

ユーザが入力した地点の近傍を優先する観光ルート推薦の一手法

中野 翔子[†] 新妻 弘崇^{††} 太田 学^{†††}

[†] 岡山大学工学部情報系学科 〒 700-8530 岡山県岡山市北区津島中 3-1-1

^{††} 大阪大学データビリティフロンティア機構 〒 565-0871 大阪府吹田市山田丘 2-8

^{†††} 岡山大学大学院自然科学研究科 〒 700-8530 岡山県岡山市北区津島中 3-1-1

E-mail: [†]p7tv0307@s.okayama-u.ac.jp, ^{††}niitsuma@ids.osaka-u.ac.jp, ^{†††}ohta@cs.okayama-u.ac.jp

あらまし 本研究ではマイクロブログに投稿された観光体験情報を利用した観光ルート推薦を行う。新井らは Twitter に投稿された観光体験情報から観光スポットの複数の評価スコアを算出し、それらを用いて観光ルートを推薦する手法を提案した。しかし、スコアの大きい観光スポットを優先的に観光ルートに含めるために移動距離が長くなるという問題があった。これを解決するために、本稿ではユーザが入力した観光ルートの出発地と到着地、およびユーザが訪問先として選択した観光スポットの計 3 地点の近傍の観光スポットを優先した観光ルート推薦手法を提案し、移動効率を改善する。さらに、提案手法はユーザの嗜好に応じて移動効率と寄り道のバランスが調整できることを実験により示す。

キーワード マイクロブログ, Twitter, 観光情報, ルート推薦

1 はじめに

マイクロブログとは、短い文章でユーザの身近に起こった出来事などを発信することができるサービスである。Twitter¹はマイクロブログの一種で、ツイートと呼ばれる 140 字以内の文章を投稿できる。投稿されたツイートの中には、ユーザが観光に訪れた際に体験した内容や景観の感想などの観光体験情報も含まれている。本研究では、この観光体験情報から算出したスコアに基づいて観光スポットを評価し、その評価を元にした観光ルートを推薦する。

新井ら [1] は Twitter に投稿された観光体験情報を含むツイートを収集し、観光スポットを訪れる時間帯の適切さを表すスコアや観光スポット同士の共に訪れやすさを表すスコアなどの評価スコアを算出した。さらに観光スポットのこれらの評価スコアを利用して、新井らは観光ルートを推薦する手法を提案した。しかし、新井らのルート推薦手法は、スコアの大きい観光スポットを優先的に観光ルートに含めるために、移動距離が長くなるという問題があった。

この問題を解決するため、金川ら [2] は観光スポットの評価スコアよりも移動距離を優先して観光ルートを生成する手法を提案した。具体的には、ユーザが訪問先として選択した観光スポットから一定の時間内に移動できる観光スポットのみを利用して観光ルートを推薦した。しかし、金川らの手法では、ユーザの入力条件によっては移動距離が長い観光ルートが推薦されたり、観光地によってパラメータの調整が必要であったりするという問題があった。

そこで本稿では、選択スポットに加えてユーザが入力した出発地と到着地の計 3 地点の近傍の観光スポットを優先してルートを生成する手法を提案し、これにより推薦ルートの移動距離

のさらなる短縮を図る。さらに、ユーザの意図に沿って移動効率を調整できる観光ルートを推薦する。

本稿の構成は次の通りである。まず 2 節で関連研究について述べる。3 節で観光スポットと観光ツイートの収集方法について述べ、4 節で観光ツイートから算出する観光スポットの評価スコアと観光ルートの評価値について述べる。5 節で観光ルート推薦システムの概要と提案手法を説明し、6 節で評価実験について述べる。最後に 7 節でまとめる。

2 関連研究

中嶋ら [3] は、旅行者のツイートで頻繁に用いられるツイートの特徴と Foursquare²や Instagram³のサービスを利用して観光体験情報を含むツイートを収集した。さらに、収集したツイートのユーザのタイムラインから観光ルートを抽出し、手がかり語や品詞の特徴から「食事」、「景観」、「行動」の 3 つのカテゴリへツイートを分類し、その結果を用いて観光ルートを推薦する手法を提案した。しかし、中嶋らの手法では、ユーザのタイムラインからのみ観光スポットを取得するため、得られる観光スポットが少ないという問題があった。

