

# ソーシャルメディアを用いた訪日外国人へのパーソナライズされたイベント情報提供の最適化ルートの推薦

今井 美希<sup>1</sup> 工藤 瑠璃子<sup>2</sup> 榎 美紀<sup>2</sup> 小口 正人<sup>1</sup>

**概要:** 近年日本を訪れる外国人観光客は急激に増加しており、2020年に開催される東京オリンピック・パラリンピックを考慮すると更なる増加が見込まれる。快適に観光を満喫できる環境整備を進める必要があり、それぞれの観光客の嗜好に合わせた多様な情報を発信すること、観光客にとって使いやすい形式で情報を提供することが重要である。観光客の増加に伴い、有名な観光スポットなどの情報を載せたガイドブックや Web サイトは見受けられるようになってきた。しかしながら、それらの媒体に載っていないようなローカルな情報や今まさに開催されているイベントを知り得ることは、現状難しい。また時間や場所における制約がある観光客にとって膨大な情報量の中から、そのような制約を満たし、興味があるイベント情報を探し出すことは困難である、非常に煩わしい。そこで我々は SNS にある情報に着目した。本研究では位置情報とソーシャルストリームに基づき推定した訪日外国人の趣向によりイベント情報の順位付けを行い、その時その場で使える情報として最適なルートの提案を行う。

## Recommendation of an optimized route for providing personalized event information to foreign visitors to Japan using social media

### 1. はじめに

近年、クルーズ船寄港数の増加や航空路線の拡充、これまでの継続的な訪日旅行プロモーションに加え、ビザの緩和、消費税免税制度の拡充等の結果、日本を訪れる外国人観光客は急激に増加している [2]。2017年には過去最高の訪日外国人数を記録し、その数は2870万人、前年比は19.3パーセントであった。2020年には東京五輪・パラリンピックを控え、更なる増加が見込まれる。日本政府は2020年までに訪日外国人4000万人を目標に掲げており、観光庁をはじめとする多くの機関が訪日外国人旅行者の受入環境の整備・充実を総合的に推進している。また、新たな観光地の発掘や観光地域を作り出す取り組みも行われている。このように、さらなる観光客増加に向けて、快適に観光を満喫

できる環境整備を進める必要があり、それぞれの観光客の嗜好に合わせた多様な情報を発信すること、観光客にとって使いやすい形式で情報を提供することが重要である。

既存の外人向けの情報配信としては、ガイドブックによる王道ルートの提示や、いくつかの訪日外国人向けの Web サイトがあげられる。しかしながら、ガイドブックによる王道ルートなどの提示は恒常的に有名な観光地に関する情報が多く、外国人向けの Web サイトにおいても、有名な観光地やイベント以外の情報が取れず多様な情報は取得しにくいのが現状である。

近年、様々な SNS が普及し、人と人との社会的な繋がりを維持・促進し、情報共有、情報伝達の手段として一翼を担っている。SNS の代表である Twitter [1] は主にツイートをする (情報発信)、ツイートを読む (情報収集) といったシンプルな機能によって作られており、その使いやすさとモバイル端末からも投稿可能という利便性から、多くの人が情報発信、情報収集の場として利用している。更に、ツイートには今その場で起きていることを配信できるリアルタイム性があり、またその情報には単発のイベントや地域特有の情報など、特定の場所にいる人にとって有益なものが含まれている。Twitter には、ホームページのような固

<sup>1</sup> お茶の水女子大学  
Ochanomizu University  
2-1-1, Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan  
miki.oguchi@ogl.is.ocha.ac.jp

<sup>2</sup> 日本アイ・ビー・エム株式会社  
IBM Research - Tokyo,  
19-21 Hakozaki-cho, Nihonbashi, Chuou-ku, Tokyo 103-8510, Japan  
rurikudo.enomiki@ibm.com

定的なメディアには載っていない情報や、ローカルイベントや地域特有の情報など大小様々な有益な情報が含まれている。特定のジャンルに興味を持つ人に向けては、そのようなニッチな情報が非常に価値ある可能性が高い。しかし、これまでそのような所をターゲットとした検索・推薦システムは存在しなかったため、ニッチな情報を取得しようとする際には膨大な情報量の中から自力で探さなければならず、容易に有用な情報を取得するのは困難であった。多様化する観光客のニーズに応えるためには、バラエティに富んだ情報の発信が必要であるため、Twitter 上の情報を利用した検索・推薦システムは有効である。

