測

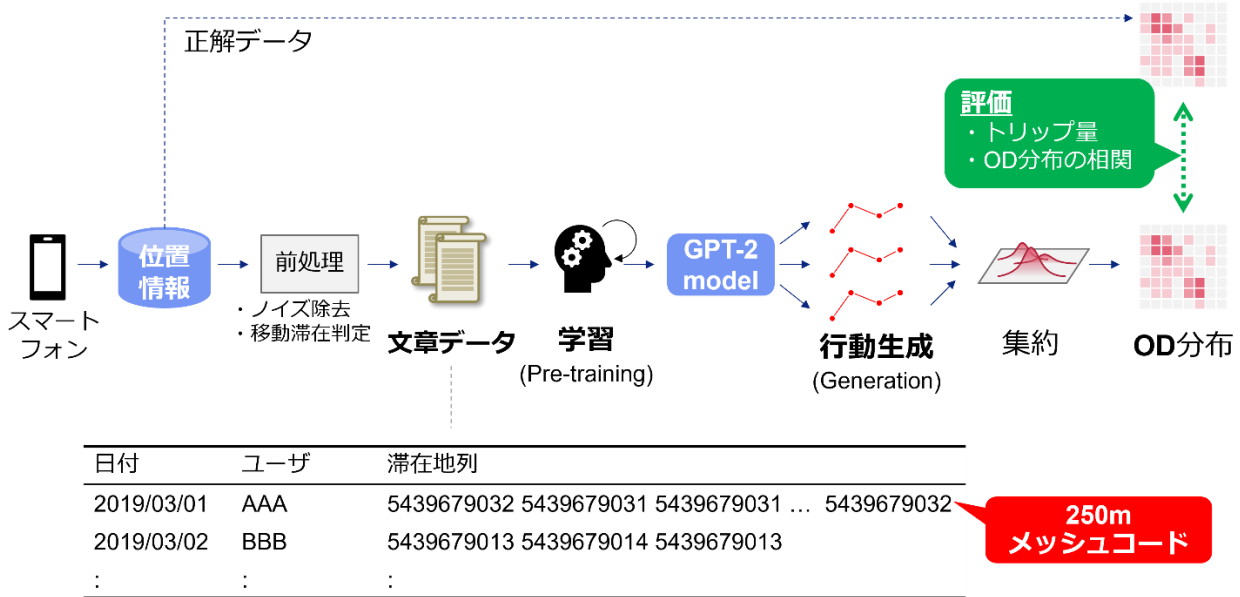


図3 交通行動シミュレーションの概要

コンテキストを用いた行動予測の概念図である。訓練データにA→B→C→Xというデータ系列が少量しか含まれていない場合であっても、ユーザが過去1週間にA→B→C→Xという行動パターンをとっていることをNNに入力することで、該当ユーザの次の1日の行動系列を高精度に予測する。このとき、長期間の行動系列をNNで取り扱うため、本研究では長期間の文脈、依存性 (long-term dependency) を表現・認識できるTransformer⁹⁾を用いてユーザの交通行動を予測する手法を提案する。

本研究では、宇都宮市を対象とし、実際に観測されたスマートフォン位置情報から行動予測モデルを訓練し、その基本的な性能を評価する。さらに、推定結果を可視化することで、提案手法の性質や限界について検討する。

2. 関連研究

Yin³⁾は、スマートフォンのCall Detail Record (CDR)からユーザの滞在地を抽出し、ユーザの一日の移動軌跡を滞在地IDの系列で離散的に表現した。この系列をLong Short Term Memory (LSTM)で訓練し、各ユーザの自宅位置を初期位置として1日の行動を生成することで、再現したサンフランシスコ湾岸地域の道路交通量が計測された道路交通量の傾向と一致することを確認した。しかし

