

テクニカルノート

GPS データの複数属性を用いた LSTM ネットワークによる 移動状態のモデル化と推定

城内 光平^{1,a)} 江口 浩二^{1,b)}

受付日 2017年6月9日, 採録日 2017年8月6日

概要: 近年, GPS 移動軌跡データを分析することにより, 移動状態推定や人流予測を実現する試みが注目されつつある. この目的のもと, 従来は隠れマルコフモデル (HMM) などが用いられてきた. 本研究では, 再帰型ニューラルネットワーク (RNN) の一種で, 短期依存性だけでなく長期依存性をもとらえることが可能な長短期記憶ネットワーク (LSTM) に着目して, GPS 移動軌跡データの複数属性による移動状態推定を実現し, さらに実データを用いて有効性の評価を行う.

キーワード: 移動状態推定, GPS 移動軌跡, LSTM, 再帰型ニューラルネットワーク

Modeling and Estimating Movement States from Multiple Attributes of GPS Data Using LSTM Networks

KOUHEI SHIROUCHI^{1,a)} KOJI EGUCHI^{1,b)}

Received: June 9, 2017, Accepted: August 6, 2017

Abstract: The problems of moving state estimation and human flow prediction have been addressed such as by analyzing GPS movement trajectory data and have taken keen interest in these years. For these objectives, Hidden Markov Models (HMM) have often been used conventionally. In this paper, we focus on Long Short-Term Memory (LSTM) networks that were extended from Recursive Neural Networks (RNN) and can capture not only short-term dependency but also long-term dependency. We apply the LSTM networks to estimate the movement states from multiple attributes of GPS trajectory data and evaluate the effectiveness through experiments with real data.

Keywords: movement state estimation, GPS trajectories, LSTM, recursive neural networks.

1. はじめに

近年, ディープラーニングの技術を用いた研究に対する注目が高まっている. ディープラーニングの特徴は, 生のデータを与えるとモデルが自動的に特徴抽出から学習までをシームレスに行うモデルであるという点であり, 画像データや音声データ, GPS (Global Positioning System) データなどをはじめとした様々なデータへの応用が考えられている.

本研究では特に時系列で変化するデータ, とりわけ GPS

移動軌跡データに対してディープラーニングの観点において検証する. 時系列データを解析する研究としては, GPS 移動軌跡から速度変化を利用して旅行者の現在の状態を推測する研究 [1], 隠れマルコフモデル (HMM) やサポートベクトルマシン (SVM) を用いて GPS 移動軌跡データからどのような手段で移動しているのかを分類する研究 [2], [3], GPS の生のデータから条件付き確立場 (CRF) や SVM を用いて移動者の状態を予測する研究 [4] など, GPS 移動軌跡データを用いて状態を予測することには様々なニーズがあり, 様々な手法での研究がなされている. しかし, ディープラーニングの手法を GPS 移動軌跡データに適用し人流データを予測, 解析した研究はまだ十分に行われていない. 本研究の目的はディープラーニングの手法を用いて GPS 移動軌跡のデータを解析し, 状態を予測すること

¹ 神戸大学大学院システム情報学研究科
Graduate School of System Informatics, Kobe University,
Kobe, Hyogo 657-8501, Japan

^{a)} shirouchi@cs25.scitec.kobe-u.ac.jp

^{b)} eguchi@port.kobe-u.ac.jp

である。時系列データを扱うモデルとして再帰型ニューラルネットワーク (Recurrent neural network: RNN) [5], [9] がある。しかし, RNN では勾配消失問題により長い系列のデータを扱うことが容易ではない。長い系列のデータを扱うために RNN を拡張したモデルとして長短期記憶ネットワーク (Long Short-Term Memory: LSTM) [6] がある。本研究では LSTM ネットワークを用いて学習を行い, 移動者の状態の推定を行う。

2. 関連研究

本章では最も基本的な順伝播型ニューラルネットワーク (Feed-forward neural network: FFNN) [7] と, その時系列データのための拡張モデルにあたる RNN および, LSTM について述べる。

