

コロナ禍イベント参加者性格特性を考慮した イベント推薦システムの提案

伊藤 桃[†] 榎 美紀^{††} 小口 正人[†]

[†] お茶の水女子大学 〒112-8610 東京都文京区大塚 2-1-1

^{††} 日本アイ・ビー・エム株式会社 〒103-8510 東京都中央区日本橋箱崎町 19-21

E-mail: [†]momo-i@ogl.is.ocha.ac.jp, enomiki@jp.ibm.com, oguchi@is.ocha.ac.jp

あらまし 昨今新型コロナウイルスが、世界中に多大な影響を与え続けている。驚異的な感染力により未だにウイルスの収束も見えず、現在も外出を控えた日常生活を送っている人が多いだろう。イベント業界においては、イベントオンラインと現地参加いずれかの方法での参加が選択できるハイブリットイベントが新しい参加形態のイベントとしてコロナ禍に現れた。それにより、従来のイベントとは異なり、同じイベント参加者の中であっても、現地参加かオンライン参加かでそれぞれ参加するユーザ層に違いが生じている可能性があり、それらを加味したシステムの構築が必要である。本論文では、高パーソナライズ性システムを構築するためにまず、ソーシャルメディアと性格推定ツールを利用して、コロナ禍の日本のイベントに参加するユーザの性格値傾向を検証する。結果として、オンラインで参加するか、現地で参加するかの性格特性に関係がありあそうな「外交的性格度」の値に差がみられることを発見した。

キーワード ソーシャルメディア, Personality Insights, 性格値

1 はじめに

昨今新型コロナウイルスが、世界中に多大な影響を与え続けている。驚異的な感染力により未だにウイルスの収束も見えず、現在も外出を控えた日常生活を送っている人が多いだろう。しかしながら、収束後はその反動から、以前よりイベントやエンタメ業界の需要が再び盛り上がりを見せることは容易に想像でき、その際に、多様なイベント推薦システムの需要も共に高まってくると考える。現在既に、有名なイベントは様々な Web サイトなどから簡単に情報を取得でき、AI を使ったイベント推薦システムなども増えてきた。これらのシステムの主流はユーザの趣味嗜好情報からイベントを推薦する。しかし、そのような既存の推薦システムは、似かよった趣味嗜好のイベントを推薦する傾向があり、ユーザにとって単一的な推薦になりがちである。

本研究では、ユーザの性格特性情報からイベントを推薦するという新たなアプローチを提案する。性格を考慮することで、単に趣味趣向情報から推薦するようなシステムでは参加することがなかったイベントの発見ができ、よりイベントの選択肢を増やすことが期待できる。また、オンラインと現地参加いずれかの方法での参加が選択できるハイブリットイベントが新しい参加形態のイベントとしてコロナ禍に現れた。それにより、従来のイベントとは異なり、同じイベント参加者の中であっても、現地参加かオンライン参加かでそれぞれ参加するユーザ層に違いが生じている可能性があり、それらを加味したシステムの構築が必要である。本論文では、システムを構築するためにまず、ソーシャルメディアと性格推定ツールを利用して、コロナ禍の日本のイベントに参加するユーザの性格値傾向を検証する。結

果として、オンラインで参加するか、現地で参加するかの性格特性に関係がありあそうな「外交的性格度」の値に差がみられることを発見した。

本論文の構成は以下のとおりである。2 章で関連研究について述べ、3 章で先行研究について述べる。4 章では提案システムの概要を紹介し、5 章で本研究で使用する PI について具体的に説明する。6 章で本論文の実験で使用するデータについて述べ、7 章で分析を行う。最後に、8 章で本稿をまとめる。

2 関連研究

本研究では、Twitter にて情報収集を行いそこで得たテキスト情報からユーザの性格値を判断する。Jalal ら [1] や, Gou ら [2] が Twitter のテキスト情報から性格を判断するその正確性について説いている。Jalal らは、空港の待ち時間などその場にはいないと分からないローカルな情報を Twitter のユーザから収集する qCrowd というシステムを構築したいと考え、Twitter の情報からローカル情報を提供してくれるユーザか否かを判断するモデルを作成した。結果としては、65%以上の確率で正しくユーザをモデルが判別した。また、Gou らも、同じように Twitter の情報から最も主流な心理学的な性格の指標である Big Five Personality に基づいてユーザの性格値を判定するモデルを作成しており、正確性を立証している。