また、観光客には時間や場所おける制約がある。様々なイベント情報の中から、そのような制約を考慮した上で興味がある情報を引きだそうとすると、時間が限られる観光客にとっては非常に困難であり煩わしい。多くの選択肢がある中から制約を満たし興味があると思われる情報を抽出することは、観光地に詳しくない観光客にとっては有効である。観光中に置いてイベントの内容や、交通状況など状況により予定は変わってくる。「その時」「その場」で役立つ情報が必要となる。

そこで本研究ではそれらの情報を使えるよう整理し、ユーザの趣向に合わせてリアルタイムに発信していこうと考えた。旅行者などの時間とともに移動していく人々に有用な情報を Twitter から抽出し、その情報をユーザの過去のツイート内容から推定した趣向と位置情報により順位づけし、その時その場で使える情報として最適なルートの提案を行うといったインバウンド対応のタイムリーな情報提示手法を提案する。

以下のような流れで本研究を進める。

- (1) ユーザの過去のツイートを 4 カテゴリに分類し、分類結果に基づいてユーザの趣向を推定する。
- (2) 位置情報と、推定したユーザの趣向によりイベントを順位づけする。
- (3) 順位づけしたイベントをその時、その場で使える情報としてユーザの趣向や状況にあわせてルートの提供を行う。

## 2. 関連研究

外国語を用いるユーザが、訪日外国人なのか、在日外国人なのかを判別する手法を佐伯ら [3] が提案している。ユーザが日本国内でツイートを投稿した期間が、一時的な滞在かどうかを推定することによって判別、または分析する期間を設定し、その期間内の日本国内でのツイートの割合に着目して判別するといった手法をとっている。

趣向に合わせた情報提供の手法については石野ら [4]、向井ら [5]、上原ら [6]、矢部ら [7] によっても提案されている。

石野らは各個人の SNS 上の画像群を解析することにより、ユーザがどのような観光地を好むかを判定する手法を提

案している。向井らの研究では、マイクロブログ特有のリアルタイム性の高い投稿を活用して効果的な情報推薦を行う手法を提案している。「リツイート」に着目し、そこから各ユーザのプロファイルを作成し、またセレンディビティのある推薦を可能にするため、プロファイルの類似するユーザのクラスタリングを行うことでユーザの嗜好を推定している。これらの研究では、ある決められたイベント情報を配信するというスタティックな情報提供になっている。本研究ではダイナミックに、リアルタイムに時間と場所を考慮した上で配信することを目指している点で、これらの研究とは異なる。

上原らは、Web から観光情報を抽出し、複数の特徴ベクトルから観光地間の類似性を評価することで、観光地を推薦するシステムを提案している。ユーザはお気に入りの観光地を入力し、知恵袋・ブログ上での共起キーワードと時系列分布、知恵袋上でのカテゴリ構造、観光地周辺施設、地図画像 から生成した観光地の特徴ベクトルから、お気に入りの観光地と類似性の高い観光地を推薦する。

矢部らの研究では、利用者の趣味・嗜好を踏まえた上で配信情報との関係性を動的に計量する位置情報に基づく情報配信システムを提案している。利用者の興味・関心というリアルタイムに変化する利用者の意図を、変動しないベクトルに変換し、地理情報・配信情報と利用者間関係性を動的に計量することにより、趣向、位置情報を考慮した情報配信に成功している。

