

隠れマルコフモデルによる観光情報アプリユーザの移動距離と滞在時間を考慮したパターン推定

A Pattern Estimation considering Moving distance and Staying time of Tourism Information Application Users using Hidden Markov Model

佐藤 啓一[†] 多田 十兵衛[†] 長橋 陽介[†]
Keiichi Sato Jube Tada Yosuke Nagahashi

1. はじめに

本研究では、山形県を訪れる観光客を対象とした観光情報提供アプリケーションのユーザがとる行動パターン推定のため、滞在状況を説明変数として隠れマルコフモデル(HMM: Hidden Markov Model)を適用する。具体的には、滞在状況の異なるグループごとに代表ユーザを選出する。次いで当該ユーザにおけるトラッキングデータに含まれる緯度経度からなる位置情報とデータ取得日時より、山形県内各地を訪れた滞在時間と移動距離情報を抽出する。それらをHMMへの学習データとして適用し行動パターン推定モデルを生成する。その結果、各ユーザが各地を回遊した際の滞在時間と移動距離に見られる行動パターン推定を試みる。

2. 観光情報アプリ Wi-Fi YAMAGATA

Wi-Fi YAMAGATA は株式会社デジコンキューブが提供する携帯端末用観光情報提供アプリである [1]。アプリ自体は無料で入手および利用できる。主なサービスとして山形県を訪れる旅行客(以下、ユーザ)が県内各地を周遊する際に、各地に設置された Wi-Fi アクセスポイントへパスワード入力無しに自動接続して無料のインターネットを利用することができる。可能となる。また、本アプリを通じて現在位置周辺の観光情報の入手や観光施設で使える割引クーポンを入手してサービスを受けることができる。観光情報については提供元が開設した観光案内サイトを通じて入手できるようになる。本研究で用いるトラッキングデータはユーザが本アプリケーションを所有する携帯端末上で実行している間に収集される。

3. トラッキングデータ概要

分析対象のトラッキングデータはユーザ ID、データ取得日時、緯度経度等の複数項目のデータから構成される。それらの中で分析対象のデータ項目は表 1 に示すものである。ユーザ ID はランダムな英数字で付与され、日時は西暦、月日、および時刻から成り、本アプリ起動時に約 15 分毎に取得される。位置情報を示す緯度経度は分析の際に地域メッシュコード(以降、メッシュ)に変換して離散的な位置情報として用いる [2]。本研究では、2019 年 2 月の期間に 64 ユーザから取得された 42,156 件のデータを分析対象とする。

表 1. 分析対象のトラッキングデータ項目

項目	内容
ユーザID	・本アプリのユーザ識別用ID ・IDはランダムな英数字で表現(匿名化済)
日時	データの取得日時
緯度	データの取得位置(緯度, 十進角表記)
経度	データの取得位置(経度, 十進角表記)

4. 移動距離と滞在時間データの生成

前述のトラッキングデータを用いて HMM のパターン推定する。既存研究では、和田らはスマートフォンから過去に取得された位置情報を用いて HMM によるユーザの位置予測が報告されている [3]。本研究では位置情報(メッシュ)を目的変数とし、それに与える説明変数としてメッシュ毎の滞在時間とメッシュ間移動距離を新たなデータとして生成する。具体的には、1 日あたり 1 ユーザが訪れたメッシュ毎の最終取得データの送信日時から初回取得データの送信日時の差分を求めた時間を求めることで滞在時間データを得る。次いで、それに伴うメッシュ間の移動距離を 2 地点間のメッシュの各桁に対してユークリッド距離に基づき差分を求め、各差分に対して距離に応じた重みづけにより距離情報を得る。

以上の滞在時間および移動距離をユーザ毎に日別かつ時系列順に配置したデータ系列を用意し、それらに該当する位置情報と共に HMM によるパターン推定を行う。

5. クラスタリングによるユーザの分類

本研究では、ユーザの滞在時間と移動距離に基づくクラスタリングを行い、各クラスタ特有に見られる代表的な行動パターンを HMM 推定する。まず Ward 法による階層クラスタリングから得られる樹形図からクラスタ数を 4 とし、次に K-means 法による非階層クラスタリングを実施する。その結果を図 2 に示す。これは滞在時間と移動距離を両軸にとった散布図に各ユーザの合計の滞在時間と移動距離をプロットしたものをクラスタリングしたものである。この結果より、クラスタ C1 は移動距離および滞在時間共に最小であり、山形県内またはその近隣の県外地域に在住するユーザのグループと考えられる。クラスタ C2 および C4 はクラスタ C1 に比べ移動距離は同程度だが滞在時間が長い。従って山形県内またはその近隣の県外地域に在住するユーザの中で、宿泊を伴ったり定期的に山形県の各地を訪れるユーザのグループと考えられる。一方、クラスタ C4 については滞在時間がクラスタ C1 および C2 と同程度だが、移動距離が他のクラスタよりも長い傾向にある。以上より、山形県から比較的遠距離に在住するユーザで旅行目的で訪れた可

[†] 山形県立産業技術短期大学校 情報システム科
Yamagata College of Industry and Technology, Department of Information Systems

^{††} 山形大学大学院理工学研究科
Yamagata University Graduate School of

^{†††} 株式会社デジコンキューブ
Digiconcube Inc.

