

システム情報工学研究科修士論文概要

年 度	平成 27 年度	学 位 名	修士(工学)
専 攻	知能機能システム専攻	専攻	著者氏名 今田 貴和
指導教員氏名 宇津呂 武仁			
論文題目 ウェブ検索における関心动向のトピックモデリングによる市場シェア分析			
論文概要 <p>インターネットの普及により、日頃からウェブサイトを開覧する機会が増えている。ウェブ閲覧者の多くは、自らの関心事項についてGoogle やYahoo!, Bing といった検索エンジンを用いてウェブ検索を行っている。これに対して、近年、検索ログを分析することにより、景気の動向、選挙結果の動向、インフルエンザの動向、等を予測する試みがなされており、検索者の関心の割合と実社会の動向を比較する取り組みが盛んである。本研究では、このような検索者の関心の割合と実社会の動向を比較する取り組みに関連して、特定商品ジャンルにおける製品・サービス等の供給者である複数の企業の間で、検索における関心の割合を比較するというタスクを設定する。そして、検索における関心の割合が、実社会における売れ行きシェア統計との間でどの程度の相関を持つのかについて分析を行う。</p> <p>本研究の方式では、検索における関心の割合を示す指標として、検索エンジン・サジェストを使用する。検索エンジン・サジェストとは、検索エンジン会社に蓄積された検索語のログを用いて提示される語である。検索エンジン・サジェストには、検索における関心が反映されている。提案手法においては、まず、分析対象の企業をクエリ・フォーカスとして検索エンジン・サジェストを最大約1,000個収集する。そして、クエリ・フォーカスに加えて一つの検索エンジン・サジェストを指定したAND検索によってウェブページを収集する。最大約1,000個の検索エンジン・サジェストに対してこの方法を適用することにより、各企業に対する大規模ウェブページ集合を収集する。さらに、複数の企業に対して収集されたウェブページから構成される混合文書集合を作成する。この混合文書集合に対してトピックモデルを適用し、トピック集合を生成する。以上の手順によって生成されたトピックにおいては、「テレビ関連製品」、「パソコン関連製品」等、商品ジャンルの話題が含まれる。そこで、それぞれの商品ジャンルに対応するトピックについて、検索エンジン・サジェストの企業別割合の統計分布を作成する。さらに、作成した検索エンジン・サジェストの企業別割合が、実際の商品売り上げ統計との間で相関を持つか否かを調べるために、価格.comにおける統計情報との間の相関について分析した結果、高い相関が見られた。以上より、本論文では、検索における関心の割合を手がかりに、実社会における市場シェア統計を推定する方式の有効性について分析を行った結果について述べる。</p>			
審査日 平成 28 年 1 月 27 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	博士(工学)	中内 靖
副査	筑波大学 教授	博士(工学)	古賀 弘樹