

平成 24 年度
修 士 論 文

利用者の主観的評価を考慮した
ページ評価アルゴリズムの提案

Proposal of Web Page Ranking Algorithm
by Considering Subjective Evaluation

米岡 毅

神戸情報大学院大学
情報技術研究科 情報システム専攻

内容梗概

近年、21世紀の到来と共にインターネットの拡大や情報技術の進歩などにより、情報が爆発的に増加する「情報爆発時代」を迎えている。これは、従来の新聞やテレビ、ラジオと言った媒体からインターネットへと情報の中心が移行したことを意味し、同時に個人による情報の取得・発信に対する敷居が低下していることが窺える。

一方、このような膨大な情報量を蓄積するインターネットから自身の求める情報を検索する仕組みとして検索エンジンがある。日本では Google や Yahoo, Bing と言ったものがその代表として挙げられる。しかし、対象となるインターネットの規模拡大に伴い、それら検索エンジンが候補として列挙するページ数も増加し、利用者はより断定的な情報を入力することでしか情報を絞り込むことができなくなってきている。また、検索結果として出力される情報の中には、スパムページやノイズページなどが含まれているため、利用者はそれらを閲覧・選定し、自身が求める情報を精査しなければならない。そのため、検索エンジンで用いられるページ評価アルゴリズムは、これらの様々なニーズを満たし、より利用者が情報を得やすくなるよう日々進化することが求められている。

一方で、利用者が求める情報と検索結果として出力されたページの情報との一致の度合いやページの内容に対する理解の容易さと言ったものは、システムによって機械的に測ることには限界があり、人間による精査が必要となる。

本研究では、検索エンジンの利用者がページに下す主観的評価を従来のページ評価アルゴリズムに取り入れることで、利用者の求める情報を提示することを目指した。その結果、検索エンジンのページ評価アルゴリズムに利用者の主観的評価を取り入れることで、検索効率の向上させることができた。また、過去に検索・評価された際の検索キーワードによって抽象的な検索キーワードを補正することが出来た。

目次

1. 序論	1
1.1. 研究背景	1
1.2. 研究目的	1
1.3. 論文構成	2
2. 検索エンジン	3
2.1. 検索エンジンとは	3
2.2. 検索エンジンの利用	3
2.3. 検索エンジンの歴史と至近の動向	4
2.3.1. 検索エンジンの歴史	4
2.3.2. ナレッジグラフ	4
2.3.3. Graph Search	5
2.4. 検索エンジンの構成	5
2.4.1. 全体構成	5
2.4.2. クローラ	6
2.4.3. サーチャ	6
2.5. ページ評価アルゴリズム	6
2.5.1. ページ評価とは	6
2.5.2. 検索キーワードの出現頻度	7
2.5.3. 検索キーワードの用途	7
2.5.4. 検索キーワードの出現位置	7
2.5.5. 検索キーワード同士の近接度	8
2.5.6. ページランク	8
3. 主観的評価指標	10
3.1. 主観的評価指標の定義	10
3.2. 主観的評価を測る手法	10
4. 利用者の評価を取り入れた検索システムに関する先行研究	12
4.1. 主観的評価収集型	12
4.2. 閲覧履歴型	12
4.3. ブックマーク型	12
4.4. 閲覧履歴・ブックマーク併用型	13
5. システム開発	14

5.1.	開発環境	14
5.2.	運用環境	14
5.3.	システム全体	15
5.4.	クローラ	16
5.4.1.	処理フロー	16
5.4.2.	テーブル定義	18
5.5.	主観的評価収集構造	22
5.6.	サーチャ	24
5.6.1.	処理フロー	24
5.6.2.	テーブル定義	25
5.7.	画面設計	26
5.8.	ページ評価	29
6.	ページ評価	30
6.1.	従来のページ評価指標	30
6.2.	主観的評価指標	31
6.3.	スコアの合成	32
7.	目的達成型検索の検証	33
7.1.	検証目的	33
7.2.	検証条件	33
7.2.1.	検証環境	33
7.2.2.	被験者	33
7.2.3.	事前準備	34
7.3.	検証対象	34
7.4.	検証手順	34
7.5.	検証結果・考察	35
8.	追加開発	37
8.1.	検索キーワードの用途	37
8.2.	スコアの合成	38
9.	情報確認型検索の検証	40
9.1.	検証目的	40
9.2.	検証条件	40
9.2.1.	検証環境	40
9.2.2.	被験者	40
9.2.3.	事前準備	40
9.3.	検証データ	40

9.4. 検証手順	41
9.5. 検証結果	42
9.5.1. 検証実施順	42
9.5.2. 検証データ 1 の実験結果	42
9.5.3. 検証データ 2 の実験結果	43
9.5.4. 検証データ 3 の実験結果	44
9.5.5. 検証のまとめ	45
10. 締結	47
10.1. まとめ	47
10.2. 今後の展望	47
謝辞	48
参考文献	49
付録 1 : 評価実験手順 (目的達成型検索の検証)	51
付録 2 : 評価実験手順 (情報確認型検索の検証)	53
付録 3 : 電子情報通信学会 総合大会 論文 (2013.3)	62
付録 4 : 修了発表会 発表資料	63