LSTMなどのRNNは、並列化が困難な再帰処理を系列の長さに応じて必用とするため、取り扱える系列長に制限があった。

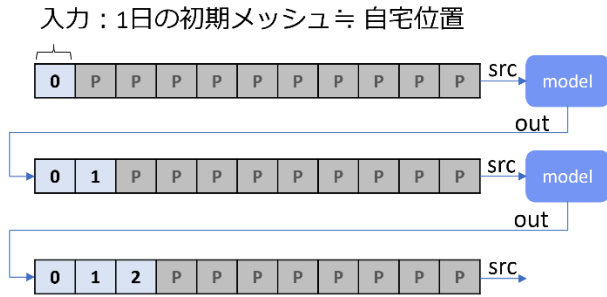
この課題に対しTransformer⁹⁾は、並列化が容易なSelf-Attentionを用いて系列のどこに注目すべきかをデータから訓練しているため、Graphics Processing Unit (GPU)を用いることで長大な系列を高速に取り扱うことが可能であり、勾配消失なしに長期間の文脈を訓練できる。Transformerは翻訳などの主に自然言語処理で用いられてきたが、近年では位置情報の表現に活用する試みが始まっている。Lin⁴⁾は、Transformerを用いた自然言語モデルBERT⁹⁾を用いて位置を埋め込みベクトルで表現した。グリッドで離散表現された移動軌跡の一部をマスクし、正解となるグリッドを当てるタスクをBERTが訓練することで位置の分散表現を獲得し、移動軌跡の予測精度を向上させた。しかし、BERTの活用は位置の表現に留まり、予測には依然としてLSTMが用いられている。

3. GPT-2を用いた交通行動のシミュレーション

(1) 交通行動シミュレーションの概要

図3に本研究が提案する言語処理モデルを活用した交通行動のシミュレーションの概要を示す。スマートフォ

■ 生成方式 I (従来法) [4][5]



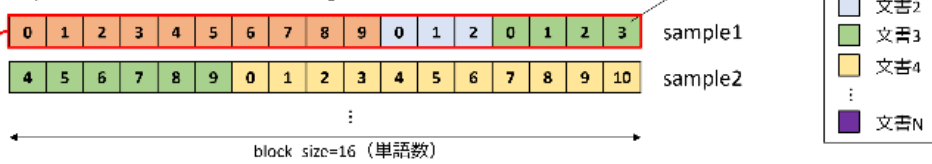
■ 生成方式 II (提案法)



図 4 コンテキストを用いた行動生成

■ 訓練タスクの生成

- ・入力された文書をバッファリングしてsampleを作成
- ・各sampleの先頭/末尾を削除してsrcとtgtを作成



■ 訓練

srcをGPT-2 Modelに投入して得たout とtgt との Cross Entropy Loss を計算し誤差を逆伝搬させる (全サンプルに対して実施)

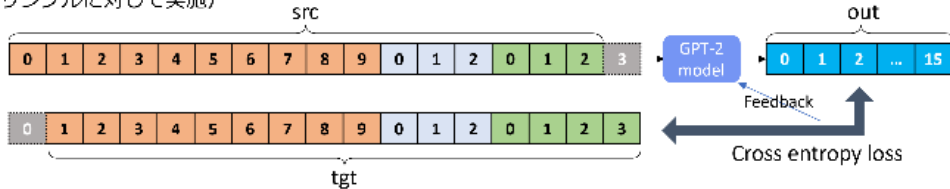


図 5 GPT-2 の訓練

ンにより計測された位置情報から、ユーザの1日の滞在位置を時刻順に抽出し、地域メッシュコードで表現する。図3下の表のように離散表現された滞在系列を、本稿では文章データと呼ぶ。個々のメッシュコードを単語と捉えた場合、文章生成課題を解くことで、ユーザの1日の移動行動を生成できる。ある都市から抽出した大量の文章データを訓練し、一定のバリエーションを持たせた文章生成を複数回実施することで、その都市の人流分布を模したデータを生成できる。