2.1 順伝播型ニューラルネットワーク

順伝播型ニューラルネットワーク (FFNN) は, 層状に並べたユニットが隣接層間でのみ結合した構造を持ち, 情報が入力側から出力側に 1 方向にのみ伝播するニューラルネットワークである。ネットワークを構成する各ユニットは複数の入力を受け取り, 1 つの出力を計算する。ここで, ユニットは 4 つの入力 x_1, x_2, x_3, x_4 を受け取るとすると, その総入力 u は各入力にそれぞれ異なる重みを掛けたものを加算し, バイアス b を加算したものとなる。

$$u = w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + b \quad (1)$$

ユニットの出力 z はユニットへの総入力 u に対する活性化関数と呼ばれる関数 f の出力となる。活性化関数にはシグモイド関数やソフトマックス関数などがよく用いられる。

2.2 再帰型ニューラルネットワーク

系列データは個々の要素が順序を持ち, その並びに意味が隠されているようなデータのことである。系列の並びが時刻の並びとなっている系列データを特別に時系列データと呼ぶ。系列データの例としては音声の波形, 動画, 文章 (単語列) などがある。それにより, 系列データの分類問題においてはある時刻, もしくは位置におけるデータの前後におけるデータとの関係性をうまく考慮することができれば, 結果が向上することが知られている。そして FFNN を系列データを扱うために拡張したモデルとして再帰型ニューラルネットワーク (RNN) がある。RNN における分類問題は入力系列 $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \dots$ から正解の系列 $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \dots$ を推定する問題として定式化することができる。このとき時刻 t における出力 \mathbf{y}^t はそれ以前の入力 $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \dots, \mathbf{x}^t$ の影響を受けていると考えることができる。

RNN は内部に有向な閉路を持つニューラルネットの総称である。RNN はこの構造により情報を一時的に記憶し,

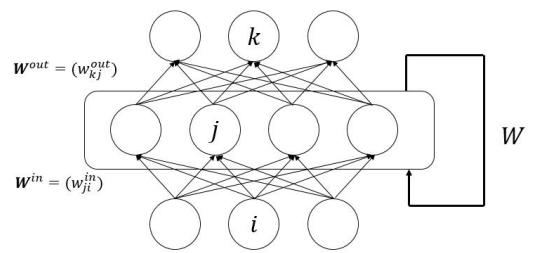


図 1 再帰型ニューラルネットワーク
Fig. 1 Recurrent neural network.

また振舞いを動的に変化させることができる。これにより, 系列データ中に存在する時間的な依存関係である「文脈」をとらえることができ, 系列データに対して有効に処理することができる。RNN のグラフィカルモデルを図 1 に示す。

図 1 にあるように RNN は FFNN と同様の構造を持ち, ただし中間層のユニットの出力が自分自身に戻される「帰還路」を持っている。RNN の動作は各時刻 t につき 1 つの入力 \mathbf{x}^t を受け取り, また同時に 1 つの出力 \mathbf{y}^t を返すというものである。RNN は理論上, 過去のすべての入力から 1 つの出力への写像を表現する。

系列データの分類問題においては, RNN の出力層は対象となるクラス数と同じ数のユニットを並べ, ソフトマックス関数を活性化関数に選ぶ。誤差関数は交差エントロピーを用いる。出力層 L の k 番目のユニットの出力はその下の層 $L-1$ 層からの入力を用いて以下のように定義される。

$$y_k \equiv z_k^{(L)} = \frac{\exp(u_k^{(L)})}{\sum_{j=1}^K \exp(u_j^{(L)})} \quad (2)$$

ここで, $u_k^{(L)}$ は出力層の 1 つ下の層からの入力である。

またこの出力 y_1, \dots, y_K は, 総和が 1 になることに注意が必要である。次に, 誤差関数は以下のように定義される。

$$E(\mathbf{w}) = \sum_n \sum_t \sum_k d_{nk}^t \log y_k^t(\mathbf{x}_n; \mathbf{w}) \quad (3)$$

