

うる覚え Web ページ再発見のための閲覧履歴における被記憶ページ特徴群の発見

Discovery of Memorized Features from Web Browsing History for Re-finding Vaguely-memorized Pages

井倉 真一[†]
Shinichi Ikura

近藤 司[‡]
Tsukasa Kondo

原田 史子[§]
Fumiko Harada

島川 博光[§]
Hiromitsu Shimakawa

1. はじめに

近年, WWW の普及により, ユーザは Web ページ群から手軽に欲しい情報を得ることができるようになった. ユーザは大量の Web ページ群を閲覧する過程で, 過去に閲覧した Web ページを再度, 閲覧したくなることもある. この場合, ユーザはブラウザの履歴検索機能を用いて, 閲覧履歴からもう一度閲覧したい Web ページを再発見する必要がある. しかし, 検索エンジンを用いて, もう一度閲覧したい Web ページを再発見するには困難な場合がある. ユーザが Web ページの内容を曖昧にしか覚えていない場合, もう一度閲覧したい Web ページを再発見するための適切なキーワードを思いつけない. そのため, ユーザは検索エンジンを使えない.

本研究では, ユーザがもう一度閲覧したい Web ページの内容を表すキーワードが分からず, もう一度閲覧したい Web ページを再発見できない状態をうる覚え状態とする. ユーザがうる覚え状態時に, もう一度閲覧したい Web ページを目的 Web ページとする. 本論文では, キーワード以外の情報を使い, 目的 Web ページを絞り込む手法を提案する.

2. 閲覧履歴における検索

ユーザは Web ページを閲覧したさいに, その Web ページの漠然とした内容を, キーワード以外にもさまざまな情報群からもう覚えしている. これらの情報群は, 閲覧履歴から目的 Web ページを探すさいの手がかりになると考えられる. またうる覚え状態であるため, これらの情報群に関してもう覚えしていることが考えられる. そして, キーワード以外の情報というのは単独では目的 Web ページが含まれると思われる Web ページ群を絞り込む力が弱いと考えられる. そのため, これらの情報群を検索に利用する場合は単独で使うのではなく, 複数の情報群を用いて検索に利用する必要があると考えられる. Web ページの閲覧と結びつく情報を条件として閲覧履歴内を調べることで, ユーザは目的 Web ページの内容を表すキーワードが分からない場合でも, 目的 Web ページを閲覧履歴から探し出せると考えられる.

3. うる覚えの Web ページの再発見支援

3.1 質問回答による目的 Web ページの再発見

本論文では, ユーザのうる覚え状態において, 目的 Web ページの再発見を支援する手法を提案する. 本手法では, ユーザへ対話的に質問しその回答を取得することを繰り返す. 質問の回答に合致する Web ページ群に得点付けをし, ランキング形式で提示することで目的 Web ページを再発見する. 本手法の全体図を図 1 に示す. ま

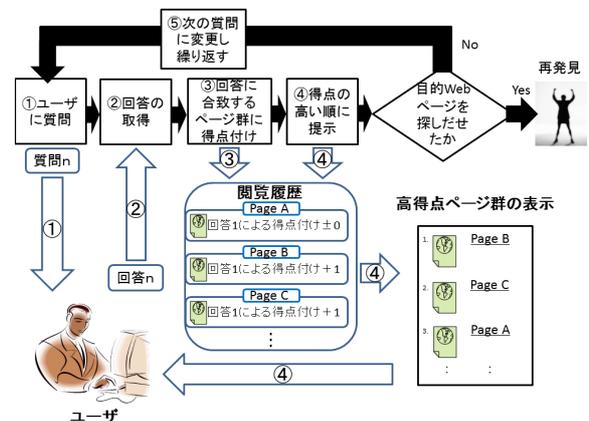


図 1: 全体図

ず, ユーザに対して「はい」または「いいえ」で答えられる質問をすることで, 閲覧履歴から目的 Web ページを再発見するために, 目的 Web ページがキーワード以外のどんな特徴を持っているかを答えさせる. 本手法では, 予めいくつかの質問が設定されている. ユーザに対して, 設定された中から 1 つ質問をする. 次にユーザは質問に対して「はい」「いいえ」「分からない」の 3 択で回答する. ユーザは目的 Web ページに対してうる覚え状態であるため, 目的 Web ページの特徴に関して詳細な情報を入力できないと考えられる. そこで 3 択で答えさせることによって, ユーザは目的 Web ページが質問にある特徴を含むか否かを判断するだけで良い! 「分からない」はユーザが質問に答えられない場合に選択する. そして質問の回答に合致する閲覧履歴中の Web ページ群に得点付けをする. これは, 質問の回答から得た目的 Web ページの特徴を使って閲覧履歴から目的 Web ページの候補を絞り込むためである. その後閲覧履歴から得点の高い順に Web ページ群をユーザに提示する. 回答による閲覧履歴の Web ページ群の得点付けがされる度に, ユーザに提示する Web ページ群を更新する. 1 つの質問による得点付けの結果をすぐに確認できることで, その度にユーザは目的 Web ページがあるかどうか確認できる. 最後にユーザが以上の手順を繰り返すと応答した場合, 次に設定された質問を使って繰り返す. 現在提示されている Web ページから目的 Web ページを発見できた場合, ユーザは繰り返さない旨を応答する.