上原らや、矢部らの研究では、自身で属性や趣向などを入力した上で情報推薦を行うのに対し、本研究では過去のツイート内容から自動的に情報推薦を行う点で異なる。

## 3. 提案システム

観光客などに有用な情報をタイムリーにインバウンド対応で提示するために、本研究で提案するシステムの概要を図 1 に示す。

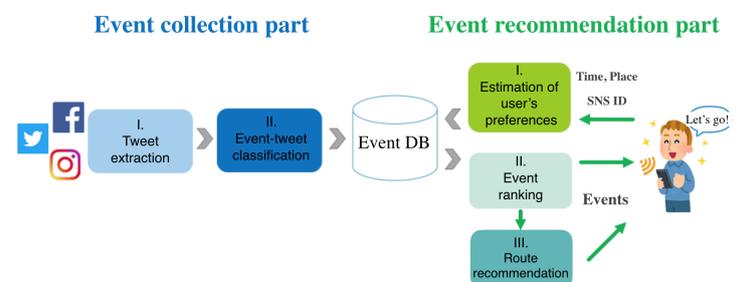


図 1 提案システムの概要

### ● Event collection part

- (1) SNS から観光客にとって有用なイベント情報を抽出する

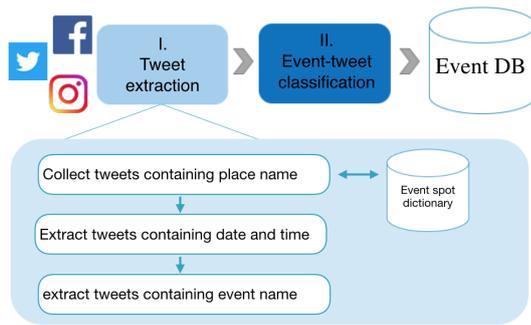


図 2 Event collection part

(2) イベントをカテゴリごとに分類し、保存する

### ● Event recommendation part

- (1) ユーザの過去のツイートをカテゴリに分類し、分類結果によりユーザの趣向を推定する
- (2) 推定した趣向、時間、位置情報によりイベントを順位づけする
- (3) 順位づけしたイベントをその時、その場で使える情報としてユーザの趣向や状況にあわせてルートの提供を行う

Event collection part は工藤ら [8] によって研究が進められている。本研究では Event recommendation part の部分に焦点をあて、先行研究で抽出したイベント情報をユーザの SNS から推定した趣向に合わせて時間と場所を考慮したルートの提案を行う。

### 3.1 Event collection part

ソーシャルメディアからのイベント収集について説明する。

#### 1 Extracting tweets

図 2 に示すようにソーシャルメディアからイベント情報を収集する。まず、Twitter API[9] のキーワード検索を用いて地名を含むツイートを収集し、次にイベントがいつ、どこで開催されるかを抽出するために正規表現を用いてテキストを分析することにより、日時を含むツイートを抽出する。

次にイベント名を抽出する。ツイートには文字数以外の制約はないため、記述の自由度が極めて高い。記述がある程度形式化された日時などの情報は、正規表現を使用して自動的に抽出することが可能であるが、「イベント名」のような形式の定まっていない記述は、自動的に抽出することが困難である。そこで、ユーザに提示する有益な情報を増やすために、外部情報を用いて情報を補足する機能を提案する。正規表現を用いてツイートから抽出するという手法 (以下、手法 1) に加え、外部情報を利用する手法として、スクレイピング

“Music Event” “Comedy/Vaudeville”  
“Stage” “Movie” “Exhibition”

を行い、提示するイベントの Web サイトの HTML から、title タグの中身を取得する手法を利用する。スクレイピングを用いる手法においては、ツイートに Web サイトの URL が記載されている場合は、正規表現で URL を抽出し HTML を取得する (以下、手法 2-a)。URL の記載がない場合は、ツイートを検索エンジンにかけた結果の上位数件の Web サイトの HTML データを取得する (以下、手法 2-b)。

まず、手法 1 が適用可能なツイートに対し手法 1 を適用し、正解とする情報が取得できたか判定を行う。手法 1 において正しく情報が取得できなかった場合、手法 2-a または手法 2-b のうち、適用可能な手法で再度、情報取得を試みる。このように順に手法を適用した結果、最終的に有用な情報がどのくらい取得できるかを次節にて評価する。

先行研究において [8]、このように順に手法を適用した結果、最終的に有用な情報がどのくらい取得できるかを次節にて評価する。

#### 2 Classifying tweet events

ユーザの趣向に合わせてイベント情報を提供するために、イベントのツイートをカテゴリに分類する。MeCab[10] を用いて文書において必要な部分 (名詞、動詞、形容詞、副詞) を抽出し、ツイートの Bag of Words のベクトルのモデルを作成した。カテゴリの定義は代表的な都市情報誌である「東京ウォーカー」[11] のイベントカテゴリを参考にした。

MeCab 用の日本語のシステム辞書 (ipadic-NEologd) を用いて、SVM とランダムフォレストによるイベント分類の評価を行なった。実験結果は精度や計算スピードの観点から見て SVM よりもランダムフォレストを使用の方が良い結果を得られたため、イベント分類にランダムフォレストを使用した。

表 1 Classifying preferences for each category random forest + mecab-ipadic-NEologd

Category	Precision	Recall	F-measure
Exhibition	0.87	0.81	0.84
Movie	1.00	0.73	0.84
Music	0.88	0.76	0.81
Stage	0.50	0.40	0.44
Comedy	0.79	0.76	0.78