本研究では Generative Pre-trained Transformer 2 (GPT-2)⁷⁾を用いることで、位置表現の獲得と移動軌跡の予測を同時に実現する。GPT-2は、Transformerをベースとしたテキスト生成モデルであり、与えられた単語の系列から、次の単語を予測するタスクを訓練(Pre-Training)する。そのため逐次的に単語の予測を繰り返すことで文章生成(Generation)が可能である。さらに、大量のテキストを訓練することで、文章判別やQ&Aなど汎用的な自然言語処理タスクに活用可能なテキストエンコーダを獲得し、人間が作成した文章と区別がつかない自然な文章を生成すると言われている⁸⁾。一方で、訓練に大量のデータと計算量を要することがGPT-2の欠点である。しかし、ス

マートフォンを通じて多数のユーザの移動行動を継続的かつ大量に観測可能となった今、GPT-2を用いた交通行動のシミュレーションに取り組む価値は高い。

(2) コンテキストを用いた行動生成

図4は行動生成部の動作を示す。図中の0, 1, 2はメッシュコード、Pはパディングを示す。GPT-2は、図4左上に示すデータ系列が入力されると、パディングを無視しメッシュコードの系列から次にユーザが滞在する可能性が高いメッシュを予測する。従来は、1日の初期位置をGPT-2に入力し(多くの場合ユーザの自宅位置となる)、逐次的に次の滞在地を予測し、1日の終了を示す特殊なコードが現れた場合に行動の生成を終了する(生成方式I)。これに対し提案手法(生成方式II)では、該当ユーザの長期間の行動系列(例えば過去1週間の滞在系列)と1日の初期滞在地をGPT-2に入力する。過去1週間の系列にはユーザの周期的な行動(例えば、水曜日が定休日)が含まれるため、ユーザの細かな特性に応じた行動が生成可能となる。

従来GPT-2は図5に示す方式で訓練を行う。訓練データ中の文書を一定の単語数のバッファに格納していき、

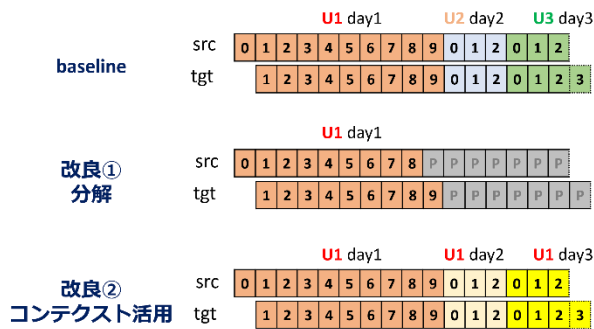


図 6 訓練タスク生成部の改良

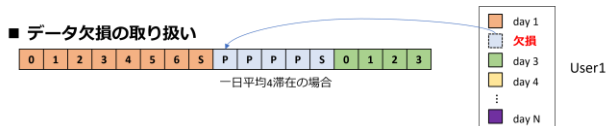


図 7 データ欠損への対応

1 サンプルの訓練タスクとする。各サンプルについて、最後の 1 単語を除いた系列(src)を GPT-2 に入力し、GPT-2 が出力した結果(out)と、サンプルの最初の一文字を除いた正解(tgt)と比較し Cross Entropy 誤差を算出する。この誤差を逆伝搬させることでモデルを訓練している。しかし、訓練タスクの生成時に文書間の順序は考慮されないため、スマートフォン位置情報から抽出した文章データに適用した際に、異なるユーザ・異なる日の行動系列が混在して連結されてしまう。

本研究は上記の課題を解決するため、訓練タスクの生成部に対し 2 種類の改良を検討した(図 6)。改良①ではユーザや日付の混在を無くするため 1 サンプルに含まれる滞在系列を 1 ユーザ 1 日分に限定した。改良②ではコンテキストを活用した行動生成の精度を向上するため、同一ユーザの系列が 1 日毎に連続するよう訓練タスクを生成する。本研究では、コンテキストとなる 7 日分の滞在系列と予測対象となる 1 日分の系列を連結し、訓練タスクとした。

ただし、スマートフォンの測位周期には粗密があり、データが 1 日中欠損することも起こりうる。ここで、欠損日を飛ばして隙間の無いサンプルを生成した場合、「月曜日にだけ習い事に通う」などのユーザ行動の周期性を正しく訓練できない。そのため、本研究では欠損日はパディングを平均的な滞在長だけ埋めることで、滞在系列に含まれる周期性を保つ(図 7)。

