

Twitterにおける宣伝情報拡散に影響するインフルエンサーの特徴分析

1X16C057-1 後藤孝輔
指導教員 大野高裕

1. 研究背景と目的

SNS が広く普及した現在、SNS 上で影響力のある人に宣伝広告を依頼するというインフルエンサーマーケティングが市場規模を拡大させている。本研究での影響力とは、宣伝情報をどれだけ拡散させられるかということとし、その拡散度合をリツイート(以下 RT)を用いて測る。インフルエンサーは、フォロワー数が多いほど影響力が強い事が経験的にも研究的にも明らかである。しかし、インフルエンサーマーケティングでは、フォロワー数に対して比例的にコストがかかるフォロワー単価制をとっているため、企業が宣伝広告にかけられるコストを考えると、自ずとフォロワー数に制限がかかってしまう。その制約の中でいかにして効果的なインフルエンサーを選択するのが重要である。

現状、企業はコスト制約の中で、宣伝したい商品・サービスのジャンルとインフルエンサーの投稿内容の親和性、インフルエンサーが持つフォロワーの性別・年齢の属性との親和性などを見てインフルエンサーを選考している。しかし、候補となるインフルエンサーは多く、宣伝効果の大きいインフルエンサーの選択要因がわからないという声が上がっている。

よって本研究では、フォロワー数を含め、インフルエンサーの投稿傾向やプロフィールといった特徴と、インフルエンサーの影響力の関係を分析するモデルを構築し、これを明らかにする。この研究によって、インフルエンサーマーケティングを試みようとする企業、またはインフルエンサーを紹介する企業が、インフルエンサーの選考する際の検討要因としてその特徴を用いることができる。また、インフルエンサー自身が自らの影響力を高めるための努力指標とすることも可能である。

2. 従来研究

Eytan Baksy(2011)[1]では、Twitter ユーザーの影響力をそのユーザーの投稿の拡散に関わった全ての人の総数であるカスケードサイズで評価し、これをフォロワー数、フォロワー数、ツイート数、Twitter 使用歴、その人の過去の影響力を用いて予測する研究を行っている。また Eva Lahuerta-Otero(2016)[2]では、Twitter における投稿の拡散度合とフォロワー数、ツイートの感情、語彙の多様性、ハッシュタグの使用、メンションの使用、URL の使用との因果関係の分析を行っている。同様に、ユーザーの影響力と他の要素との因果関係について分析した研究は多数存在する。これらの研究は Twitter 上の一般的な投稿とユーザーを対象として分析を行っており、宣伝投稿やインフルエンサーに絞って分析を行っているものは見られない。また、多くの研究は 1 投稿ごとに分析を行っており、ユーザーの過去の投稿から抽出した特徴を使って分析を行うような研究も見られない。

3. 本研究の提案

3.1. 研究概要

本研究では、研究対象となるインフルエンサーの条件として二つ設定する。一つ目は、「#PR」と付けた投稿をしていることである。この理由は、本研究では宣伝広告情報を拡散することを対象とし、ステルスマーケティングになることを避けるためである。二つ目は、フォロワー数が 1,000 人から 10 万人の範囲であることを設定する。この理由は、従来研究からフォロワー数とリツイートの関係には減衰の効用があることが言われており、企業の一般的な広告予算現状から見て、フォロワー数 1,000 人から 10 万人ほどの人が重視されているからである。そして上記に該当するインフルエンサーの特徴をプロフィールと投稿から抽出し、これと影響力との因果関係を明らかにする。

3.2. 分析するインフルエンサーの特徴要素の決定

まずインフルエンサーが RT を得るまでの流れとその中で RT に影響を与える要因について整理した結果を図 1 に示す。今回はインフルエンサーについての研究であるので、まず投稿する側に

着目する。ここで 3 つの要因を挙げた理由を述べる。まず、RT されるためにはその投稿が見られなければならないと考え、被接触率を挙げた。次に、RT することは自らもその情報の発信を行うということで、RT する人は怪しい情報を発信することを避けるために宣伝広告の情報源が信頼できるか気に掛けると考え、信頼度を挙げた。最後に、RT によって自分のフォロワーに宣伝の内容を発信したいと思うためにはその内容に共感が持たなければならないと考え、共感度を挙げた。

