

氏 名	濱崎 雅弘
学位（専攻分野）	博士（情報学）
学位記番号	総研大甲第 860 号
学位授与の日付	平成 17 年 3 月 24 日
学位授与の要件	複合科学研究科 情報学専攻 学位規則第 6 条第 1 項該当
学位論文題目	不均質な情報源間での情報共有支援
論文審査員	主 査 教授 武田 英明 教授 山田 誠二 教授 相澤 彰子 特任教授 中小路 久美代（東京大学） 教授 角 康之（京都大学）

本論文は多種多様な情報源が存在する World Wide Web（以下、Web）における情報収集を支援する技術およびシステムを提案するものである。

近年、ネットワーク技術の発達と普及に伴い、オンライン上に膨大な情報が蓄積されるようになった。それらの情報の多くはネットワークに接続さえすれば誰でも容易にアクセス可能であり、私たちは日々溢れんばかりの情報に囲まれている状態にある。そのような傾向は Web において特に顕著に見られ、その結果、多くの情報の中から必要なものを取り出すのが困難である「情報過多」と呼ばれる状況になっている。この問題を解決するために様々な情報検索や情報推薦技術が提案されているが、Web 上での情報発信量は増す一方で情報過多問題を解決するには至っていない。

そこで本研究では、Web 上の情報源の多くが Web 上の情報を収集し同時に発信もしている「情報発信者」である点に着目し、情報収集支援の対象となる利用者が一つの情報源となり、そのような情報源間での情報共有を支援することによって情報収集の支援を行うというアプローチを取る。これは情報収集の支援に積極的に人の知識を利用するアプローチであり、品質にばらつきのあり内容分類だけでなく価値判断が重要な Web 情報には適していると考えられる。

本論文は 6 章からなるが、大きく分けると 2 部からなる。前半では、どのようにして互いに異なる概念体系を持つ不均質な情報源間での情報共有を実現すればよいかということについて述べる。後半では、そのような不均質な情報源が複数存在し、ネットワーク状に情報共有を行う場合において、どのようにして適切な共有相手を発見するか、また、そのようなネットワーク構造を用いた情報共有はどのような特性を持つのかについて述べる。

1 章では、本論文の目的と問題の分析、および本研究のアプローチを述べる。この章において、本論文での情報収集支援とは情報を収集する主体的な存在の間での情報共有を支援することによって為されることを示す。その上で本研究で取り組むべき課題として情報共有の問題点を示し、それに関係する既存の研究を概観して位置づけることにより、本論文が扱う領域について明確化を行う。章の最後に本論文の構成を述べる。

2 章では、異なる概念体系を持つ不均質な情報源間での情報共有の方法について議論する。これは情報共有を行う情報源をそれぞれどのように表現するかという問題と密接に関係する。本研究では、情報源をそれぞれ概念階層で表現し、共有対象となる情報はその概念階層の中に格納されているものとした。この手法の有効性を被験者を用いた実験により検証する。

3 章では、4 章で概念階層を用いた情報共有の有効性が示されたのを踏まえ、類似する概念階層の発見手法について議論する。概念階層内の情報を利用した類似概念階層の発見手法はいくつか提案されているが、本研究では概念階層内の情報の内容的類似性と、概念階層の構造自体の特徴を用いた類似概念階層発見システム WebHical を提案する。この手法の有効性を実データを用いた実験により検証する。

4 章では、情報共有の相手をどのようにして見つけるかという点について議論する。より良い情報共有を実現するためには共有相手の選択が重要である。しかし、一般に膨大な数の共有相手となる情報源の中から適切なものを見つけるのは困難である。そこで本研究

では、友人に友人の友人を紹介してもらおうという私たちが日常生活において行っている新しい知り合いの獲得手法を模倣した、近傍仲介法を提案する。近傍仲介法はボトムアップな共有相手発見手法であり、トップダウンなマッチングシステムとは異なった特徴を持つ。この手法の有効性をシミュレーション実験により検証する。

5章では、情報源である人がもともと知り合いである相手をシステムに登録し情報共有を行った場合にどのような振る舞いが見られるか、また、そのような情報源のネットワーク（情報源が人に限定されるのであればパーソナルネットワーク）が、情報共有にどのように応用することができるか、これら点について実際にシステムを運用し、その利用ログの分析から検証する。

6章では、本論文の結論と展望を述べる。

本論文は、多様な情報源が情報共有を行うための技術に関する研究をまとめたものである。インターネットにおいては、多様な情報源が存在している。さらに、インターネットにおいては情報源は情報を提供するだけでなく、他の情報源のもつ情報を収集したり、関係付けたりすることも一般に行われており、情報源同士がインタラクションをもつようになっている。この研究ではそのような多様な情報源がいかにほかの情報源との関係を見つけ、情報を共有・交換するかについて、2つの視点から研究を行っている。第1点は情報源のもつ情報の類似性をどう計算するかというものである。第2点は情報源間の関係が形成するネットワークをどのように扱うか、という点である。それぞれの視点に対してまた2つのアプローチで、技術的可能性を追求している。

1. 情報源のもつ情報の類似性の計算

ここでは扱う情報を階層的な構造と仮定し、その情報間での類似性の計算について論じている。

1a. 共通話題ネットワークに基づいた情報共有

ここでは、個人のもつ WWW ブックマークを利用した情報共有について述べている。ブックマークの持つ階層性を利用して、サブツリー単位での類似性の発見がユーザに効果的であることを検証した。また、ユーザ間での情報共有の受け入れやすさは相互のブックマークの構造的類似性によることも示した。

1b. 階層的知識と内容的類似性を用いた異なる概念体系の統合

ここでは構造だけを利用した類似性発見アルゴリズム HICAL と個々の情報のもつ類似性を併用する仕組みを提案している。これにより構造的な類似性が少ない場合でも、内容的な類似性から部分的に補えることを示した。

2. 情報源間のネットワークの利用

情報源間に関係が存在すると、全体としてネットワークになる。インターネットではこのようなネットワークは大規模かつ複雑になりがちである。2つめの視点ではこのネットワークという視点から、その改善手法の提案や実際の利用でのネットワークの分析などを行っている。

2a. 仲介によるネットワークの拡張

インターネットのような状況では全ノードを一元的に管理することは難しく、分散的に処理するしかない。このような状況での情報源のネットワークを改善するアルゴリズムとして近傍仲介法というものを提案している。これは各情報源の直接つながっている相手同士の関係の追加、削除だけを行うだけで全体のネットワークを改善を行うというものである。この手法に基づくシミュレーション実験によって、本手法が一元管理手法における最適値に十分近い値をだすようなネットワークを構築できることを示した。

2b. パーソナルネットワークを用いた会議支援

ここでは実際の会議支援をケーススタディに情報源間のネットワーク（ここでは参加者がつくるネットワーク）の性質の分析や情報共有のために利用について論じている。会議支援システムの中に知人関係を入力する仕組みを作り、それを用いて情報アクセス経路の生成や情報推薦を行い、このようなネットワークが情報共有で有用であることを示した。イ

インターネットにおいては多様性は必須の性質である。

本研究は、各々の情報源の持つ多様性を損なわずにむしろそれを利用して情報共有を実現しようとしているという点で独創的な研究である。そのような新しい仕組みを現実のデータを利用して研究を行っており、この点は大いに評価できる。以上の点をもって、この論文は博士（情報学）の学位論文として十分な価値があるものとして認められる。