

ソーシャルメディアでのコミュニティとユーザの性格特性の関係

松香研究室 16L1005U 飯嶋孝浩

1. はじめに

【ビッグファイブ推定】

ビッグファイブは、協調性、誠実性、外向性、神経症傾向（情緒不安定性）、および開放性からなり、それらの小分類が存在する。IBM Watson は、ユーザの投稿データからビッグファイブとそれぞれの小分類の 35 種類の性格特性のスコアを推定する Personality Insights というモデルを公開している。

【集団と性格の関係性】

Barrick ら (1998): チームの性格スコアの平均値・分散などとチームでのパフォーマンスとの相関係数を計算した結果、誠実性 (Conscientiousness)、協調性 (Agreeableness)、外向性 (extraversion)、および情緒安定性 (emotional stability) が平均的に高いチームほど仕事は優秀である傾向が見られることや、チーム内の誠実性のばらつきがパフォーマンスを低下させることを発見した。

【Twitter 上のコミュニティ形成と性格特性】

鈴木ら (2018): 「ユーザの属性」や「ユーザの感情・思想」による影響を取り除き、ユーザの性格とコミュニティ形成の関係を考察した。その結果、Gregariousness (社交性) や Conscientiousness (誠実性) についてホモフィリー現象を発見し、チーム内でのこれらの性格特性のばらつきが人々の結束を低下させると考えた。

2. 調査

2.1. 目的

ランダムに選んだユーザの投稿データから性格特性スコアを計算し、トピックの種類による性格特性スコアの差について検討した。

2.2. 方法

データ: 3 種類のハッシュタグ (#おしゃれさんと繋がりたい #DTMer と繋がりたい #絵師さんと繋がりたい) の付与されたツイート

装置: R、Python、Tweepy、Personality Insights

手続き: ハッシュタグの付与されたツイートを投稿したユーザから 10 人のユーザをランダムに選択し、ユーザにつき直近の 100 個のツイートを取得し、IBM が提供するモデル Personality Insights を用いてユーザの性格特性を計算した。

2.3. 結果と考察

Self-consciousness (自意識過剰) 相対的に性格スコアの標準偏差が小さい性格特性については、表 1 の結果となった。

表 1. 標準偏差 (平均= 0.248 最大値=0.357) が相対的に小さい性格特性

トピックネットワーク	性格特性	平均値	標準偏差
#おしゃれさんと繋がりたい	Activity level(活発度)	0.779	0.0778
	Sympathy(共感度)	0.752	0.0957
#DTMer と繋がりたい	Artistic interests(芸術的関心度)	0.833	0.1049
	Emotionality(情動性)	0.625	0.1253
#絵師さんと繋がりたい	Uncompromising(強硬さ)	0.476	0.0465
	Self-consciousness(自意識過剰)	0.595	0.0891

「#おしゃれさんと繋がりたい」は Activity level(活発度)の標準偏差が相対的に小さく、平均値も相対的に高い値のため、活発度が高いユーザが集まっていると推測された。また、「#DTMer と繋がりたい」では Artistic interests(芸術的関心度)の標準偏差が相対的に小さく、かつ平均値も相対的に高い値のため、芸術的関心度が高いユーザが集まっていると推測された。トピックネットワーク同士の平均値の差分が相対的に大きくなるような性格特性については、「#おしゃれさんと繋がりたい」は「#絵師さんと繋がりたい」と比較すると大胆性が高く、悲観性が低いと考えられた。同様に、「#おしゃれさんと繋がりたい」よりも「#DTMer と繋がりたい」の方が芸術的関心度と明朗性が高く、「#絵師さんと繋がりたい」よりも「#DTMer と繋がりたい」の方が思考力と明朗性が高いことがわかった。

改善点としては、指定されたハッシュタグの付与された投稿をリツイートした人たちの一部をランダムに取り出したので、分析の正確さに欠けるものとなった。また、本調査ではデータ取得したのは特定のハッシュタグを使用したユーザというだけであったため、リツイートの関係や返信の関係をリンクとして持つようなネットワークを作成することで、トピックによらず見られるような性格特性の特徴について検討することができると考えられる。

3. 実験 1

3.1. 目的

特定のトピックに関してのツイートをリツイートしたユーザの集団をトピックネットワークと見なした際に見られるホモフィリーな性格特性やコミュニティの規模と性格特性との関係について分析した。

3.2. 方法

データ : 4 種類のハッシュタグ(#映画好きと繋がりたい #おしゃれさんと繋がりたい #絵師さんと繋がりたい #メンヘラさんと繋がりたい)を付与されたツイートと、それをリツイートしたユーザの投稿データ

