

豊橋技術科学大学長 殿

平成 17 年 11 月 30 日

審査委員長 青野 雅樹



論文審査及び最終試験の結果報告書

このことについて、下記の結果を得ましたので報告いたします。

記

学位申請者	Xu Yinghui	学籍番号	第 0 1 9 3 0 5 号	
申請学位	博士(工学)	専攻名	電子・情報工学専攻	
論文題目	A Study on Empirical Models for Large Scale Information Retrieval (大規模情報検索のための経験的モデルに関する研究)			
公開審査会の日	平成 17年 11 月 15 日			
論文審査の期間	平成17年 9 月 26 日~平成17年	11月30日	論文審査の結果	合格
最終試験の日	平成 17 年 11 月 15 日		最終試験の結果	合格
論文内容の要旨	<p>本研究は、大規模な情報検索のためのモデルを示し、それを応用した検索システムについて論じている。第1章では、情報検索システムで用いられる潜在的意味索引、ページランクアルゴリズム、Katzの単語分布モデルについて、本論で扱う問題点を述べ、本研究の意義について述べている。第2章では、従来のベクトル空間モデル、Webページの参照関係のモデル、単語の分布モデルについてまとめている。第3章では、潜在的意味索引における意味を構成する最適な次元数が、質問ごとの意味空間を生成するという着想を用いることで、1あるいは2になることを論じ、それをもとにした情報検索システムが構築できることを述べている。そして、従来の潜在的意味索引では大規模な文書を対象に実行するには計算量の問題があることを示したあと、提案する着想では大規模な文書対処でも実行できることを明らかにしている。第4章では、従来のWebの情報検索で用いられている解析で、ページランクアルゴリズムによるリンク解析について、実際には性質の異なるリンクを区別しないという問題を示し、リンクに含まれる参照元の情報を含んだ仮想文書をつくり、そこでの単語の出現頻度からリンクを類別する着想を述べている。そして、その着想を実際のWebを対象とした情報検索の課題について適用し、効果があることを示している。第5章では、広くKatzのモデルとして知られるモデルで、単語の繰り返し出現数のモデルについて内容を表す単語や固有名詞などの単語での経験的な値とのモデルに従った値において系統的な差が生じている問題を示した。そのあとに、その差を是正し、十分に自然な背景を説明できる新しいモデルを提案し、それが経験的に与えられる出現数とより一致することを詳細な実験で示している。第6章では、第3章から第5章までの学術貢献をまとめて、情報検索におけるモデルの重要性を確認し、今後の展望を述べている。</p>			
審査結果の要旨	<p>本研究での貢献は、情報検索において重要とされており、よく使われる概念を改良したことである。まず、潜在的意味索引において、意味を構成する次元を100以上とすることが多いが、質問ごとの空間において、桁違いに小さい次元で効果があることを示したことは、興味深い着想である。これを、小規模に実験するだけでなく、通常潜在的意味索引の作成が困難なほどに大規模な文書集合に適用し、いままで実現できなかったことが可能とするような貢献にしていることは評価できる。また、Webの検索について、リンク解析は基本的な手法であり、ページランクは確立されたかのように思われる方法であるが、リンクを区別するという着想をもち、それを実際に実装し、効果を測定したことは評価できる。Webの検索は、検索対象の文書のサイズが大きく、それゆえ、着想をもったとしてもそれを実装して検証するのは困難な作業であるが、本研究ではそれを実現している。また、この準備として、広く有効と知られている方法を実装して、Web検索の競争的なワークショップにおいて、優れた結果を得たことから、検証作業が緻密に行われていると判断できる。また、Katzの単語分布モデルは、情報検索システムで重要な単語のモデルとして優れていることが知られているモデルである。本研究では、Katzのモデルと実際の単語分布の間の系統的な「ずれ」の存在を指摘し、その「ずれ」を是正するモデルを提案し、大規模な文書データでの精密な実証実験で、その有効性を示した。本研究で扱っている潜在的意味索引、ページランクアルゴリズム、Katzの単語分布モデルのいずれもが、広く知られており、よく利用されるものであるため、これらのもつ問題を示し、それを解決する方法を示したことは学術的な貢献が大きいと判断できる。また、それぞれの提案のどれもが、大規模なデータを用いて、緻密に効果を検証している。それゆえ、提案したモデルと結果は情報検索システムを実現するときに有益な情報を示し、この分野の発展に大きく貢献するものと判断できた。これらの貢献は2件の学術論文と4件の国際会議での発表にまとめられている。よって、本論文は博士(工学)学位論文に相当するものと判断した。</p>			
審査委員	青野 雅樹	中川 聖一	増山 繁	梅村 恭司

(注) 論文審査の結果及び最終試験の結果は「合格」又は「不合格」の評語で記入すること。