

# 論文内容の要旨

博士論文題目 WWW 文書検索支援のための対話的な情報検索手法に関する研究

氏名 仲川 こころ

## 論文内容の要旨

1989年に誕生したインターネット上のWWW (World Wide Web) サービスは、いまやテレビやラジオに並ぶ主要メディアとして現代社会に広く浸透し、爆発的にその利用が進んでいる。それに伴い、膨大な情報に誰もが自由にアクセスできる環境が実現された一方で、WWW上の情報検索システムがきわめて重要な役割を担うようになった。現在数多くのWWW検索サービスが提供されているが、未だ問題点も多く、専門的な知識や技術に乏しくても必要な情報を簡単に探し出せるような、使いやすい検索手法の必要性が高まっている。

従来の検索支援に関する研究が主に検索の効率や精度(再現率・適合率)の改善を目的としているのに対し、本研究では特に情報検索作業に対するユーザのメンタルモデルに着目し、それに沿ったユーザにとって無理のない検索手法の考案を目的としている。本論文では、上記の目的に沿って提案した2種類のWWW検索支援手法(可変なカテゴリ構造を用いた支援手法、調整可能な文書ランキングを用いた支援手法)について、その詳細と評価実験の結果を報告する。

可変なカテゴリ構造を用いた検索支援手法では、検索のたびに、ユーザの検索目的を反映した小規模なディレクトリを対話的に構築することで、検索作業を支援する。この支援手法は、キーワード入力型の検索手法では適切なキーワードを考えることが難しく、またディレクトリ探索型の検索手法ではユーザの想定する情報の分類方法が検索目的に応じて様々に変化することに対応できない、という問題の解決を図り、ユーザにとってより自然な検索手法を提案することを目的としている。既存の検索手法であるキーワード検索手法および統計的分類手法(クラスタリング)との比較を行った評価実験では、この支援手法が特に難しい検索目的に対して有効であることが確認できた。更に使いやすさに対する被験者の主観的評価は、クラスタリングに大きく勝り、キーワード検索手法に対して遜色ない程度であった。

一方、調整可能な文書ランキングを用いた検索支援手法では、キーワード検索手法の出力である文書のランキングをユーザが直観的な方法で調整できる手段を提供することで、検索作業中に生じる心理的な負担の軽減を図る。既存のキーワード検索システムにこの支援手法を適用し、調整によるランキングの改良効果を調べた評価実験では、調整を行うことによって適合文書を選択的に上位に持ち上げていることが観測できた。これにより、下位に埋もれた未知の適合文書を発見し易くするという点で有効な手法であることが確認できた。また、調整によって全体的に適合文書の順位を上げ、不適合文書の順位を下げていることも観測でき、本研究で提案するランキング調整のアルゴリズム(クエリーベクトルの推測手法)が妥当であることも確認できた。

本研究で提案する2つの支援手法は互いに独立しており、どちらも既存の検索手法および検索支援手法に対する改良として利用・併用できる技術である。本論文では最後に今後の展望として、2つの提案手法の統合や、従来の各検索支援技術と提案手法との統合の可能性について述べる。

## (論文審査結果の要旨)

WWW情報の急速な増大に伴い、キーワード検索を中心とするWWW検索サーバが多数実用化されているが、一般ユーザにとっての使い易さに関する問題がいまだに残っている。本論文では、ユーザの検索意図を反映し、ユーザと検索システムが対話的に検索結果を精練してゆく、以下のような二つの手法を提案している。

1. 可変なカテゴリ構造を用いた検索手法：従来のキーワード検索法には、一般ユーザにとっては、入力キーワードを精練して検索結果の精度を上げるのが難しいという問題点があり、一方ディレクトリサービスには、巨大なディレクトリ中で適切なカテゴリを選択するのが難しいという問題点がある。本論文では、検索ごとにその検索目的に応じて小規模なカテゴリ構造を提供することにより、上記の両方の問題を解決する検索手法を目指している。具体的に、あらかじめ分類観点と呼ばれるサブカテゴリ集合を多数用意しておき、システムは、初期検索結果の文書集合に対して、明確さとエントロピーという二つの尺度によって、各分類観点の適切さを評価し、ユーザに提示する。ユーザが分類観点を選択すると、システムはそれに基づいて文書集合を分類する。これを再帰的に繰り返すことによって、検索目的に適合した小規模なカテゴリ構造が構成される。
2. ユーザによる文書ランキングの調整が可能な検索手法：キーワード検索によって得られた文書集合は通常、一定の計算規則で各文書に与えられたスコアの降順に表示される(文書ランキング)。しかし、ユーザにとって有用な文書がランキングの下位に埋もれていることも多く、そのような文書を手作業で探すのはユーザにとって負担が大きい。そこで本手法では、ユーザが文書ランキング中の任意の文書を直接上げ下げできるようなインタフェースを提供する。システムは、ユーザによる調整結果からユーザの検索意図を推測してスコア計算規則を変更し、新しい計算規則によるランキングを提示する。

本研究では両提案手法に基づく検索システムが試作されており、論文ではベンチマークテストならびに被験者を用いた評価実験を行なった結果が詳細に報告されている。一つ目の手法は、BMIR-J2 テストコレクション、二つ目の手法は、実在のWWW文書を検索対象にして、実用レベルの規模の検索実験が行なわれている。実験では従来の単純なキーワード検索法や統計的分類法との比較がなされ、提案手法が検索精度や使い易さの面ですぐれていることが実証されている。

これらの結果は、情報検索、とりわけ対話的WWW検索支援法における重要な知見を与えており、その発展に寄与するところが大きい。従って、博士(工学)の学位論文として価値あるものと認める。