

ソーシャルストリームを用いた場所と時間に対応するリアルタイムなイベント検索

工藤 瑠璃子¹ 丸 千尋¹ 榎 美紀² 中尾 彰宏³ 山本 周³ 山口 実靖⁴ 小口 正人¹

Real Time Event Search Corresponding to Place and Time Using Social Stream

RURIKO KUDO¹ CHIHIRO MARU¹ MIKI ENOKI² AKIHIRO NAKANO³ SHU YAMAMOTO³
SANEYASU YAMAGUCHI⁴ MASATO OGUCHI¹

1. はじめに

近年, 様々な SNS(ソーシャル・ネットワーキング・サービス) が人々の間に普及し, 一種の情報伝達手段として一翼を担っている. SNS の代表である Twitter[1] は, ツイートを読む (=情報収集) とツイートする (=情報発信) というシンプルな機能を軸に作られており, その使いやすさとモバイル端末からも投稿できるという利便性から, 多くの人が気軽に情報を発信している. 例えば, 訪れた観光地の感想や, 映画や舞台を観に行ったら感想をその場でツイートしたり, お祭りやイルミネーションのようなイベントでは写真や感想, 混雑具合をツイートすることも多い. 実際に行った感想ではなく, これから行くイベントなどの予定をツイートするユーザもよく見られる. また, 「リツイート」機能を利用してツイートを拡散し, 多くの人に情報を知ってもらうことも可能であるため, イベントの告知などにも使われている. そのような目的で使用する際には, ツイートの文字数が 140 字に制限されていることから, ツイート本文には重要な情報が簡潔にまとめられているという特徴もある. このように Twitter を代表とする SNS には, ローカル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情報が含まれている. しかし, 有名な観光スポットなどの情報はガイドブックや WEB サイトから取得できるが, 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取得するのは, 膨大な情報の中から自力で探し出さなければならず, 非常に困難である. 特に, 場所と時間の制約がある旅行者などが必要

とする「その時」「その場所」に関する情報を発信している媒体は少なく, その場で役立つスポット的な情報が配信可能になれば, 観光に来た土地勘の無い旅行者などに有益だと考える.

また, 東京オリンピック開催決定前の 2012 年の訪日外国人数は 8,358,105 人であったのに対し, 2013 年は 10,363,904 人, 2014 年は 13,413,467 人, さらに 2015 年は 19,737,409 人に上り, 東京オリンピック開催が確定した 2013 年以降, 訪日外国人数は大幅に増加している [2]. 急増しているインバウンドへの情報配信手段を充実させるため, IoT デバイスの開発も進んでいる. 東京大学では 2015 年にネットワーク仮想化に対応したアクセスポイントを用いて, 様々な情報配信サービスを行う実証実験を実施している [3]. この実証実験では, Wi-Fi の無線ビーコンに情報をのせユーザのスマートフォンに情報を配信する BeaconCast 技術を実装したアクセスポイントを, 空港と都内を結ぶリムジンバスに設置しており, インターネットにアクセスできない場合のメッセージ配信を実現している. リムジンバスの乗客は, 専用アプリケーション「LimoCast」をインストールすることで, 車内に配置されたアクセスポイントを利用して, 通常の無線 LAN アクセスによるインターネット接続を楽しむと同時に, バス到着地周辺のイベント情報, 店舗のクーポン, リアルタイムで更新されるお得な情報を受け取ることが可能になる. これは, 通信契約をしていないインバウンドへの情報配信手段としても有用である.

本研究では, このような IoT デバイスを基盤と考え, 旅行者などの時間とともに移動していく人に有用な情報を SNS の代表である Twitter から抽出し, 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.