また本研究では、テキスト情報から性格値を判断するのに独自のモデルではなく、IBM Personality Insights(以下 PI) [3] という IBM のサービスを用いる。このサービスは Twitter などのユーザが書いたテキスト情報をインプットとして、性格の特性を、Big five(個性), Needs(要求), Value(価値感) の 3 つの次元に分割して出力する。PI を用いた結果を検証している研究

論文として、富永 [4] の論文を挙げる。この論文では、Twitter のユーザの時間とともに変化する人格を、PI を用いて分析している。結果として、人格特性の変化が感覚的に理解のできる要因によって変化していて、PI の正確性を裏付けるような結果であると解釈している。先述した従来の研究から、Twitter から取得したデータから正確性の比較的高い性格値の判断ができることが言える。

また、Twitter から収集したデータからコロナウイルス禍における人々の感情分析を行なった研究は既に複数存在するので紹介したい。例えば鳥海 [5] らはコロナ関連のキーワードを含んだツイートを長期的に収集し、感情表現辞書を用いて時系列的にどのような感情の変化があるかを分析している。その結果、コロナに関するニュースや報道を受けて、恐怖の感情数値が上下している事が確認された。これにより、ソーシャルメディアテキストから比較的適切に感情分析ができる事わかる。四方田も [6] 同様に、コロナウイルスに関連する単語を複数グルーピングし、どの要因からコロナウイルスに関するツイートを行なったのかを長期的に分析している。本研究とこれらの従来研究の相違点として、Twitter のようなソーシャルメディアから得られるテキストから感情分析や要因分析といった分析を行っている点においては同じであるが、PI を用い分析を行っている点において本研究と違う点である。

3 先行研究

訪日外国人の趣向に合わせたイベント情報提供のためのソーシャルメディア活用を検討 (DEIM 2019) [7] という題で、今井らが発表している。図 1 にシステム概要図と流れを示す。

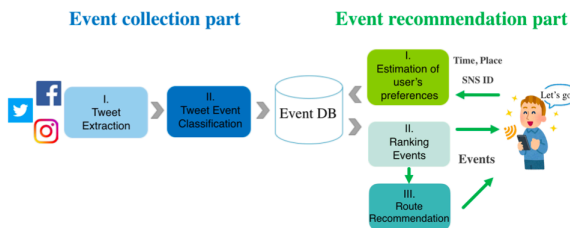


図 1 先行研究システムの概要

システムはイベント収集パートとイベント推薦パートの二つから成っている。まず、イベント収集パートでは、観光客に有益なイベント情報をソーシャルメディアから抽出し、データベースに保存する。推薦パートについては、まず各ユーザの SNS のアカウントを分析することでユーザの趣味趣向を検出する。そして、それらをユーザの位置情報を元に、リアルタイムにその時その場のイベント情報を提供するというものである。ソーシャルメディア情報から収集したイベント情報を持つことで一般的なイベントまとめサイトから得られるイベントだけでなく、マイナーなイベントも収集されており、またイベント推薦パートでは位置情報と趣味趣向データからよりリアルタイムなパーソナライズ性の高い推薦に成功している。研究では、こ

ちらの収集パートと推薦パートから成るシステムの形を引き継いで研究を進めている。

4 提案システムの概要

本研究で提案する概要図を図 2 に示す。

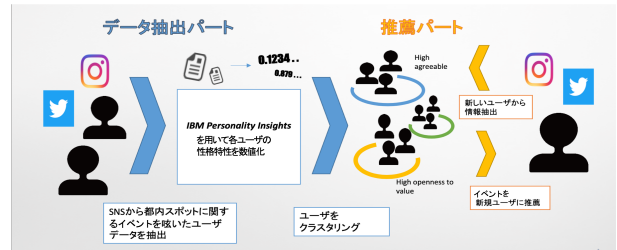


図 2 提案システムの概要

- SNS から事前に大量のユーザテキスト情報を抽出
 - PI によりテキストデータをパーソナリティ数値配列に変換
 - PI 数値を入力とするクラスタリングモデルを作成し、いくつかのクラスグループに分割
 - 問合せユーザのテキストデータを抽出し、そのデータを PI を通じてパーソナリティ数値配列に変換
 - 問合せユーザが所属するクラスタリンググループを決定し、クラスタリンググループが頻繁に参加するイベントを推薦
- なお本論文では、推薦パートではなく、データ抽出パートに焦点を当てていく。

