（ 続紙 1 ）

京都大学	博士（　　情報　　学）	氏名	山本 祐輔
論文題目	ウェブ情報の信憑性分析に関する研究		

（論文内容の要旨）

本論文では、ユーザが安心してウェブ情報の検索・閲覧を行うために、社会心理学分野における信憑性の基礎理論と情報検索分野における検索・データマイニング技術を横断し、信憑性指向の情報検索・閲覧のための信憑性分析手法について提案を行っている。

社会心理学の分野では、1950年代から信憑性が重要な研究トピックとして扱われてきた。これらの分野では、信憑性に影響する要因や信憑性の認知メカニズムの解明が行われてきた。しかし、社会心理学の分野における信憑性研究は説得理論の構築が目的であり、ウェブ情報の信憑性を検証するための具体的・工学的な方法の提案はほとんど行われていない。情報検索の分野ではウェブ情報の評価に関する研究が盛んに行われてきた。しかし、そのような研究の多くは、検索要求（クエリ）と文書との類似度計算や検索ユーザの行動分析による適合性評価、リンク解析による人気度評価に基づくデータのランキング手法に焦点があてられており、ウェブ情報の信憑性評価手法に焦点があてられた研究はほとんど存在しない。

そこで、本論文では、信憑性指向のウェブ情報の検索・閲覧を実現するために、信憑性指向の検索ランキングモデルおよび閲覧情報の信憑性を判断するための情報抽出に関する提案を行っている。具体的には、本論文では「様々なウェブ情報の信憑性を評価するための汎用モデル」、「ユーザの信憑性判断モデルの推定による検索結果の最適化」、「ウェブ情報の集約による不確かなファクト型知識の信憑性を判断するための情報検索」に関する研究を行い、提案した手法の評価を行っている。

本論文は全6章から構成されている。その概要は以下の通りである。

第1章は序論であり、本論文の研究の背景、本論文の研究を行うに至った動機、および、本論文の研究の全体の概要について述べている。

第2章では、第3～5章において説明するそれぞれの研究課題に特に深く関連する従来研究について整理し、ウェブ情報の信憑性との関連および本研究との位置づけを明確にしている。

第3章では、ウェブ情報の種類や評価観点によって異なる信憑性を集合知的アプローチによって評価するためのモデルを提案しこれについて論じている。提案モデルでは特定のウェブ情報の信憑性を与えられた観点から評価するために、評価対象である

ウェブ情報をデータ対として表現し、関連するデータ対間の支持関係の強さを分析する。信憑性の分析観点に応じてデータ対の構成、データ対間の支持関係の定義を行うことで様々なウェブ情報の信憑性を評価することが可能となる。提案モデルの適用例として、コミュニティ質問応答（Q&A）サイト上の回答の信憑性を評価するために質問・回答対間の支持関係を分析するケース、および画像信憑性を評価するために画像・テキスト対間のサポート関係を分析するケースを取り上げ、実際に提案モデルによる信憑性評価の有効性を示している。

第4章では、ユーザが行った信憑性判断情報をもとにユーザの信憑性判断モデルを推定し、それをウェブ検索結果のランキングに反映させるシステムを提案しこれについて論じている。また、実験によって、提案システムによってユーザは信憑性の高いウェブページを効率良く検索できるようになったことを示しており、同時にユーザの信憑性判断モデルが検索トピックに関するユーザの知識量によって異なることを明らかにしている。

第5章では、膨大なウェブ情報を効率的に集約することによって様々な信憑性判断情報を収集するシステムについて論じている。提案システムでは、信憑性検証対象として「エジソンは電気を発明した」といったようなフレーズで表現されるファクト型知識に焦点をあて、通常の検索エンジンでは収集するのが困難な様々な信憑性判断情報をウェブからリアルタイムに集約し提示する。これらの情報を提供しユーザの信憑性判断を支援するシステムを提案し、実際にプロトタイプシステムを用いた評価実験を行っている。評価実験によって既存の検索エンジンでは収集しづらい信憑性判断情報を、提案システムでは効率良く収集できることを示している。

第6章では、本研究で得られた研究成果をまとめ、さらに今後の展開について述べている。

（続紙　２　）

（論文審査の結果の要旨）

本論文は、信憑性指向のウェブ情報検索およびウェブ情報の信憑性判断支援の実現を目指し、ウェブ情報の信憑性を評価するモデル、ユーザの信憑性判断モデルに基づくウェブ情報検索、ウェブからの信憑性判断情報の抽出手法について提案を行い、その有効性の検証を行ったものである。

情報検索の分野では従来、ユーザの検索要求に対して適合するコンテンツの検索技術に焦点があてられてきたため、情報の信憑性を評価・分析する工学的な手法は殆ど提案されてこなかった。情報の質に関して玉石混淆の状態にあるウェブにおいて、ユーザが信憑性の高い情報を取得することは困難な課題である。

学位申請者は、社会心理学分野における信憑性の基礎理論と情報検索分野における検索・データマイニング技術を横断し、これまで情報検索分野で議論されてこなかった信憑性指向のウェブ情報検索および情報閲覧支援を実現するための手法を提案している。さらに、それら提案手法について、有効性検証のための評価実験を行い、興味深い結果を提示している。

具体的には、本論文の研究によって得られた成果は以下のように要約される。

1. データ対間のサポート関係分析に基づくウェブ情報の信憑性評価モデルを提案した。提案モデルは、評価対象となっているウェブ情報をデータ対として表現し、信憑性の評価観点に応じてデータ対間のサポート関係を定義することで、集合知的アプローチによって様々なタイプの信憑性を評価することができる汎用性の高いモデルとして提案されている。信憑性評価の例題として、質問・回答対に着目したQAサイト上の回答の信憑性評価、および画像・テキスト対の信憑性評価を取り上げ、提案モデルが様々なウェブ情報の信憑性評価に有用であることを実験的に示した。
2. 信憑性フィードバック情報からユーザの信憑性判断モデルを推定し、それに基づきウェブ検索結果の再ランキングをするシステムを提案した。提案システムを評価するために大規模なオンライン実験とユーザ実験を行っており、実験によって提案システムを用いることで、ユーザは膨大なウェブ検索結果リストから信憑性の高いウェブページを効率よく検索できるようになることを明らかにしている。また、検索トピックに関するユーザの知識量に応じて信憑性判断モデルは異なり、トピックによっては知識が乏しいユーザが信憑性判断に有用であると考える指標を、知識の豊富なユーザは信憑性判断に有用ではないと考えているケースがあるという興味深い知見を得ている。
3. 不確かなファクト型知識の信憑性判断支援に関するシステムを提案した。提案システムは、不確かなファクト型知識がフレーズ形式で入力されると、膨大なウェブ情報を集約することで、信憑性を判断する際に比較すべきファクト型知識、信憑性を判断するために重要な観点など様々な信憑性判断情報を提示するものである。本研究では比較的低コストな言語処理を検索エンジンの索引情報に繰り返し適用することで、リアルタイムに効率良く判断材料を抽出する手法を提案している。評価実験では、大規模なユーザ評価を行い、信憑性判断情報収集システムとしての提案システムの有用性を明らかにした。

以上、本論文は、ウェブ情報の信憑性分析を行うための新たなフレームワークを提案し、その中で新たな情報の評価モデルや従来はなかった目的に対する情報検索手法を提案しており、学術上、実際上寄与するところが少なくない。よって、本論

文は、博士（情報学）の学位論文として価値あるものと認める。また、平成23年2月9日に論文内容とそれに関連した事項について諮問を行った結果、合格と認めた。