この問題を解決するために、新井ら [1] は Google Places API⁴を用いて観光スポットの候補となる施設を収集し、Yahoo! 知恵袋 質問検索 API によりそれらの施設が観光スポットとして適切であるかを判定した。また、収集した観光スポット名を手がかり語として観光体験情報を含むツイートを収集し、それらのツイートから各観光スポットについての 3 種類のスコアを求め、これらのスコアの和が最大となる観光ルートを推薦する手法を提案した。この手法では観光スポットのスコアを利用す

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

3 : Instagram, <https://www.instagram.com/>

4 : Google Places API, <https://cloud.google.com/maps-platform/places/>

るため、観光スポットを訪問する時間帯の適切さなどを考慮した観光ルートを推薦することができたが、スコアの高い観光スポットが離れた位置に存在する場合に移動効率の悪い観光ルートが推薦されるという問題があった。

中川ら [4] は観光ルートの移動効率の改善を目的に、観光スポットのスコアとルート生成アルゴリズムを改良した。具体的には、共に訪れやすい観光スポットを表す共起スコアを改良し、観光スポットを訪問する順番が近いほどそれらのスポットの共起スコアが大きくなるようにした。また遺伝的アルゴリズムを利用した推薦ルート生成を提案した。

松本ら [5] は写真共有サービス Flickr⁵から取得した写真データに付与された位置情報を利用し、ユーザが指定した観光時間内に観光可能な観光ルートを推薦する手法を提案した。この手法では、まずユーザが指定した出発地から一番近い観光スポットを選択し、さらに選択した観光スポットから一番近い観光スポットを選択する。これをユーザが指定した観光スポットを全て含むまで繰り返すことで最適ルートを生成する。最適ルートに必要な観光時間を算出した後に、それがユーザの指定した観光時間よりも短ければ、写真データの情報から最適ルートに含まれる観光スポットと共起回数が高い観光スポットを算出し、それらの観光スポットを上限の観光時間に達するまでルートに追加することで、推薦ルートを生成する。

中里ら [6] は Twitter に投稿された位置情報付きツイートを利用して、イベントに参加しているユーザがイベントの前後によく訪れるスポットを抽出する手法を提案した。スポットの抽出はツイート数とユーザ数を基準に行い、イベント参加者が多く集まるスポットや時間帯などのユーザ行動を把握した。

3 観光スポットと観光ツイートの収集

本節では新井ら [1] により提案された観光スポットの収集方法と収集した観光スポットを用いた観光体験情報を含むツイートの収集方法について説明する。

3.1 観光スポットの収集

本研究での観光スポットの収集は、新井ら [1] の方法を利用する。まず観光スポットの収集の中心となる任意の地点を定め、その地点の周辺の施設の情報を Google Places API を用いて 20 件取得する。さらにその取得した施設の周辺の施設も同様に取得していく、新しい施設が出現しなくなるまで取得を続ける。ここで取得した施設には、病院など観光スポットとして適当でない施設も含まれるため、Yahoo!知恵袋の質問を用いて、観光スポットとして適切な施設かどうかを判定する。具体的には、収集した施設の名称が Yahoo!知恵袋の「地域、旅行、おでかけ」カテゴリ下にある「国内」カテゴリでの出現回数が閾値以上の施設を観光スポットの名称と判定する。

3.2 観光ツイートの収集

3.1 節の手順で収集した観光スポット名を手がかり語とする

5 : Flickr, <https://www.flickr.com/>

ことで、各観光スポットについて観光体験情報を含むツイートを収集する。具体的には Twitter API⁶を用いて、以下の手順でツイートを収集する。

(1) 観光スポット名と収集した観光スポットが存在する県名を含むツイートの投稿ユーザの ID を取得する。この際、取得日から最大 10 日間のユーザ ID を取得することができる。

(2) 取得したユーザ ID から 1 ユーザあたり最大 3,200 件のツイートを取得する。

(3) 取得したツイートから本文に観光スポット名と県名を含むものを抽出し、ツイートの本文や投稿日時等の情報を得る。

なお、手順 (3) においてツイートが得られなかった観光スポットは、本研究の観光ルート推薦には利用しない。

4 観光スポットとルートの評価

本節では、3 節の方法で収集した観光体験情報を含むツイートから算出する観光スポットの評価スコアについて説明する。次に評価スコアを用いて算出する観光ルートの評価値について説明する。

4.1 観光スポットの評価スコア

3.2 節の手順で収集した観光体験情報を含むツイートから、各観光スポットについて評価スコアを算出する。本研究で使用する評価スコアは「時間帯スコア」と「共起スコア」と「人気スコア」の 3 つである。