4. 対象都市における実験

本研究では、対象都市で観測されたスマートフォン位置情報から図 6 に示す 3 種の方法(baseline, 改良①, 改良

②) で構築されたモデルを用い、図 4 に示す 2 種の方法(生成方式 I, 生成方式 II) で移動行動を生成し、スマートフォンで実際に観測された移動行動と比較することで提案手法の有効性を評価した。

(1) データセット

本研究では、携帯電話の GPS データのうち、KDDI 株式会社が au スマートフォン利用者より同意を得て取得する GPS データ(以下、「スマートフォン位置情報データ」という。)を使用して検証を行った。スマートフォン位置情報は測位間隔に粗密があるが、集計した結果は各種人口統計⁹⁾や PT 調査¹⁰⁾と関連するだけでなく、交通需要の絶対数も精度良く推計可能¹¹⁾なデータである。

本研究では、2019 年 3 月から 7 月において、宇都宮市に外接する矩形上で検出された、スマートフォン位置情報データを抽出し、文章データ化した。滞在地の抽出は手法¹²⁾を用い、250m メッシュコードの系列でユーザの 1 日の移動行動を表現した。なお、2019/3/1-7/17 を GPT-2 の訓練に用い、2019/7/19-7/31 を評価に用いた。

(2) GPT-2 の訓練と移動軌跡の生成

GPT-2 は単語を ID に変換した上で訓練している。ID への変換処理は、言語非依存の ID 変換処理を実現した Sentencepiece¹³⁾を用いた。生成時に無効なメッシュコードが生成されることを避けるため、出現頻度が低い単語を分割する機能は用いず、1 メッシュ IID に変換した。変換した ID 系列から図 6 に示す 3 種の方式でタスクを生成し、それぞれ GPT-2 を訓練した。報告⁷⁾におけるパラメータ数 117M のモデルに従い、レイヤ数: 12, 隠れ層: 768 次元, 自己注目ヘッド数: 12 を設定した。GPT-2 の実装は Hugging Face transformers¹⁴⁾を用いた。

ユーザの行動系列は、確率的に様々なパターンの文章を生成する方式である Top k, Top p サンプルリング¹⁵⁾を用いた。Top k は次の滞在先として、尤度が上位何件までのメッシュから選択するかを示し、Top p は Top k で選ばれた候補からさらに尤度の合計でどれだけ絞り込むかを示し、いずれも値が大きいほどバリエーションに富んだ行動を生成する。一方で文章生成においては、beam search と呼ばれる、深さ優先探索を用いて可能な限り尤もらしい文章を 1 つ生成しようとする手法がある。Welleck ら¹⁶⁾によれば Top p, Top k サンプルリングに比べて beam search 方式は人間が作る文章に近い文章を生成すると言われている。しかし、事前実験では、都市全体で集約した場合、Top p, Top k サンプルリングの方が beam search に比べて実際の分布に近いことが分かったため、前者の方式を用いて生成した。具体的には、Top p は 0.8/0.85/0.9/0.95 の 4 種のパラメータを用い、Top k は 100 に設定した。

