

# 第1章 背景と関連領域

## 概要

本章では、本研究の背景と関連領域について述べる。

## 1.1. はじめに

近年、情報爆発により、膨大な量の情報にアクセスすることが可能になってきた。このような膨大な量の情報にアクセスするための手法として、様々な情報検索手法、情報フィルタリング手法、情報提示手法が提案されてきた。ここでは、これらの情報検索手法や情報フィルタリング手法の発展と共に本研究の背景について述べる。

## 1.2. 情報検索

従来、Web 検索エンジンは、人手で検索結果を作成する、ディレクトリ型の検索エンジン [1] が主流であった。例えば、初期の頃の Yahoo の検索エンジンなどがこれにあたる。だが、Web ページの急激な増加に伴い、人手だけによる Web ページの選択・分類が難しくなってきた。そこで、近年では、ロボットが自動的に Web ページを巡回して、検索結果を生成するロボット型の検索エンジンが主流となっている [2]。例えば、現在の Google 検索 [3] や Bing [4] などの検索エンジンなどがこれにあたる。

だが、現在のロボット型の検索エンジンでは、検索結果が大量に取得されるため、検索キーワードの選択が重要であるが、その適切な選択は容易でないことも多い [5]。また、検索者にとって不要な情報も大量に検索結果としてひっかかってしまい、検索結果の中から所望の Web ページを見つけるのも容易ではない場合がある [6]。そのため、近年では、情報フィルタリング\*1 や情報レコメンデーション\*2 の重要性が問われている。

また、近年では、パーソナライズド検索 ?? の研究も多くなされている。パーソナライズド検索とは、検索ユーザの過去の検索履歴などから、ユーザの興味や関心を推定し、それぞれのユーザに合わせた検索結果を提示する検索手法のことである。パーソナライズド検索の研究例として、例えば [7] があげられる。

一般的に、Web 検索の問題を解決するために、主に以下 2 種類のアプローチが取られている。

- アルゴリズムの改善による解決
- 人力利用による解決

近年では、アルゴリズムのみによる検索手法の提案ではなく、人力情報も利用した検索手法が注目されてきている [8]。その主な理由として、以下 2 つの理由があげられる。

1 つ目の理由は、現在成功している Web 情報検索システムや Web 情報フィルタリングシステムの多くが、人力情報を利用しているからである。例えば、Google 検索で利用されている PageRank [9] も被リンクを利用しているため、人力情報を存分に利用している。

2 つ目の理由は、近年、ソーシャルなシステムが広く普及してきたからである。このため、多くのユーザから発信された様々な情報を容易に取得できるようになってきた。

こういった理由からも筆者は、機械的なアルゴリズムだけでなく人力も利用した検索手法の方が、有用な検索システムが作れる可能性が高いと考えている。そのため本研究では、機械的なアルゴリズムだけでなく、人力情報も利用したアプローチを行う。

## 1.3. 情報フィルタリング ??

情報フィルタリングとは、大量の情報の中から、ユーザにとって必要な情報を取り出し、不要な情報を排除する処理を自動的に行う技術のことである [10]。情報検索と情報フィルタリングの違い

いは、検索質問が動的であるか静的であるかという点である。情報フィルタリングでは、ユーザの興味や関心などをプロファイルと呼ばれるファイルに記述しておき、これを用いて新規に生み出される情報を選別し、不要な情報を削除することが主な目的となる[?]。情報フィルタリングには、様々な手法が存在するが、情報推薦や収集に関しては、「内容に基づくフィルタリング??」と「協調フィルタリング??」という手法がよく利用される。

情報フィルタリングや情報推薦の技術を用いることによって、各ユーザにあった情報を見つけることを目的としている研究例は多い[?][?][?]。また、今流行の情報を見つけるための研究やシステムも多い[?][?][?]。また、人気の高い情報を集めるシステムは多く存在する。

これに対して、長期的に利用されている情報を見つけることを目的としている研究例はほとんど存在しない。このため、長期的に利用される情報に特化して見つけたり、収集したりすることは難しいのではないかと考えた。さらに、こういった情報の中で、あまり人気の高くない情報を探すが、特に難しいのではないかと考えた。こういった背景もあり、本研究では、長期的に利用されている情報に焦点を当てて研究を行うことにする。

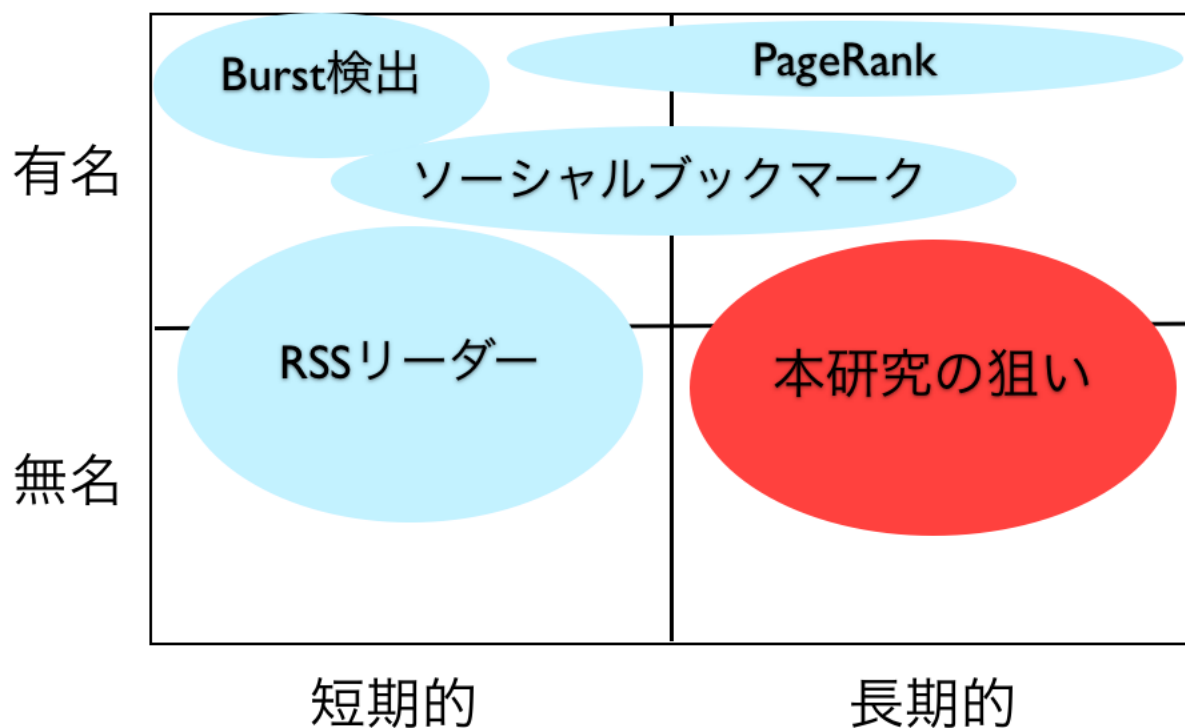


図 1.1: 狙いとする情報

図 1.1

#### 1.4. まとめ

以上のような、情報検索と情報フィルタリングの研究背景に基づき、本研究では長期的に利用される情報を検索する手法に焦点をあてて研究を行う。