5 PI から分かる性格値

性格特性を数値化するため、先述した通り IBM Personality Insights を用いる。PI が推定する性格値は表 1 に示されるような性格の基本的な次元が 5 つであるという Big Five Model の特性項目と、Kevin Ford の Universal Needs Map に沿った Needs (欲求) 分析、Schwartz の価値概説 (Schwartz Value Survey) に沿った Values (価値観) 分析の特性項目からなる。

表 1 Personality traits estimated by PI

Big Five(個性)	数値の大小による意味
協調性	人当たりの良い・温情のある vs 冷たい・不親切
誠実性	勤勉・マメな人 vs 楽観的・不注意
外向性	外交的・エネルギーッシュ vs 孤独を好む・控えめ
感情起伏	繊細・神経質 vs 情緒安定な・自信家の
知的好奇心	好奇心が強い・独創的 vs 着実・警戒心が強い

6 各ハイブリットイベント参加ユーザデータ

本論文では、以下の手順でハイブリットイベント参加ユーザデータを取得する。データ取得期間は 2021 年の 11 月である。

- (1) Twitter の公式 API [8] を用いて過去に行われていて、かつ現地と配信に参加したと思われるユーザ (それぞれ 10 人

以上) のデータ取得が見込めるイベントを表 2 の通り 7 件取得

- (2) イベント名を含んだツイートを行っているユーザの過去ツイート全てを取得
- (3) 取得した過去のツイートをテキストデータとして PI にて性格特性値を算出

性格特性の数値は, 具体的には 0 から 1 の連続値である. なお, PI は 100 単語以上からなるテキストからしか性格特性を算出しない制限がある. よって 100 単語以上のツイートをしているユーザに限定している. 正しく性格特性を判断できないと思われる箇所 (URL 表記など) は適宜ツイートを削除するなどして加工している.

表 2 取得ハイブリットイベント一覧

	イベント名
1	プロレスリング・ノア試合
2	バンドリ Zepp 札幌・ツアーライブ
3	SHIORoom
4	ダンガンロンパ 10 周年イベント
5	ESPICE 発表会
6	AMUSE VOICE ACTORS CHANNEL FES 2021
7	第 5 回京都アニメーションファン感謝イベント

7 ハイブリットイベント分析

目視で各イベントに現地参加したと思われるユーザとオンライン参加したと思われるユーザに分け, まず各性格項目ごとに箱髭図に表した. 全ての性格項目においてイベント種別あるいは現地参加かオンライン参加かの参加形態ごとに傾向が見られないかを確認した. 例として, Big Five Model の特性項目の一つで, 社交的な性格かどうかを表す big5_extraversion を図 3 に, 同じく Big Five Model の特性項目の一つであり知的好奇心度を表す big5_openness を図 4 に示す. これらの性格特性を始めとするほとんどの性格項目において傾向は特に確認ができなかった. 図中末尾 ‘_real’ は現地参加を表し ‘_online’ はオンライン参加を表す.

図 5 の熱意や意欲欲求を追求する性格値である need_excitement については, ほとんどのイベントについて現地参加ユーザの方が高い値を示す結果が得られた. オンライン参加の方が参加ハードルが低いことから, 現地参加ユーザの方が熱意を表す性格値が高くなったこの結果も感覚的に腑に落ちるものである.

また, Needs (欲求) 分析の 1 つの指標である, 外交的性格の指標をあらわす need_love が図 6 で示す通り, 比較的現地参加の方がオンライン参加より比較的高い結果となった.

