

Twitter を用いた潜在的所望スポットの提示による観光ルート推薦の検討 Examination of tour recommendation by presenting potential desired spots using Twitter

竹下 知宏

Chihiro Takeshita

岡山大学 太田研究室

Ohta Laboratory, Okayama University

概要 本稿では、Twitter で観光スポットに関するツイートが投稿されていることに着目する。各観光スポットに関するツイートを 4 つのジャンルに分類し、各観光スポットにおけるジャンルを決定する。そして、ジャンルごとにそのジャンルのツイート数が多い上位 10 件をおすすめスポットとして表示し、ユーザに情報の提供を行うシステムを提案する。さらに、おすすめスポットの中から 3 つまで選択できるようにし、それらが組み込まれた観光ルートの生成を検討する。

1 はじめに

近年、インターネットの利用率は 8 割を超え^{*1}、多くの人がウェブ上の情報にアクセスするようになった。特に Twitter のようなソーシャルネットワークワーキングサービスには多数の情報が投稿されている。その結果、インターネット上で様々な知識や経験、価値観を共有できるようになった。例えば図 1 のように観光に関係する知識を共有するツイートなどが多数 Twitter には存在している。

観光客が観光するスポットを決める際、歴史的な城や寺社仏閣などの特定のジャンルだけでなく、食事をしたり、お土産を買ったりするスポットを選出することが考えられる。また観光客は、初めて行く地域ではその地域に対する知識がなく、その周辺で人気の観光スポットが何であるかわからない場合も想定される。

そこで本稿では、Twitter から観光スポットに関するツイートを取得し、「食事」、「景観」、「行動」、「土産」の 4 つのジャンルに分類を行い、各観光スポットのジャンルを決め、各ジャンルで人気のスポットをおすすめスポットとして提示することでユーザの潜在的なニーズに応えた観光プランの立案の支援を行う。

本稿ではまず 2 節でおすすめスポットを提示する手法について述べ、3 節でそれを導入した観光ルート推薦の手法を説明する。そして 4 節で実際にシステムを使用した結果について記述し、5 節で考察を記す。

2 おすすめスポットの提示

本節では、本稿で提案する観光スポットのジャンルの分類方法を説明する。提案手法では、最初にツイートのジャンル分類を行い、次にその結果を利用して観光スポットのジャンルの分類を行う。

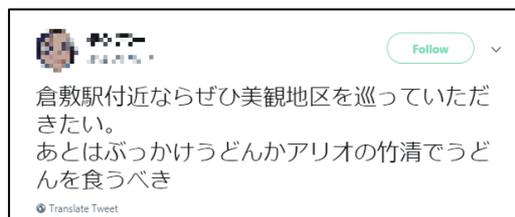


図 1：観光に関するツイートの 1 例

2.1 ツイートのジャンル分類

ツイートを、手掛かり語が出現するかどうかに基づいて「食事」、「景観」、「行動」、「土産」の 4 つのジャンルに分類する。手掛かり語とは事前に Wikipedia と Weblio 類語辞典から収集したものである。例えばジャンル「食事」では「から揚げ」や「コロケ」などを手掛かり語として利用する。

2.2 観光スポットのジャンル分類

次に各観光スポットのジャンル決定を各ジャンルのツイート数を用いて行う。例えば、観光スポット「岡山後楽園」の場合、ジャンル「食事」に関するツイートの数は 284、ジャンル「景観」に関するツイートの数は 1361、ジャンル「行動」に関するツイートの数は 250、ジャンル「土産」に関するツイートの数は 30 となる。このとき、「岡山後楽園」のジャンルはツイート数が最も多い「景観」となる。このようにツイートの多数決で観光スポットのジャンルを決定する。

2.3 おすすめスポットの表示

おすすめスポットはジャンルごとにそのジャンルのツイート数が多い上位 10 件を表示させる。例えば、観光スポット「岡山後楽園」はジャンル「景観」で、そのジャンルとなるスポット内でジャンル「景観」に関するツイートの数が 9 番目に多い。このことから「岡山後楽園」はおすすめスポットのジャンル「景観」の 9 番目に表示される。なお、ジャンルのスポットが 10 件に満たない場合はそのジャンルのスポットを全件表示させることにする。

3 観光ルート推薦

本稿では前節で作成したおすすめスポットから 3 つ選択しそれを実際の観光ルートに組み込む方法を提案する。詳細は 3.2 節で述べる。また選択したスポット以外のスポットも組み込み観光ルートの生成を行う。詳細は 3.3 節で述べる。