次に投稿する側の要因を構成する要素として考えられるものを実物の Twitter を見ながら検討し、表 1 に整理した。この表にあるすべての要素について、データとして取得できるか、取得したデータから定量的に表現できるか検討を行った。その結果、本研究で扱える要素をフォロワー数、フォロワー数、Twitter 使用歴、画像・動画(メディア)を使っているか、ハッシュタグを使っているか、URL を載せているか、ツイートの文字数の 7 つとする。

3.3. 提案モデル

まず目的変数は、どれだけ拡散されたかを見るための RT 数と、フォロワーにどれだけ RT されたかを疑似的に確認できる RT 率 (RT 数/フォロワー数)として、それぞれに対して分析を行う。次に説明変数は、前述した 3.2 項の要素を用いる。フォロワー数、フォロワー数、Twitter 使用歴以外は平均使用率として、インフルエンサーの傾向を見るためにリプライ投稿を除いた最新の 100 件のデータの平均値をとった。ここまでで設定された変数に対する本研究の仮説を表 2 に示す。また説明変数と目的変数の関係を図 2 に示し、二つのモデルのモデル式を以下に示す。モデル式では α と β は各モデル式における変数の推定係数を示し、 x は各説明変数の値を示している。(添え字の対応は図 2 を参照) RT 率のモデル式では目的変数の計算にフォロワー数を使用しているため、フォロワー数が説明変数に含まれていない。

$$(RT \text{ 数}) = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4 + \alpha_5 x_5 + \alpha_6 x_6 + \alpha_7 x_7 + \alpha_0 \quad (1)$$

$$(RT \text{ 率}) = \beta_1 x_2 + \beta_2 x_3 + \beta_3 x_4 + \beta_4 x_5 + \beta_5 x_6 + \beta_6 x_7 + \beta_0 \quad (2)$$

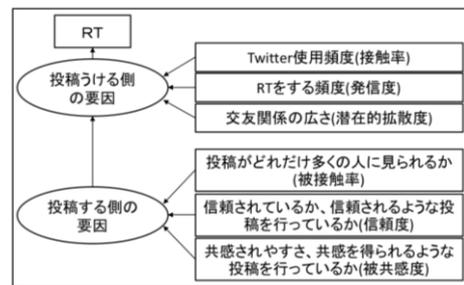


図 1.RT 数に影響を与える要因を整理した図

表 1.投稿する側の要因を構成する要素

要素	投稿する側の要因		
	被接触率	信頼度	被共感度
	フォロワー数	デモグラフィックデータ(年齢、性別など)	画像、動画(メディア)の使用頻度
	ツイート頻度	ユーザー名(本名かどうかなど)	URL を投稿・転載する頻度
	ハッシュタグの使用頻度	紹介文(詳しく書かれているかなど)	ツイートの文字数
		アイコン画像(本人画像かなど)	ツイートの内容
		Twitter使用歴	語彙力
		フォロワー数	メンション機能の使用頻度
		フォロワー数	ハッシュタグの使用頻度
		相互フォロワー数	
		ツイート頻度	
		リプライへの対応	