ここで, d_{nk}^t は n 番目のサンプルの時刻 t での目標出力, y_k^t はそれと比較される RNN の出力, \mathbf{x}_n は n 番目のサンプルの入力系列である。

RNN では, 中間層に十分な数のユニットがあれば, 任意の系列から系列への写像を, 任意の精度で近似することができることが証明されている。

2.3 Long Short-Term Memory (LSTM)

RNN は系列データの文脈をとらえて推定を行うことができる。このとき, とらえることのできる文脈の長さ, すなわち現時刻からどれだけ遠い過去の入力を出力に反映させることができるかは重要である。しかし, RNN で実際に出力に反映することができるのはたかだか過去の 10 時刻分程度であるといわれている。この限界は勾配消失問題により生じている。層の数の多い深いネットワークにおい

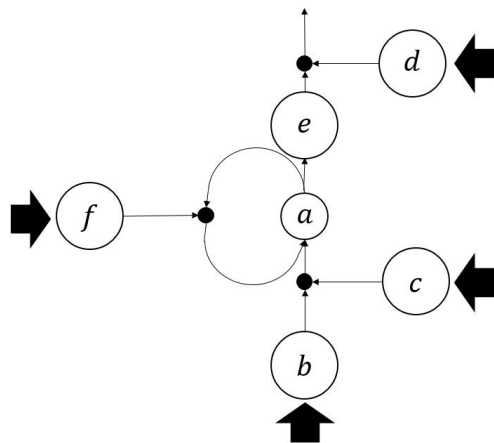


図 2 メモリユニット
Fig. 2 Memory unit.

て、誤差逆伝播法によって勾配を計算するとき層をさかのぼるにつれて勾配の値が爆発的に大きくなるか、あるいは 0 に消滅してしまう。長期にわたる記憶を実現するために提案されたモデルが長短期記憶ネットワーク (LSTM) である。LSTM では上で述べたような RNN に対し、その中間層の各ユニットをメモリユニットと呼ぶ要素で置き換えた構造を持っている。その他の構造は通常の RNN と同様である。メモリユニットの構造を図 2 に示す。

中央にメモリセル (図中記号 a) があり、その周囲に 5 つのユニットが配置されている。メモリセル a は状態 s_j^t を保持し、これを 1 時刻隔ててメモリセル自身に帰還させることで記憶を実現している。この帰還路には途中で忘却ゲートが挿入されており、ユニット f の出力がゲートの値 $g_j^{F,t} \in [0, 1]$ となる。 s_j^t に $g_j^{F,t}$ をかけたものが伝えられ、 $g_j^{F,t}$ の値が 1 に近ければ現状態がそのまま記憶され、0 に近ければリセット (忘却) される。

メモリユニットへの外部からの入力にはユニット b が受け取り、その出力がメモリセルに入力される。その間には入力ゲートがあり、ユニット c の出力がゲートの値 $g_j^{I,t} \in [0, 1]$ となる。ユニット b は通常の RNN の中間層のユニット 1 つに相当しており、これにゲートの値 $g_j^{I,t}$ をかけたものがメモリセルに伝えられる。

3. LSTM ネットワークによる時系列データの分類

LSTM ネットワークに対して GPS 移動軌跡より得られる種々のデータ (マルチモーダルデータ) を入力として与え、対応する状態ラベルを推定する。ここで、状態ラベルは入力として与えたあるデータと対応する移動状態のラベルを指す。GPS 移動軌跡は多くの外れ値や欠損値を持っており、また広範囲に分布しているため生のままの GPS 移動軌跡を用いるとスパース性が問題となる。外れ値や欠損値の影響を少なくするため GPS 移動軌跡から速度を計算

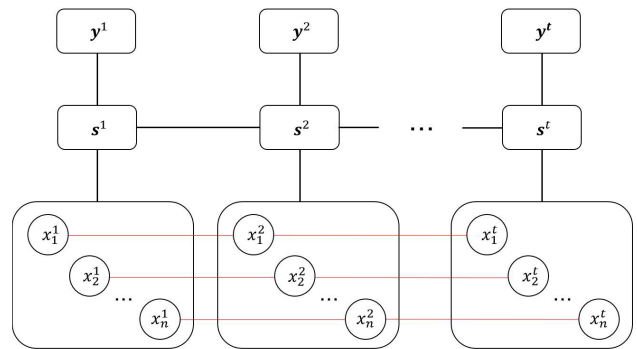


図 3 マルチモーダルデータのための LSTM ネットワーク
Fig. 3 LSTM network for multimodal data.