時間帯スコアは、観光スポットを訪れる時間帯の適切さを表すスコアであり、新井ら [1] と同じく 24 時間を 3 時間ごとで区切った 8 つの時間帯ごとにスコアをつける。例えば、ある観光スポットについてのツイートが全部で 10 件あり、その内 5 件が 12 時～15 時の間に投稿されたツイートの場合、この観光スポットの 12 時～15 時の時間帯スコアは $5/10=0.5$ である。

共起スコアは、観光スポット同士が共に訪れられやすいかを表すスコアであり、ツイートを投稿した同じユーザの同日のツイートに出てくる他の観光スポットの共起確率を表す。例えば、観光スポット A を訪れたユーザが 10 人いて、その内 2 人が同じ日に観光スポット B も訪れていた場合、観光スポット A における観光スポット B の共起スコアは $2/10=0.2$ である。

人気スコアは、観光スポットの人気を表すスコアであり、各観光スポット名を含むツイート数の 5 段階評価により算出する。具体的には、総ツイート数の下位 20% の観光スポットの評価を 1、上位 20% の観光スポットの評価を 5、その間を 20% 毎に 2, 3, 4 とする。

4.2 ルートの評価値

本研究の評価実験では、観光スポットの評価スコアを用いた以下に示すルートの評価値 v を観光ルートごとに算出する。

$$v = \frac{F}{D} = \frac{\sum_{i=1}^n f_i}{D} \quad (1)$$

ここで

6 : Twitter Developer, <https://developer.twitter.com/>

$$f_i = f_i^P (f_i^T + f_i^{CO}) \quad (2)$$

である。ここで n は観光ルートに含まれる観光スポットの数, f_i は観光ルートにおいて i 番目に訪れる観光スポット $spot_i$ のスコアを表す。 f_i^T , f_i^{CO} , f_i^P はそれぞれ $spot_i$ の時間帯スコア, 共起スコア, 人気スコアを表す。また D は観光ルートの総移動距離 (km) を表す。この評価値 v は観光ルート内の全ての観光スポットの評価スコアの和を観光ルートの総移動距離で割った値であるため, 観光スポットの満足度と移動距離の両方を考慮した評価値といえる。

5 観光ルート推薦システムの概要

本研究では, 新井ら [1] が提案した観光ルート推薦システムを用いる。このシステムは, 以下の手順でユーザに観光ルートを推薦する。

- (1) ユーザが出発地, 出発時刻, 到着地, 到着時刻を入力する。
- (2) システムが表示する観光スポットからユーザが訪問したい観光スポットを 1 つ選択する。
- (3) 選択した観光スポットを含み, 出発時刻から到着時刻の間に収まる観光ルートをシステムが推薦する。

ユーザの入力に対する推薦ルートの生成手法を以下で説明する。

5.1 擬似全探索による推薦ルート生成手法

観光ルートを生成する手法として, 観光スポットの全探索によりルートを生成する手法がある。しかし, 全探索ではルートの生成に時間がかかりすぎるため, 本研究では探索時間に制約を設けるカット処理を行う。このカット処理を含む全探索により生成された観光ルートの中で, 最も式 (1) の F が大きい観光ルートをユーザに推薦する。以降, この手法を擬似全探索と呼ぶ。

この手法では, スコアが大きい観光スポットを優先的に推薦ルートに含める傾向にあるために移動距離が長くなることがあるという問題がある。

5.2 金川らの推薦ルート生成手法

金川らは擬似全探索による推薦ルートにおいて移動距離が長くなるという問題を解決するために, ユーザが選択したスポット(選択スポット)から一定時間内に移動できる観光スポットのみを利用してルートを生成する手法を提案した。以下に金川らが提案した手法の手順を示す。

- (1) ユーザの選択スポットから 15 分以内に移動できる観光スポットを取得する。
- (2) 手順 (1) で取得した観光スポットの数が 5 個未満の場合, 観光スポットを取得する範囲を 15 分延ばして, 観光スポットを再取得する。
- (3) 取得した観光スポットの数が 5 個以上になるまで手順 (2) を繰り返す。
- (4) 取得した観光スポットの数が 5 個以上になれば, それ

らのみを用いて観光ルートを生成し, 最も式 (1) の F が大きい観光ルートを推薦する。

しかし, この手法では出発地や到着地と選択スポットが離れている場合や選択スポットの周りに観光スポットが少ない場合に, 移動距離が長いルートが生成されるという問題があった。また「15 分」という時間や「5 個」という取得する観光スポットの数は恣意的に決められている。

