

ソーシャルメディア上のコミュニティ形成に対して個人の性格が与える影響

Analyzing the influence of human personality on SNS community formation

鈴木 凱亜^{*1}
Gaia Suzuki

大知 正直^{*1}
Masanao Ochi

榎 剛史^{*1}
Takeshi Sakaki

坂田 一郎^{*1}
Ichiro Sakata

^{*1}東京大学 工学系研究科
School of Engineering, The University of Tokyo

With its huge spread in daily life, social media is now a big platform to make new human relationships. While recent research identified the relationship between users' personality traits and how individual users behave on social media, knowledge about how personality traits affect users' mutual interactions is limited. This research proposes a method to relate human personality to community formation on social media. By applying this method to Twitter data, it was inferred that personality factors such as gregariousness, conscientiousness, cheerfulness shows a "homophilous" tendency. This means that similarity in these personality traits breeds user connection. Furthermore, we investigated the relationship between community's composition of personality and community dynamics(size, activeness). Analysis by the proposed method may contribute to creating a social matching system with better user satisfaction.

1. はじめに

ソーシャルメディア上の人々の関わり合いは、今後ますます発展し多様化すると見込まれる。そのような環境下において、ソーシャルメディア上のコミュニティ形成のメカニズムを把握することは、人々の関わり合いを通してユーザが得られる効用や、チームワークが生み出すパフォーマンスを最大化するために必要不可欠である。本研究では、ユーザの特徴がその所属するコミュニティの挙動とどのように関係するかを分析する。SNS ユーザは、投稿の文面に表出される互いの性格を見て交流の取捨選択を行う、という仮定のもと、「ユーザの性格」がコミュニティ形成にどのような影響を与えるか考える。

そもそもソーシャルメディア上の人々のつながりに影響を与える要因は「ユーザの性格」以外にも、「ユーザの属性」、そして「ユーザの思想・感情」があると考えられる。McPherson らはネットワークにおいて「共通の属性を持つ人同士が繋がりがやすい」、Homophily と呼ばれる傾向に着目した。そして、様々な社会ネットワークで民族や人種、宗教といった「ユーザの属性」が Homophily な傾向を示す、すなわち似た者同士でコミュニティを形成しやすいことを確認した [McPherson 01]。他方で「ユーザの思想・感情」によるコミュニティ形成の例として、ソーシャルメディア上で偏った意見を持つ人々が集まり、そのグループの中でその声が反響されることで意見が過激化・極端化する、Echo Chamber と呼ばれる現象も知られている。「ユーザの性格」によるコミュニティ形成については、他の 2 つの要因と比較して小さなスケールでのネットワークダイナミクスであること、そして定量化された性格データの取得が困難であったことから、研究例が少ない。

本論文の目的は、SNS の会話ネットワークにおいて、どのような性格同士がユーザが活発に会話を行うグループを形成するか、コミュニティの形成や挙動と構成員の性格との関係を分析することである。これにより得られる知見は、SNS 上のチームビルディングの最適化や、より相性の良い人となつながら SNS プラットフォームの設計などの応用が期待される。

2. 関連研究

心理学の分野では、人間の性格と行動との相関を分析した研究が広くなされている。そういった研究において性格を捉える枠組みとしてスタンダードになっているのが、ビッグファイブと呼ばれる性格因子である。これは、性格を「Openness (知的好奇心)」、「Conscientiousness (誠実性)」、「Extraversion (外向性)」、「Agreeableness (協調性)」、「Emotional range (感情起伏)」の 5 つの性格特性のスコアの組み合わせにより座標的に表現する枠組みであり、これは文化圏などによらないユニバーサルなものである [McCrae 97]。性格特性のスコアは、被験者への心理テストによって得られる。

性格の異なるユーザの SNS 上の行動の違いを捉える研究方法として、分析対象を定量的な指標で表し、上記の性格特性のスコアとの相関係数を計算するというものがある。Ryan らは Openness (知的好奇心) のスコアが高い人ほど SNS への積極的なコミットが見られることを示し [Ryan 11]、Skues らもそのような人ほど友人を多くもつ傾向にあるとした [Skues 12]。また、Seidman らは Agreeableness (協調性) のスコアが高い人ほど返信など他人に向けた積極的なアプローチを行う傾向にあることを示した。

複数人の性格の組み合わせを捉えるとき、性格スコアの単純な平均値に加え、分散や最大値・最小値などもグループの性格の分布を見るための重要なパラメータとなる。Barrick らは、チームによる仕事のパフォーマンスとメンバーの性格の関係を分析するために、チームごとの性格スコアの平均や分散、最大値・最小値と、チームパフォーマンスの相関係数を計算し、0.23 以上の高い相関を示す関係を有意なものとして抽出した。その結果、Conscientiousness (誠実性) が平均的に高いチームほど仕事が優秀な傾向が見られるが、チーム内で Conscientiousness (誠実性) のばらつきが大きいとパフォーマンスは低下してしまうことを発見した [Barrick 98]。

