

バースト解析を用いた Twitter ユーザの興味・関心情報の探索

佐藤 健太[†] 荒川 達也^{††}

[†]群馬工業高等専門学校専攻科生産システム工学専攻 ^{††}群馬工業高等専門学校電子情報工学科

1. はじめに

マイクロブログサービスの特徴の1つとして、各ユーザが考えたことや感じたことを即座に投稿するリアルタイム性が挙げられる。そのため、マイクロブログへの投稿を分析することによりユーザの興味や関心の推移を追跡できると考えられる。

本研究では、Twitter のユーザを対象とし、時系列テキストデータの中で特定のキーワードを含むテキストが急激に増加する「バースト」に注目して、個人のツイートデータに対してバースト解析を行うことで個人の興味・関心の探索を行う。また、個人のバーストと社会全体でのバーストの比較による個人の興味と社会との関係の分析を試みる。

Twitter を対象とした研究は数多くあり、その中に Twitter 全体でのトピックを探索・追跡する研究がある [1][2]。それらに対し、本研究では個人のツイートデータに対するバースト解析およびそれと全体との比較を行うことにより、ユーザー一人一人の興味・関心情報の抽出を行うことを目標とする。

2. 提案手法

本研究では「個人のツイートデータ内でバーストしている事象はその個人が興味を持つ事象である」と考え、バースト解析を用いてユーザの興味情報の探索を行う。また、個人のツイートデータ内で検知されたバーストに対し、それらを同時期に社会全体のツイートデータから検知されたバーストと比較することにより個人の興味の傾向のより詳しい推測を試みる。

3. 実験と考察

提案手法の有効性を検証するため、社会のツイートと、ある2名のユーザ A, B のツイートに対してそれぞれバースト解析を行い、その結果を比較することによってより詳しい情報が得られるか検証・考察を行った。今回は「ゆるゆり」という漫画についてのツイートであることを示すハッシュタグを用いて解析を行うこと

表1 実験 解析設定

解析期間	2015年10月1日 ～ 2015年11月30日
解析に用いる単語	#yuruyuri
1期間の単位	1日

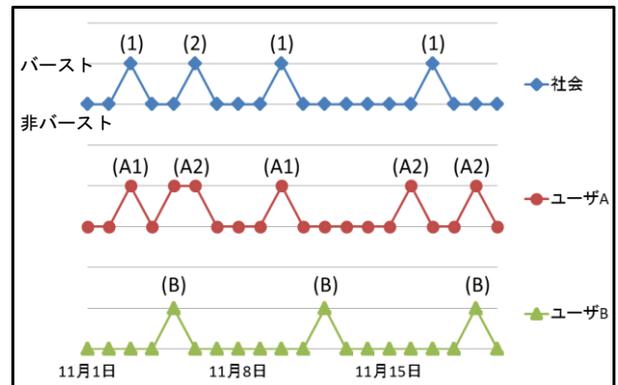


図1：個人と社会全体のバースト比較（抜粋）

として表1のように実験の設定を行った。この設定で解析結果をグラフにしたものの抜粋を図1に示す。

図1のグラフから各ユーザの状況について考えると、次のようなことが推測できる。

- (1)と(A1)のバーストが同時期に起きていることから、ユーザAはこの期間は広く注目されていた事象に対して同じように注目した。
- (A2)や(B)は、社会でバーストしていない期間でバーストをしていることから、ユーザA, Bはこの期間に独自の興味を持っていた。
- (B)は(1)から2日遅れて定期的にバーストをしているとも見ることができる。これより、Bは何らかの理由で社会的な事象に対して遅れて反応をした。

その後実際のツイート内容を目視で確認したところ、これらの推測が概ね正しかったことを確認できた。

4. 今後の予定

より様々な条件での実験と考察を重ねて多くのデータを収集し、提案手法の有効性の検討を進める。

参考文献

- [1] M. Mathioudakis, N. Koudas. Twittermonitor: trend detection over the twitter stream. Proceedings of the 2010 ACM SIGMOD International Conference on Management of data p.1155-1158.
- [2] 北田剛士, 風間一洋, 榎剛史, 鳥海不二夫, 栗原聡, 篠田孝祐, 野田五十樹, 斎藤和己. Twitter のトピック変遷の可視化法の提案. DEIM Forum 2015 E2-6, 2015.