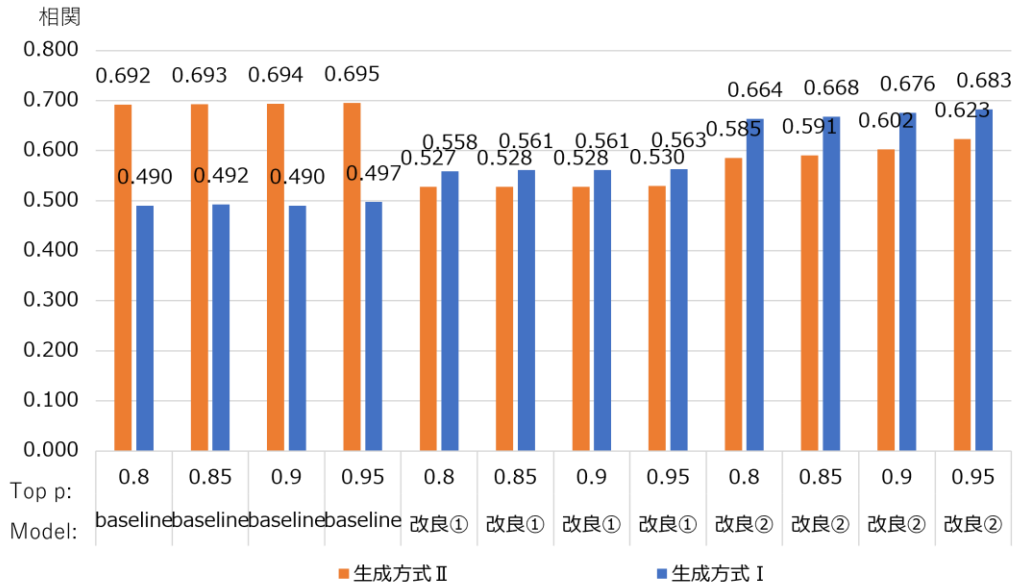


図 8 GPT-2 が生成した OD と実際の OD との相関

行動生成にあたり、図 4 の 2 種類の方式を入力し結果を比較した。生成方式 I では 3/29-31 に観測された滞在系列の最初の 1 メッシュのみを入力とし、生成方式 II では上記 1 メッシュに加えて直前の 1 週間分の系列を入力した。生成方式 II においては、直前の 1 週間に滞在が全く存在しない日が 3 日以上存在する場合は評価対象外とした。また、1 日の行動生成の限界として訓練データ中における 1 日の滞在数の 3 四分位を設定し、同じ入力に対し 100 回乱数を変えて 100 パターンの行動を生成した。

(3) 生成した移動軌跡の評価

生成した移動軌跡は図 3 に示すように、集計した OD がログの分布と合致するかといった観点で評価した。滞在地間の遷移をトリップと捉え都市全体で OD 表を集計し、生成した OD 表と観測した OD 表の相関を取り、分布の傾向が実際の人流とどの程度一致するか確認した。また、1 日におけるトリップ数の平均を生成した系列と実際の系列で比較し、生成量が実際の人流と一致するか確認した。また、各メッシュの集中交通量を算出し可視化することで、生成された交通量の傾向について分析した。

5. GPT-2 が生成した行動の分析

図 8 は、GPT-2 が生成したトリップの OD 分布と、観測した OD 分布の相関を示す。Baseline モデルで比較した場合、生成方式 I は相関が 0.497 (Top p = 0.95) に対し、生成方式 II では相関が 0.695 (Top p = 0.95) と、40% も精度が

向上している。一方で、モデル間を比較すると、生成方式 I で行動を生成する場合は改良②→改良①→Baseline の順に高精度だが、生成方式 II で行動生成した場合は想定と異なり、Baseline → 改良② → 改良① の順の精度となっている。Top p に関しては、生成のバリエーションが増えるよう大きな値を設定した方が高精度となり、事前実験と同様の傾向を示した。

表 1 は、GPT-2 が生成した 1 日の平均トリップ数をまとめたものである。前節で述べた通り、生成方式 II で行動生成をする場合は、生成日の直前の 1 週間にデータ欠落が 3 日以上ある場合は評価対象から除いているため、生成方式 I と II でトリップ数の真値が異なる。生成方式 I においては、真値と比較すると 1 日のトリップ量が下振れする傾向があり、生成方式 II では真値に対して上振れするトリップを生成する傾向がある。モデル間で比較した場合、改良②モデルが生成方式 I・II いずれの場合においても最も真値に近いトリップ量を生成しており、

表 1 1 日のトリップ数平均

| model | Top p | 生成方式 I | 生成方式 II |
|----------|-------|--------|---------|
| baseline | 0.8 | 1.19 | 28.34 |
| | 0.85 | 1.25 | 28.04 |
| | 0.9 | 1.38 | 27.86 |
| | 0.95 | 1.62 | 27.83 |
| 改良① | 0.8 | 5.53 | 13.75 |
| | 0.85 | 5.54 | 13.50 |
| | 0.9 | 5.56 | 13.20 |
| | 0.95 | 5.59 | 12.77 |
| 改良② | 0.8 | 5.16 | 9.38 |
| | 0.85 | 4.99 | 8.63 |
| | 0.9 | 4.91 | 7.99 |
| | 0.95 | 4.89 | 7.54 |
| 真値 | | 5.92 | 4.63 |