5.3 提案手法

本研究では, 金川らの手法の問題の解決とさらなる移動距離の短縮を目指し, 出発地と到着地と選択スポットの計 3 地点の近傍の観光スポットを優先した観光ルートを生成する手法を提案する。

提案手法では, 出発地, 到着地, 選択スポットの 3 地点のいずれかからの移動時間が短い X 個の観光スポットを候補観光スポットとして取得する。この X は,

$$X = \alpha \cdot \frac{T_G - T_S}{\frac{1}{m} \sum_{j=1}^m V_j} \quad (3)$$

として求める。ここで T_G は到着時刻, T_S は出発時刻, m は 3 節で収集した全ての観光スポットの数, V_j は $spot_j$ の滞在時間を表す。すなわち式 (3) の右辺の分母は全ての観光スポットの滞在時間の平均を表す。出発から到着までの観光時間が長いほど X が大きくなり多くの観光スポットを取得するため, 提案手法は長時間の観光ルートにも対応できる。また α は移動効率を調整するための係数であり, α が小さければ取得する候補観光スポットが少なくなるために移動を制限した観光ルートが生成され, α が大きければ取得する観光スポット数が多くなるために寄り道ができる観光ルートが生成される。さらに α が十分に大きければ, これは 5.1 節で述べた擬似全探索と等しくなる。

この X 個の観光スポットのみを用いて観光ルートを生成し, 最も式 (1) の F が大きい観光ルートをユーザに推薦する。

6 評価実験

本稿では 2 つの評価実験を行う。1 つ目は, 擬似全探索と金川らの手法と提案手法の 3 つ手法により推薦された観光ルートを評価値 v 等で比較する。2 つ目は, 式 (3) の係数 α の値を変更して様々な推薦観光ルートを生成し比較する。

本実験では新井ら [1] が取得した岡山県内の観光スポットの内 51 スポットを使用する。またこれらの観光スポットの評価スコアは, 2019 年 11 月 11 日から 11 月 15 日の期間に収集した 945 ユーザによる 12,172 件のツイートから算出した。

6.1 推荐ルート生成手法の比較

実験においてユーザの入力とする, 出発地, 出発時刻, 到着地, 到着時刻, 選択スポットを表 1 にまとめる。これらの入力条件に対して推薦された観光ルートを比較する。ここで入力条件 (i) と (ii) は入力の 3 地点間の距離が 2km 程度と近く, 3 地点の周辺に観光スポットが多い地点を選んだ。入力条件 (iii) は出発地から選択スポットへの距離が 20km 程度と遠く, 3 地点の周辺に観光スポットが多い地点を選んだ。入力条件 (iv) は出

表 1 観光ルート推薦システムへの入力条件

	出発地	出発時刻	到着地	到着時刻	選択スポット
(i)	岡山駅	10:00	岡山駅	18:00	岡山城
(ii)	岡山駅	12:00	岡山駅	18:00	岡山城
(iii)	岡山駅	10:00	倉敷駅	18:00	美観地区
(iv)	岡山駅	10:00	倉敷駅	18:00	渋川動物公園
(v)	岡山空港	9:00	岡山空港	21:00	井倉洞

発地から選択スポットまで、および選択スポットから到着地までの距離が 25km 程度と遠く、周辺に観光スポットが少ない選択スポットを選んだ。入力条件 (v) は周辺に観光スポットが少ない 3 地点を選んだ。また、入力条件 (ii) は観光時間が 6 時間、(i) と (iii) と (iv) は 8 時間、(v) は 12 時間のルートを推薦する。提案手法において式 (3) の α の値は、本節の実験では全ての入力条件において $\alpha=2$ と設定した。

表 1 で示した入力条件によって 3 つの手法で推薦された観光ルートに含まれる観光スポット数と観光スポットのスコア合計 F と移動距離 D と評価値 v を表 2 にまとめる。

すべての入力条件において、提案手法は擬似全探索と金川らの手法に比べて移動距離が短かった。擬似全探索と比べると入力の 3 地点の距離が近い入力条件 (i) と (ii) において提案手法の方が評価値が大きくなつた。しかし、提案手法はユーザが入力した 3 地点の周辺の観光スポットしか観光ルートに含められないため、ほぼ全ての観光スポットからルートを生成する擬似全探索と比べるとスコア合計は小さくなつた。また金川らの手法と比べると、入力条件 (ii) 以外において提案手法の方が評価値が大きくなつた。

