

# Web ページの重ね合わせを利用した効率的な情報アクセス手法

## An Approach for Effective Information Access with Overlapping Operation on Webpages

IDS07193T 富永一成 Kazushige TOMINAGA

### 1. はじめに

現在、Web 上には多種多様な情報が存在し、日常生活の様々な場面において、Web ページによる情報取得が必要とされることが多い。近年、Web 上の情報を利用する環境としては、パーソナルコンピュータばかりではなく、携帯電話やスマートフォン等の情報端末による利用が広まりつつある。一般的に、Web 上に存在する特定の情報にアクセスするためには検索エンジンを利用することが多い。

検索エンジンにおける情報取得は、既知情報の検索と未知情報の検索に大別できる。未知情報の検索では、様々な検索キーワードを用いた検索と検索結果の閲覧を繰り返しながら、複数の候補を見つけ、それらを比較しながら、意思決定を行う際に利用される。しかし、現在の検索エンジンでは、ユーザが指定したキーワードを含む Web ページを重要度順に提示することにより、既知情報を検索するには適しているが、未知情報を検索して意思決定を行う作業を効率的に行うために、十分な機能を提供しているとは言い難い。特に、携帯電話等の小型モバイル端末では、貧弱な入力インタフェースしか提供しないため文字入力がつづらく、こうした作業を、より一層困難にしている。

本研究の目的は、同一の対象に関する複数のページに存在する情報を利用して、総合的に比較検討する作業を支援することである。本研究では、ユーザの欲しい情報に関連する複数の Web ページを重ね合わせることにより、ユーザがクエリを入力等の冗長な操作を必要とせず欲しい情報を効率的に取得できる検索方式を提案する。

### 2. アプローチ

武田ら [1] は、実空間上のコンテンツを携帯電話のカメラによって拾い上げ (Pick)、他の実空間上のコンテンツに重ね合わせる (Lap) ことにより、複数の実空間上のコンテンツを仮想的に融合し、ユーザの目的に適した形式で情報を提示するという Pick and Lap 手法を提案している。Pick and Lap 手法では、複数のコンテンツの情報を融合するためのメタファとして重ね合わせという操作を利用している。

本研究では、未知情報を比較検討する際に、ある一つのサイトにおける情報が、別のサイトにおいてどのように表現されているかを調べる操作が多く行われることに着目し、ユーザが発見した Web ページに含まれる対象に関する、別の情報を取得するために重ね合わせ操作を利用する。具体的には、ユーザが興味を持つ Web ページ A を、関連情報へのリンクを含む Web ページ B に重ね合わせたとき、Web ページ B から数クリックで到達可能な Web ページの中に含

まれる Web ページ A に関連が強い Web ページ C を提示する。(図 1)。

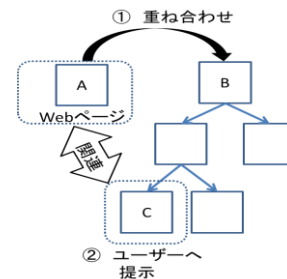


図 1: 重ね合わせによる Web ページの提示

### 3. 実現方法

二つの Web ページ間の関連性を見つけるために、本研究では TF・IDF 値を利用して、ユーザの求める情報を含む Web ページのキーワードを取得する。TF・IDF 値とは、文書中の単語の重要度を表す代表的な特徴量である。次に、Web ページ A と、Web ページ B がリンクしている Web ページの類似度を求める。類似度の計算にはコサイン相関値を利用する。ここでは重ね合わせる Web ページに含まれる単語の重みベクトルを  $a$ 、重ね合わせられる Web サイトの Top ページがリンクしている Web ページのうち、一つに含まれる単語の重みベクトルを  $b$  とする。コサイン相関値は次の式で表される。

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|} = \frac{\sum_{i=1}^n a_i \cdot b_i}{\sqrt{\sum_{i=1}^n a_i^2} \times \sqrt{\sum_{i=1}^n b_i^2}}$$

この計算式によって Web ページ A と、重ね合わせられる Web ページがリンクしている Web ページ B とのそれぞれの類似度を求める。そしてその結果を元に、Web ページ A に関連の高い順に Web ページ B がリンクしている Web ページを表示する。

### 4. まとめ

本研究では、二つの Web ページを重ね合わせることでユーザがクエリを入力等の冗長な操作を必要とせず欲しい情報を効率的に取得できる検索方式を提案した。今回は、重ね合わせる Web ページを単位としたが、一つの Web ページに複数の情報が含まれることがある。Web ページに含まれる情報のまとまりを単位として、関連情報を抽出することが重要であると考えている。

### 参考文献

- [1] 武田 十季, 鶴野 玲治, 牛尼 剛聡: “Pick and Lap: カメラ付き携帯電話を用いた実空間コンテンツの重ね合わせによる効果的な情報閲覧”, DEIM Forum 2010