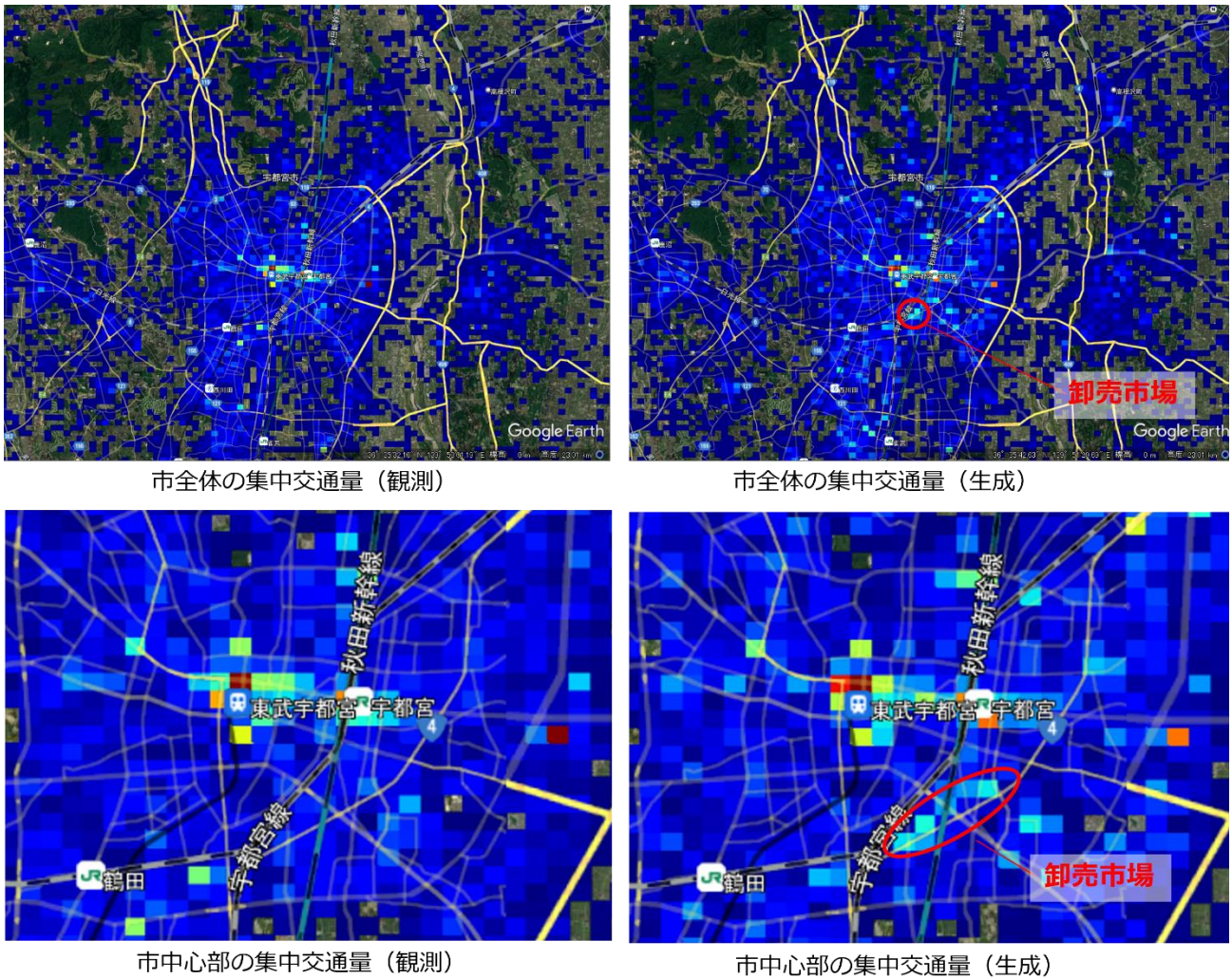


図9 GPT-2 が生成した集中交通量と観測値の比較

交通量予測などのアプリケーションにおいては有用と考えられる。

図 9 は、観測されたメッシュ毎の集中交通量と、OD 分布の相関が最も高い、生成方式 II (baseline, Top p = 0.95)