抽出したイベントはイベントデータベースにためておく。

従来の手法である Web サイトからの情報取得に比べ、

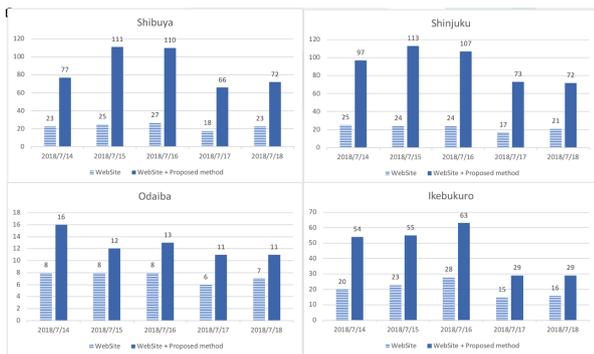


図 3 Experimental results 1

## Event recommendation part

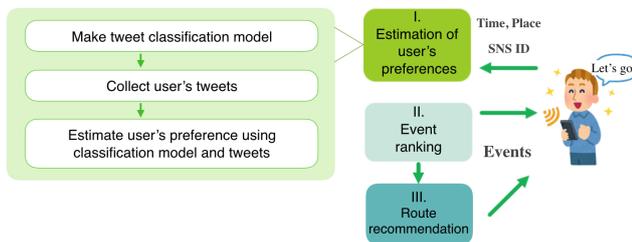


図 4 Event recommendation part  
 本手法を用いることでどれくらいのイベントが取得できるか検証を行ったところ、取得可能なイベント情報はどの日付、どの地域においても増加することを示した。図 3 に検証結果を示す。2018 年 7 月 14 日から 7 月 18 日まで、渋谷、新宿、お台場、池袋で開催されたイベントを収集した。

## 3.2 Event recommendation part

### 1 本研究の概要

次に、ソーシャルメディアを利用したイベント推薦について簡単に紹介する。

図 4 に示す。まず、ツイートを 4 つのイベントカテゴリに分類する分類モデルを作成する。次に、位置情報とツイートの分類結果から推定したユーザの好みによるイベントの順位づけ方法について述べる。最後に、イベントのランキングに従って最適化されたその時、その場で使えるルートの提案を行う。

### 2 Estimating user's preferences

趣向の解析方法について説明していく。

趣向の判定にはユーザの過去のツイートをを用いる。過去のツイートをあらかじめ設定したイベントのカテゴリに分類し、分類結果をもとに各カテゴリの点数を求める。

カテゴリは、イベント情報の収集を行なっている先行研究に合わせて展示、映画、ライブ、舞台の 4 カテゴリとした。過去のツイートをカテゴリーに分類するた

めの分類モデルを作成する。

### (1) 分類モデル作成のデータ

Twitter API[9] のキーワード検索で各カテゴリごとのキーワードを設定、また言語を英語に設定することで英語のツイートを取得する。ツイートは各カテゴリ 500 ツイートずつ取得し、キーワードは以下のように設定する。

展示会 art, museum, gallery, anime, manga, cartoon

映画 movie, cinema, film, Academy-Awards, Oscars, Golden-Globe-Awards

ライブ concert, gig, recital, song, sing, melody

舞台 musical, opera, theater, drama, show, stage  
 上記以外のツイート キーワードなし

取得する際、ツイートをする人が多くいる時間帯を見計らうことで多くのツイートを取得することが可能となった。

### (2) 分類モデル作成

3.2 で取得したツイートをを使い、ユーザのツイートを分類するための学習器を作成していく。まず特徴語辞書を作成する。各カテゴリのツイートに関する特徴を捉えるためである。ツイートを単語に分割、更に小文字に直す。そこから is, a, RT など出現回数が多い単語を予め設定しておき取り除く。

特徴ベクトル (出現頻度のベクトル) に変換するため、単語と ID, ID と頻度にマッピングした後、ベクトルにする。続いて学習を行なっていく。

先行研究においてイベント分類で代表的な手法である SVM とランダムフォレストを比較したところ、精度、処理速度の観点からみてランダムフォレストの方が研究では優れていた為、本研究においてもランダムフォレストを使用して学習を行う。

英単語には過去分詞 (ed), 現在分詞 (ing), 複数形があり、分類の妨げとなると考えた。そこで、Tree Tagger[12] と呼ばれるパッケージを用いて形態素解析を行い、英単語を過去分詞 (ed), 現在分詞 (ing), 複数形などを標準形でまとめた。

また、scikit-learn のランダムフォレスト [13] には多くのパラメタがある。調整することで、よりランダムフォレストで正確な分類をできるようになる。そこでグリッドサーチという自動的な最適化ツールを使い、与えたパラメタの中で最も精度の良いものを選ぶ。学習データを用いてランダムフォレストによって学習を行った。