本論文の構成は以下の通りである. 2 章で関連研究につ

¹ お茶の水女子大学

² 日本アイ・ビー・エム株式会社

³ 東京大学

⁴ 工学院大学

いて述べ、3章で提案システムの概要を紹介する。提案システムの詳細を、4章、5章、6章に分けて説明する。最後に、7章で本稿をまとめる。

2. 関連研究

Twitter から観光情報や地域特徴を抽出する手法は、前田ら [4]、長谷川ら [5]、中川ら [6]、佐伯ら [7] によっても提案されている。前田らの手法では、Twitter 位置情報を用いて、人の高頻度滞在地と低頻度移動地を求め、移動件数と距離から地域の魅力を算出し、その魅力が固有性によるものなのか、利便性によるものなのかを区別することで観光地を抽出している。長谷川らは、Twitter 上に投稿されたコンテンツの中から、地域の特徴を表す特徴語を抽出して地域特徴語辞書の構築を行い、構築された辞書を利用して Twitter からユーザの観光体験を検索する手法を提案している。中川らは、ツイートされた時刻に着目して観光ルートを推薦する手法の問題点を改善し、観光スポット間の距離を考慮した観光ルートを推薦するための手法を提案している。また、佐伯らは、外国人のツイートに出現する地名に注目し外国人向けの観光情報抽出の手法を提案している。これらの研究では、蓄積された情報を解析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、リアルタイムに情報を収集して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。

Twitter を利用したイベント検知に関連する研究として、Chao ら [8]、山口ら [9] の手法があげられる。イベント検知手法の多くは、今まさに起こっているイベントの検知を対象としており、類似した内容を持つ投稿がある地域から局所的に多く投稿された時、その地域で何らかのローカルイベントが発生している可能性が高い、というアイデアに基づき提案されている。また、これらの手法では、地理的局所性のある地震やデモなどのローカルイベントを検知対象としている場合が多い。本研究では、観光者に対して情報提示を行うことを想定しているため、今まさに起きているイベントではなく、少し未来の情報が抽出対象となる。検知対象は、祭りやライブなどの地理的局所性を持たないイベントも対象とする。

3. 提案システム

観光者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本研究では以下のシステムを提案する。提案システムの概要を図 1 に示す。

(1) ツイートの抽出

1. Twitter API[10] のキーワード検索で地名をキーワードに設定し、地名に紐付いたツイートを取得
2. ツイートテキストを解析し情報を整理する。

(2) ツイートのイベント分類

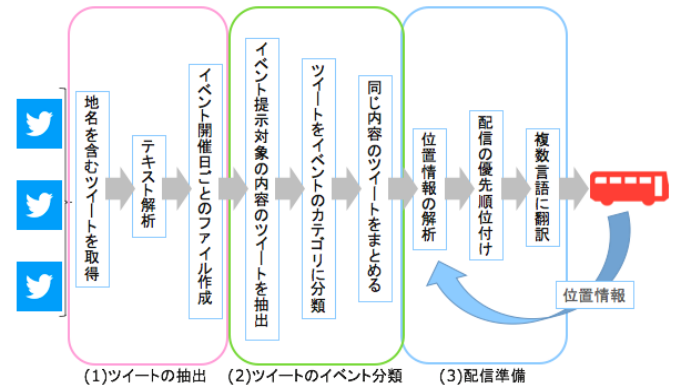


図 1 提案システムの概要

1. イベント提示対象の内容のツイートを抽出
2. 抽出されたツイートをイベントのカテゴリごとに分類
- (3) 配信準備

1. 提供する情報を複数言語に翻訳
2. ユーザと提供するイベントの位置情報を取得
3. 提供する情報を並び替え

4. テキスト解析

本章では、提案システムのツイート抽出機能の一部であるテキスト解析について説明する。

テキスト解析では、取得した JSON ファイルからテキスト、アカウント名、ID、RT 数、公式アカウントかどうか、フォロワー数を抜粋し、JSON ファイルにまとめ直す。さらに、イベント開催日時、開催場所をテキストを解析することにより取得する。

4.1 イベント開催日時の取得

キーワード検索で取得したツイートの本文に日付と時間が記載されているかを正規表現を使用して判定する。日付と時間が記載されているツイートを抽出し、イベント開催日ごとのファイルにまとめる。

4.2 イベント開催地の取得

本システムでは、山手線の駅名、東京 23 区の地名をキーワードに設定しツイートを取得している。イベント開催地として、キーワード検索で設定した地名を使用することは、地名の粒度が非常に大きいため望ましくない。そのため、本システムでは、イベントスポット辞書と正規表現を使用し、建物名や住所でイベント開催地を取得する。

イベントスポット辞書は、東京 23 区内のアミューズメント施設、ミュージアム、ショッピング施設、エンターテインメント施設、温泉、劇場、ホールを東京ウォーカー [11] とナビタイム [12] から抜粋し作成する。ツイートテキストにイベントスポット辞書内の要素が存在する場合は、その要素をイベント開催地として JSON ファイルに保持する。また、イベント開催地の住所がテキストに記載されている場合は、

