

アテンションリカレントネットワークを用いた人流データ解析の評価と考察

Evaluation and Discussion of Human Flow Data Analysis Using Attention Recurrent Networks

服部 まどか¹⁾ 来住 伸子¹⁾ 小口 正人²⁾ 中野 美由紀¹⁾
Madoka Hattori Nobuko Kishi Masato Oguchi Miyuki Nakano

1 要旨

近年、人流データは災害対応や観光、まちづくりなど様々な分野における活用が期待されている [1, 2, 3]. 本研究では、DeepMove: Predicting Human Mobility with Attentional Recurrent Networks[4] のアルゴリズムを用いた人流解析を行った. 異なる移動パターンや異なる行動をとるユーザーを検出することを目的に人工人流データを用いた実験を実施し、精度の評価及び人流制御への有効性について検討を行う.

人流データとはある場所から別の場所へ人が移動する動きをスマートフォンなどの機器を介して観測・収集された位置情報のことを指す [5]. 計測方法や集計加工方法によって、カウントデータ、滞留データ、OD データ、移動軌跡データの 4 種類に分けられる. 本研究では、ユーザーの移動経路を把握できる移動軌跡データを用いて人流解析を行なっている.

2 人流解析アルゴリズム

本研究では、先行研究のアルゴリズム [4] をベースに MacBookPro 上で GPU を用いて動作するように実装した.

図 1 に示すように、モデルは (1) マルチモーダル埋め込みモジュール (2) リカレントモジュールと履歴アテンションモジュール (3) 予測モジュールの 3 つの主要要素から構成されている. リカレントモジュールでは人間の移動の変遷の規則性を支配する複数の要因を捉えるために、一方で、履歴アテンションモジュールでは人間の移動の多段階周期性をモデル化するためにそれぞれ用いられている. これら 2 つのモジュールは並列に存在する.

現在の移動状態を表すクエリベクトルがリカレントモジュールで算出され、予測モジュールと履歴アテンションモジュールに渡される. 履歴アテンションモジュールでは、現在の移動状況に関連する過去の移動パターンを表す候補地ベクトルが生成され、クエリベクトルと比較して最も関連性の高いものを選択し予測モジュールに入力する.

3 人工人流データ作成

3.1 人工人流データ

人流データはプライバシー保護などの理由からオープンにはならないことが多い. そのため本研究は人工的に人流データを生成し、実験を行う. GitHub[6] 上に掲載されているサンプル人流データ (tweets_clean_sample.txt) を元に人工的な人流データを約 8000 個作成した.

サンプルデータの内容は位置情報 SNS, Foursquare API[7] から収集した数個の移動軌跡データとなっていた. タブで区切られた 9 つのデータが格納され、ユーザ

1) 津田塾大学 Tsuda University

2) お茶の水女子大学 Ochanomizu University

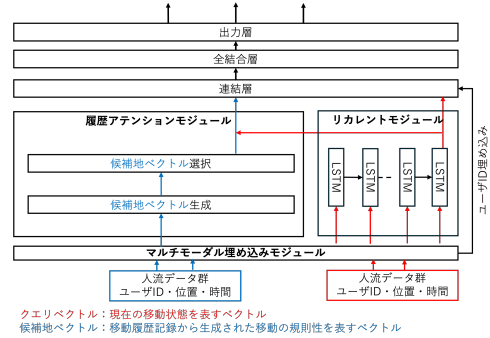


図 1 モデル概要

ID・緯度・経度・日付時刻・場所 ID が含まれていることを確認できた. 残りのデータは詳細不明であるが、モデルが学習に利用するのは前述の 5 つだけなので、人工データ作成の際には残りのデータには適当な文字列を作成し格納した. 具体的な作成方法を次の節で紹介する.

3.2 ユーザ設定

移動軌跡データは、誰が・いつ・どんな移動をしているのかという情報を含む. そのためデータを生成するにあたって、まずユーザーの基本情報そして平日・休日の移動設定を行なった. 具体的には、ユーザーが住んでいるエリアや職場のエリア、最寄り駅をはじめ、休日に訪れるスポットなどあらかじめ定めてからデータを作った. 本研究では著者を含む実在する人物を参考に 4 人のユーザーを作成し、その設定を行なった.

