

Twitterにおけるバーストの検出と生起要因に関する分析

Detection of bursty status on Twitter and analysis of its occurrence factors

学籍番号: 201221601

氏名: 水沼友宏

Yuhiro MIZUNUMA

近年、ソーシャルメディアを通じた情報の収集・発信、コミュニケーションが一般的なものとなっている。ソーシャルメディアの一種である Twitter は 2006 年の登場以降ユーザ数をのぼし、2012 年時点でそのユーザ数は 5 億人を突破した。

Twitter の「速報性」、「情報拡散性」、「簡便性」といった特徴により、あるイベントが生じた際にツイート数が平常時と比較し大きく増加することがある。本研究では、これを「バースト」と定義し、どのようにバーストの検出を行うのが妥当であるか、どのようなイベントによってバーストが生起するのか、バースト時のツイートにはどのような特徴があるのかを明らかにする。Twitter Search API を用いて収集した 2011 年 11 月 16 日から 2013 年 2 月 15 日までのデータ、5,285,607,227 件を対象に分析と考察を行った。

バーストの検出手法については、種々の外れ値検出手法やこれまで行われたバースト検出手法を比較することにより、 3σ 法によるバースト検出が妥当であると判断した。

また、どのようなイベントによってバーストが生起するのかについては、様々な要因によってバーストが生起するが、特に地震などの災害によってバーストが生起しやすいことや、他のメディア、とりわけ速報性の高いメディアと関係があることが明らかになった。また、バーストは、「怖」、「哀」などのネガティブな感情と関連するものが全体の 30%以上を占めていること、バーストした際にツイート数が多くなるのは、ポジティブな感情が見られるバーストであることが分かった。さらに、地震バーストに影響を与える要因のうち、最大震度と都心から被災地までの距離の短さがバースト生起の有無に影響を及ぼすこと、特に都心から被災地までの距離の短さがバースト生起により強い影響を及ぼすことが明らかとなった。

バースト時と非バースト時の比較により、バースト時はツイートの平均文字数が短く、リツイート (RT) の比率が高く、リプライ (@) の比率が低いといった傾向が明らかとなった。このことからバースト時には平常時と比較して、情報入手、拡散ツールとしての側面が強くなることが示された。また、バーストの類型化を行い、各クラスごとにイベントの特徴を推測することで、各々のバーストは、「小さなイベント型」、「既知イベントピーク型」、「既知イベント準備期間型」、「突発的イベント型」、「情報拡散型」の 5 つに類型化された。

研究指導教員: 池内淳

副研究指導教員: 宇陀則彦