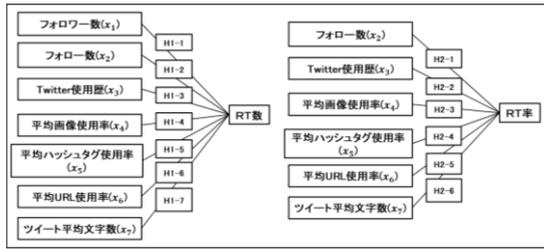


図 2.本研究のモデル

表 2.本研究の仮説

仮説	仮説
H1 RT数に対する仮説	H2 RT率に対する仮説
H1-1 フォロワー数はRT数に正の影響を与える	H2-1 フォロワー数はRT率に正の影響を与える
H1-2 フォロワー数はRT数に正の影響を与える	H2-2 Twitter使用歴はRT率に正の影響を与える
H1-3 Twitter使用歴はRT数に正の影響を与える	H2-3 平均メディア使用率はRT率に正の影響を与える
H1-4 平均メディア使用率はRT数に正の影響を与える	H2-4 平均ハッシュタグ使用率はRT率に正の影響を与える
H1-5 平均ハッシュタグ使用率はRT数に正の影響を与える	H2-5 平均URL使用率はRT率に正の影響を与える
H1-6 平均URL使用率はRT数に正の影響を与える	H2-6 ツイート平均文字数はRT率に正の影響を与える
H1-7 ツイート平均文字数はRT数に正の影響を与える	

4. 検証

4.1. 使用データと分析方法

本研究は、Twitter server 上のデータから個人用 Twitter API を用いてデータ取得を行い、次の 2 つの step で取得を行った。

step1: 「#PR」がついている日本の投稿を Twitter タイムラインより取得する。その中でフォロワー数 1,000 人から 10 万人のユーザーが本研究の対象となる。データ取得期間中に複数投稿を行っているユーザーが多数見られたが、本研究ではユーザーの頭数がサンプル数になるため、取得期間中のユーザーの RT 数の平均値を目的変数とした。

step2: 次に各ユーザーの直近 100 件の投稿を取得し、これより各説明変数の算出を行った。最後に各ユーザーに対して算出した変数に最初に取得した目的変数のデータを結び付けた。この手順により取得したデータは 2019 年 7 月 19 日から 1 週間前までに投稿を行ったユーザー 2,081 名分となった。

分析手法は、RT 数は負の二項分布、RT 率は対数正規分布に従うと仮定し、説明変数を正規分布へ標準化した値で重回帰分析を行った。RT 率について分析を行う際、RT 率が 0 のものは対数変換できないため、そのようなデータを除いて分析を行った。

4.2. 分析結果

分析結果と仮説の検討を整理したものを表 3、表 4、表 5 に示す。RT 数についてはフォロワー数、平均メディア使用率、平均ツイート文字数の仮説が支持される結果となり、RT 率については全ての仮説が支持されない結果となった。

5. 考察

5.1. RT 数の仮説に関して

まずフォロワー数が正の影響を持つ理由として、フォロワー数が多いほど多くの人に投稿が見られ、RT される機会が増えるからだと考えられる。この結果より、インフルエンサーマーケティングにおいてフォロワー単価制が取られる根拠が確認できる。

次に平均メディア使用率、ツイート平均文字数が正の影響を持つ理由として、投稿の内容は文字数が多いほど商品・サービスがどのようなものか、インフルエンサーが何を伝えたいのかが伝わりやすく、それに付け加えて画像・動画があることでその商品・サービスがイメージされやすくなり、共感が得られやすい投稿となるため RT につながるということが考えられる。

次にフォロワー数が負の影響を持つ理由として、フォローを多くしていることがフォロワー獲得のためであると投稿を見る人に思われることによって、信頼を失ってしまうことが考えられた。

次に Twitter 使用歴が負の影響を持つ理由として、投稿を見る人はより新しい情報を好むため、使用歴が長い人の情報は古いと判断され共感が得られにくいことが考えられた。

表 3. RT 数の分析結果

変数	推定係数	P値	有意性
切片	-0.25029	9.02e-05	***
フォロワー数	1.62514	<2e-16	***
フォロー数	-0.52946	<2e-16	***
Twitter使用歴	-0.27924	0.000139	**
平均メディア使用率	0.23941	0.000859	**
平均ハッシュタグ使用率	-0.47526	2.80e-10	***
平均URL使用率	0.06392	0.556691	
平均文字使用率	0.61022	3.48e-10	***

表 4. RT 率の分析結果

変数	推定係数	P値	有意性
切片	-7.56027	<2e-16	***
フォロワー数	-0.32455	5.86e-12	***
Twitter使用歴	-0.06072	0.345	
平均メディア使用率	-0.04895	-0.706	
平均ハッシュタグ使用率	0.04939	0.725	
平均URL使用率	0.09460	0.354	
平均文字使用率	-0.15841	0.117	