3.3 レコード生成

次に、レコードを生成する. レコードには、ユーザー ID、位置情報 (経度緯度, POI)、時間情報 (日付・時刻) などが含まれる. 作成したレコードの内訳は、ユーザーのモデルとなった人物達の実際の過去 2 ヶ月間の移動を基に手動で作ったレコードが 1032 個、平日と週末で移動のパターンを固定し、日付時刻はランダムに設定し自動で生成したレコードが 7080 個となっている. なお、後者の時刻についてはユーザーが移動を行う時間帯として最も相応しい時間帯 (例: 出退勤ならば朝と夕方) からランダムに設定した.

3.4 セッション生成

レコードをそのまま分析や予測に用いると不便なことが多い. 人の移動軌跡データはスパースであり、長時間にわたって断続的に記録されるので扱いづらいためである. そこで本研究ではセッションを作成し、レコードを分割し、扱いやすくする.

本研究において、セッションとは人の 1 日を想定した一連の移動の単位を指す. セッション作成によって、一貫した移動パターンが捉えやすくなり、異常値が排除さ

49180	2023_all	136_46392	35_17930	2023-01-01 09:51:54	Sun 11:16	rhdwv	4a_fe37
49180	2023_all	139_46598	35_72339	2023-01-01 21:48:10	Sun 11:16	rhdwv	4a_fe24
52448	2023_all	139_46598	35_72237	2023-01-02 07:14:29	Mon 12:02	eouq1	4a_fe1
52448	2023_all	139_46598	35_72237	2023-01-02 07:29:29	Mon 12:02	eouq1	4a_fe2
52448	2023_all	139_75284	35_68598	2023-01-02 08:14:29	Mon 12:02	eouq1	4a_fe3
52448	2023_all	139_75284	35_68598	2023-01-02 17:32:42	Mon 12:02	eouq1	4a_fe3
52448	2023_all	139_46598	35_72237	2023-01-02 18:32:42	Mon 12:02	eouq1	4a_fe1
52811	2023_all	139_46598	35_72237	2023-01-03 07:22:43	Tue 10:00	ftndp	4a_fe1
52811	2023_all	139_46598	35_72237	2023-01-03 07:22:43	Tue 10:00	ftndp	4a_fe2
52811	2023_all	139_75284	35_68598	2023-01-03 08:22:43	Tue 10:00	ftndp	4a_fe3
52811	2023_all	139_75284	35_68598	2023-01-03 19:53:32	Tue 10:00	ftndp	4a_fe3
52811	2023_all	139_46598	35_72237	2023-01-03 20:53:32	Tue 10:00	ftndp	4a_fe1
78328	2023_all	139_46598	35_72237	2023-01-04 08:33:35	Wed 10:00	trcha	4a_fe1
78328	2023_all	139_46598	35_72237	2023-01-04 08:33:35	Wed 10:00	trcha	4a_fe2
78328	2023_all	139_75284	35_68598	2023-01-04 09:33:35	Wed 10:00	trcha	4a_fe3
78328	2023_all	139_75284	35_68598	2023-01-04 17:42:31	Wed 10:00	trcha	4a_fe3
78328	2023_all	139_46598	35_72237	2023-01-04 18:42:31	Wed 10:00	trcha	4a_fe1
39580	2023_all	139_46598	35_72237	2023-01-05 07:55:53	Thu 10:00	odlma	4a_fe1
39580	2023_all	139_46598	35_72237	2023-01-05 07:55:53	Thu 10:00	odlma	4a_fe2
39580	2023_all	139_75284	35_68598	2023-01-05 08:48:53	Thu 10:00	odlma	4a_fe3
39580	2023_all	139_75284	35_68598	2023-01-05 17:32:12	Thu 10:00	odlma	4a_fe3

図 2 作成した人工人流データの一部

れることで効率的な学習と予測が可能になる。以下にレコードをセッションに格納する際のルールを示し、そのイメージを図 3 に示す。

① 1つのセッションに格納できるレコード数は最大 11 個

② 前後に並ぶレコードの時間間隔は 10 分以上必要

③ 前後に並ぶレコードの時間間隔が 4 日以上開くと、後者は新しいセッションに格納される

④ 1度だけしか行かなかった場所への移動レコードは格納されない

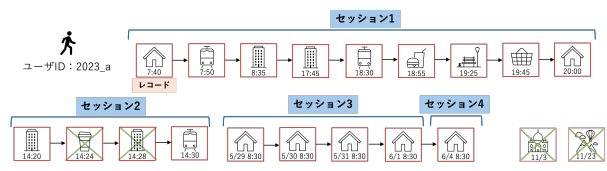


図 3 セッション生成

4 人工人流データ学習結果

作成したデータを用いて 2 つの実験を行なった。どの実験でもデータの 70 パーセントを訓練データ、残り 30 パーセントをテストデータとして用いた。なおデータ分割はセッション単位で行われている