### (3) カテゴリ再考

作成した分類モデルの分類結果に偏りがある場合はカテゴリを再考する必要がある。そのため、事前に以下のようにカテゴリを定義しておく。各カテゴリは図5に示すように木構造で定義しておく。分類の結果、全体の半分以上を占めるカテゴリがある場合には木構造の真ん中でさらにカテゴリを分割し、各分割した木構造以下の単語をカテゴリのキーワードとして設定する。図5は展示会カテゴリが全体の半分以上を占めている場合に、展示会カテゴリを再考する例を示す。

展示カテゴリが全体の半分以上を占めている場合は、展示の木構造を真ん中で分割したときその次には美術とアニメがある。美術のキーワードは art, gallery, museum. アニメのキーワードは anime, manga, cartoon となる。

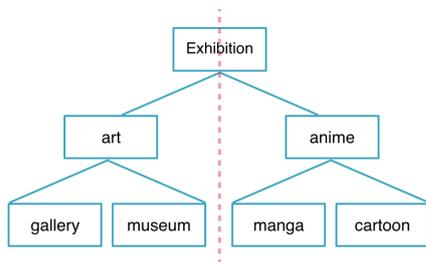


図5 Reconsideration of categories

### 3 Event ranking

時間、行動範囲が限られる旅行者には、「その時」「その場」で役立つ情報を配信していく必要がある。そのため、イベントの順位付けにおいて、位置情報や時間を考慮していくべきと考える。また、ユーザが興味を示す、ユーザの趣向に沿ったイベント情報の提供が必要であり、イベント自体の質も考慮すべきである。それらを組み合わせると、位置情報を考慮した上で、趣向を考慮し、更にイベント自体の質を考えたイベント情報の提供ができる。

#### (1) 順位づけ手法

##### (a) 位置情報のみを考慮した場合

地図上の標準地域メッシュを利用し、現在地を考慮してメッシュ毎に Event score(a) をつけ順位付けを行う。

図6のピンが立っているところを現在地 x とすると、Event score (a) = mesh(x) で図の該当する値を返す。

0.3	0.5	0.3
0.5		0.5
0.3	0.5	0.3

図6 地図メッシュ

#### (b) 趣向を考慮した場合

Event score (b) = (分類されたジャンルのツイート数)/(全体のツイート数) × Event score (a) とすることで、位置情報に趣向を加えて順位付けを行う。

#### (c) メジャーなイベント、マイナーなイベント

より、ユーザそれぞれにあった情報を配信するため、ユーザの好みに合わせてメジャーなイベント、マイナーなイベントを選択する。

follower 数によるスコアを (b) に掛け合わせた。Event score (c) = Event score (b) × follower-score. RT などメジャー、マイナーを判断する基準と考えたが、今回扱っているイベントはローカルなイベントや単発のイベントのため、RT をされているイベントは少なく、基準とならないと考えた。

メジャー、マイナーによる順位付けにおいて、follower/max-follower によりスコア付けを行うと、差が大きくなりすぎてしまうため、予め閾値を定めスコアを求めた。メジャー、マイナーによる順位付けは趣向を重視した研究のため、重み付けを小さくした。表2、にスコアの求め方を表した。

表2 follower-score

x-major	x-minor	follow
x = 0.03	x = 0.15	100 ≤ follower < 500
x = 0.05	x = 0.13	500 ≤ follower < 1000
x = 0.07	x = 0.11	1000 ≤ follower < 5000
x = 0.09	x = 0.09	5000 ≤ follower < 10000
x = 0.11	x = 0.07	10000 ≤ follower < 15000
x = 0.13	x = 0.05	15000 ≤ follower < 20000
x = 0.15	x = 0.03	20000 ≤ follower

### 4 Route recommendation

#### (1) Route recommendation method

順位づけしたイベント情報をその時、その場で使える情報としてユーザの好み、状況にあわせてルートの提供を行う。

Event collection part においてどの日付、どの地域においても多くのイベント情報を取得することが可能であった。そのため、多くのイベント情報の中からユーザにとって最適なルートの探索を行うには計算時間が莫大になる。そこで、3.2 節に置いて順位づけしたイベントのうち上位のイベントのみ抽出し、そのイベント情報の中からルートの推薦を行うこととする。ユーザの趣向に合わせてながらヒューリスティックにルートの探索を行う。

ユーザはあらかじめスタート、ゴールの時間とイベントにおける予定滞在時間を入力し、抽出したイベント情報の中から、順位が高いイベントを極力多くいけるルートの探索を行う。ユーザはイベントに参加し終わると実際の滞在時間を入力し、予定滞在時間と異なる場合には再び最適なルートの探索を行い、提示する。ユーザの予定が変わったとしても、状況に合わせてダイナミックに情報の提供を行う。

