

## タグの重要度を指定したコンテンツ検索手法の提案

宮嶋 清人

通信技術の進歩により、文字データだけでなく、画像や音声、動画など大容量のファイルを Web 上でやり取りすることが可能となってきた。こうしたファイルなどのコンテンツを検索しようとする際、検索者の意図を検索文字列として表現し検索システムに伝えることは困難な場合が多い。

そのため、ユーザーによるタグ付けを採用し、検索に利用するサービスが増加している。しかし、こうしたサービスでみられるタグ検索システムを利用して不便さを感じることもある。たとえば、検索時にタグの組み合わせを変化させ何度も検索を行っても検索できないコンテンツが多々あったり、複数のタグを組み合わせで検索を行う際に自分の意図と全く異なる内容のコンテンツが上位に多数表示されたりすることがある。

本研究で提案する検索システムは、検索を行う際、タグを指定して検索を行うだけでなく、タグの重要度を設定し、ユーザーがどのタグを重視しているかを指示することができるシステムとする。指定されたタグの重要度と、コンテンツに付与されたタグの付与率を利用してコンテンツの要求への合致度を求め、それに基づきランキングを行い、検索結果として表示する。ユーザーは重要度を操作することで、同じタグでも検索結果の順位を変動させることができ、それらの組み合わせで様々な検索結果が得られる。また、タグの付与率を用いて検索を行うことで、付与数では下位に位置づけられるようなコンテンツを大幅に見つけやすくすることができる。

検索結果のランキングを作成する方法として、タグの重要度と付与率の距離によってコンテンツのユーザー要求への合致度を評価する方法と、タグの重要度と付与率の類似度によって評価する方法を採用した。また、好ましくないタグが付与されたコンテンツの順位を低下させるよう、重要度に負の値も設定できる方式を導入した。併せて、タグの完全一致のみではなく、部分一致した場合についても適切に合致度を評価できる方式とした。

実験により、指定したタグを含むコンテンツを単純に検索する方式や、タグの付与数やコンテンツのブックマーク数などで評価する既存のタグ検索システムと比較して、検索可能なコンテンツの範囲の増加と適切な順位づけが確認できた。

しかし、指定したタグを含まないコンテンツに対してどのような補正をかけるか、部分一致するタグをどのように評価するかといった点については、さまざまな方法が考えられ、何度も実験を行い、経験的に最適な方法を見つけていく必要がある。

(指導教員 谷口祥一)