各入力条件で推薦された観光ルートについて詳しく説明する。入力条件 (i) では、提案手法により推薦された観光ルートの移動距離が、擬似全探索と金川らの手法により推薦された観光ルートよりもかなり短くなつたため、スコア合計は小さいが評価値が大きくなつた。提案手法のスコア合計が小さい主な原因として、他の 2 つの手法で推薦された観光ルートでは 7 つの観光スポットを訪問するが、提案手法で推薦された観光ルートでは 2 つ少ない 5 つの観光スポットしか訪問しないことが挙げられる。

入力条件 (ii) で提案手法と擬似全探索を比較すると、入力条件 (i) と同様に提案手法により推薦された観光ルートの方が移動距離が短くなつたため、スコア合計は小さいが評価値が大きくなつた。しかし金川らの手法と比べると、移動距離は同じであるが、スコア合計が小さいために評価値が小さくなつた。提案手法のスコア合計が金川らの手法のスコア合計よりも小さい理由として、提案手法により推薦される観光ルートでは、滞在時間の長い観光スポットが含まれるために、金川らの手法により推薦される観光ルートよりも訪問する観光スポットの数が 1 つ少ないことが挙げられる。提案手法において滞在時間の長い観光スポットを訪問する原因には、入力条件 (ii) の観光時間が 6 時間と短いことがある。金川らの手法では、選択スポットから 15 分以内の観光スポットを候補観光スポットとして取得するため、周辺に観光スポットの多い「岡山城」を選択スポット

表 2 推薦された観光ルートのスコアと移動距離

入力条件	手法	ルート内のスポット数	スコア合計 F	移動距離 D (km)	評価値 v
(i)	擬似全探索	7	10.22	10.9	0.93
	金川らの手法	7	10.27	11.6	0.88
	提案手法	5	8.66	7.1	1.21
(ii)	擬似全探索	5	9.34	19.5	0.47
	金川らの手法	5	9.14	8.2	1.11
	提案手法	4	6.11	8.2	0.74
(iii)	擬似全探索	5	8.32	29.8	0.27
	金川らの手法	4	3.59	33.1	0.10
	提案手法	4	4.88	25.0	0.19
(iv)	擬似全探索	5	7.40	84.6	0.08
	金川らの手法	4	5.38	77.2	0.06
	提案手法	4	4.67	62.0	0.07
(v)	擬似全探索	6	8.25	191.4	0.04
	金川らの手法	4	3.39	199.9	0.01
	提案手法	6	7.69	183.8	0.04

とする入力条件 (ii) では 17 個の候補観光スポットを取得できた。一方、提案手法では観光時間に応じた時間の数だけ観光スポットを取得するために、6 時間の観光時間の場合は 8 個の候補観光スポットしか取得しなかつた。そのため、提案手法により推薦される観光ルートでは滞在時間の長い観光スポットを訪問せざるを得なくなつた。

入力条件 (iii) で推薦された観光ルートを比較すると、手法ごとの違いが明確に表れている。図 1～図 3 にこの条件で各手法により推薦された観光ルートを示す。図 1 の擬似全探索では、入力の 3 地点から訪問する観光スポットへの距離をあまり考慮しない観光ルートを生成するため、出発地と到着地の間に位置する「吉備津彦神社」と「吉備津神社」を訪問する。図 2 の金川らの手法では、選択スポットである「美観地区」の近傍の観光スポットしか観光ルートに含めないため、訪問する観光スポットが選択スポットの周辺に限られている。図 3 の提案手法では、入力の 3 地点の近傍の観光スポットを観光ルートに含めるため、出発地の周辺にある「クレド岡山」と「イオンモール岡山」を訪問し、到着地および選択スポットの周辺にある「アリオ倉敷」を訪問する観光ルートを推薦した。

入力条件 (iv) では、3 つの手法で評価値の差がほとんどない。擬似全探索により推薦された観光ルートに含まれる観光スポットは 5 つ、金川らの手法と提案手法により推薦された観光ルートに含まれる観光スポットは 4 つと、訪問する観光スポットが擬似全探索では 1 つ多いことがスコア合計が最も大きくなつた主な理由といえる。また金川らの手法では選択スポット周辺の観光スポットのみを観光ルートに含めるため、共起スコアが大きくなるような観光スポットを複数訪問し、提案手法よりもスコア合計が大きくなつた。

