

Twitter を利用した岡山における観光ルート推薦に関する検討

Examination of tour recommendation in Okayama using Twitter

中川 智也

Tomoya Nakagawa

岡山大学 太田研究室

Ohta Laboratory, Okayama University

概要 今日, Twitter に代表されるマイクロブログの普及に伴い, 一般ユーザが体験した事, 感じた事を簡単に発信できるようになった. その中には, 観光客が観光スポットで見たり, 感じたりしたことを投稿したものがあある. 本稿では岡山県内のスポットを収集し, そのスポットが観光スポットとして適当かどうか判定した. さらに, 観光スポットについてのツイートを収集, 分析し, それを利用した観光ルート推薦を行った.

1 はじめに

近年, マイクロブログと呼ばれる短い文章で, ユーザの身近で起こった出来事を発信するサービスの利用者が増えている. Twitter^{*1}はマイクロブログの一種で, ツイートと呼ばれる 140 文字以内の文章を発信できるサービスである. 2016 年 1 月現在, 世界で 3 億 2000 万人が利用しており [1], ツイートの中には何気ない発言からニュース速報に至るまで幅広い情報が含まれている. また, Foursquare^{*2}のように Twitter と連携して自分の位置情報を共有できるサービスや, Instagram^{*3}のような画像共有サービスも登場し, ユーザの位置情報や体験情報が発信されている.

新井ら [2] は, ツイートから各観光スポットがどの時間帯に訪れられやすいか, どのような目的で訪れられているかをスコア化し, そのスコアを使って観光ルートを推薦した.

本稿では, 新井らの手法を用いて岡山の観光スポットの収集, ツイートの収集と分析をし, DEIM2016 で提案した手法 [3] を用いて観光ルート推薦を行った.

2 観光スポットの収集

2.1 岡山県内のスポットの収集

本節では岡山県内のスポットを収集手法を説明する. スポットの収集には Google Places API^{*4}を用いる. Google Places API はある一点の経緯度を入力すると, その周辺のスポット群を返す. 本研究では得られたスポットの周辺のスポットを新たに収集し, 新しいスポットが出現しなくなるまでスポットの収集を繰り返す.

2.2 観光スポットの選定

2.1 節で収集したスポットには, 病院や学校など観光スポットとして適当ではないスポットも含まれている. そこで, 本研究では Yahoo!知恵袋質問検索 API^{*5} を用いて観光スポットかどうか判定する. Yahoo!知恵袋質問検索 API はカテゴリを指定して質問を検索できる. そこで, スポット名で検索したときのすべてのカテゴリの質問数に対する, 「地域, 旅行, おでかけ」のカテゴリの下の「国内」カテゴリにおいてスポット名で検索したときの質問数の割合が, 閾値以上のものを観光スポット候補とする. この閾値は新井ら [2] と同じ 0.27 を用いた.

3 観光ツイートの収集

次に Google Places API と Yahoo!知恵袋質問検索 API を用いて収集した観光スポット名を含むツイートを Twitter API^{*6}を使って収集する. この時に, 位置情報共有サービスの URL を含んでいるか, 画像が添付されているか, 実際に訪れたとみなせる表現がされているか, などの特徴も収集する. この情報を使うことで実際の観光体験に基づくツイートかどうか判定する. またツイートの投稿時刻も記録する. この時刻情報は後述する各観光スポットの観光所要時間と時間帯スコアを算出するために必要となる.

4 観光ツイートの集計

観光と関係するツイートを収集することで, それぞれの観光スポットの平均的な観光所要時間を推定することができる. 具体的には実際の観光体験に基づく判断されたツイートの投稿時刻, そのユーザのタイムラインで次に訪れたスポットでされたと思われるツイートの投稿時刻及びそのスポット間の移動時間から滞在時間を推定する.

さらに, 観光に関連したツイートを収集すると, 観光スポットの良さを評価する次の 3 つのスコアを計算することもできる [2].