正規表現を使用して住所を取得し、保持する。上述のどちらも記載されていない場合に対応するため、キーワード検索で設定した地名も JSON ファイルに保持する。

5. ツイートのイベント分類評価

本章では、提案システムの一部であるツイートのイベント分類について説明する。

場所と時間を含むツイートには観光者などに有用でないツイートも多く存在する。そこで、有用でないツイートを排除するためにツイートの分類を行う。また、観光者などに有用なツイートには、様々なジャンルが混在しており、ユーザーの嗜好にあった情報提供ができない。そのため、抽出した有用なツイートをカテゴリごとに分類する。

機械学習を用いた有用判定とカテゴリ分類の精度の評価実験の詳細を以下に示す。

5.1 カテゴリの設定と実験環境

本論文における「有用なツイート」とは以下のカテゴリのイベントに関するツイートとする。

音楽イベント	寄席・お笑い	舞台
映画	展覧会	ポケモン GO その他

カテゴリは、都市情報雑誌の代表格である「東京 Walker」[15] のイベントカテゴリを参考にし、予備実験で地名と日時を含むツイートの種類を分析した際に、一定数のツイートが得られると判明したカテゴリを設定した。

学習データは、2016 年 9 月 10 月に取得したツイートを人手で有用か有用でないか、と各カテゴリに分類したものを使用する。実験データには、2016 年 12 月 6 日～2016 年 12 月 13 日に取得したツイートの中で、イベント開催日が 12 月 13 日のツイート 1000 件を使用した。

機械的にツイートを分類するためには、ツイート本文を数値化する前処理が必要である。MeCab[17] を用いて必要な品詞を取り出し、Bag-of-words により数値化を行い、ツイートのベクトルモデルを作成する。ここで、MeCab により取り出す品詞は「名詞、動詞、形容詞、副詞」とした。MeCab の形態素辞書には、IPADIC[18] を使用している。

評価実験の検証方法としては、ツイートを正解のツイートと不正解のツイートに人手で分け、機械学習による判定結果の正答率と、正解のツイートと機械学習による判定結果との適合率、再現率、F 値を求める。

5.2 実験結果

5.2.1 Support Vector Machine

2 クラスの分類を行うことができる学習機械である SVM(Support Vector Machine) を用いて、ツイートのイベント分類を行った結果について説明する。SVM のソフトウェアには SVM-Light[16] を使用した。

SVM によるツイートの分類は、まず有用か有用でないか

の一对一分類を行う。次に有用なツイートをカテゴリごとに分類する一对他分類を行う。SVM は 2 クラスの分類であるため、カテゴリ分類では、あるカテゴリとその他のカテゴリ全部で SVM 判別モデルを構築していき、多カテゴリに分類する。今回はカテゴリを 6 個設定したので、6 個の SVM 判別モデルを作成し判定を行った。どのカテゴリにも判別されなかったツイートはその他のカテゴリとする。

有用か有用でないかに分ける一对一分類 (1)、カテゴリごとに分ける一对他分類 (2) の結果を表 1, 2 に示す。(1) の分類では、1000 件のうち 200 件で検証を行った。(2) の分類では、(1) で有用であると判定された 482 件のうち、カテゴリごとに 100 件で検証を行った。

表 1 Performance of classification(1).

Accuracy	Precision	Recall	F 値
0.8650	0.9091	0.8333	0.8857

表 2 Performance of classification for each category(2).

カテゴリ	Accuracy	Precision	Recall	F 値
音楽イベント	0.8700	0.8867	0.8703	0.8784
寄席・お笑い	0.9100	0.6500	0.8666	0.7428
舞台	0.8100	0.1363	1.0000	0.2399
映画	0.9300	0.4000	0.8000	0.5333
展覧会	0.9600	0.4285	1.0000	0.5999
ポケモン GO	1.0000	1.0000	1.0000	1.0000

(1),(2) の分類ともに高い Accuracy と Recall の結果が得られたが、表 2 の Precision と F 値には偏りが出てしまった。

「舞台」の Precision の値が極めて低いのは、「舞台」以外のカテゴリの SVM 判別モデルでも重みが大きくなる素性が、「舞台」の SVM 判別モデルの重みの大きな素性の中にも存在していることが原因であると考えられる。以下に「舞台」と「音楽イベント」の各々の SVM 判別モデルにおいて、重みの大きかった素性の一部を挙げる。

