

ユーザの Web 探索履歴におけるキーワード遷移に基づく Web 文書推薦システム

枝 隼也

Web には誰でも情報を発信することができるため、新たなページが生まれるたび、古いページが残ったまま別の場所に新しい情報が書かれ、重複した内容のページが様々な場所に点在している。また、ページの著者が匿名である、内容の信憑性が疑わしいページも多く存在する。

このため、ユーザは情報を得るために、興味や関心に依じて何度も検索を繰り返し様々なリンクを辿って複数のページを見比べるという探索行動が必要となり、情報探索におけるユーザ負担は決して少なくない。この問題を解決するために、ユーザの情報探索における興味を的確に捉えた情報推薦を行おうとする様々な研究がなされている。

本論文では、ページを移動することによって変化する話題からユーザの興味を推定する手法を提案する。提案法を実装した Web 文書推薦システムでは、ユーザが閲覧したページのキーワードを抽出し、ページを移動することによって変化する話題のキーワード遷移からユーザの興味を推定する。ここでキーワード遷移とは、各ページに出現するそのページの話題を表すキーワードをノード、キーワード間の閾値付き *Simpson* 係数の値をエッジの重みとした非循環有向グラフによって表わされる。ユーザが新しいページを閲覧するたびに、グラフを再構成するとともに、グラフを解析して、ユーザの興味を推定し、次に閲覧するのに適したページを推薦する。本研究は、ユーザの Web 探索履歴を用いて Web 文書の推薦を行い、URL やページの移動順の他にそのページの内容まで利用することに特徴がある。また、ユーザの Web 探索行動の過程で変化する興味を推薦に随時反映することも特徴である。ユーザの Web 探索行動をキーワードのリストの積み重ねで表現することと、*Simpson* 係数によって構築されるキーワード遷移グラフの有効性を検証するために、ユーザの Web 閲覧行動を 4 つのパターンに分けて利用者実験を行った。

その結果、キーワードのリストの積み重ねからでもユーザの Web 探索行動の特徴を知ることができ、またシステムがキーワード遷移グラフを用いて提示するキーワードとユーザが選択するキーワードも重なるものがあることが確認できた。

今後は、ユーザの Web 閲覧行動を精緻化するとともに、パターンを考慮した推薦法に拡張していく。

(指導教員 佐藤 哲司)