- 時間帯スコア・・・全ツイート数に対する 3 時間ごとの各時間帯におけるツイート数の割合

*1 Twitter, <http://twitter.com/>

*2 Foursquare, <https://ja.foursquare.com/>

*3 Instagram, <http://instagram.com/>

*4 Google Places API, <https://developers.google.com/places/>

*5 Yahoo!知恵袋質問検索 API, <http://developer.yahoo.co.jp/webapi/chiebukuro/chiebukuro/v1/questionsearch.html>

*6 Twitter API, <https://apps.twitter.com/>

表 1: 市町村別のスポット数 (S_1), API で判定した観光スポット数 (S_2), 人手で判定した観光スポット数 (S_3)

市町村名	S_1	S_2	S_3
岡山市	37455	179	37
倉敷市	20610	85	22
津山市	6342	26	7
真庭市	2902	36	18
玉野市	2698	18	5
その他	23378	156	65
計	93385	500	154

- カテゴリスコア・・・ツイートを「食事」, 「景観」, 「土産」, 「行動」の4つのカテゴリに分類した時の割合
- 共起スコア・・・ツイートした同じユーザのタイムラインに出てくる他の観光スポットの共起確率

5 観光ルートの生成

収集したツイートから、それぞれの観光スポットの良さの評価値や平均滞在時間を求めると、スコアの合計がより大きくなるような観光ルートを計算することができる。本研究では出発地、到着地、出発時刻、到着時刻、最も訪れたいスポットを指定し、この指定された条件の中でより良いスコアとなる観光ルートを求める実験を行った。

6 評価実験

6.1 観光スポット収集

2.1 節で示した方法でスポットの収集を行った場合の評価実験を行った。岡山駅を最初のスポットとして収集した岡山県内のスポット数 (S_1), 2.2 節で説明した API を用いて観光スポット候補と判定したスポット数 (S_2), S_2 のスポット候補からさらに人手で観光スポットを判定したスポット数 (S_3) を市町村別に表 1 に示す。

表 1 に示す通り、API を用いて観光スポットと判定した 500 件を人手で観光スポットかどうか判定すると、適合率は 0.31 と低く、正確な選定はできていなかった。そのため以下のルート推薦の実験では人手で判定した 154 スポットを実験に用いる。具体的には、この 154 スポットに対して 3 節で説明した方法で収集した 2016 年 1 月 1 日から 2016 年 6 月 30 日の期間に投稿されたツイートを用いて滞在時間やスコアを計算する。この計算結果を用いてルート推薦を行う。

6.2 ルート推薦

10:00 岡山駅出発, 17:00 岡山市内ホテル到着, 最も訪れたいスポットを岡山後楽園, 移動手段を自動車と



図 1: 推薦プログラムが推薦したルート

表 2: 推薦プログラムが推薦したルート及び出発時刻, 到着時刻, 滞在時刻

スポット名	出発時刻	到着時刻	滞在時間
A:岡山駅		10:00	
B:岡山城	10:08	11:10	62
C:岡山後楽園	11:15	12:36	81
D:石山公園	12:40	13:21	41
E:吉備津彦神社	13:46	14:46	60
F:吉備津神社	14:51	16:36	105
G:岡山市内ホテル	17:00		

して推薦したルートを図 1, 表 2 に示す。岡山後楽園以外に岡山後楽園と隣接している岡山城や、桃太郎のモデルとなった吉備津彦命を祀る吉備津神社など歴史を感じられるスポットが推薦された。

7 考察

観光スポット選定において 2.2 節で説明した手法で良い結果を得られなかったのは、岡山県内にある「清水寺」のように、選定するスポットと同じ名称の観光スポットが他県に存在しており、この手法では区別できなかったことが一因と考えられる。

8 まとめ

本研究では、岡山県内のスポットを収集し、スポットの中から観光スポットとして適当なものを選別した。さらに、判定した観光スポットを用いて観光ルート推薦を行った。

参考文献

- [1] Twitter, Inc. について, <http://about.twitter.com/ja/company>
- [2] 新井晃平, 新妻弘崇, 太田学, “Twitter を利用した観光ルート推薦の一手法”, 第 7 回データ工学と情報マネジメントに関するフォーラム (DEIM2015), G7-6, pp. 1-8, 2015.
- [3] 中川智也, 新妻弘崇, 太田学, “マイクロブログを利用した観光ルート推薦における移動効率の改善”, 第 8 回データ工学と情報マネジメントに関するフォーラム (DEIM2016), H1-3, pp. 1-8, 2016.