公演	席	舞台	シアター	演劇	朗読
上演	枚	ミュージカル	劇	演出	開演

ライブ	LIVE	演奏	ワンマン	予約
料金	コンサート	ピアノ	開演	アイドル

上記に列挙した素性の中には、「開演」という素性が共通して含まれている。このように複数のカテゴリの判別モデルに共通して出現する、重みの大きな素性が存在する場合、あるツイートが複数のカテゴリにおいて正解と判別されてしまい、分類の精度が下がってしまうと考えられる。そこで、カテゴリ「舞台」に多く混在していた [音楽イベント] との一对一分類の SVM 判別モデルを作成、分類し、「舞台」でないツイートを排除した。一对一分類で有用であると判定された 482 件のうち、一对他分類のカテゴリ分類で「舞台」に振り分けられた 106 件に、新しく作成した SVM 判別モ

表 3 Performance of classification(3).

カテゴリ	Accuracy	Precision	Recall	F 値
舞台	0.9439	0.3809	0.9411	0.5516

デルを適用した結果を表 3 に示す.

表 3 に示されているように,Recall の値が少し下がってしまったが,Precision と F 値の値を改善することができた.これより,カテゴリ「舞台」に多く混在していた「音楽イベント」を排除できたことが分かる.

次に,MeCab の形態素辞書を変更し,前述の実験と同じ条件で実験を行った.

前述の実験で使用していた形態素辞書である IPADIC は,2007 年 3 月以降更新されておらず,2007 年 3 月以降に流行った単語などは収録されていない. 実際に,カテゴリ「映画」の判別モデルにおいて重みが大きい素性である「六本木」「ヒルズ」や,「TOHO」「シネマズ」といった素性は,「六本木ヒルズ」「TOHO シネマズ」と分割されるべき素性であるが,IPADIC では,このような固有表現に対応しておらず,望ましくない形態素に分割されてしまっている. また,「舞台」のツイートに多く出現する「赤坂 ACT シアター」という固有表現は,2008 年 3 月に完成した劇場であるため,IPADIC では「赤坂」「ACT」「シアター」の 3 つの形態素に分割されてしまう. このように,IPADIC はしかるべき機関によって作成された信頼性の高い辞書であるが,収録されている単語が長い期間更新されていないため,流行語や新語,芸能人の名前や,新製品の名前,ランドマークの名称や通称が多く含まれる SNS に投稿されるような文章の解析には不向きな点もある. そこで,LINE の Data Labs で開発された辞書である,mecab-ipadic-NEologd[19] を併用することにした.

mecab-ipadic-NEologd は,多数の Web 上の言語資源から得た新語を追加することでカスタマイズした MeCab 用のシステム辞書であり,以下のような特徴があげられる.

- (1) IPADIC では複数の形態素に分割されてしまう固有表現を採録
- (2) 毎週 2 回以上更新
- (3) Web 上の言語資源を活用して更新時に新しい固有表現を随時追加

MeCab の形態素辞書に mecab-ipadic-NEologd を併用して実験を行った結果を表 4,5 に示す.

表 4 Performance of classification.

Accuracy	Precision	Recall	F 値
0.8750	0.9175	0.8396	0.8768

有用判定,カテゴリ分類ともに,前述の実験結果に比べ,全体的に F 値が高い結果が得られた.特に,表 2 では極端に低い値が出てしまっていたカテゴリ「舞台」の Precision と F 値については,一回の分類で表 3 の値に近づけることができ,mecab-ipadic-NEologd を併用したことによ

表 5 Performance of classification for each category.

カテゴリ	Accuracy	Precision	Recall	F 値
音楽イベント	0.8400	0.8484	0.9032	0.8750
寄席・お笑い	0.8600	0.5357	0.9375	0.6818
舞台	0.7200	0.3500	0.8750	0.5000
映画	0.9400	0.5454	0.8571	0.6667
展覧会	0.8300	0.3703	1.0000	0.5405
ポケモン GO	1.0000	1.0000	1.0000	1.0000

る効果が見られた.この先の実験では,MeCab の辞書には mecab-ipadic-NEologd を併用することとする.

