

スマホ・アプリ型回遊調査データを用いた SVM による行動状態推定

熊本大学 学生会員 ○井村祥太郎

熊本大学 学生会員 濱澤憲駿
熊本大学 正会員 円山琢也

1. はじめに

2012 年度に熊本都市圏パーソントリップ調査 (PT 調査) と連携したスマートフォン (以下, スマホ) 調査, 2013 年度に PT 調査の補完として熊本都心部回遊調査が実施された. PT 調査は近年郵送型紙面調査によるものが主流であるが, スマホの普及によりプローブパーソン調査 (以下, PP 調査) の展開が期待されている.

この調査における基礎分析については野原¹⁾らによって行われている. しかしこの研究では加速度データ利用した分析は行われていない. 位置情報の記録が正確に取れていない店舗内での動きについて, 加速度から移動や滞在を判別することが出来れば, 店舗内の滞在についてはストップポイントとして見ることが出来るようになる. これが出来ればより詳細な回遊データの取得が可能となり, 空間設計などに有用な分析が可能となりうる.

そこで本研究では, 事後調査として行動状態を記録しながら行った再調査の結果を基に, 機械学習サポートベクトルマシン (以下, SVM) を用いて行動状態についての判定, 精度の検証及び GIS での可視化を試みる.

2. スマホ・アプリを用いた調査

調査用に開発を行ったスマホのアプリである「スマくま」は GPS による位置の記録と加速度を取得することが出来る.

(1) 第一回事後調査

2014 年 7 月に機械学習判定を行うためのデータ取得を行う目的でアプリを用いた調査を行った. Nexus 7 で 15 台, スマホ端末 3 台で位置情報及び加速度情報を取得. 記録を行う際に二人一組で行動, 一人が記録作業に専念しもう一人はアプリの入った端末を複数台持ち, まちなかを歩く. 記録する内容として, 回遊時の行動状態を状態ごとにフラグ付けを行う. 表 1 に調査概要, 表 2 にフラグ付けを行った項目を示す.

表 1 一回目事後調査概要

調査日	平成 26 年 7 月 17 日 (水) 晴れ
調査時間	午前 10 時～午後 12 時
調査エリア	熊本都心部 (上通り, 下通, 新市街, 交通センター)
調査対象	熊本大学学生 6 名

表 2 行動フラグ一覧

フラグ	行動種類
1	通常移動
2	店舗内移動
3	階段
4	エスカレーター
5	エレベーター
6	信号停止
7	休憩 椅子に座る
8	その他立ち止まり

(2) 第二回事後調査

2014 年 12 月 3 日に第二回目の事後調査を行った. 目的として, 一回目調査でタブレット端末の記録不具合などにおける教師データサンプル数の不足があったため同様の調査を Nexus 7 23 台, スマホ 4 台で行った.

表 3 二回目事後調査概要

調査日	平成 26 年 7 月 17 日 (水) 晴れ
調査時間	午後 2 時～午後 4 時
調査エリア	熊本都心部 (上通り, 下通, 新市街, 交通センター)
調査対象	熊本大学学生 8 名

3. 機械学習 SVM について

本章では移動滞在判定に用いる SVM について説明する. SVM は教師あり学習を用いる識別手法の一つであり, 1995 年に V. Vapnik によって提案された機械学習理論である. 識別能力が高い理由として 3 点挙げられる.

a) 学習データの中の, サポートベクトルと識別境界との距離であるマージンを最大化するように識別境界 (超平面) を構築し 2 クラス分類を行うため汎化能力が高い.

b) 線形識別できない問題においても、カーネルトリックを用いることによって、高次元特徴空間に写像することにより、特徴空間上での線形識別が可能となる。

c) 学習が二次計画問題に帰着されることから、局所最適解が広域最適解と必ず一致する。

このような特徴や工夫を用いることによって高い識別能力を有することになる。今回この SVM を用いるに当たって、フリーの統計ソフトである R の "e1071" パッケージを用いてデータを学習させ判定の精度を検証するものとする。

4. SVM による行動状態判定

本章では移動滞在判定に用いる SVM を用いた移動滞在判定の結果について示す。調査時にフラグ付けを行ったものから、通常の移動「move」(表 3 の 1, 6), 店舗内回遊時の移動「shop」(表 3 の 2, 3, 4, 5), 停止や休憩の「stop」(表 3 の 7, 8) の 3 つに分類し機械学習を行う。

(1) 教師データ

判定に用いるデータとして、12 月調査時に得られた 5 サンプルより 3 分類したものから均等かつランダムに 2000 データを取り出す。それらを組み合わせたものを今回の機械学習における教師データとして扱う。また同様に 10 サンプルから 3 分類の各 1000 データずつ取り出した教師データも用いて判定率の違いを確認する。判定するデータは 7 月調査時のデータを用いて精度の検証を行う。

(2) 判定結果

教師データサンプル数 5 と 10 のもので SVM の判定を行った結果を表 4, 5 に示す。判定率として店舗内回遊である shop に関してはどちらも低い判定結果となっている。stop に関しては高い判定率が出ている。結果が顕著に表れたのは move の判定率が 5 サンプルの教師データのものより 10 サンプルで学習した判定結果が高い結果となった。このような結果になった原因としてはサンプル数を増やしたことにより、加速度取得状況の均一化が図れたのではないかと考えられる。また図 1 に判定用データの stop 状況、図 2 に 10 サンプルの教師データで判定された stop 状態の密集具合を表したヒートマップを示す。ここでいうヒートマップとは点の

集積分布をカーネル密度関数で数値化し、集積の度合いを可視化したものである。集積が大きいと濃く、小さいと薄く表示される。通常時の移動も stop 判定されているところが多いあるためほぼ歩いた軌跡を描写してしまっているが、密度は小さくなっている。図 1 で示した実際立ち止まった箇所付近では stop が集中していることが確認され、滞在箇所の推定に繋がる可能性が示唆された。

表 4 5 サンプル判定率 表 5 10 サンプル判定率

行動状態	判定率(%)
move	18.34
shop	0.11
stop	98.89

行動状態	判定率(%)
move	52.96
shop	0.02
stop	97.07



図 1 stop データ

図 2 10 サンプル stop ヒートマップ

5. おわりに

今回教師データに用いるサンプル数を増やした際に判定率の向上がみられた。また GIS 上に可視化したことによって、上手く判定できていない箇所に関しても移動と滞在の違いを確認することが出来た。

今後の展開として、教師データの精度向上の更なるまとめ方の手法検討及び、データを用いた 2013 年度に行われた本調査への適応、誤判定データの修正方法の検討などを考えていく必要がある。

【参考文献】

- 1) 野原浩大朗, 福所誠也, 井村祥太郎, 円山琢也: スマホ・アプリを利用した熊本都心部回遊調査の分析, 土木計画学研究・講演集, Vol.49, 2014.
- 2) 井村祥太郎, 佐藤貴大, 円山琢也: スマホアプリ型回遊調査のデータ特性と SVM による移動滞在判別の基礎分析, 土木計画学研究・講演集, Vol.50, 2014
- 3) 金森敬文: R で学ぶデータサイエンスパターン認識 5, 共立出版, 2009
- 4) Christopher M. Bishop: パターン認識と機械学習 下, 丸善出版, 2008