

旅行者の購買行動プロセスに関連するツイートの自動抽出

藤村 雄偉[†] 石野 亜耶[†] 難波 英嗣[‡] 竹澤 寿幸[‡]

[†] 広島経済大学 ビジネス情報学科 〒731-0138 広島県広島市安佐南区祇園 5-37-1

[‡] 広島市立大学大学院 情報科学研究科 〒731-3194 広島県広島市安佐南区大塚東 3-4-1

E-mail: [†] ay-ishino@hue.ac.jp, [‡] {nanba, takezawa}@ls.info.hiroshima-cu.ac.jp

あらまし 一般的な商品を購入する際の消費者の購買行動プロセスは、AIDMA モデルや AISAS モデルが知られている。これまでも、旅行商品の購買行動プロセスも研究されているが、主に旅行先を決定するためのプロセスに限定されている。しかし、個人旅行においては、旅行先が決定した後も、どこに宿泊するか、何を食べるかなど、旅行に関連する商品の購買行動の連続である。本研究では、個人旅行者がどのようなプロセスを経て旅行関連商品を選択しているのかを明らかにするために、旅行を計画中のユーザが Twitter に投稿した購買行動プロセスに関連するツイートを自動で抽出する手法を提案する。

キーワード 観光情報処理, 旅行者, 購買行動, Twitter

1. はじめに

観光地のマーケティングを効果的に行うためには、旅行者が、旅行先などの旅行に関連する商品をどのように選択しているのかという購買行動プロセスを明らかにする必要がある。一般的な商品を購入する際のユーザの購買行動プロセスのモデルとして、AIDMA モデルや AISAS モデルが知られている。このようなモデルを、旅行に関連する商品の購買行動プロセスに応用した研究が行われているが、主に旅行先を決定するためのプロセスに限定されている。しかし、個人旅行においては、旅行先以外にも、どこに宿泊するか、どこで何を食べるかなど、旅行に関連する商品の購買行動の連続である。本研究では、個人旅行者がどのようなプロセスを経て旅行関連商品を選択しているのかを明らかにするための情報源として、Twitter に注目する。

Twitter には、旅行中の体験や感想だけではなく、旅行の計画中の情報を投稿するユーザも存在する。このようなユーザを発見し、旅行の計画から実行までの一連のツイートを収集することができれば、旅行者の購買行動プロセスを明らかにすることが可能であると考えられる。

そこで本研究では、旅行者の購買行動プロセスを明らかにするための第一歩として、旅行を計画中のユーザが Twitter に投稿した購買行動プロセスに関連するツイートを自動で抽出する手法を提案する。

本論文の構成は以下の通りである。2 節では関連研究, 3 節では提案手法, 4 節では実験結果について述べ, 5 節で本稿をまとめる。

2. 関連研究

本節では、本研究に関連する研究を紹介する。本研究の最終的な目標は、Twitter に投稿された旅行者の購買行動プロセスに関連するツイートを利用することで、旅行者の購買行動プロセスを明らかにすることである。2.1 節では、購買行動プロセスに関連する研究を紹介する。2.2 節では、Twitter などのソーシャルメディアを利用した購買行動プロセスに関連する研究、2.3 節では、ソーシャルメディアを情報源とした観光情報の抽出または分析に関連する研究を紹介する。

2.1. 購買行動プロセスに関連する研究

一般的な商品を購入する際の消費者の購買行動プロセスの古典的なモデルとしては、AIDMA モデルが知られている。AIDMA モデルは、消費者の購買行動を、「注意 (Attention)」⇒「興味・関心 (Interest)」⇒「欲求 (Desire)」⇒「記憶 (Memory)」⇒「行動 (Action)」というプロセスで表している。

近年では、インターネットの普及や、Google や Yahoo! などの検索エンジンの登場により、消費者にとって情報検索が容易になった。またソーシャルメディアの利用が一般的になり、商品を購入した消費者自身が、ブログや Twitter など、使用感や感想などの情報発信や情報共有を行うようになった。このような、新しい消費者の購買行動として、電通が AISAS モデル¹を提案している。AISAS モデルでは、「注意 (Attention)」⇒「興味・関心 (Interest)」⇒「検索 (Search)」⇒「行動 (Action)」⇒「共有 (Share)」という 5 段階のプロセスが想定されている。