5.2.2 ランダムフォレスト

本節では,分類器にランダムフォレストを用いてツイートのイベント分類を行った結果について述べる.

ランダムフォレストとは,決定木を弱学習器とする集団学習アルゴリズムである.2 クラスの分類を行う SVM は,各クラスごとに学習を行う必要があり,本研究では有用判定を行った後に,6 個の SVM を作成してカテゴリ分類を行っていたが,ランダムフォレストは多クラスの分類を行うことが可能であるため,有用判定を行わず,「有用でない」「音楽イベント」「寄席,お笑い」「舞台」「映画」「展覧会」「ポケモン GO」の 7 クラスの分類を一度に行うことにした.

分類は,Python の機械学習ライブラリである scikit-learn[20] を用いて行った.

学習データ,実験データ,検証方法の条件は SVM を用いた実験の条件と変えていない.検証件数は,各カテゴリごとに実験データ 1000 件のうち 400 件で行った.

ランダムフォレストによるカテゴリ分類の結果を表 6 に示す.

表 6 Performance of classification for each category.

カテゴリ	Accuracy	Precision	Recall	F 値
音楽イベント	0.8950	0.8750	0.7583	0.8125
寄席・お笑い	0.9550	0.7948	0.7561	0.7750
舞台	0.9875	0.5000	0.4000	0.4444
映画	0.9925	1.0000	0.7272	0.8421
展覧会	0.9875	0.8666	0.8125	0.8387
ポケモン GO	1.0000	1.0000	1.0000	1.0000

表 6 に示されている通り,カテゴリ「舞台」以外では,F 値が 0.8 前後になり,高い精度で分類が行われている.カテゴリ「舞台」においても,SVM の場合と同等の結果を得ることができた.

SVM とランダムフォレストのどちらの分類器を用いて分類を行った場合でも,最終的には高い精度が得られることが分かったが,両者を比較すると,ランダムフォレストの方が F 値が全体的に高い結果が得られ,精度がより優れていた.この精度は,学習データを増やすこと,高頻度すぎる単語,低頻度すぎる単語を削除する等の手法によって,更に高めることができると考える.

また,処理速度についても,SVM のように総当たりでカテゴリ分類を行う必要のないランダムフォレストの方が,

処理にかかる時間が少なく優れていた。

本研究では、情報をリアルタイムに処理してユーザに提示することを想定しているため、ツイートのイベント分類にはランダムフォレストを適用することにする。

6. 配信準備

本章では、システムの一部である配信準備について、情報の多言語化と配信の優先順位付けに分けて説明する。

6.1 情報の多言語化

インバウンドに対応するため、ツイート本文を多言語に翻訳する必要がある。しかし、ツイート本文は不完全な日本語で書かれていることも多く、ツイート本文を原文のまま翻訳すると正しく翻訳されないことが多い。そのため、ツイート本文中の情報を予め整理し、必要な情報のみを提供する。整理した情報を情報通信研究機構が開発した自動翻訳エンジン「みんなの自動翻訳@ TexTra」[13]を使用して翻訳する。翻訳する言語は、訪日外国人の国籍別の割合[14]を基に、英語、中国語(簡体字、繁体字)、韓国語を選んだ。情報を整理する前処理を行わなかった英語翻訳の例を図2に示す。ツイートの情報を整理する前処理を施した例を図3に示す。図3を英語に翻訳した例を図4に示す。

〈紙わざ大賞26〉12/13(火)~17(土)10:00~19:00(最終日18:00まで)
東京交通会館 東京都千代田区有楽町2-10-1 私制作の紙わざ大賞 歴代入
選作品 今年26回〈南天雄鶏像〉25回〈飛翔-白鷹〉24回〈ヴェ…
<https://t.co/zhblqq79ZH>

<Paper's Awards 26> 12/13 (Tue) 17 (sat) 10:00 to 19:00 (until the final day 18:00) 2-10-1 I make Tokyo Traffic Hall in Chiyoda Ward, Tokyo, Yurakucho Awards past winners of this year and 26 times southern rooster image> 25 times flying-Shirataka> 24 times ヴェ… <https://t.co/zhblqq79ZH>

図2 Result of translation(full text).

イベント名: 〈紙わざ大賞26〉
日付: 12/13
時間: 10:00
場所: 東京都千代田区有楽町2-10-1
移動時間(driving): 16分
カテゴリ: 展覧会

図3 Result of organizing data.