で生成した集中交通量のヒートマップである。GPT-2 が生成した集中交通量は、市中心部や外延部において観測結果と類似した形状に分布しており、GPT-2 は宇都宮市の都市交通を概ね再現できていると言える。一方で、卸売市場の周辺で観測結果には見られないトリップが生成されており、考察において検討する。

6. 考察

生成方式 II を用いた場合、改良②が想定と異なり Baseline に対して高い精度が出なかった。Baseline では GPT-2 のデフォルト設定に従い、サンプルのバッファサイズを 1024 に設定している。表 1 より 1 日のトリップ数

の平均は 5.92 回であるため、GPT-2 は約 169 日分のトリップ系列を同時に訓練していると言える。Baseline の訓練時にはユーザや日時をランダム化してバッファリングする処理は実施していなかったため、意図せずしてユーザのコンテキストが系列中に含まれていた可能性がある。

生成した OD 分布から近距離トリップを除くと異なる結果となる。文献⁹⁾によるとスマートフォン位置情報は PT 調査に比べて近距離のトリップを検出しやすく、本実験においても観測されたメッシュ内移動の割合は 45% と大きい。メッシュ内移動を除いて評価した場合、Baseline モデルが生成した OD 表と観測結果との相関は、0.457 (生成方式 II : Top p = 0.95) に対して、改良②の相関は 0.479 (生成方式 II : Top p = 0.95) と向上している。そのため、長距離の移動に関しては改良②のモデルの方が人流を適切にモデル化していると言える。長距離トリップの割合が多い宇都宮市より人口が少ない都市では、改良②モデルを用い生成方式 II で都市交通のシミュレーションを行うことが望ましいと思われる。

トリップの生成数という観点からは、生成方式 II は真

値を大きく上振れているため改善の必用がある。一方で、GPT-2をそのまま用いた場合、つまり、Baselineモデルを用い生成方式Iで行動を生成した場合、平均トリップ数2以下になっている。これは生成時に「1日に1回も外出しない」と誤って推定されたケースが多いことが原因と考えられる。一方で改良②によりユーザ毎に過去の行動系列をまとめてモデル化することで、この傾向が是正され、真値に近いトリップ数が生成されたものと考えられる。

図9では、卸売市場の近辺で真値に比べて過剰にトリップが生成されている。前述の通り、そもそもメッシュ内移動45%と多いデータセットで訓練し、生成結果もメッシュ内移動が62%（生成方式II，baselineモデル，Top p = 0.95）と多くなる傾向があるため、ユーザの初期滞在地がトリップ数に大きな影響を与える。卸売市場は、他のメッシュと異なり夜間に人口が集中する傾向があり、人口が集中した状態で行動生成がはじまっているため、過剰にトリップが生成されたと考えられる。生成方式IIは原理的にはどの時刻から行動の生成を開始しても良いため、コンテキストに含まれる人それぞれの生活形態を考慮して行動を生成することで、この傾向は是正されると期待できる。

7. おわりに

本論文では、テキスト生成で用いられているGPT-2を用いて、スマートフォン位置情報から都市の人流をモデル化し、シミュレーションする技術を提案した。特に人間の周期的な行動に着目し、ユーザ毎のコンテキストを用いて行動を生成する手法を提案した。提案した手法により生成した都市のOD分布と観測結果とを比較したところ、最大で40%相関が向上した。一方で、生成トリップ数は真値に比べて上振れる傾向が見られ、課題が残った。