#### 4. 検証結果

##### 1 ユーザの趣向性考慮の効果の検証実験

###### (1) 分類モデルの制度

形態素解析、チューニングを行い、ランダムフォレストによって学習器を作成し、ツイートを分類を行った。

適合率 (Precision), 検出率 (Recall), F 値 (f-score) などから見ても、高い結果が得られた。

表 3 Classification accuracy

	precision	recall	f-score
Exhibition	0.79	0.83	0.81
Movie	0.96	0.82	0.89
Music	0.87	0.84	0.86
Stage	0.89	0.82	0.86
avg/total	0.84	0.83	0.83

###### (2) 全体の分類傾向 (外国人と日本人別)

作成した分類モデルを用いて、趣向の解析を行った。一定数の日本人と外国人のツイートをそれぞれ抽出、分類し、それらの結果からグラフを作成した。結果を図 7 に示す。双方のグラフ (図 7) を比

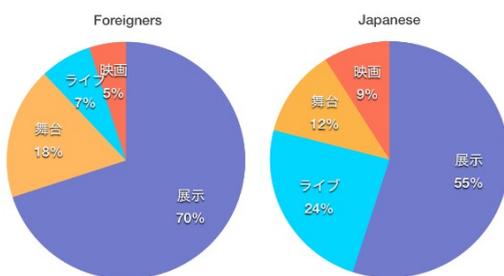


図 7 Trends of Japanese and foreigners

べると、日本人は外国人よりライブのツイートの割合が多いことが見て取れる。また、どちらのグラフからも展示の割合が大きくなっていることがわかる。これは展示として分類しているツイートが、アニメから美術など多岐にわたっていることが要因と考える。

###### (3) カテゴリ再考

展示に偏って分類される傾向がある。分類結果の半数以上を占めるカテゴリがある場合は、そのカテゴリを更に詳細に分ける事とする。

今回は展示が半数以上を占めていた為、アニメと美術に分類した。キーワードは

アニメ anime, manga

美術 art, museum, gallery

とした。

4-1-(2) 節と同様に日本人と外国人に対してツイートの分類を行ない、グラフにまとめた。これまでと比べるとバランスはよくなったが、アニメに多く分類される傾向がある。また、分類結果が 4-1-(2) 節と比べ展示以外のカテゴリも割合が変わっているが、これは分類モデルの精度が 8 割ほどである事が要因と考えられる。

今回は元の展示、映画、ライブ、舞台の 4 カテゴリのまま研究を進めることとする。

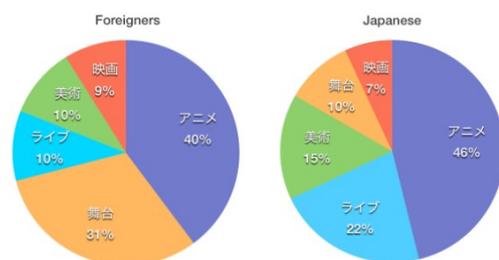


図 8 Reconsidering categories

###### (4) 趣向推定の分類モデルの評価

作成した分類モデルによって正しくユーザの趣向を推定できているのか検証を行なっていく。今回は映画に興味を持つユーザの映画に関するアカウントを用いてツイート分類し、趣向を推定する。3 人のユーザに対して分類を行なった結果を表 4 に示す。

表 4 映画用アカウントの分類

	ユーザ 1	ユーザ 2	ユーザ 3
展示会	13 件	246 件	21 件
映画	113 件	336 件	37 件
ライブ	1 件	85 件	15 件
舞台	8 件	72 件	16 件

どのユーザを見ても映画に最も多くツイートが分類されており映画に興味があることがわかる。趣向を推定できていると考えられる。

続いて、推薦されているイベントを実際に訪れたユーザに対してツイートの分類を行う。推薦して

いるイベントを実際に訪れていた外国人のツイートを探し出すことが困難であった為、日本人でイベントを訪れたユーザのツイートを英語に自動翻訳してツイートの分類を行う。今回はある展示イベントを訪れたユーザの過去のツイートを分類した。結果を以下の表5に示す。

表5 展示イベントに訪れたユーザツイートの分類

	ユーザ4	ユーザ5	ユーザ6
展示会	783件	393件	128件
映画	172件	64件	52件
ライブ	374件	187件	166件
舞台	98件	47件	25件

ユーザ9, ユーザ10は展示のイベントに最も多くツイートが分類されたが、ユーザ11はライブが最も多い結果となった。ユーザ11に関しては、ユーザ11の総数(371)に対して34.5%が展示会、44.8%がライブ、ライブが1位だが展示もそこそこ関心が高いことが見受けられる。