Event Name: <Paper's Award (26)>
Date: 13/12
Time: 10 00
Location: Yurakucho, Chiyoda-ku, Tokyo 2-10-1
Travel Time (driving): 16 minutes
Category: exhibition

図4 Result of translation(英語).

これより、図2のように原文のままでは正しく翻訳できず、複数箇所で誤った翻訳結果になる場合でも、図3のような前処理を施すことで必要な情報を簡潔に正しく提示できることが示された。

6.2 配信の順位付け

場所と日付と時間の条件に適合する情報を、ユーザにとって有益な順で提示するために、情報の順位付けを行う。どのような順で情報提示すればユーザにとって有益なのかは、ユーザによって異なり、例えば、順位付けの条件として、イベントの開催地や人気度、開催時間、金額が考えられる。

今回は、その中の一例として、Google Maps API の Distance Matrix API[21] を使用し、イベント開催地に着目した順位付けを行った。Google Maps API の Distance Matrix API[21] は、出発地点と到着地点間の推奨ルートに基づいた時間と距離の値が取得できる API である。本システムでは空港と都内を結ぶリムジンバスの乗客をユーザに想定しているため、ユーザの位置情報や目的地が取得可能である。このリムジンバスより取得する地点を出発地点、4章において各ツイートごとに JSON ファイルに格納したイベントスポット辞書の要素、住所、キーワード検索で設定した地名のいずれかを到着地点に設定する。Distance Matrix API から各イベントについて取得した移動時間の小さい順に情報を並びかえる。

ユーザの位置、目的地に「新宿駅」を選択し、カテゴリ「寄席/お笑い」の情報に順位付けを行った結果の一部を図5に示す。なお、実験データとして、2017年2月18日から2月26日の期間で、山手線の駅名をキーワードに Twitter API のキーワード検索で取得したツイートのうち、イベント開催日が2017年2月26日のツイートを使用した。

7. まとめと今後の課題

ガイドブックや検索エンジンなどから容易に取得できない情報、場所と時間を考慮した「その場」「その時」に観光者などが利用出来る有用性の高い情報を Twitter から抽出し、ユーザの嗜好に合った情報提示を視野に入れ、ツイートのイベント分類を行った。ツイートのイベント分類では、SVM とランダムフォレストで比較を行い、多カテゴリに分類できるランダムフォレストが、分類精度、処理速度どちらにおいても優れていることが分かった。また、インバウンド向け多言語で配信を実現するために、情報を整理し、高い精度での翻訳を可能にした。そして、ユーザにとって有益な順で情報を提示するために、提示する情報の順位付けを行った。情報の順位付けでは、ツイートテキストを解析した結果を用い、より正確な位置情報に基づいた移動時間を算出可能にした。

今後の課題としては、イベント分類、イベント名の抽出の精度向上があげられる。また、現在のシステムでは、地名と日付と時間の3点をツイートの本文に含むもののみを対象としており、解析対象のツイートの種類が少ないため、この3点の情報の一部が欠けている場合に対処する手法を考え、解析するツイートの種類を増やしたい。

イベント名: 【弱井龍免・再任反対運動のしめくり集会を開催します】
日付: 2/26
時間: 13時30分
場所: 東京都渋谷区代々木2丁目5-5
移動時間(driving): 7分
カテゴリ: 寄席・お笑い

イベント名: 『トータルテンボスの「大貧民」
日付: 2/26
時間: 19:00
場所: ルミネtheよしもと
移動時間(driving): 9分
カテゴリ: 寄席・お笑い

イベント名: 「real/dummy」
日付: 2/26
時間: 15:30
場所: 原宿
移動時間(driving): 9分
カテゴリ: 寄席・お笑い

日付: 2/26
時間: 13:30
場所: 新宿パッショ
移動時間(driving): 9分
カテゴリ: 寄席・お笑い

⋮

日付: 2/26
時間: 19:30
場所: 新大久保
移動時間(driving): 11分
カテゴリ: 寄席・お笑い

イベント名: 『事務所対抗 先輩後輩 漫才チーム戦』
日付: 2/26
時間: 18:30
場所: ラスタ原宿
移動時間(driving): 11分
カテゴリ: 寄席・お笑い

イベント名: 「さんびーち泣かせ」
日付: 2/26
時間: 15:30
場所: ラスタ原宿
移動時間(driving): 11分
カテゴリ: 寄席・お笑い

イベント名: 【夜はこちらだよ!】
日付: 2/26
時間: 18:30
場所: ラスタ原宿
移動時間(driving): 11分
カテゴリ: 寄席・お笑い

⋮

図 5 Result of prioritized information.