装置 : R、Python、Tweepy、Personality Insights

手続き：Twitter ユーザをノード、リツイートした人とリツイートされた人の関係をリンクとしたトピックネットワークを作成し、トピックネットワークの中で関係性が密となっているコミュニティを抽出した。その後、コミュニティを構成するユーザの直近のツイート 100 個をもとに、IBM が提供するモデル Personality Insights を用いてユーザの性格特性を計算した。

3.3. 結果と考察

表 2 構築したネットワーク

トピック	トピックネットワーク	ノード数	コミュニティ数
ファッション	#おしゃれさんと繋がりたい	209	14
映画	#映画好きと繋がりたい	39	2
イラスト	#絵師さんと繋がりたい	72	1
他者依存	#メンヘラさんと繋がりたい	198	13

トピックネットワークに関わらず標準偏差が相対的に小さい値になった性格特性は Conscientiousness で、「#映画好きと繋がりたい」以外のトピックネットワークについては Agreeableness の標準偏差も相対的に小さくなった。これは、Noe ら(2016)による、Agreeableness や Conscientiousness が似ている人たちが繋がりやすいという研究結果を支持する結果となった。また、性格特性コミュニティの規模が大きくなるほど Emotional range(神経症傾向)と Trust(信用度)の標準偏差が大きくなることから、コミュニティの規模が大きくなるほどユーザの神経症傾向や信用度は多様になることが推測された。

改善点としては、本実験で取得したデータは他者の投稿のリツイートをしたユーザであったが、あるトピックに対して関心度の低いユーザである可能性があり、返信の関係をリンクとしたネットワークの再構築の必要性があった。また、Personality Insights で使用するデータには自身のツイートではない投稿のリツイートが含まれており、同じコミュニティのユーザ達が同じ投稿をリツイートしていた可能性が高く、ホモフィリーな傾向が出やすくなってしまったと考えられる。そのため、ユーザ自身の投稿だけを性格分析用のデータとして使うことが必要であると考えられた。

4. 実験 2

4.1. 目的

トピックネットワークの種類に関わらずホモフィリーな性格特性や、コミュニティの規模と性格特性との関係について検討した。

4.2. 方法

データ：4 種類のハッシュタグ(#映画好きと繋がりたい #おしゃれさんと繋がりたい #絵師さんと繋がりたい #メンヘラさんと繋がりたい)を付与されたツイートと、そのユーザに返信したユーザの投稿データ

装置 : R、Python、Tweepy、Personality Insights

手続き : 4種類のハッシュタグによって検索し、その中からリプライが4つ以上あるツイートを投稿したユーザをランダムに1つ抽出した。抽出したユーザの直近のツイートを取得し、そのユーザが返信した先のユーザを抽出した。そのツイートに含まれる返信先のユーザのIDを取得し、そのユーザの直近のツイート30個を取得した。その後、ユーザの返信の関係をリンクとしたトピックネットワークを作成し、その中で関係性が密となっているコミュニティを抽出した。その後、コミュニティを構成するユーザの直近のツイート100個をもとに、IBMが提供するモデルPersonality Insightsを用いてユーザの性格特性を計算した。

4.3. 結果と考察

表3 構築したネットワーク

トピック	トピックネットワーク	ノード数	コミュニティ数
ファッション	#おしゃれさんと繋がりたい	259	18
映画	#映画好きと繋がりたい	82	5
イラスト	#絵師さんと繋がりたい	158	12
他者依存	#メンヘラさんと繋がりたい	79	8

標準偏差が相対的に小さい値になった性格特性はAgreeablenessであった。これは、実験1と同様にNoeら(2016)の研究を支持する結果となった。また、Seidman(2013)が指摘した、Agreeableness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という研究結果を踏まえると、返信をして繋がろうとするユーザのコミュニティではAgreeablenessのスコアが同程度である人が集まりやすいことが予測された。

また、Immoderation(利己的)の平均値はどのトピックネットワークにおいてもユーザ数と負の相関を持ったことから、ユーザの性格特性が利己的でないほどコミュニティの規模が大きくなる、すなわち利己的であるほどコミュニティの規模が小さくなることが推測された。

5. 総合考察

本研究では、ソーシャルメディアでの性格とコミュニティの関係性について、トピックネットワークの種類によらず共通して見られる性格特性や、トピックネットワークによって異なる性格特性があるのかについて分析した。

今後の展望としては、より広い分野のコミュニティを抽出できるようなハッシュタグの選択が望まれる。トピックネットワークの種類にかかわらず一貫して見られたような傾向が少なかった。これは、ユーザが関心のあるトピックの種類の違いがコミュニティ抽出に与える影響が除ききれてない可能性がある。また、Twitter上の投稿と投稿者本人の性格の違いがある場合もあるため、Personality Insightsではなく質問紙から得られた性格特性スコアを使用することが挙げられる。そのほかには、Instagramなどの他のソーシャルメディアのデータを使用した分析をすることで、より多角的なソーシャルメディアの利用者についての知見が得られると考えられる。