し、その系列を実験に用いた。また GPS 移動軌跡にはそのデータのユーザ ID や時刻印など位置情報以外にも複数の情報が含まれている。マルチモーダルな入力データを扱うことで推定をより正確にできることが期待される。実験で用いた LSTM ネットワークを図 3 に示す。

ここで、 \mathbf{x} は入力データであり、インデックス n は入力データの次元数を表している。入力データとしては速度やユーザ ID、位置情報や角度などが考えられ、たとえば速度とユーザ ID を入力とするならば $n = 2$ となる。 s^t は時間 t における中間層の状態であり、中間層はメモリユニットで構成されている。 y^t は時間 t における LSTM ネットワークの出力を表している。移動状態のラベル集合を $\mathcal{L} = \{l_1, l_2, \dots, l_\ell\}$ とすると目標出力は正解と対応するラベルの値のみが 1 をとり、その他の要素が 0 となる ℓ 次元のベクトルである。

4. 実験

本章では GPS データから求めた複数属性を用いた LSTM ネットワークによる状態推定の実験について述べる。複数属性を用いて状態の推定を行い、その精度を測定し、性能を評価する。

4.1 データセット

本研究では、Microsoft Research Asia Geolife project により集められた、GPS センサを持った人間の移動の GPS 軌跡に関するデータ*1 [4], [5], [6] を使用する。データはおおよそ 1-5 秒おき、もしくは 5-10 メートルおきに取得されている。データの内容としてはユーザ ID、経度、緯度、高度、取得時間の情報を持っている。また 1 部のデータにはラベルが付与されており、本実験ではラベルが付与されているデータを抽出し使用している。データの不均衡性を軽減するため、ラベルの種類として、車、自転車、徒歩、バスの 4 種類のみに着目した。また実験には、速度、ユーザ ID、観測された位置、出発地点からの角度の 4 つの属性を用いた。

*1 <http://research.microsoft.com/en-us/downloads/b16d359d-d164-469e-9fd4-daa38f2b2e13/>

表 1 実験結果
Table 1 Experiment results.

	Accuracy
LSTM (速度)	0.4676
LSTM (速度+ID)	0.5040
LSTM (速度+ID+角度)	0.5165
LSTM (速度+ID+位置)	0.4963
LSTM (速度+ID+角度+位置)	0.5010
ランダム	0.2500

4.2 パラメータ設定

実験は単一の間層を持つ LSTM ネットワークで行った。LSTM ネットワークに対して速度データ、あるいは速度データとユーザ ID を入力として与え、移動状態を推定しその正解率 (accuracy) を評価する。入力層は速度だけを用いる場合には、ユニット数が各タイムステップにおいて 1 であり、速度とユーザ ID を用いた場合にはユニット数は 2 となる。予備実験の結果として、中間層のユニット数は 200 とし、活性化関数は tanh とした。出力層はラベルの種類が 4 種類のためユニット数を 4、活性化関数をソフトマックス関数とした。誤差関数は交差エントロピーを用い、学習は確率的勾配降下法により行った。