- 動モデルの構築と観光地推薦手法の提案”, 人工知能学会
合同研究会 2015 第 9 回 データ指向構成マイニングとシ
ミュレーション研究会
- [5] 長谷川 馨亮, 馬 強, 吉川 正俊, ”Twitter からの地域
特徴語辞書の構築と その観光情報検索への応用”, DEIM
Forum 2014 B3-4.
- [6] 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”, DEIM
Forum 2016 H1-3.
- [7] 佐伯 圭介, 村山 敬祐, 遠藤 雅樹, 横山 昌平, 石川
博, ”外国人向け観光プロモーション支援のためのソーシャ
ルメディア分析手法の研究”, DEIM Forum 2014 B4-2.
- [8] Chao Zhang, Guangyu Zhou, Quan Yuan, Honglei
Zhuang, Yu Zheng, Lance Kaplan, Shaowen Wang, and
Jiawei Han, ”GeoBurst: Real-Time Local Event Detec-
tion in Geo-Tagged Tweet Streams”, Proceedings of the
39th International ACM SIGIR Conference on Research
and Development in Information Retrieval.
- [9] 山口 祐人, 伊川 洋平, 天笠 俊之, 北川 博之, ”ソーシャ
ルストリームからのイベント検出とユーザ位置推定の統合”,
DEIM Forum 2013.
- [10] ”Twitter Serch API,” [https://dev.twitter.com/
rest/public/search](https://dev.twitter.com/rest/public/search)
- [11] ”Walker+ 東京都のおでかけスポット一覧,” [http:
//www.walkerplus.com/spot_list/ar0313/](http://www.walkerplus.com/spot_list/ar0313/)
- [12] ”NAVITIME,” [https://www.navitime.co.jp/
category/0106003/13/](https://www.navitime.co.jp/category/0106003/13/)
- [13] ”みんなの自動翻訳 @ TexTra” [https://
mt-auto-minhon-mlt.ucrjgn-x.jp](https://mt-auto-minhon-mlt.ucrjgn-x.jp)
- [14] 日本政府観光局 「国籍/月別 訪日外客数 (2003 年～2016
年)」: [http://www.jnto.go.jp/jpn/statistics/since2003-
tourists.pdf](http://www.jnto.go.jp/jpn/statistics/since2003-tourists.pdf)
- [15] ”Walker+” <http://www.walkerplus.com>
- [16] ”SVM light” <http://svmlight.joachims.org>
- [17] ”MeCab,” <http://mecab.sourceforge.net/>
- [18] ”IPADIC,” <https://ja.osdn.net/projects/ipadic/>
- [19] ”mecab-ipadic-NEologd,” [https://github.com/
neologd/mecab-ipadic-neologd/blob/master/
README.ja.md](https://github.com/neologd/mecab-ipadic-neologd/blob/master/README.ja.md)
- [20] ”scikit-learn Machine Learning in Python,” [http://
scikit-learn.org/stable/index.html](http://scikit-learn.org/stable/index.html)
- [21] ”Distance Matrix API,” [https://developers.google.
com/maps/documentation/distance-matrix/?hl=ja](https://developers.google.com/maps/documentation/distance-matrix/?hl=ja)

8. 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 「年別訪日外客数、出国人数の推移」:
[http://www.jnto.go.jp/jpn/statistics/marketingdata-
outbound.pdf](http://www.jnto.go.jp/jpn/statistics/marketingdata-outbound.pdf)
- [3] 訪日外国人向け O2O マーケティングも視野に iPhone・
Android 両スマートフォン向け新規技術「BeaconCast」
アプリを用いて 高速リムジンバス内での情報配信を行
う実証実験を共同で実施 [http://www.u-tokyo.ac.jp/
public/public01-261023-j.html](http://www.u-tokyo.ac.jp/public/public01-261023-j.html)
- [4] 前田 高志ニコラス, 吉田 光男, 鳥海 不二夫, 大橋
弘忠, ”Twitter 位置情報・テキスト情報を用いた人の移