

Twitterにおける炎上・バズ事象の特性分析手法の提案に向けた事前分析

Preliminary analysis for proposal of characteristic analysis method of flame and buzz events on Twitter

東 優介[†] 高野知佐[†] 前田香織[†]

Yusuke Azuma[†] Chisa Takano[†] Kaori Maeda[†]

[†] 広島市立大学大学院 情報科学研究科

1. 概要

近年、ソーシャル・ネットワーキングサービス (SNS) の普及により、年齢、職種、地理的制約を超えた人と人とのコミュニケーションが活性化している。SNSは、直接の友人だけでなく、共通の趣味を持った人とネット上でリアルタイムに情報交換できるといった「交友関係を広める媒体」として有用であるが、ネット炎上やバズといった予測のつかない事象が起こり大きな社会問題となっている。ネット炎上やバズはどちらもユーザのダイナミクスが爆発的に活発化する事象であるが、その原因は不祥事やいじめ、ニュース、災害情報など多岐にわたっており、個々の事象の原因を本質的に理解することは難しい。本稿ではTwitterの実データを利用して、ネット炎上やバズの発生過程および収束過程を調査する方法を示す。最終的には、Twitterのネットワーク構造特性と関連付けて、発生過程/収束過程の特性を明らかにし、炎上検知やその対策技術、さらにはバズマーケティング技術等に発展させる予定である。

2. はじめに

近年、情報通信ネットワーク技術の発展や情報通信手段の発達、スマートフォンのような通信デバイスの普及を背景に、TwitterやFacebook、Instagramといった、様々なソーシャルネットワーキングサービス (SNS) が世に広まっている。これまでのコミュニケーションは、人と人が直接つながることにより形成されるものであったが、SNSの普及によって、間接的な繋がりの上に形成される「年齢」、「職業」、「地理的」制約を超えた、これまでにない新しい形態のコミュニケーションが生まれている。また、SNSの普及によって、「ネットを介した口コミによる流行の爆発現象 (バズ)」のような経済効果につながる事象がある一方、個人の軽率な行為によって多数のインターネット利用者が誹謗中傷のツイートを行う (炎上) といった予測不可能な事象も発生している。このような炎上事象およびバズ事象は、インターネットの情報取得や伝達のための情報基盤という側面だけでは説明ができず、その背後にあるユーザ間の相互作用を表す社会ネットワークの構造が影響を与えているはずである。

本研究では、SNS上の炎上およびバズといった爆発的事象の特性を調査し、この特性と社会ネットワーク構造との関係を明らかにする。この知見により、社会

ネットワーク構造と情報特性 (炎上またはバズのキーワード) の関係性によって爆発的事象の予兆を検知することができ、最終的には炎上対策技術およびバズマーケティング技術に発展させる予定である。ただし、本稿は事前分析として爆発的事象の特性を調査する方針を示すものであり、多くの課題が残っている。

3. 炎上事象における関連研究

この章では、炎上分析に関する関連研究について簡単に説明する。

文献 [1] は、過去の炎上事例 50 件について「内容」、「投稿時刻」、「炎上発生日時」等の関係性を調査したものである。ただし、Twitterのデータは炎上が発生すると、早期に削除される可能性があるため、Webアーカイブスや2ちゃんねる等の掲示板機能、キュレーションサービスを用いて情報収集を行い、炎上の詳細を分析している。また、文献 [2] では、炎上事例をできるだけ早く検出、対策するために、SNSの各コメントに対して機械学習を用い感情極性 (肯定/否定/中立の感情) 予測を行っている。感情極性予測を用いることによって、否定的なコメントと判断されたコメントの増減を監視し、否定的なコメントが急激に増加している部分を炎上発生部分と定義して炎上検知を行っている。

これらの研究は、今までに起きた炎上事例を機械学習を用いて分析し、炎上の傾向を予測するものであり、精度の高い学習を行うためには膨大なサンプルデータや分析する学習期間が必要、炎上あるいはバズが起きた因果関係がわからないといった問題点がある。

4. 炎上・バズ事象の特性分析方法

SNSの中でも、特にネットユーザに占める利用率も高く、情報発信や情報拡散に強い影響力を持つサービスであるTwitterの実データを利用する。炎上、バズ事象は、「あるキーワードに関するツイート数の急激な増加」、「あるツイートに関するリツイート数の急激な増加」、「画像や動画が入ったツイートのリツイート数の増加」といった様々な特性変化によって判断される。我々は「あるキーワードに関するツイート数の急激な増加」を事象発生の判断要素とし、使用するTwitter APIのエンドポイントはGET search/tweetsである。