入力条件 (v) では、擬似全探索と提案手法の評価値には差がないが、金川らの手法の評価値はそれより小さくなつた。図 4 と図 5 にそれぞれ金川らの手法と提案手法により推薦された観光ルートを示す。この入力条件では選択スポットである「井倉洞」の周辺の観光スポットが少ないので、図 4 の金川らの手



図 1 擬似全探索による観光ルート（入力条件 (iii)）

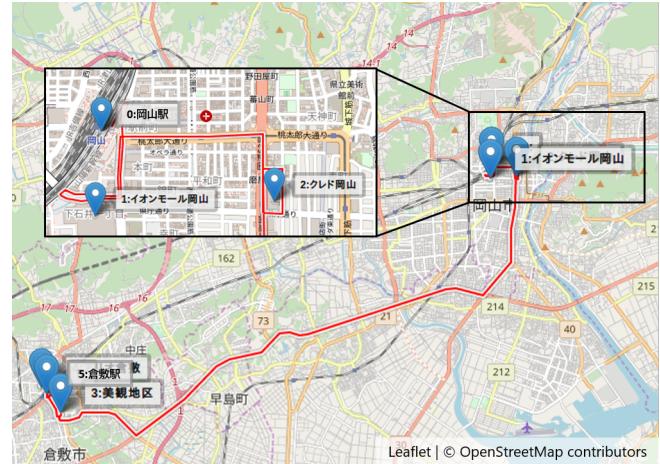
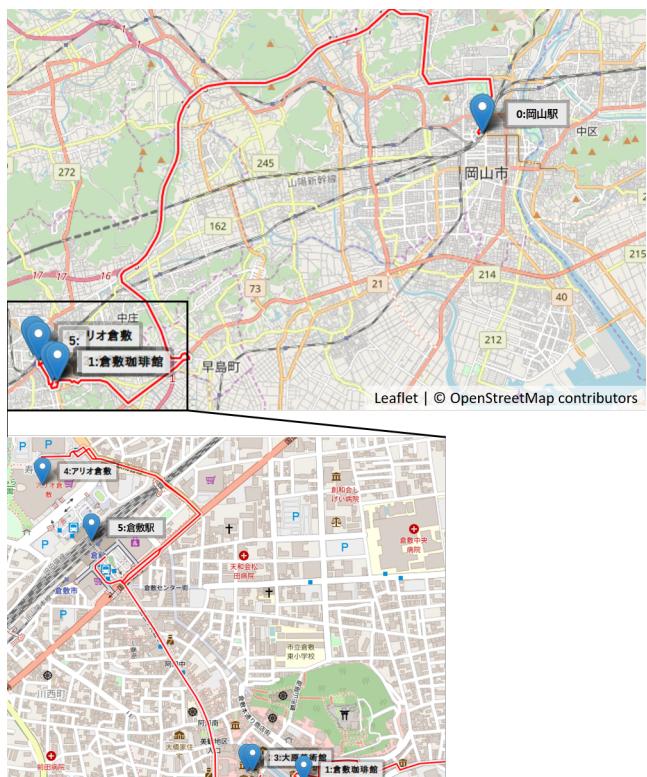


図 3 提案手法による観光ルート（入力条件 (iii)）

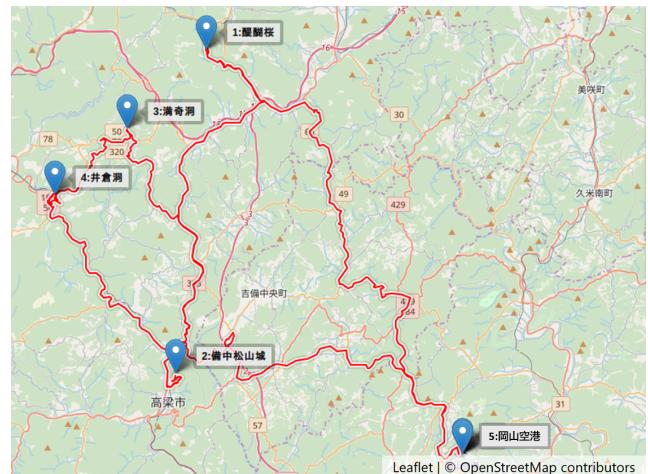
	到着時刻	出発時刻
0 岡山駅		10:00
1 岡山城	10:04	11:39
2 吉備津彦神社	12:03	12:44
3 吉備津彦神社	12:49	14:22
4 倉敷珈琲館	14:52	15:52
5 美観地区	15:53	17:55
6 倉敷駅	17:57	