旅行商品の購買行動プロセスも研究されているが、

¹ 商標登録番号第 4874525 号

主に旅行先を決定するためのプロセスに限定されている。旅行先を商品としてとらえた購買行動プロセスのモデルとして、Schmoll のモデル[1]がある。Schmoll のモデルでは、「旅行欲求 (Travel desire)」⇒「情報収集 (Info search)」⇒「選択肢の検討 (Travel alternatives)」⇒「決定 (Decision)」というプロセスで説明されている。このモデルは、旅行者の旅行先決定までの購買行動プロセスとそれらに影響する要因を示すフレームワークとして多くの研究で引用されているが、そのプロセスを具体的に説明するには至っていないことが指摘されている[2]。このように、旅行先を商品ととらえ、購買行動プロセスを明らかにするための研究が行われている。

観光庁の調査によると、訪日外国人の 28.6%が団体ツアー、71.4%が団体ツアーではない（個人旅行）形態での来訪である[3]。個人旅行においては、旅行先以外にも、どこに宿泊するか、どこで何を食べるかなど、旅行に関連する商品の購買行動の連続である。そのため、旅行先だけでなく旅行に関連する様々な商品の購買行動プロセスを明らかにする必要があるが、我々が調査した範囲では、このような研究は行われていない。そこで本研究では、個人旅行者がどのようなプロセスを経て旅行関連商品を選択しているのかという旅行に関連する商品の購買行動プロセスを明らかにするために、Twitter に着目し、旅行者の購買行動プロセスに関連するツイートを自動抽出する手法を提案する。

近年、Twitter ユーザの属性（性別、年齢など）を自動的に推定する研究が活発に行われている[4, 5]。このような技術を利用し、Twitter ユーザの属性と、そのユーザの購買行動プロセスを組み合わせて分析することで、より詳細なマーケティングが可能になると考えられる。

2.2. ソーシャルメディアを利用した購買行動プロセスに関連する研究

岩田ら[6]は、Twitter などのソーシャルメディアに、特定の商品やサービスに対して自身の購買行動に関する状態（〇〇欲しい、××買った等）を投稿するユーザが増えていることに着目し、広告配信や意見分析のために、ツイートからユーザの購買行動プロセスのフェーズを自動推定する手法を提案している。購買行動プロセスのフェーズは、AIDMA モデルの「注意 (Attention)」を「認知フェーズ」、「興味・関心 (Interest)」、「欲求 (Desire)」、「記憶 (Memory)」を「感情フェーズ」、「行動 (Action)」を「行動フェーズ」としている。提案手法では、同一ユーザの投稿を時系列に追跡し、各フェーズに特徴的な単語を手掛かり語とし、分類器 CNB (Complement Naïve Bayes) を用いてフェーズを

推定している。Twitter を利用してユーザの購買行動プロセスを明らかにする点では本研究と類似しているが、岩田らは、通信端末や車などの一般的な商品を対象としており、本研究は旅行に関連する商品を対象としている点で異なる。

ソーシャルメディアから、「~したい」のようなユーザの購買意図を抽出するための研究がある。購買意図を抽出するための情報源として Twitter を利用した研究[7, 8]、商品レビューを利用した研究[9, 10]がある。購買意図は、購買行動プロセスの一部であると考えられる。

2.3. ソーシャルメディアを情報源とした観光情報の抽出または分析に関連する研究

Twitter を観光情報の情報源として利用する研究は、数多く行われている。佐伯ら[11]は、外国語で投稿している Twitter のユーザが、訪日外国人なのか、在日外国人なのかを判別する手法を提案している。また、日本人ユーザ、訪日外国人ユーザ、在日外国人ユーザがツイートしている場所を明らかにし、その結果に差異があったことを報告している。

土田ら[12]は、Twitter から収集した都市・地域やランドマークの名称を含むツイートに対して、Word2Vec を用いることで、ある地域のある観光スポットに対する他の地域の観光スポットとの関係性を抽出する手法を提案している。