以下の方針 (2つの調査項目) で、爆発的事象の特性を調査し、社会ネットワーク構造との関係性を明ら

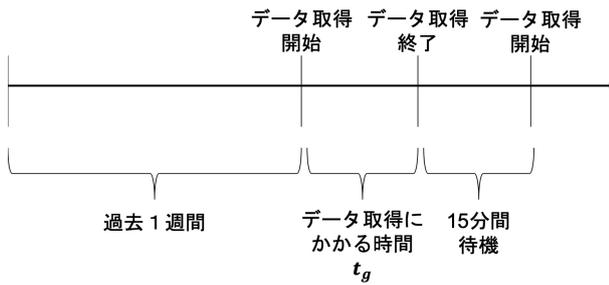


図 1: データ取得時間に関する模式図

かにする。

調査 1 炎上、バズりそうなキーワードを監視しておくことにより、炎上、バズの振る舞いを調査し、発生原因となったツイートの特性を調査

調査 2 炎上しそうなユーザを監視しておくことにより、炎上が発生する前と後で、炎上ユーザのユーザ特性（フォロワー数、リツイート数、リプライ数等）、ユーザ間の特性相関（次数相関）の変化を調べ、社会ネットワーク構造の観点から爆発的現象の発生要因を調査

本稿は 調査 1 に着目し、事前分析を行う。調査 1 において、はじめに検討しなければならない項目は

- (1) どんなキーワードを監視するのか
- (2) どれくらいの頻度でデータ取得するか

でありこれらについては詳細に検討をすすめる。

5. 事前分析

本節では事前分析として、ある特定のキーワードに対するツイート数を調査し、その結果を示す。

5.1. 分析条件

選択したキーワードは、バズ事象が発生すると予想される「バルス」と「地震」である。また、Twitter のキーワード検索によるツイート数と比較するために、Googleトレンド [3] で同じキーワードの検索回数の推移も調査した。観測区間は、2019年8月27日から9月6日である。

データを取得する頻度は図 1 に示すとおりである。Twitter API の仕様で、データ取得開始時から過去 1 週間のツイートデータしか取得できない。事前分析として、この仕様に従い過去 1 週間遡ってツイート数をカウントする。また過去 1 週間の全データを一度に取得できるわけではなく、データリクエスト回数が 15 分に 180 回という制限があり、ツイート数に応じたデータ取得時間 t_g (分) を要する。調査の結果、キーワードによらずツイート数 n に対して以下のような線形的な関係があることがわかっている。

$$t_g = 0.0011 n + \alpha$$

ただし n が大きい場合、 α は無視できるほど小さな値である。

また、リクエスト制限によりデータ取得終了後に次のデータ取得開始時間まで 15 分間待機する必要がある。

5.2. 分析結果と考察

図 2, 図 3 はそれぞれキーワード「バルス」と「地震」のデータ取得結果を示す。また、各図の上 (a-1, b-1) は Google トレンド、下 (a-2, b-2) は Twitter API で取得した結果である。各グラフの横軸は取得した日時、縦軸は Twitter の場合、過去 1 週間のツイート数であり、Google トレンドの場合、キーワードが検索された回数の相対的な値 (観測区間の最大値で割った値) である。

図 2 の (a-1) を見ると、2019 年 8 月 30 日に急激に値が増加していることが確認できる。これは、同日、テレビ放送で「天空の城ラピュタ」が放映されたため、Google での検索回数が増えたためである。実際にテレビの中で「バルス」と叫んだ時間は、2019 年 8 月 30 日 23 時 20 分ごろである。しかし、下の Twitter の結果 (a-2) を見ると、Google トレンドのピークの時間より遅れてツイート件数が増加している。これは、5.1 節で述べたように、ツイート件数が多ければ多いほどデータ取得に時間がかかってしまうためである。具体的には、バルスに関して 57 万件以上のツイートがあり、取得時間は約 10 時間以上を要した。

また、図 3 のキーワード「地震」の方を見ると、Google トレンド、Twitter どちらも 1 日の平均で考えると時間的な変動は少ない。特にこの区間に大きな地震はなかったからである。ただし、データ取得頻度の違いだけでなく、「Twitter でツイートすること」と「Google で検索すること」の使用目的 (日常をつぶやく、詳細な情報を知りたい等) が異なるため、Twitter ツイート数と Google 検索数のデータを直接比較することはできない。

また、今回の分析では、2 つのキーワードのツイート数の時間変化しか取得しておらず、データ取得にかかる時間がそれぞれ違うためグラフにプロットする間隔が統一されてない。

今後は、炎上やバズを誘発するキーワードについて検討し、データ取得時間を削減するためにツイート数の観測区間を 1 週間ではなく、1 日あるいは 1 時間にすることを予定である。

6. まとめと今後の課題

本研究では、社会問題となっている SNS における炎上やバズといった爆発的事象の特性を調査するための方針を示し、2 つのキーワード (バルス, 地震) において事前分析を行った。今後の課題としては、分析するキーワードやデータ取得の頻度を考察する。また、炎上やバズといった爆発的なダイナミクスが発生している状態とそれ以外の定常状態の比較や、それらの事象の発生過程や収束過程の特性を分析をしていく予定である。

