

## 審査の結果の要旨

氏 名 佐藤 進也 (さとう しんや)

本論文は「Web における情報生産及び利用状況の解析とその応用に関する研究」と題し、Web を、情報のみならず、それを生み・利用する者からなるシステムとしてとらえ、その「振る舞い」を観測・解析し、それを真似、応用することで、情報の意味や価値を把握する方法について論じている。本論文は全7章より構成されている。

第1章は序論であり、研究の背景や目的、方針について述べている。まず、Web において、必要とする情報の効率的獲得が難しくなっているという問題を提起している。情報の取捨選択のためには、ことばの意味や情報の価値の判断処理が必要であるが、そのためには、情報（データ）そのものの特徴を把握するだけでなく、情報生産者や利用者の振る舞いをも把握・理解し、応用すべきであることを主張している。

第2章は、人々が情報を利用している状況を Web 検索の履歴から自動的に汲み取る方法を示している。検索履歴から、ことば（検索語）と情報（検索結果の中から検索者が選択し閲覧したページ）の組合せを抜き出し、それらを相互関係に基づいて構造化することにより意味体系を自動生成するアルゴリズムを記述している。また、そのアルゴリズムを実装した GUI の明解な情報システムの例を示している。

第3章は文脈理解の新たな手法を提案している。実社会において人物の特徴が協働・交友関係などの他者との関係により把握し得るという知識を情報処理に採り入れ、人名が用いられている文脈を理解するために、その他の人名との共起の状況を調べる手法を詳細に記述している。さらに、Web ページに出現する人名を個々の人物に対応付けるというタスクを通じ、この方法の有効性を定量的に示している。これは従来の一般的なキーワード共起を用いる手法に比べ、同姓同名人物の識別に関して有意に優れた性能を達成している。

第3章で示された手法は、人名の相互関係をネットワークにより表現し、その構造の解析により人物の同一性を識別している。第4章は、識別の精度がネットワークの構成方法に依存するという問題に着眼している。その上で、高い識別精度をもたらすネットワークの構造的特徴を定量的に明らかにした。この特徴を用い、第3章での人名の相互関係を適切に表現するための、自動化可能な方法を具体的に示した。

第5章は、Web における情報生産者としての Web サーバの振る舞いについて述べている。サーバのアクセスログ解析により、情報需要の低い状態から高い状態へ変化する過程を明らかにし、情報需要の状態を定量的に示すために、アクセスログから得られるパラメタのゆらぎのスペクトルを用いた斬新な計量法を提案し、膨大なアクセスログを使ってそれを検証した。

第6章は、情報提供者である Web サーバ群の協調により情報流通を効率化させる手法を展開している。現状では複数の Web サイトが独立に情報を生産する、いわば無秩序な状況であることを指摘した上で、この状況を改善する手段として、Web サーバが互いに自己のプロファイルを交換し、関連性のあるもの同士で自律的にグループを形成するという情報流通のアーキテクチャを提案している。こうやって形成されたグループが、利用者にとって、情報の価値評価や必要とする情報の効率的収集に役立つ。さらに、このアーキテクチャを実現する分散アルゴリズムを提案し、その妥当性をシミュレーションによって検証している。

第7章は、本研究の貢献内容を整理して、わかりやすくまとめたものである。

以上のように、本論文は、情報過負荷あるいは情報洪水と呼ばれている本質的問題を解決するには、情報というものの捉え方を見直す必要があることを指摘し、情報のみならず、それを生み出す者、利用する者までを含めて1つのシステムとして捉え、そのダイナミクスを理解、応用するという一貫した方法論を論じている。この方法論は、情報システムの新しい方向性を示した示唆に富むものである。さらに、従来の情報システムで扱うことが困難であったことばの意味の理解や情報の価値判断という問題に対して解決の方策を与えるものであり、現実の問題を解くための有用な指針となることを、第2章から第6章にわたって複数の側面から実証している。

すなわち、本研究の成果は、Web における情報過負荷問題に対処可能な情報システムアーキテクチャの設計に関するさまざまな指針を、具体的かつ定量的な裏付けを伴って示している点にある。よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。