

## システム情報工学研究科修士論文概要

年 度	平成 24 年度	学位名		修士(工学)
専 攻	知能機能システム	専攻	著者氏名	高橋 佑介
指導教員氏名 宇津呂 武仁				
論文題目  時系列トピックのバーストの検出				
論文概要 <p>Kleinberg のバースト解析モデルは、一般的に、時系列文書中に出現するキーワードが、バーストしているか否かを同定する手法である。バースト解析は、例えば通常と異なる特異な出来事が発生し、ある時からある話題に関する記述が急激に増加するような現象を、自動で同定するために用いられる。しかし、キーワード単位のバースト解析を行った場合、同じバーストの原因を指す似通ったキーワードが、バースト解析の結果として複数現れるため、結果として得られる情報が冗長になりやすい。</p> <p>そこで本論文では、時系列文書集合に対して時系列トピックモデルを適用することにより、まず、トピックの系列を求める。そこで得られたトピックを単位として Kleinberg のバースト解析モデルを適用することにより、キーワード単位のバースト解析における冗長性を集約し、話題のまとまりの単位でのバーストを同定する。Kleinberg のバースト解析モデルは、各日におけるキーワードの関連記事数を定義し、時系列文書におけるその出現確率の期待値からコスト関数を決定することで、解析期間におけるバースト期間・非バースト期間を切り分けるものである。そこで本研究では、各日におけるトピック <math>z_n</math> の関連記事数を、文書 <math>b</math> ごとのトピック <math>z_n</math> の確率 <math>p(z_n   b)</math> の総和とすることにより、トピックの関連記事数を定義した。これにより、時系列文書集合に対して、トピックのバーストを同定する方式を確立した。さらに、密接に関連しあう 2 種類の情報源から生成される時系列文書集合から推定した時系列トピックを対象として、トピックのバーストを同定する方式を確立した。特に、この方式においては、情報源ごとに独立にトピックのバーストを同定する機能を実現した。時系列ニュース記事およびツイッターの 2 種類の情報源を対象として、この方式を適用し、その有用性を評価した結果、時系列ニュース記事、および、ツイッターの双方において、最大約 90% の精度でトピックのバーストを同定できることを実証した。また、時系列ニュース記事とツイッターの 2 つの情報源間で同一トピックについてのバーストが同日に発生した場合に、実際に 2 つのバーストが密接に関連する割合を評価した結果、87.5% と高い精度を達成した。一方、その他の多数のトピックにおいて、ツイートでのみのバーストを容易に観測することができた。</p>				
審査日	平成 25 年 1 月 31 日			
審査員	(大学名 職名)	(学位)	(氏名)	
主査	筑波大学 教授	博士 (工学)	宇津呂 武仁	
副査	筑波大学 教授	工学博士	白川 友紀	
副査	筑波大学 准教授	博士 (工学)	中内 靖	