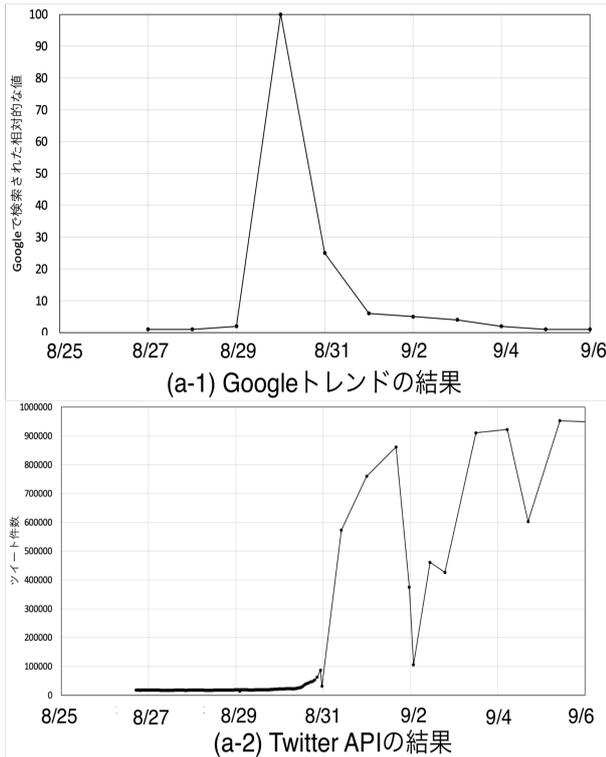


図 2: 「バルス」に関するデータ取得結果

キーワード検索による今回の方法では、Twitter における社会ネットワーク構造と Twitter の爆発的事象を関連づけることはできない。なぜならば、炎上の発端となったユーザを特定することは難しく、炎上ユーザを中心としたネットワーク構造はわからないからである。よって、今後炎上しそうなユーザ（インフルエンサー）を監視しておき、実際に炎上した段階で、炎上前後のフォロワー数やリプライ数の変化を分析することで、ネットワーク構造の変化と炎上事象特性の関連性を調査することも必要である。

謝辞

本研究は JSPS 科研費 17H01737, 17K06438, 18K11271, 19K11929 より研究費の支援を受けて実施したものです。

参考文献

- [1] S. Yamagata, H. Kawamura and K. Suzuki, "Trend analysis for flaming of SNS," Proceedings of the 2013 IEEE/SICE International Symposium on System Integration, 2013. (DOI: 10.1109/SII.2013.6776721)
- [2] S. Ozawa, S. Yoshida, J. Kitazono and T. Sugawara and T. Haga, "A sentiment polarity prediction model using transfer learning and its application to SNS flaming event detection," IEEE Symposium Series on Computational Intel-

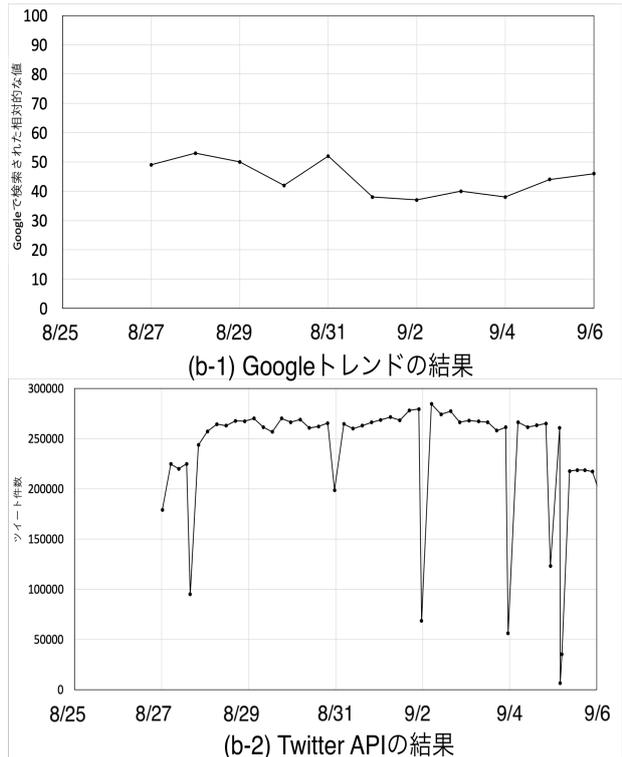


図 3: 「地震」に関するデータ取得結果

ligence(SSCI), 2016.

(DOI: 10.1109/SSCI.2016.7849868)

- [3] Google Trends - Googleトレンド, <https://trends.google.co.jp/trends/?geo=JP>