3. 提案手法

提案手法の概略

SNS 上のネットワークダイナミクスに対する性格の影響を見る際に考慮しなければならないのが、「ユーザの性格」以外の要因によるつながりの存在である。純粋に性格が与える影響を捉えるためには、上述した「ユーザの属性」、「ユーザの感情・思想」による影響を区別し取り除いたネットワークを考える必要がある。

そこで本論文では、同じトピックに属するユーザを投稿のクエリ検索によって抽出し、興味分野やトピックに対する見解が共通したユーザによる会話ネットワーク（以下、トピックネットワークと呼ぶ）を構築する。適切なクエリの選択により、「ユーザの属性」と「ユーザの感情・思想」をできる限り統制したトピックネットワークの構築が可能であると考えられる。そして、このトピックネットワークの中で、ユーザの性格推定、クラスタリング、分析指標との相関分析などを行うことを提案する（図 1）。

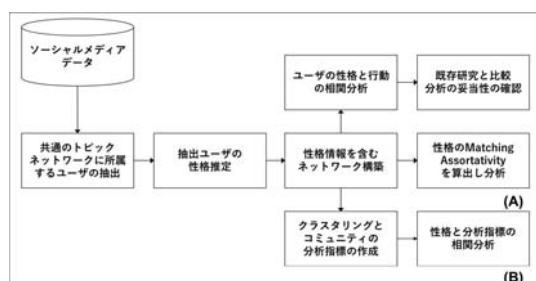


図 1: 分析のフレームワーク

3.1 ユーザの性格推定

近年、ビッグファイブを心理テストでなく SNS の投稿データのテキスト分析により推定する研究が進んでいる。たとえば IBM Watson は、ユーザの SNS 投稿データからビッグファイブと、それぞれの小分類（ファセット）にあたる計 35 種類の性格特性のスコアを推定するモデルを API として公開している。それぞれの性格特性について、具体的な説明は脚注の URL を参照されたい^{*1}。この Personality Insights API は日本語の投稿にも対応しており、伝統的な心理テストにより計算された実スコアと 0.2 程度の比較的高い相関を示すことから、本論文の実験でもこのサービスを用いる。取得可能なデータから性格が計算できるようになると、被験者を集め心理テストを実施する必要がなくなり、より多様なユーザデータを用いた分析が可能になる。

3.2 クラスタリングによるコミュニティ抽出

構築したユーザ間会話ネットワークから、活発に会話を行うグループを抽出するために、Louvain 法 [Blondel 08] を用いたモジュラリティ最大化によるネットワークのクラスタリングを行う。この手法により、分割数を指定することなく重み付きネットワークをつながりの密なグループに分割することができる。本研究では、得られたクラスタのうち 10 人以上のものを分析対象のコミュニティとした。

3.3 Homophily な性格特性の確認

スコアの近いもの同士がつながりやすいような、Homophily な性格特性が存在するかを確認するため、ネットワーク上でリンクを結ぶノード同士の属性値がどのくらい似ているかを捉える指標である Matching Assortativity [Newman 03] を計算する。これは 0 から 1 の値を取り、大きくなるほどノードの属性の近いもの同士がつながったネットワークであることを示す。

3.4 分析指標とユーザの性格の相関分析

本研究で用いる分析指標を表 1 に示す。ユーザ行動の分析指標はユーザの 35 種類の性格特性との相関係数を、コミュニティの分析指標は所属するユーザの性格特性の平均値、標準偏差との間の相関係数を計算する。前述した Barrick らの研究で 0.23 以上の相関があるものを取り上げていたことや、Personality Insights API の推定結果と実スコアとの間の相関が 0.2 程度であることをふまえると、本研究においてもあまり大きな相関係数は期待されない。そこで、他の性格指標と比較し相対的に大きな相関を示し、作成したどのトピックネットワークにおいても一貫してその相関が見られるような関係を取り上げる。

表 1: 本研究で使用する分析指標

分析指標の種類	指標	説明
ユーザ行動の評価	投稿の頻度	1 日あたりの投稿数
	リプライ率	投稿内の返信の割合
コミュニティの評価	ユーザ数	コミュニティの人数
	活性度	コミュニティ内の会話の増加率

4. 実験

4.1 トピックネットワークの抽出

本実験では、トピックネットワークの抽出に Twitter のハッシュタグ検索機能を用いた。ハッシュタグの種類を適切に選択することで、同一の話題に関して、お互いに積極的に話しかける密なトピックネットワークの抽出が可能となる。そこで、表 2 に示す 4 種類のトピックに関するハッシュタグ検索により取得されたツイートから、それぞれのトピックネットワークを構成するユーザを抽出した。表 2 にあるようなハッシュタグのユーザは、共通してトピックに対する興味を持ち、共通して「同じ趣味を持つ人間と繋がりたい」というモチベーションを持つため、繋がりを形成する最終的な決め手が性格的な相性になりやすいと考える。