今後は、訓練時におけるサンプルデータ抽出期間の制御によるデータ効率の向上、今回平日3日間のみであった評価期間の増大、単語のID変換における単語分割の活用による訓練効率の向上が必要である。さらに、今回単なるメッシュ系列であったユーザ行動への時刻情報の付与、政策変数の追加による未知の環境におけるシミュレーションなどが課題である。

参考文献

1) 東京都市圏交通計画協議会、「新たなライフスタイル

- を実現する人中心のモビリティネットワークと生活圏 — 転換点を迎えた東京都市圏の都市交通戦略 —」, 2021年（令和3年3月）
- 2) 国土交通省都市局 都市計画課都市計画調査室, 「総合都市交通体系調査におけるビッグデータ活用の手引き」, 2018年（平成30年6月）
 - 3) Yin, M., Sheehan, M., Feygin, S., Paiement, J.F. and Pozdnoukhov, A. 2018. A Generative Model of Urban Activities from Cellular Data. *IEEE Transactions on Intelligent Transportation Systems*. 19, 6 (Jun. 2018), 1682–1696.
 - 4) Lin, Y., Wan, H., Guo, S., Lin, Y.: Pre-training Context and Time Aware Location Embeddings from Spatial-Temporal Trajectories for User Next Location Prediction. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35 (5), 4241–4248. (2021).
 - 5) Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. ukasz and Polosukhin, I. 2017. Attention is All you Need. *Advances in Neural Information Processing Systems* (2017).
 - 6) Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *NAACL-HLT (1) 2019*: 4171–4186
 - 7) Radford, Alec, et al. "Language models are unsupervised multitask learners." *OpenAI blog* 1.8 (2019): 9.
 - 8) Brown, Tom, et al. "Language models are few-shot learners." *Advances in neural information processing systems* 33 (2020): 1877–1901
 - 9) 石井良治, 末成浩嗣, 越智健吾, 関信郎, 大塚賢太, 酒井幸輝, 會田優磨, 南川敦宣, 「携帯電話 GPS ビッグデータの都市交通分野における活用に向けた信頼性検証」, 土木計画学研究・講演集, vol. 58, 2018.
 - 10) 三瀬遼太郎, 井原雄人, 小林亮博, 上坂大輔, 南川敦宣, 森本章倫: 「将来の交通環境の変化が交通環境負荷に与える影響に関する研究」, 土木学会論文集 D3 (土木計画学), 78(to Appear), (2022)
 - 11) 吉羽崇, 小林亮博, 中管章浩, 南川敦宣, 富岡秀虎, 森本章倫. 「スマートフォン位置情報データを活用したバス需要予測に関する研究」. 土木学会論文集 D3 (土木計画学), 76(5), (2021)
 - 12) 小林直, 石塚宏紀, 南川敦宣, 村松茂樹, 小野智弘, 「携帯電話通信履歴に適した移動滞在状態推定手法の提案」, 情報処理学会論文誌データベース (TOD), Vol.10, No.1, pp.13–23, 2017.
 - 13) Kudo, Taku. "Subword Regularization: Improving Neural Network Translation Models with Multiple Subword Candidates." *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 2018.
 - 14) <https://huggingface.co/transformers/>
 - 15) Welleck, S., Kulikov, I., Roller, S., Dinan, E., Cho, K., & Weston, J.: Neural Text Generation With Unlikelihood Training. In *International Conference on Learning Representations*. (2019).

(XXXX.XX 受付)

Traffic Demand Generation Technology Using Location-Based Big Data

Akihiro KOBAYASHI, Daisuke KAMISAKA, Naoto TAKEDA,
Atsunori MINAMIKAWA and Akinori MORIMOTO

Attempts have begun to predict traffic demand by modeling large-scale location information (big data) acquired on smartphones using deep learning technology, which has been remarkably developed in recent years. However, most of them are based on RNNs, and it is difficult to predict behaviors by utilizing the long-term contexts of each user. In this study, we propose a training task setting method and a context-based behavior generation technique to apply the language model GPT-2, which is used for natural language generation, to human flow modeling. We modeled the human flow in Utsunomiya City based on actual smartphone location data and compared the generated OD table with observed data.