

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 久野 雄一郎
指導教員氏名 鬼沢 武久			
論文題目 Twitter における極大部分文字列の時系列相関を用いた類似単語抽出とその検索支援への応用			
論文概要 インターネットおよび Web サービスの普及により、流通情報量が爆発的に増えたことで、ユーザは膨大な情報の中から必要な情報のみを見つけ出すことが難しくなっているため、検索支援は重要になってきている。欲しい情報を見つける場合にキーワードを用いた検索は日常的に行われるが、検索するためのキーワードとなる固有名詞は日々のニュースや流行に伴って増加し続けている。さらに、近年のソーシャルメディアの登場により誰もが容易に情報発信できるようになったため、ソーシャルメディアの中で固有名詞の類似単語である俗語や略語等が数多く使われている。そのため、効率的な情報検索を行うためのポイントであるキーワードをする際に、これら固有名詞と類似単語の増加によって、この選択自体が難しい問題となってきている。本研究では、この問題を解決するために、固有名詞と類似単語の対応関係について着目する。本研究では、Twitter 検索支援を目的とした類似単語を考慮したキーワード推薦のために、ある話題に出現する類似単語の判定手法を提案する。今回の手法は Twitter の即時性の高さに注目した新たな類似単語の判定方法で、テレビ番組などの多人数が同時に共有する話題におけるツイート群では類似単語同士の出現傾向が類似する性質を利用している。より具体的には、ある話題に関するハッシュタグ付きのツイート群から固有名詞を抽出し、それらの時系列相関と共起確率から単語間の類似度を算出、閾値処理により類似単語の判定を行う。検証実験として、2011 年 12 月 31 日および 2012 年 12 月 31 日に行われた紅白歌合戦とに関するハッシュタグを含むツイートに対して提案手法を適用することによって、その有効性を検証する。			
審査日	平成 25 年 1 月 30 日		
審査員	(大学名 職名)	(学位)	(氏名)
主査	筑波大学 教授	工学博士	鬼沢 武久
副査	筑波大学 講師	博士(工学)	延原 肇
副査	筑波大学 教授(連携大学院)	博士(工学)	野田 五十樹