^{*1}<http://www.soumu.go.jp/johotsusintokei/statistics/statistics05.html>

3.1 初期入力

おすすめスポットを提示する前にユーザはルート生成に必要な情報（出発地、出発時刻、到着地、到着時間）の入力を行う。さらに、ユーザは観光するにあたって観光の目的となるスポットが1つはあることが想定されるため、訪れたい観光スポットを1つ入力できるようにする。これを選択スポットと呼び、選択スポットのジャンルを興味ジャンルと呼ぶ。

3.2 おすすめスポットの選択

先の入力が終了した後、前節で作成したジャンルごとのおすすめスポットを表示しその中から3つまで選択できるようにする。行きたいスポットがその中に無い場合には選択しなくてもよいことにする。そしてそれをルートに組み込む。

3.3 おすすめスポット以外のスポットの導入

おすすめスポット内からスポットを選択したあとで設定した観光時間に他のスポットを入れる余裕がある場合には以下の3つのスコアの和が最大となる観光スポットを探し、導入する。

- ▶ 時間帯スコア
1日を8等分した3時間ごとの各時間帯におけるツイート数の割合で、その観光スポットを訪れるのに適した時間帯を表すスコアである。
- ▶ 共起スコア
ツイートした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で、ルート内のスポットと共に訪れられやすいスポットを表すスコアである。
- ▶ ジャンルスコア
2.1節で述べた手法でツイートを4つのジャンルに分類したときの各ジャンルの割合である。興味ジャンルの割合が高いほどスコアが高くなる。

4 実験

本節では提案したシステムの実行結果を示す。出発地、到着地を岡山駅、出発時間、到着時間をそれぞれ10時、19時とし、選択スポットを岡山後楽園にしたところ、表1のようなおすすめスポット一覧が生成された。また、表1の一覧の中から文の助茶屋、岡山城、大原美術館を選択したところ、表2のような観光ルートが生成された。

5 考察

本節では提案したシステムを動かした結果の考察について述べる。

5.1 おすすめスポットの提示の導入

今回作成したおすすめスポット一覧はジャンル別でスポット名を表示しただけのものである。スポット名でどういふ場所であるかを認知できるものもあるが、多くはそれができないためユーザはそのスポットについて調べなければいけなくなる。この手間を省くためにはおすすめスポット

表1：おすすめスポット一覧

	食事	景観	行動	土産
1	石川屋	清水寺	テラー	お土産センター
2	伊勢屋	美観地区	鷺羽山ハイランド	
3	ひさご	貴船	大原美術館	
4	浜屋	藤公園	明治池	
5	さくら亭	白糸の滝	宝鏡寺	
6	そばの館	鷺羽山		
7	ひるぜん	岡山城		
8	チロリン村	花やしき		
9	吉祥院	諏訪神社		
10	文の助茶屋	千光寺		

表2：生成された観光ルート

START(岡山駅)	10:00発	
大原美術館	10:43着	12:02発
岡山後楽園	12:44着	13:48発
岡山城	13:55着	15:30発
吉備津彦神社	15:57着	16:38発
文の助茶屋	17:04着	18:52発
GOAL(岡山駅)	19:00着	

にスポット名だけでなく出発地、到着地からの距離やスポットの説明文などの具体的な情報を付与する必要がある。

また、おすすめスポット「景観」の一番先頭には「清水寺」というスポットが表示されるが、データを見る限りこれは岡山県にある「清水寺」ではなく京都の「清水寺」についてのツイートであることが推測される。このような曖昧性についても中川らの研究[1]を参考にすることで回避することを検討しなければならない。

5.2 観光ルート推薦の導入

提案する観光ルート推薦システムでは、おすすめスポットから3つまで選択し、それを観光ルートに組み込み、ルートを生成することができたが、3つ選択したときに選択したスポットが全部ルートに組み込まれない場合もあった。これは観光する時間の都合上選択したスポットを入れられなかったと考えられる。対策として、スポットがすべて入らないとわかった段階で警告文を出したり、再度選択してもらったりすることなどの考慮が必要となる。

また、近くにあるスポット同士が同一ルートに含まれている場合にそれらを連続で行かずに間に距離の離れたスポットが入るといふ大変非効率なルートを生成してしまうことがあった。この問題を解決するためにもスポット間の距離を考慮した新たなスコアを設けるなどして対策を行わなければならない。

6 まとめ

本報告ではおすすめスポットを提示し、ユーザに情報を与え、それを踏まえたうえで観光ルート推薦を行うシステムを提案した。今後は5節で述べた問題点や解決策を熟考し、改善することでよりユーザの満足度の高い観光ルート推薦を目指していく。

参考文献

- [1] 中川智也, 新妻弘崇, 太田学, “Yahoo!知恵袋を利用した施設名の曖昧性解消手法の提案”, 第9回データ工学と情報マネジメントに関するフォーラム, C4-4, pp. 2-3, 2017.