表 2: 作成したトピックネットワーク

トピックネットワーク	検索ハッシュタグの例
ファッション	#おしゃれさんと繋がりたい
映画	#映画好きと繋がりたい
イラスト	#絵師さんと繋がりたい
他者依存	#メンヘラさんと繋がりたい

4.2 ネットワークの構築とクラスタリング

ハッシュタグにより抽出したユーザの全ツイート（最大で 2000 ツイート）を取得し、その投稿データをもとに Personality Insights API によって 35 種類の性格特性のスコアを計算した。また、取得したツイートからネットワーク内のメンション関係を抽出し、性格特性のスコアを属性として持つノードと、会話

*1 <https://console.bluemix.net/docs/services/personality-insights/models.html>

の回数を重みとして持つリンクによるトピックネットワークを構築した。

構築したネットワークから提案手法のクラスタリングを用い分析対象のコミュニティを抽出した。各ネットワークの基本情報、抽出されたコミュニティの数は表3の通り。

表 3: 構築したネットワーク

トピック	ノード	リンク	コミュニティ数
ファッション	923	742	12
映画	490	623	8
イラスト	947	662	16
他者依存	916	1575	19

4.3 分析の妥当性の検証

そもそも日本語のツイートから、外部モデルにより性格を推定し、それを前提に議論を行うことに妥当性はあるか、実際の分析を行う前に検証した。関連研究で述べたように、SNSへの積極的なコミットには Openness (知的好奇心) の高さが、他人に向けた投稿をする傾向には Agreeableness (協調性) の高さが関わると考えられる。

そこで、それぞれを表す指標として、「1日あたりの投稿数 (投稿の頻度)」と「全投稿に対するメンション投稿の割合 (リプライ率)」をユーザごとに計算し、Openness (知的好奇心)、Agreeableness (協調性) との相関を計算した。結果は表4の通りである。

相関の値は高くないが、4つのネットワーク共に一貫して0.1付近の相関係数を示し、また他の性格特性と比較すると相対的に高い相関を示した。大きな相関は確認されなかったが、既存研究と矛盾しないため、本分析を行うことは有意義であると考えられる。

表 4: 性格特性とユーザ行動の相関関係

性格の種類	Openness	Agreeableness
ユーザ行動の分析指標	投稿の頻度	リプライ率
相関係数の平均	0.12	0.13
35 性格中のスコアの順位	1 位	7 位

4.4 分析指標の作成とコミュニティ形成の分析

本研究では二種類の分析を行う。

(A) コミュニティの形成過程と性格の関係の分析

会話ネットワークの Matching Assortativity を各性格特性について計算することで、コミュニティの形成過程において、「ユーザの属性」や「ユーザの思想・感情」で見られる Homophily な傾向が、性格特性についても見られるか分析した。

(B) 形成されたコミュニティの挙動と性格の関係の分析

コミュニティのふるまいを測る指標として「コミュニティの規模 (ユーザ数)」と、「コミュニティの活性度 (一人当たりの会話の増加率)」を計算し、コミュニティ内の性格特性の平均値、標準偏差との相関を算出した。

5. 結果および考察

(A) コミュニティの形成過程と性格の関係の分析

それぞれの性格特性の各トピックネットワークにおける Matching Assortativity の平均値を計算した。Matching Assortativity が相対的に大きい3つの性格特性を表5に示す。

分析結果から、Gregariousness (社交性) や, Conscientiousness (誠実性), Cheerfulness (明朗性) にはスコアの高い人は高い人同士で、低い人は低い人同士で繋がりがやすい Homophily な傾向があると思われる。これらの性格特性は、自らと反対の傾向を持つ人との交流がストレスを生むのではないかと考えられる。人々がそういったストレスを回避しようとした結果、コミュニティ形成において Homophily な傾向が見られていると推測される。

この研究結果は、関連研究で述べたチーム内の Conscientiousness (誠実性) のばらつきがパフォーマンスの低下を招くことと整合する。誠実性の差は、互いの仕事への姿勢や自制力の差に対する不満やストレスにつながり、人々の結束を低下させると考える。

表 5: Matching Assortativity の大きな性格特性

性格特性の種類	Matching Assortativity
Gregariousness (社交性)	0.260
Conscientiousness (誠実性)	0.225
Cheerfulness (明朗性)	0.195

(B) 形成されたコミュニティの挙動と性格の関係の分析

「コミュニティの規模」、「コミュニティの活性度」のそれぞれのコミュニティの挙動について、相関関係が0.2以上であり、各トピックネットワークについて一貫しているような性格特性の平均値や標準偏差を表6, 7に示す。