藤井ら[13]は、観光地の特徴を明らかにするために、旅行ブログエントリを投稿したブロガーの属性（「年齢」と「使用言語」と各旅行ブログエントリの属性（観光タイプ）を自動判定する手法を提案している。観光タイプは、表 1 に示す「見る」、「体験する」、「買う」、「食べる」、「泊まる」、「その他」に分類している。本研究では、旅行者の購買行動プロセスに関連するツイートは、藤井らの設定した観光タイプに分類する予定である。

表 1: 観光タイプとその内容

タイプ	内容
見る	観光名所などの見て楽しめる物やイベントについての情報。
体験する	〇〇体験やスキューバダイビングなど、自分の体を使って楽しめる物についての情報。
買う	土産物に関する情報。
食べる	飲食に関する情報。
泊まる	宿泊施設に関する情報。
その他	「見る」、「体験する」、「買う」、「食べる」、「泊まる」に該当しない場合。例として広告ページや巻末の交通情報。

3. 旅行者の購買行動プロセスに関連するツイートの自動抽出手法

本研究では、旅行者の購買行動プロセスに関連するツイートを自動抽出する手法を提案する。提案手法の流れを図 1 に示す。まず、旅行計画者の候補となる Twitter ユーザを収集する (図 1 の①)。次に、旅行計画者であるユーザを自動判定する (図 1 の②)。最後に、旅行計画者であるユーザのツイートから、購買行動プロセスに関連するツイートを自動抽出する (図 1 の③)。

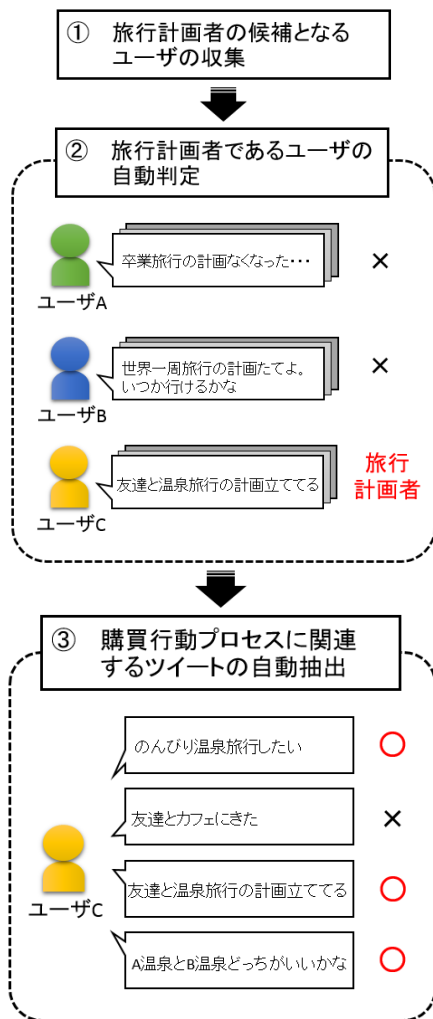


図 1: 旅行者の購買行動プロセスに関連するツイートの自動抽出手法の流れ

本研究では、旅行計画者の候補となるユーザの収集と、旅行計画者であるユーザの自動判定を行う。それぞれの手法については、3.1 節、3.2 節で説明する。

3.1. 旅行計画者の候補となるユーザの収集

まず、旅行を計画していると考えられる Twitter ユーザを収集する。旅行を計画しているユーザは、「旅行」や「計画」といった単語を含んだツイートをすると考

えられる。そのため、まず、「旅行」や「計画」というキーワードを含むツイートを収集する。ただし、リツイートは除外することとする。具体的には、「旅行 計画 -RT」を検索キーワードとしてツイートを検索し収集する。また、これらのツイートを投稿したユーザの情報も収集する。ツイートとユーザの情報は、Twitter API²を利用して収集する。