(5) 訪日外国人に対する検証

実際に日本へ観光に来た訪日外国人のツイートを用いて検証を行った。

(a) 検証データ

ツイート取得の際、ツイートの位置情報を日本国内の緯度経度に合わせ、更に言語を英語とすることで訪日外国人のツイートを取得した。今回はツイートをした人が在日外国人か、観光目的での訪日外国人か目視で判断した。外国人観光客と思われる5人分の過去のツイートを取得した。この5人をユーザと見立て、ツイートの分類を行い趣向の解析をしていく。

(b) ツイート分類

取得したツイートを作成したユーザの趣向性を判定するための分類モデルによって分類した。分類の結果を以下に示す。

表6 訪日外国人のツイート分類

	ユーザ7	ユーザ8	ユーザ9	ユーザ10	ユーザ11
展示会	274件	2048件	663件	448件	807件
映画	40件	66件	53件	48件	76件
ライブ	78件	80件	60件	61件	135件
舞台	144件	490件	236件	63件	141件

実際にユーザのツイートを目視で確認する。ユーザ7について過去のツイートを確認すると、アートに関するツイートやリツイートが多く確認で

き、またミュージアムへ行ったツイートも確認できたため展示には興味があるのではないかと推測できる。また、実際に日本に滞在し観光していた際のツイートを確認すると、日本のミュージアムを観光していたツイートを確認することができたため、情報の推薦に役立つとみなすことができる。表6の結果から、訪日外国人の趣向の傾向が展示会にあることが見受けられる。この結果を用いてツイッターの情報が無い観光客にも、訪日外国人の傾向を反映してイベントの配信を行えると考ええる。

2 イベントの順位づけの検証

(1) イベントの順位づけの検証

4-1-(5)節でツイートの分類を行ったユーザ7に対してイベントの推薦を行っていく。このユーザのその他に分類されたツイート以外の全ツイート数は536件であった。またユーザ7は渋谷駅にいると仮定してイベントの推薦を行う。

表7 趣向に対する点数付け

	分類結果	score
展示会	274件	$274/536 = 0.51$
映画	40件	$40/536 = 0.07$
ライブ	78件	$78/536 = 0.14$
舞台	144件	$144/536 = 0.26$

表8 場所, 趣向, メジャー, マイナー

	(1):場所	(2):(1)+趣向	(3):(2)+メジャー	(4):(2)+マイナー
1	ライブ1(1.0)	展示1(0.51)	展示1(0.0561)	展示2(0.0561)
2	映画1(1.0)	展示2(0.51)	舞台1(0.0286)	展示3(0.0382)
3	ライブ2(1.0)	舞台1(0.26)	展示2(0.0255)	展示1(0.0255)
4	ライブ3(1.0)	展示3(0.255)	ライブ2(0.0154)	展示4(0.021)
5	ライブ4(1.0)	ライブ1(0.14)	映画2(0.0151)	ライブ5(0.021)
6	展示1(1.0)	ライブ2(0.14)	舞台2(0.0065)	舞台2(0.0169)
7	展示2(1.0)	ライブ3(0.14)	映画3(0.00525)	舞台1(0.013)
8	舞台1(1.0)	ライブ4(0.14)	展示3(0.00255)	映画1(0.0105)
9	舞台2(0.5)	映画2(0.13)	映画4(0.00245)	舞台5(0.0091)
10	展示3(0.5)	映画1(0.07)	舞台3(0.00245)	映画5(0.00315)

表8は、左から(1):位置情報のみを考慮した場合、(2):(1)に加え趣向を考慮した場合、(3):(2)に加えメジャーなイベントを抽出した場合、(4):(2)に加えマイナーなイベントを抽出した場合を比較する表である。

位置情報のみを考慮したものと、それに加え趣向を考慮したものを比べると、位置情報のみだった場合は同じスコアのもの10件中8件を占め、その中で適切に順位付けが行われていなかったもの

が、趣向を考慮することで本人の趣向により合ったリストになっているので趣向の考慮が有効である。また、趣向を考慮することで、位置情報だけでは遠いため下位にあったイベントが中位に表示されるようになった。

趣向を考慮した上で、メジャーなイベント、マイナーなイベントを考慮したものは、趣向のみでは上位 10 件に推薦されなかったイベントも推薦可能となった。例えば、表 8 の下線部のイベントである。