	到着時刻	出発時刻
0 岡山駅		10:00
1 イオンモール岡山	10:01	11:51
2 クレド岡山	11:58	13:31
3 美観地区	14:12	16:14
4 アリオ倉敷	16:23	17:41
5 倉敷駅	17:41	



	到着時刻	出発時刻
0 岡山駅		10:00
1 倉敷珈琲館	10:42	11:42
2 美観地区	11:42	13:44
3 大原美術館	13:44	15:46
4 アリオ倉敷	15:55	17:13
5 倉敷駅	17:14	

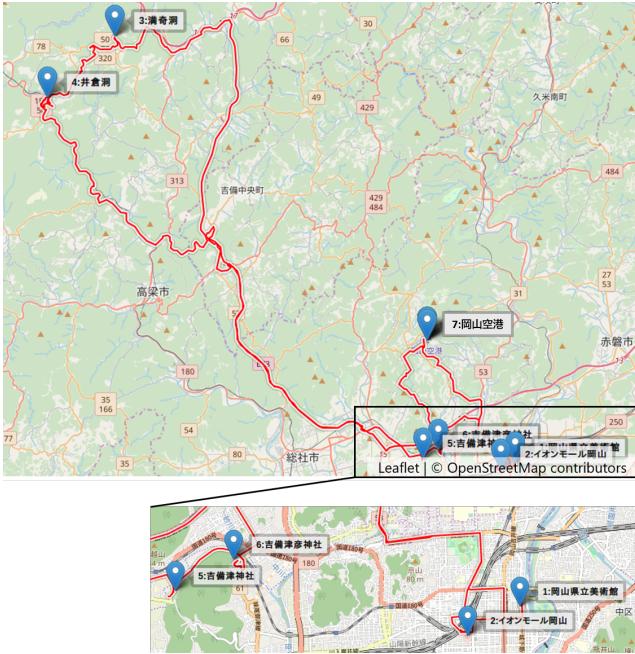
図 2 金川らの手法による観光ルート（入力条件 (iii)）



	到着時刻	出発時刻
0 岡山空港		9:00
1 醍醐桜	10:36	11:36
2 備中松山城	12:41	14:48
3 满奇洞	15:38	16:38
4 井倉洞	16:55	19:06
5 岡山空港	20:43	

図 4 金川らの手法による観光ルート（入力条件 (v)）

法では、滞在時間の長い「備中松山城」や評価スコアが小さい「醍醐桜」を訪問せざるを得なくなり、スコア合計が小さくなつた。しかし図 5 の提案手法では、出発地であり到着地でもある「岡山空港」から少し距離が離れているが岡山駅周辺の観光スポットも観光ルートに含めることができたため金川らの手法と比べるとスコア合計が大きくなつた。



	到着時刻	出発時刻
0 岡山空港		9:00
1 岡山県立美術館	9:32	10:32
2 イオンモール岡山	10:36	12:26
3 満奇洞	13:39	14:39
4 井倉洞	14:56	17:07
5 吉備津彦神社	18:13	19:46
6 吉備津彦神社	19:51	20:32
7 岡山空港	20:56	

図 5 提案手法による観光ルート（入力条件 (v)）

6.2 移動効率の調整

提案手法における式(3)の係数 α の値を 2, 4, 6 の 3 つの値で変更し、表 1 と同様の入力条件に対して推薦された観光ルートを比較する。表 3 には α の値と推薦された観光ルートに含まれる観光スポット数と観光スポットのスコア合計 F と移動距離 D と評価値 v をまとめた。

入力条件 (iv) 以外において、 $\alpha=2$ の場合に最も移動距離が短い観光ルートを推薦し、入力条件 (iii) 以外において、 $\alpha=6$ の場合に最も移動距離が長い観光ルートを推薦した。以上のことから、より移動距離の短い観光ルートを推薦する場合には α の値を小さく設定することが適切であり、反対に α の値を大きく設定することにより入力地点から少し離れた観光スポットへ寄り道をする観光ルートが推薦できるといえる。

しかし、 α の値を大きくすれば必ず移動距離が延びる訳ではなく、入力条件 (ii) の $\alpha=2$ と $\alpha=4$ の場合のように移動距離に変化が無かったり、入力条件 (iii) の $\alpha=4$ と $\alpha=6$ や入力条件 (iv) の $\alpha=2$ と $\alpha=4$ の場合のように移動距離が短くなったりすることがあった。これらの入力条件では、入力の 3 地点の周辺に観光スポットが多く存在するため、 α の値が大きくなってしまっても入力地点から近い観光スポットを候補観光スポットとすることことができた。そのため、同じ移動距離や短い移動距離でも、より観光スポットのスコア合計が高くなる観光ルートを推薦できた