4.3 結果

実験の結果を表 1 に示す。ここで、ランダムは 1–4 までの数字をランダムに生成し、その正解率について求めたものである。表 1 において、LSTM (速度) は LSTM ネットワークに対して速度の系列のみを与えて予測を行った結果を示しており、それぞれ、LSTM ネットワークに対して入力として用いた属性を示している。結果としては、LSTM ネットワークを用いて予測した結果、完全にランダムに予測した結果よりも良い結果を残すことができた。また、速度の系列のみを与えた場合よりも速度の系列に加えてユーザ ID を与えた場合により良い結果を残すことが確認される。さらに角度を与えると正解率が改善されることが確認できる。しかし、位置情報を含めて予測を行った場合に精度が少し下がっている。これはデータスパースネスに起因すると考えられる。

5. おわりに

本論文では再帰型ニューラルネットワークの拡張である LSTM ネットワークを用いて、GPS 移動軌跡データによる時系列で与えられた速度データを用いた状態推定を行った。実験においては、LSTM ネットワークを用いて予測を行った結果がランダムにラベルを選択した場合よりも精度が高い結果となった。また、速度だけではなく速度にユーザ ID を付与した場合の方がより高い精度を得た。これにより、LSTM ネットワークが速度データのパターンラベルによる違いを学習することができたことを示すことができ

た。また速度データとユーザ ID、さらには角度データをもとに用いたことで精度が上がったことは、今回のデータセットにおいては、人によるデータの数の違いやラベルの種類の違いが大きかったため、ユーザ ID を付与したことであるユーザ ID とラベルとの間の関係を学習したのではないかと考えられる。また、位置情報を加えると精度が下がったことは、位置を限定し、限られた区域の中のデータを用いることで改善することが考えられる。

謝辞 本研究の一部は科学研究費補助金基盤研究 (B) (15H02703) の援助による。

参考文献

- [1] Kasahara, H.: *Tourist Status Estimation from GPS Trajectory using Environmental Constraints*, The Japanese Society for Artificial Intelligence (2015).
- [2] Kim, Y.-J. and Cho, S.-B.: A HMM-Based Location Prediction Framework with Location Recognizer Combining k-Nearest Neighbor and Multiple Decision Trees, *Lecture Notes in Computer Science*, Vol.8073, pp.618–628 (2013).
- [3] Bolbol, A., Cheng, T., Tsapakis, I. and Haworth, J.: Inferring hybrid transportation modes from sparse GPS data using a moving window SVM classification, *Computers, Environment and Urban Systems*, pp.526–537 (2012).
- [4] Zheng, Y. and Xie, X.: Learning Transportation Mode from Raw GPS Data for Geographic Application on the Web, *WWW* (2008).
- [5] Zheng, Y., Li, Q., Chen, Y., Xie, X. and Ma, W.-Y.: Understanding Mobility Based on GPS Data, *Proc. ACM Conference on Ubiquitous Computing (UbiComp 2008)*, Seoul, Korea, pp.312–321, ACM Press (2008).
- [6] Zheng, Y., Xie, X. and Ma, W.-Y.: GeoLife: A Collaborative Social Networking Service among User, location and trajectory, *IEEE Data Engineering Bulletin*, Vol.33, No.2, pp.32–40 (2010).
- [7] Gers, F.A.: Learning to Forget: Continual Prediction with LSTM, *Neural Computation*, Vol.12, No.10, pp.2451–2471 (2000).
- [8] Goller, C. and Kuchler, A.: Learning task-dependent distributed representations by backpropagation through structure, *Neural Networks*, IEEE (1996).
- [9] Rosenblatt, F.: The perceptron: A probabilistic model for information storage and organization in the brain, *Psychological Review*, Vol.65, No.6, pp.386–408 (1958).
- [10] 岡谷貴之: 深層学習, 講談社 (2015).
- [11] Mikolov, T., Karafiat, M., Cernocky, J. “Honza” and Khudanpur, S.: Recurrent neural network based language model, *Proc. Interspeech* (2010).



城内 光平

平成 28 年神戸大学工学部情報知能工学科卒業。同年より、同大学大学院システム情報学研究科情報科学専攻博士前期課程に在学中。



江口 浩二

神戸大学大学院システム情報学研究科
准教授。博士（工学）。情報検索，統
計的機械学習，データマイニングの研
究に従事。

(担当編集委員 若宮 翔子)