3.2. 旅行計画者であるユーザの自動判定

3.1 節で収集したツイートには、「旅行」や「計画」という単語は含まれるが、実際には旅行を計画していないツイートも多く含まれる。旅行を計画中でであると判定されるツイートを図 2、旅行を計画中でないと判定されるツイートを図 3 に示す。

- ・ 大学院の 3 人での卒業旅行計画(笑)バリ島 6 日間の旅!! トランジットで 8 時間韓国で観光できるし、大自然と海で豪遊する予定
- ・ 九州旅行を計画中で、〇〇と別府温泉に泊まりたいと思っていますが、どの温泉宿が良いですか? おすすめの宿とか、ご存知の方、アドバイスお願いします
- ・ 温泉旅行計画してるけど熱海か伊豆どちらにしようかな?

図 2: 旅行を計画中でであると判定されるツイート

- ・ 隣に座ってるカップルがドイツ旅行に行く計画立ててるの聞いてわくわくしてる
- ・ 早いとこ卒論終わらせて旅行の計画とか立てたい
- ・ 年単位で...ヨーロッパ旅行計画を立てることを誓う。早くて 25 歳前後で行けたらな……

図 3: 旅行を計画中でないと判定されるツイート

図 3 に示すツイートのように、ツイートしたユーザ自身が旅行を計画していない場合や、旅行の計画を始めていない場合、漠然とした夢として旅行を計画している場合は、旅行を計画中でないと判定する。

本研究では、機械学習により、旅行を計画中であるツイートを自動判定する。機械学習には、単語と以下の手掛かり語の有無を特徴量として与える。

- ・ 地名
- ・ 人を表す単語
- ・ 旅行の種別に関連する単語 (卒業旅行など)

地名や人を表す単語は、CaboCha³により判定する。上記の手法により、旅行を計画中でであると判定されるツイートを投稿したユーザを、旅行を計画中のユーザであると判定する。

² <https://code.google.com/archive/p/python-twitter/>

³ <http://chasen.org/~taku/software/cabocho/>

4. 実験

本研究で提案した手法の有効性を確認するため、実験を行った。

4.1. 実験方法

【実験に用いるデータ】

3.1 節で説明した方法で、2017 年 1 月 8 日から 1 月 16 日の期間にツイートとそのユーザ情報を収集した。その結果、5,148 件のツイートと 4,201 件のユーザ情報を収集できた。これらのデータから、以下の条件に当てはまるユーザはボットである可能性があるため、実験に使用するデータから除くこととした。

- ・ ユーザ名の始まりや終わりが「bot」であるユーザ
- ・ 同一のツイートを繰り返しツイートするユーザ

上記のデータから 2,000 件のツイートを選択し、人手により、旅行を計画中のツイートであるかどうかを判定した。その結果を表 2 に示す。機械学習には、旅行計画中心であると判定されたツイート、旅行計画中心でないと判定されたツイートをそれぞれ 500 件ずつ使用した。

表 2: 旅行を計画中のツイートであるかどうかを人手で判定した結果

旅行計画中心であると判定されたツイート	旅行計画中心でないと判定されたツイート	合計
692	1,308	2,000

【機械学習と評価尺度】

旅行を計画中心であるツイートの自動判定の機械学習には SVM を用いた。SVM には、機械学習ライブラリの scikit-learn⁴ を用いた。2 次の多項式カーネルを使用し、5 分割交差検定を行った。評価尺度として、以下に示す精度、再現率、F 値を用いた。

$$\text{精度} = \frac{\text{システムが検出した正解件数}}{\text{システムが検出した件数}}$$

$$\text{再現率} = \frac{\text{システムが検出した正解件数}}{\text{人手で判定した正解件数}}$$

$$\text{F 値} = \frac{2 \times \text{精度} \times \text{再現率}}{\text{精度} + \text{再現率}}$$

【比較手法】

単語の有無を特徴量とした場合を比較手法、単語と 3.2 節で説明した手掛かり語の有無を特徴量とした場合を提案手法として実験を行った。

4.2. 実験結果と考察

実験を行った結果を表 3 に示す。比較手法に比べ、提案手法では、精度は 0.07 ポイント、再現率は 0.03 ポイント、F 値は 0.06 ポイントとわずかではあるが改善することができた。