また, いくつかのイベントにおいてこの性格値が一番高かった現地参加ユーザと一番低かったオンライン参加ユーザ両者の具体的なツイートの中身を比較した. すると, どのイベントにおいても, 性格値の高いユーザは同じイベントに参加する予

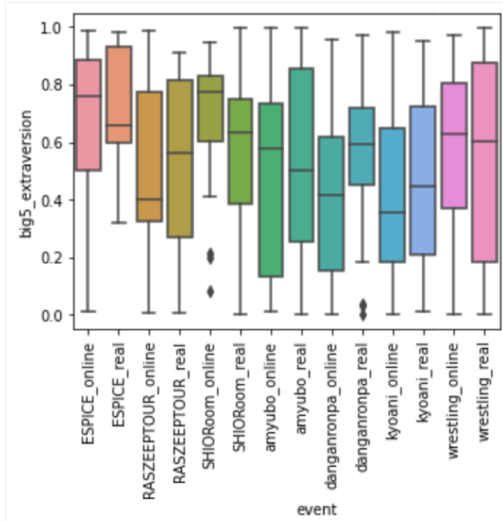


図 3 各イベント毎 ‘big5_extraversion’ 性格項目箱髭図

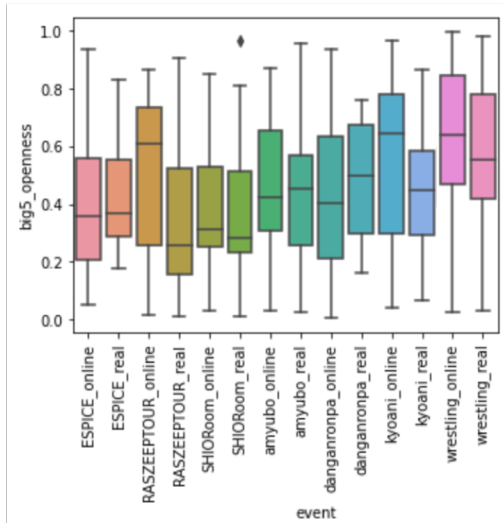


図 4 各イベント毎 ‘big5_openness’ 性格項目箱髭図

定の複数の他のユーザとやりとりをする傾向があり, やり取りの中で現場で一緒に会う約束をしていたりと, 現地でイベントに参加をし他のユーザと交流をする様子が伺えた. 一方で, 性格値の低いユーザは他の参加者とあまり交流する様子は見られなかった. このような傾向は非常に感覚的に腑に落ちるものであり, また ‘need_love’ の性格値傾向に反映されていると考える.

7.1 マンホイットニーの U 検定

先ほどの箱髭図にて, 視覚的に現地参加ユーザの ‘need_love’ 性格値がオンライン参加ユーザのものよりも高い傾向が見えたが, それらを統計的に数値で示すことにする. 検定に使用するのマンホイットニーの U 検定である. これはノンパラメトリックな検定の一つであり, PI の性格値分布は正規分布を仮定していないためこの検定方法を選択した. 具体的に検定をするにあたり使用したライブラリは, scipy [9] の stats.mannwhitneyu である.