### 3 Route recommendation

#### (1) Route recommendation method

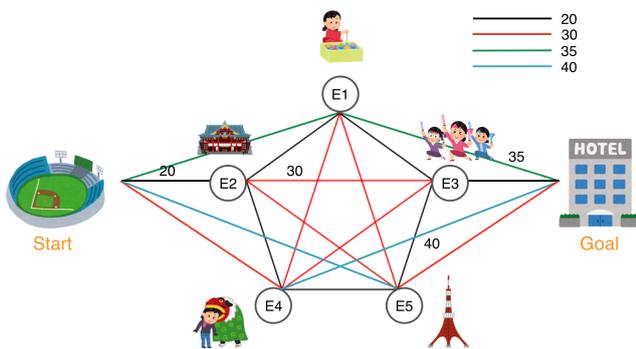


図 9 Route recommendation

- 14:00 にスタート
- 20:00 にゴール
- イベントの滞在予定時刻はユーザが設定し、今回は 60 分とする
- E1~E5 はイベントであり、E1 から E5 の順にイベントの順位は低くなっている

(i) 以下のルートをもっとも適したルートとして推薦

**start - E2 - E4 - E1 - E3 - goal**  
E2 に予定より 40 分長く滞在した

(ii) E4 には行かないルートを紹介

**E2 - E1 - E3 - goal**  
E1 に予定より 30 分短い滞在であった

(iii) 当初推薦された E4 には行けないが、E4 より優先度は低いが E5 には行けるルートを紹介

**E1 - E5 - E3 - goal**

### 5. まとめと今後の課題

増加する外国人観光客に向けて、ガイドブックや Web サイトなどから容易に取得できないような、ローカルな情報や今まさに開催されているイベントを、趣向を考慮してイベントを順位づけすることで、ユーザに合わせてその時その場で使える情報として提供することを目的として研究を進めた。イベントの順位付けにおいて位置情報のみと比べ、

趣向を考慮した場合は、本人の趣向により合ったリストになっているので趣向の考慮が有効である。更にメジャーなイベント、マイナーなイベントをユーザに選択できるようにすることでよりユーザの望むイベントの推薦が可能になると考えられる。

ルート推薦においては、順位づけしたイベント情報をその時、その場で使える情報としてユーザの予定が変わった場合にも対応したダイナミックなルート推薦を行なった。

今後は、ルートの推薦における検証、検証手法を提案していきたい。

### 6. 謝辞

本研究は一部、JST CREST JPMJCR1503 の支援を受けたものである。

#### 参考文献

- [1] "Twitter," <http://twitter.com/>
- [2] 日本政府観光局  
[https://www.jnto.go.jp/jpn/statistics/data\\_info\\_listing/pdf/170117\\_monthly.pdf](https://www.jnto.go.jp/jpn/statistics/data_info_listing/pdf/170117_monthly.pdf)
- [3] 佐伯 圭介, 遠藤 雅樹, 廣田 雅春, 倉田 陽平, 横山 昌平, '外国人 Twitter ユーザの観光訪問先の属性別分析', DEIM Forum 2015 C4-3.
- [4] 石野 淳一, 中田 洋平, 日吉 久礎, 'SNS 上の画像群からのユーザー嗜好の抽出と観光広告への応用', 情報処理学会第 76 回全国大会 6T-5.
- [5] 向井 友宏, 黒澤 義明, 目良 和也, 竹澤 寿幸, "マイクロブログの分析に基づくユーザの嗜好とタイミングを考慮した情報推薦手法の提案", 言語処理学会 第 17 回年次大会 発表論文集.
- [6] 上原 尚, 嶋田 和孝, 遠藤 勉, 'Web 上に混在する観光情報を活用した観光地推薦システム', 電子情報通信学会.
- [7] 矢部 竜太, 倉林 修一, 清木 康, "配信情報と利用者の関係性を動的に計量する位置情報に基づく情報配信システムの提案", DEIM Forum 2012.
- [8] 工藤 瑠璃子, 榎 美紀, 中尾 彰宏, 山本周, 山口実靖, 小口 正人, "場所と時間を考慮した SNS データを用いる訪日外国人観光客へのタイムリーな情報配信", 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 (DEIM2018), B1-3.
- [9] "Twitter Search API,"  
<https://dev.twitter.com/rest/public/search>
- [10] "MeCab," <https://taku910.github.io/mecab/>
- [11] "Walker+ 東京都のおでかけスポット一覧," <http://www.walkerplus.com/>
- [12] "Tree Tagger,"  
<http://www.cis.uni-muenchen.de/schmid/tools/TreeTagger/>
- [13] <http://scikit-learn.org/stable/modules/generated/sklearn.RandomForestClassifier.html>