表 3: 実験結果

	精度	再現率	F 値
提案手法	0.65	0.82	0.73
比較手法	0.58	0.79	0.67

本研究では、1 件のツイートに対し、旅行を計画中のツイートかどうかを自動判定した。旅行を計画中のユーザは、対象としたツイートの前後にも、「ガイドブック買った」や「飛行機を予約した」など、旅行の計画に関連するツイートを行う場合も多々ある。そのため、今後は、対象となるツイートの前後のツイートの情報も考慮することで、更なる精度向上を目指す。

5. おわりに

本研究では、旅行者がどのようなプロセスを経て旅行関連商品を選択しているのかを明らかにするために、Twitter に着目し、旅行者の購買行動プロセスに関連するツイートを自動抽出する手法を提案した。旅行を計画中のユーザの自動判定について実験を行い、精度 0.65、再現率 0.82、F 値 0.73 で判定することができた。

今後は、購買行動プロセスに関連するツイートに対し、「旅行先を検討している」や「宿泊ホテルを予約した」など、購買行動プロセスの具体的な段階を自動判定する手法を検討する。

謝 辞

本研究は、JSPS 科研費 JP16K16679 の助成を受けたものです。

参 考 文 献

- [1] Gottfried A. Schmol, “Tourism promotion: marketing background, promotion techniques and promotion planning methods”, Tourism International Press, 1977.
- [2] 大方優子, “旅行先選択行動に関する考察”, 東海大学福岡短期大学紀要, 2015.
- [3] 観光庁, “訪日外国人の消費動向”, 平成 25 年 7-9 月期報告書, 2013.
- [4] 池田和史, 服部元, 松本一則, 小野智弘, 東野輝夫, “マーケット分析のための Twitter 投稿者プロフィール推定手法”, 情報処理学会論文誌 コンシューマ・デバイス & システム (CDS), Vol.2, No.1,

⁴ <http://scikit-learn.org/stable/index.html>

pp.82-93, 2012.

- [5] John D. Burger, John Henderson, George Kim, Guido Zarrella, “Discriminating Gender on Twitter”, Proc. of the Conference on Empirical Methods in Natural Language Processing, 2011.
- [6] 岩田麻佑, 服部元, 池田和史, 滝嶋康弘, “変化点に着目した Twitter ユーザの購買行動フェーズ推定手法”, 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(DEIM Forum 2015), 2015.
- [7] Bernd Hollerit, Mark Kröll, Markus Strohmaier, “Towards Linking Buyers and Sellers: Detecting Commercial Intent on Twitter”, Proc. of the 22nd International Conference on World Wide Web, 2013.
- [8] Jinpeng Wang, Gao Cong, Wayne Xin Zhao, Xiaoming Li, “Mining User Intents in Twitter: A Semi-Supervised Approach to Inferring Intent Categories for Tweets”, Proc. of the 29th AAAI Conference on Artificial Intelligence, 2015.
- [9] 石野亜耶, 村上浩司, 関根聡, “商品レビューからの購買意図の抽出とそれを用いた商品検索システムの構築”, 言語処理学会第 20 回年次大会, 2014.
- [10] 金子貴美, 村上浩司, 石野亜耶, “商品レビューからの購買意図の抽出と自動階層化”, 2016 年度人工知能学会全国大会, 2016.
- [11] 佐伯圭介, 遠藤雅樹, 廣田雅春, 倉田陽平, 横山昌平, 石川博, “外国人 Twitter ユーザの観光訪問先の属性別分析”, 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(DEIM Forum 2015), 2015.
- [12] 土田崇仁, 遠藤雅樹, 加藤大受, 江原遥, 廣田雅春, 横山昌平, 石川博, “Word2Vec を用いた地域やランドマークの意味演算”, 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(DEIM Forum 2016), 2016.
- [13] 藤井一輝, 難波英嗣, 竹澤寿幸, 石野亜耶, “旅行プロダクトの属性情報に基づいた旅行者の行動分析”, 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(DEIM Forum 2015), 2015.