表 3 推薦された観光ルートのスコアと移動距離

入力条件	α	ルート内のスポット数	スコア合計 F	移動距離 D (km)	評価値 v
(i)	2	5	8.66	7.1	1.21
	4	6	10.39	23.7	0.43
	6	6	10.39	23.7	0.43
(ii)	2	4	6.11	8.2	0.74
	4	5	9.14	8.2	1.11
	6	5	9.34	19.5	0.47
(iii)	2	4	4.88	25.0	0.19
	4	5	8.10	35.9	0.22
	6	5	8.32	29.8	0.27
(iv)	2	4	4.67	62.0	0.07
	4	5	7.67	54.4	0.14
	6	5	7.40	74.7	0.09
(v)	2	6	7.69	183.8	0.04
	4	6	8.25	191.4	0.04
	6	6	8.25	191.4	0.04

と考えられる。

また推薦された観光ルートを見ると、入力条件 (ii) と (iii)において $\alpha=6$ の場合に擬似全探索により推薦された観光ルートと同じルートが推薦され、入力条件 (v) においても $\alpha=4$ 以上の場合に擬似全探索と同じルートが推薦された。

7 まとめ

新井らが提案した観光ルート推薦は、観光体験情報を含むツイートから算出した評価スコアの大きい観光スポットを優先して観光ルートに含めるため移動距離が長くなるという問題があった。そのため金川らは、観光ルートに必ず含めるスポットとしてユーザが選択した観光スポットから一定時間内に移動できる観光スポットのみを利用して観光ルートを生成することにより、移動距離の短縮を図った。本稿では金川らの手法を改良し、ユーザが入力した出発地、到着地、観光ルートに必ず含める選択スポットの計 3 地点周辺の観光スポットを優先的に観光ルートに含める推薦ルート生成手法を提案した。

本稿では岡山県内の 51 の観光スポットを用いた 2 つの評価実験により提案手法を評価した。1 つ目の評価実験は、擬似全探索と金川らの手法と提案手法の 3 つの推薦ルート生成手法の比較である。この実験の結果、すべての入力条件において、提案手法の推薦した観光ルートが他の 2 つの手法による観光ルートよりも移動距離を短くすることができた。また、観光スポットの評価スコアと移動距離を考慮したルートの評価値で比較すると、入力の 3 地点が近い場合は提案手法が擬似全探索よりも優れており、選択スポットと出発地または到着地との距離が遠い場合は提案手法が金川らの手法よりも優れているという結果となった。2 つ目の評価実験は、移動効率の調整の評価である。この実験の結果、移動効率を調整する係数を小さくすることにより移動距離の短い観光ルートを推薦し、反対に係数を大きくすることによりユーザが入力する 3 地点から離れた観光スポットへ寄り道をする観光ルートを推薦するように調節できた。

今後の課題としては、岡山県内の観光スポットの分布や評価スコアを精査して、様々な入力条件との関係を分析することや、より多くの観光スポットや他県の観光スポットを使って提案した観光ルート推薦の有用性を確かめることなどが挙げられる。

文 献

- [1] 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, G7-6, 2015.
- [2] 金川元紀, 新妻弘崇, 太田学, “ユーザが選択したスポットの近傍を優先する観光ルート推薦の一手法”, 第 11 回データ工学と情報マネジメントに関するフォーラム, F1-5, 2019.
- [3] 中嶋勇人, 新妻弘崇, 太田学, “位置情報付きツイートを利用した観光ルート推薦”, 情報処理学会研究報告, データベース・システム研究会報告, Vol. 2013-DBS-158, No. 28, pp. 1-6, 2013.
- [4] 中川智也, 新妻弘崇, 太田学, “分散表現を用いたパーソナライズド観光ルート推薦”, 第 10 回データ工学と情報マネジメントに関するフォーラム, H1-3, 2018.
- [5] 松本絢香, 中村健二, 小柳滋, “制限時間内の観光ルート推薦システム”, 第 77 回全国大会講演論文集, Vol. 2015, No. 1, pp. 589-590, 2015.
- [6] 中里主哉, 小林えり, 斎藤和巳, 風間一洋, 吉田光男, “Twitter の位置情報付きツイートを利用したイベントにおけるユーザの行動分析”, 第 78 回全国大会講演論文集, Vol. 2016, No. 1, pp. 401-402, 2016.