表 5. 仮説の検討結果

仮説	結果	仮説	結果
H1 RT数に対する仮説		H2 RT率に対する仮説	
H1-1 フォロワー数はRT数に正の影響を与える	0.1%有意	H2-1 フォロワー数はRT率に正の影響を与える	負の影響 0.1%有意
H1-2 フォロワー数はRT数に正の影響を与える	負の影響 0.1%有意	H2-2 Twitter使用歴はRT率に正の影響を与える	棄却
H1-3 Twitter使用歴はRT数に正の影響を与える	負の影響 0.1%有意	H2-3 平均メディア使用率はRT率に正の影響を与える	棄却
H1-4 平均メディア使用率はRT数に正の影響を与える	0.1%有意	H2-4 平均ハッシュタグ使用率はRT率に正の影響を与える	棄却
H1-5 平均ハッシュタグ使用率はRT数に正の影響を与える	負の影響 0.1%有意	H2-5 平均URL使用率はRT率に正の影響を与える	棄却
H1-6 平均URL使用率はRT数に正の影響を与える	棄却	H2-6 ツイート平均文字数はRT率に正の影響を与える	棄却
H1-7 ツイート平均文字数はRT数に正の影響を与える	0.1%有意		

次に平均ハッシュタグ使用率が負の影響を持つ理由として、ハッシュタグが付くと広告感が感じられ、投稿を見る人に不信感を与えるため負の影響となったと考えられる。

最後に平均 URL 使用率が有意にならなかったことについて二つの理由が考えられる。一つ目は、宣伝広告に関する投稿を RT するかどうかの判断は投稿の情報のみでされるため URL 先の情報が影響を与えにくいことである。二つ目は、URL 先は企業ページであることが多く、第三者の情報ではないため見られることが少ない可能性があることである。

5.2. RT 率の仮説に関して

フォロワー数が負の影響を持つのは RT 数と同様の理由と考えられる。また他の変数が有意とならなかった理由として二つが考えられた。一つ目は、RT 率が 0 の人とそれ以外の人で特徴に大きな差があったことである。本研究では、RT 率が 0 のものは対数化ができないという理由からそのデータを除いて分析を行ったことにより有意とならなかったことが考えられる。二つ目は、得られたデータには複数の商品・サービスのジャンルが混在しており、それぞれのジャンルごとに影響力と因果関係を持つ要素が異なっている可能性があることである。

6. 結論と今後の課題

本研究では、インフルエンサーの特徴としてフォロワー数、画像・動画の使用頻度、投稿の平均文字数が RT に対して正の影響を持ち、フォロワー数、Twitter 使用歴、平均ハッシュタグ使用率が負の影響を持つ結果となった。つまり、インフルエンサーが日頃からハッシュタグを多用せず、文字を使った十分な説明と、画像・動画の使用によって共感が得られやすい投稿を行っていることが重要だと結論付ける。

この研究の実務的貢献として、企業にはコスト制約に応じたフォロワー数を持つインフルエンサーの中で、比較的新参者でフォロワー数が少なく、投稿ではハッシュタグを多用せず、文字を使った十分な説明をしており、画像・動画多用している人を採用すべきと提案できる。またインフルエンサーに対しても、ハッシュタグの多用を避ける事、文字での十分な説明、画像・動画の積極的な使用が効果的だと提案することができる。

本研究を行って明らかとなった限界と今後の課題として、投稿内容の解釈が挙げられる。本研究では定量的な変数とするために投稿の内容を文字数のみで判断したが、その内容の語彙力や表現方法が影響を与える可能性が考えられる。これを解決するために語彙力・表現方法の定量的な評価方法の検討が必要である。

7. 参考文献

- [1] Eytan Bakshy, Jake M. Hofman, Winter A. Mason, Duncan J. Watts. "Everyone's an Influencer: Quantifying Influence on Twitter". WSDM '11, February 9-12, 2011.
- [2] Eva Lahuerta-Otero, Rebeca Cordero-Gutierrez, "Looking for the perfect tweet. The use of data mining techniques to find influencers on Twitter", ELSEVIER, 31 July 2016.