「コミュニティの規模」については、Prone to Worry (心配性), Melancholy (悲観的) のような性格特性が正の相関を示すことが分かった。これらの性格特性のスコアが高い人は、他人との関わりによって心の安定を図るため、できる限り多くの相手と関わりを作ろうとし、その結果コミュニティの規模が大きくなると推測される。

「コミュニティの活性度」については、Emotional Range (感情起伏) や Uncompromising (強硬さ) の標準偏差が小さく、Dutifulness (忠実さ), Self-efficacy (自己効力感), Melancholy (悲観的) の標準偏差が大きいほど、コミュニティ内で活発な会話がなされることが分かった。

この結果から、他人への接し方に関わる性格特性は画一的で、自分のあり方に関わる性格特性は多様性に富んでいるコミュニティであることが、活発な会話につながっていると示唆された。

表 6: コミュニティのふるまいと正の相関を示す性格特性

コミュニティの規模	コミュニティの活性度
Prone to Worry (平均値)	Dutifulness (標準偏差)
Melancholy (平均値)	Self-efficacy (標準偏差)
	Melancholy (標準偏差)

表 7: コミュニティのふるまいと負の相関を示す性格特性

コミュニティの規模	コミュニティの活性度
該当なし	Emotional Range (標準偏差)
	Uncompromising (標準偏差)

6. おわりに

本研究では、日本語の SNS 投稿データとテキスト性格推定モデルを用い、SNS 上のコミュニティ形成に性格がどう関わるか分析した。興味を統一したユーザの抽出という提案手法により、既存研究と矛盾しない分析ができることが確認され、また Homophily な作用が見受けられる性格の存在を確認した。提案した分析のフレームワークを用い、コミュニティ形成のメカニズムに関し新たな知見が得られる可能性を提示した。

今後の展望としては、ユーザの抽出方法の改善と、性格推定モデルの精度の向上が挙げられる。

ユーザの抽出方法について、本研究で抽出した twitter ネットワークは条件を満たすハッシュタグの選択が難しかったため、twitter コミュニティ全体の中で非常に限定的なものであった。そこで、より多くのトピックネットワークを構築できるユーザ抽出方法の考案、もしくはユーザの「属性」や「思想・感情」も特徴量に含めた包括的なコミュニティ形成モデルの議論が望まれる。

性格推定モデルについては、ユーザの投稿テキストのみでなく、プロフィールデータや投稿画像データなども特徴量とすることで、より高い精度の性格推定が可能であることが英語の SNS データでは明らかになっている [Azucar 18]。本研究で行った相関分析では、0.2 以上の相関を示す関係は非常に少なく、より説得力のある議論を行うためには、日本語のデータにおいても性格推定の精度を高める必要があると思われる。

参考文献

- [McPherson 01] McPherson, M., Smith-Lovin, L., and Cook, J. M.: Birds of a feather: Homophily in social networks, *Annual review of sociology*, Vol. 27, No. 1, pp. 415-444 (2001)
- [McCrae 97] McCrae, R. R., and Costa Jr, P. T.: Personality trait structure as a human universal, *American psychologist*, Vol. 52, No. 5, pp. 509 (1997)
- [Ryan 11] Ryan, T., and Xenos, S.: Who uses Facebook? An investigation into the relationship between the Big Five, shyness, narcissism, loneliness, and Facebook usage, *Computers in human behavior*, Vol. 27, No. 5, pp. 1658-1664 (2011)
- [Skues 12] Skues, J. L., Williams, B., and Wise, L.: The effects of personality traits, self-esteem, loneliness, and narcissism on Facebook use among university students, *Computers in Human Behavior*, Vol. 28, No. 6, pp. 2414-2419 (2012)
- [Seidman 13] Seidman, G.: Self-presentation and belonging on Facebook: How personality influences social media use and motivations, *Personality and Individual Differences*, Vol. 54, No. 3, pp. 402-407 (2013)
- [Barrick 98] Barrick, M. R., Stewart, G. L., Neubert, M. J., and Mount, M. K.: Relating member ability and personality to work-team processes and team effectiveness, *Journal of applied psychology*, Vol. 83, No. 3, pp. 377 (1998)

[Blondel 08] Blondel, V. D., Guillaume, J. L., Lambiotte, R., and Lefebvre, E.: Fast unfolding of communities in large networks, *Journal of statistical mechanics: theory and experiment*, Vol. 2008, No. 10 (2008)

[Newman 03] Newman, M. E.: Mixing patterns in networks, *Physical Review E*, Vol. 67, No. 2 (2003)

[Azucar 18] Azucar, D., Marengo, D., and Settanni, M.: Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis, *Personality and Individual Differences*, Vol. 124, pp. 150-159 (2018)