3.2 質問のための指標群

本手法ではまず, ユーザに対して質問をする. ユーザに対する質問の回答から目的 Web ページの特徴を把握することで, 閲覧履歴からユーザの目的 Web ページを絞り込む. 2 章で述べたように, ユーザが目的 Web ページを閲覧した事実, あるいは漠然とした内容を表す情報が当該 Web ページを閲覧履歴から探す手がかりになると考えられる. ここで, 質問が対象とする内容を指標と定

[†]立命館大学大学院情報理工学研究所

[‡]立命館大学大学院理工学研究所

[§]立命館大学情報理工学部

義する。指標が満たすべき条件として、うろ覚え状態でもユーザが記憶している可能性の高い内容であるということが考えられる。本手法は質問に対するユーザの回答に基づき閲覧履歴から目的 Web ページを発見する。ユーザが記憶していない内容を質問した場合、「分からない」と回答され、目的 Web ページの特徴を取得できない。そのため本手法はこの条件を満たす指標に基づいた質問を設定する必要がある。

4. 被記憶ページ特徴群の発見と検証

4.1 予備実験

提案手法では、目的 Web ページの特徴をユーザから取得するために、予め質問群を設定する必要がある。3.2 節で述べた通り、質問は指標に基づいて設定される。指標とは、ユーザがうろ覚え状態でも、記憶している可能性の高い Web ページの内容である。ユーザがうろ覚え状態でも、記憶している可能性の高い内容を調査するための実験 1 を実施した。被験者は 10 名で、出題者と被験者の 2 名 1 組で以下の手順に従って実施した。まず、被験者に対して、ランダムにテーマを与え、そのテーマに関する Web ページを検索エンジンから検索し、閲覧してもらった。テーマを変更し、再び検索、閲覧してもらうことを計 3 回繰り返した。このさいの Web ページ閲覧履歴を収集した。次に、被験者の閲覧した Web ページ群から、被験者に分からないように出題者がランダムに Web ページを 1 つ選択した。ここで、出題者の選択した Web ページを選択 Web ページと定義する。被験者が選択 Web ページを特定するために、出題者に対して選択 Web ページの特徴に関して質問をしてもらう。このとき、被験者は選択 Web ページが閲覧履歴中のどの Web ページか分からない。また質問をするさいには、「はい」または「いいえ」で答えられる質問をしてもらう。被験者が質問を思いつかなくなる、または、選択 Web ページを特定できるまで質問を繰り返してもらう。表 1 は実験のさいに、被験者から得られた質問をまとめたものである。質問事項とは、被験者が質問した内容を主観的に何を表しているかを分類したもので、質問人数とは、その質問事項に対して質問をした人数を表している。表 1 より、画像と背景色に関する内容の質問が多かった。つまり、画像、背景色は指標として有効である可能性が高い。背景色を条件として Web 閲覧履歴を検索することで、Web ページを表すキーワードが分からない場合でも、目的 Web ページを再発見できることが報告されている [1]。そこで、今回は画像について考察する。画像はユーザによって記憶されている可能性が高いことは分

表 1: 各質問事項別データ

質問事項	質問人数
画像	9 名
背景色	8 名
テキスト	6 名
リンク	5 名
動画	4 名
メニューバー	4 名
広告	4 名
ブログ	4 名
フラッシュ	3 名

表 2: 印象に残っている Web ページの順番

被験者	グループ A	被験者	グループ B
1	b e c' d a	6	b c'' d e a
2	b c' e a d	7	e b c'' d a
3	c' a e d b	8	b e c'' a d
4	c' e b a d	9	b e a c'' d
5	b c' e a d	10	e b c'' a d

かったが、被験者から取得できた質問は、「人の画像がありますか?」や「ブランドのロゴ画像がありますか?」など人によって定義の基準が異なるものばかりであった。質問はユーザに対して「はい」または「いいえ」で回答できるように質問基準を明確化する必要がある。また既存研究より、Web ページの内容を初めからじっくり読むユーザは少なく、79%のユーザはざっと目を通すだけであるということが分かっている [2]。短い閲覧時間で画像があることをユーザが認識するためには、画像の大きさが重要であると考えられる。

4.2 本実験

本実験では画像について質問の判断基準を設定するために、被験者 10 名を対象として、画像の大きさと記憶されやすさについて検証した。まず、Web ページ a, b, c, d, e の 5 つ用意した。ここで、Web ページ c の画像の 1 つを拡大編集した Web ページを Web ページ c' 、縮小編集した Web ページを Web ページ c'' とする。被験者 10 名を 2 つのグループ A, B に 5 人ずつに分け、グループ A の被験者には、Web ページ a, b, c', d, e の順に Web ページを見てもらった。グループ B の被験者には、Web ページ a, b, c'', d, e の順に Web ページを見てもらった。その後、被験者に閲覧した Web ページ群から印象に残っている Web ページの順番を答えてもらった。グループ A の被験者とグループ B の被験者で、印象に残っている Web ページの順番に違いがあるのかを検証した。本実験の結果を表 2 に示す。表 2 より画像を拡大した Web ページの方が、画像を縮小した Web ページよりも印象に残りやすいことが確認できる。つまり、画像の大きさという指標を基に、質問を設定することで、有効な質問になる可能性が示された。

5. おわりに

本論文では、うろ覚えの Web ページを再発見する手法を提案した。今回は、有効な指標に成りうるものを複数用意できなかったため、本手法自体の有用性を検証できなかった。今後は、質問を設定するために有効な指標群を得る実験をさらに行い、本手法の評価を試みたい。

参考文献

- [1] 石川幹直, 細川宜秀, 高橋直久: 色とその配置位置に基づいた視覚的印象による Web ページ検索手法の実現方式, 電子情報通信学会 第 16 回データ工学ワークショップ (DEWS2005), March 2005, 4B-i6 (2005).
- [2] Jakob Nielsen: How People Read on the Web, Jakob Nielsen's AlertBox, October 1, 1997