可能性が高い。

本研究では、各クラスタ内で両軸の第二四分位数に位置するユーザを当該クラスタの代表ユーザに選び、そのデータ系列を HMM 推定に用いる。

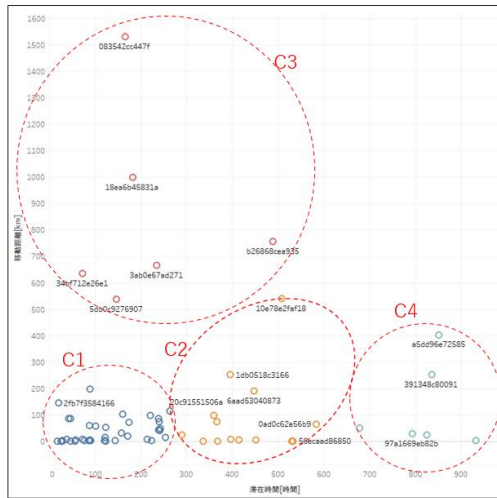


図 1 滞在時間と移動距離に基づくユーザの非階層クラスタリングの結果

6. HMM 推定

前述の各クラスタのセントロイドに該当するユーザの挙動を HMM により推定する。本研究では HMM に与えるパラメタである状態数を 3 とし、滞在時間と移動距離により得られる行動パターン抽出を試みる。図 2 に各クラスタのセントロイドユーザの HMM 推定結果を示す。

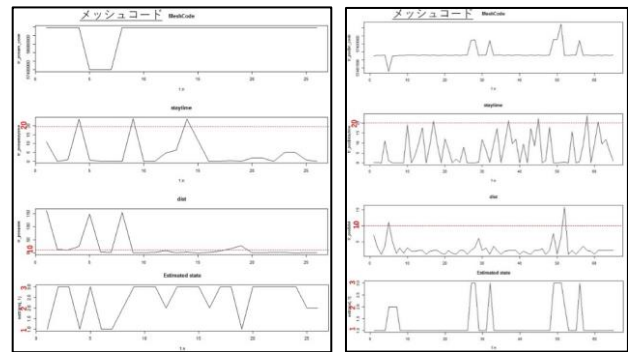
各図の構成として、X 軸はユーザが訪れたメッシュを 1 日単位で時系列順に並べたデータの取得順であり、上からメッシュ、滞在時間、移動距離、および状態遷移の変化を示している。

滞在時間の分布がほぼ同じ傾向にあるクラスタ C1, C2, および C3 で見ると比較的滞在時間の長いクラスタ C2 の滞在時間と移動距離の変動が大きい。また、それらの中で移動距離の長いクラスタ C3 については移動距離が 50[km]以上の頻度が多い(図中の赤線部)。一方、滞在時間が最も多いクラスタ C4 については移動距離が高々10[km]と比較的近距离間の移動に留まるが滞在時間が約 20 時間滞在する頻度が多い(図中の赤線部)。

次に状態遷移について見ると、明確に滞在時間と移動距離しているタイミングに連動して状態遷移していることが分かる。これについて、例えば状態 1 をユーザが拠点となる地域に滞在している状態、状態 3 を目的地に到達した状態、および状態 2 を拠点の周辺地域に滞在している状態とみなす。メッシュの変動が大きく、滞在時間と移動距離共に大きいタイミングで状態 1 から状態 3 へ、またはその逆の状態遷移が多く見られる。これに対し、メッシュの変動が小さく、滞在時間と移動距離のいずれかの変動が大きいタイミングで状態 1 と状態 2、状態 2 と状態 3 の間で状態遷移が確認できる。

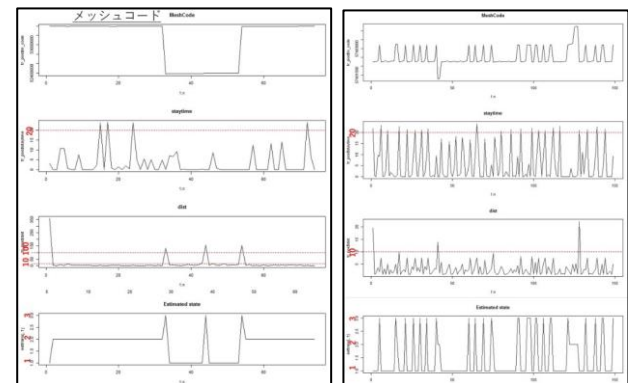
7. おわりに

本研究では、観光情報アプリから得られるトラッキングデータの位置情報について、それが示す地域ごとの滞在時間と移動距離に対して HMM によるユーザの行動パターン推定を行った。その結果、HMM の推定結果が示すパターンにはメッシュにおける移動距離と滞在時間に応じて状態遷移が決まったパターンとして観測可能であり、ユーザの挙動を推定可能であることを確認した。



a) クラスタ C1

b) クラスタ C2



c) クラスタ C3

d) クラスタ C4

図 2 各クラスタのセントロイドユーザの行動履歴における HMM のパターン推定結果

謝辞

本研究を着手するにあたり、トラッキングデータを提供頂きました株式会社デジコンキューブ様に感謝申し上げます。

参考文献

- [1] 山形県デジタルコンテンツ協議会 : Wi-FiYAMAGATA, https://www.archive.gr.jp/wifi/pdf/wifi_help.pdf (2021年2月18日にアクセス)
- [2] A.-H. Sato, S. Nishimura and H. Tsubaki, "World Grid Square codes: Definition and an example of World Grid Square data", 2017 IEEE Int. Conf. on Big Data (BIGDATA), pp. 4156-4165, Dec. 11-14, 2017
- [3] 和田 計也, 増田 早紀, 梅林 泰孝: 行動クラスタリングと隠れマルコフモデルを用いたユーザの位置予測モデル, 人工知能学会全国大会論文集 29, pp.1-4, (2015)