上記のライブラリでマンホイットニーの U 検定を行うと, p 値 0.055 という結果が出た. 検定の有意水準を 0.05 とすると

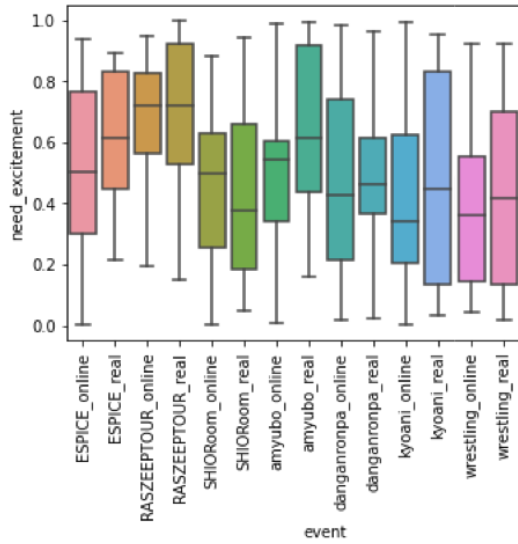


図 5 各イベント毎 ‘need_excitement’ 性格項目箱髷図

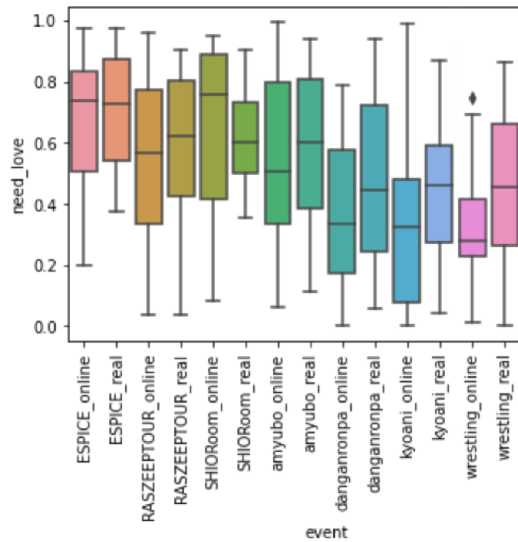


図 6 各イベント毎 ‘need_love’ 性格項目箱髷図

「現地参加とオンライン参加ユーザとで性格特性に違いがある」という仮定は棄却できない。

7.2 効果量

前節で検定を行なったが、 p 値はサンプルサイズに依存するものであるため、次式の通り Hedges の g の効果量を算出する。式中の \bar{x}_1 , s_1^2 , n_1 は本研究においては、それぞれ現地参加ユーザの need_love 性格特性値の平均、分散、サンプルサイズを表し、 \bar{x}_2 , s_2^2 , n_2 は同様にそれぞれオンライン参加ユーザの need_love 性格特性値の平均、分散、サンプルサイズを表す。

$$g = \frac{\bar{x}_1 - \bar{x}_2}{s_d}$$

$$s_d = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

効果量を計算すると効果量 0.20 という結果が出た。先ほどの検定では、仮定は棄却できなかったが、Cohen(1988) の提唱し

た基準により、0.02 より効果があると一般的に言われていることを加味すると、わずかに統計的に傾向があると言える。

8 まとめと今後の取り組み

本研究では、コロナ禍における特にハイブリットイベントに着目し、オンライン・現地両者の参加ユーザの性格特性の違いについて分析した。結果として、オンラインで参加するか、現地で参加するかの性格特性に関係がありあそうな「外交的性格度」の値に差がみられることを発見した。また、今回収集したハイブリットイベントのイベント種別は偏りがあった。収集したイベント 7 種類中、5 種類がゲームやアニメのカテゴリに相当するイベントであった。イベントカテゴリ種類を増やし分析を試みる事を今後の課題としたい。今後これらを活用しコロナ禍のような平時でない状況下にも対応する高パーソナライズイベント推薦システム構築に話を発展させたい。

9 謝 辞

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

文 献

- [1] Jalal Mahmud, Michelle X Zhou, Nimrod Megiddo, Jeffrey Nichols, and Clemens Drews. Recommending targeted strangers from whom to solicit information on social media. pp. 37–48, 2013.
- [2] Liang Gou, Michelle X Zhou, and Huahai Yang. Knowme and shareme: understanding automatically discovered personality traits from social media and user sharing preferences. pp. 955–964, 2014.
- [3] Personality Insights. <https://www.ibm.com/blogs/smarter-business/business/personality-Insights/>. (2021 年 12 月 31 日閲覧).
- [4] 富永登夢, 土方嘉徳. Twitter ユーザの受け取るフィードバックと人格特性の変化の関係に関する調査と分析. 知能と情報, Vol. 31, No. 1, pp. 516–525, 2019.
- [5] 鳥海不二夫, 榎剛史, 吉田光男. ソーシャルメディアを用いた新型コロナウイルス禍における感情変化の分析. 人工知能学会論文誌, Vol. 35, No. 4, pp. F-K45.1, 2020.
- [6] 四方田健二. 新型コロナウイルス感染拡大に伴う不安やストレスの実態: Twitter 投稿内容の計量テキスト分析から. 体育学研究, Vol. 65, pp. 757–774, 2020.
- [7] 今井美希, 工藤瑠璃子, 榎美紀, 小口正人. 訪日外国人の趣向に合わせたイベント情報提供のためのソーシャルメディア活用の検討. *DEIM Forum 2019*.
- [8] Twitter Search API. <https://dev.twitter.com/rest/public/search>. (2021 年 12 月 31 日閲覧).
- [9] SciPy. <https://scipy.org/>. (2021 年 12 月 31 日閲覧).