4.1 実験①イレギュラーな移動を検出

実験①では移動パターンが異なるデータの学習結果を比較した。作成したデータは移動パターンによって固定とイレギュラーの 2 種類に分けられる。固定データとは、4.3 節において自動で作成した 7080 個のレコードを指す。平日は家⇄駅⇄職場の往復、週末はどこかへ出かけるという固定化された移動を行うデータである。

一方でイレギュラーデータは 4.3 節における手動で作成した実データに近い 1032 個のレコードを指す。基本の移動は固定データと同じであるが、出退勤前後の買い物や食事、平日の出張や遠方への旅行などイレギュラーな移動が多数含まれている。

4.2 実験②異なる移動をするユーザを検出

実験②では同一人物のデータで訓練・テストを行った場合と、訓練とテストでそれぞれ異なる人物のデータを用いて行なった場合の精度を比較した。

4.3 結果・考察

実験①②の結果をそれぞれ表 1 と表 2 に示す。実験①では固定データの方がイレギュラーデータよりも精度が高かった。実験②では同じユーザのデータで訓練・テ

トした場合の方が精度が圧倒的に高かった。

表 1 実験①の結果

移動パターン	セッション数(個)	精度
固定	556	0.8302
イレギュラー	95	0.6371

表 2 実験②の結果

訓練・テストデータ	セッション数(個)	精度
同じユーザ	556	0.8302
異なるユーザ	556	0.2664

実験①ではイレギュラーデータは固定データより精度が下がったものの、実験②の結果よりその差は小さかった。この理由としては、学習データ数が少ないことによる過学習が原因であると考えられる。今後はイレギュラーデータの数を増やして同様の実験を行いたい。いずれの結果からも精度を下がる事が確認できたので、このアルゴリズムを用いて予想されない行動をする人を探し出せることができると考えられる。

5 まとめ

本研究では、人工人流データを先行研究 [4] のモデルによって解析し、移動パターンが異なるデータの精度比較そして、訓練・テストで用いるデータを変化させた場合の精度比較を行なった。結果として、固定化された移動パターンデータの予測と訓練とテストで同じユーザのデータを用いた予測の精度が高かった。今後はアルゴリズムの改良の検討を行い、より大規模な人流データや実データ [8] を用いた解析に取り組みたい。今後の解析に用いる実データとして、2016 年に神戸で開かれた国際会議 International Semantic Web Conference 2016 [9] の参加者の移動履歴データを検討している。

参考文献

- [1] 藤井高森. 都市空間での人流解析における深層学習の応用. AI・データサイエンス論文集, Vol. 2, No. J2, pp. 113-120, 2021.
- [2] 岡山中心市街の人流分析. https://www.dbj.jp/topics/investigate/2024/html/20240411_204773.html.
- [3] 地域課題解決のための人流データ利活用の手引き. https://www.mlit.go.jp/report/press/tochi_fudousan_kensetsugyo17_hh_000001_00017.html.
- [4] Y.; Zhang C.; Sun F.; Meng F.; Guo A.; Feng, J.; Li and D. Jin. Deepmove: Predicting human mobility with attentional recurrent networks. *Proc. of WWW*, pp. 1459-68, 2018.
- [5] Kddi location analyzer. https://k-locationanalyzer.com/column/about_peoples_flow.
- [6] Jie Feng. Deepmove. <https://github.com/vonfeng/DeepMove>.
- [7] Foursquare. <https://foursquare.com/>.
- [8] 古崎晃司; 横山輝明; 深見嘉明; Gps 移動履歴の収集とオープンデータを用いた移動軌跡の lod 化 - 国際会議 iswc2016 における実証実験を例として - デジタルプラクティス, Vol. 9, No. 1, 2018.
- [9] International semantic web conference 2016. <https://iswc2016.semanticweb.org/index.html>.