

|         |                                       |       |
|---------|---------------------------------------|-------|
| 称号及び氏名  | 博士（工学）                                | 佐賀 亮介 |
| 学位授与の日付 | 平成20年3月31日                            |       |
| 論文名     | 「社会的ネットワークと履歴を用いた<br>知的システム構成法に関する研究」 |       |
| 論文審査委員  | 主査                                    | 辻 洋   |
|         | 副査                                    | 松本啓之亮 |
|         | 副査                                    | 石渕 久生 |
|         | 副査                                    | 宮本 貴朗 |

## 論文要旨

知識は、文書や画像、動画など様々なコンテンツの中に埋もれており、知的システムを通してその知識を正しく見つけ出し、抽出し、共有することが求められている。ここでいう『知識』は厳密に情報を昇華してできた知識だけでなく情報そのものを含めたものを指し、『知的システム』はその知識を取り扱うシステムのことを指す。知的システムに実装された情報検索技術は、大量のコンテンツに埋もれた、知識や情報にたどりつくための方法として開発され、複数のキーワードなどを指定することによってコンテンツへのアクセスを可能にする。

昨今、知識は人の手では処理しきれないほど溢れてきている。知識を内包するコンテンツの供給形態は変化し、コンテンツが消費者によって生成される時代、つまり **Consumer Generated Media** へと変化している。インターネット利用者は **Blog**、**RSS**、ソーシャルブックマークなどを駆使して知識を生成し、その結果、個人が手に入れることができる情報は膨大になっている。そして、コンテンツとそこに埋もれる知識を管理することは、ますます重要となってきた。

しかしながら、知的システムにおける知識の共有と獲得、利用において次にあげる3つの問題が存在する。

- ・ 提供される知識が同質のものに偏り、多様性が失われ、イノベーションが起きにくくなる。
- ・ 変化する知識に対して、周辺知識を手に入れることができたとしても、利用者が必要な知識を取り出すことが難しい。
- ・ 一時的な出来事などから得られる知識を、現状に対応する形で再利用することが困難である。

第1の問題は知識へのアクセスの方法に原因があると考えられる。知識探索の入口である検索エンジンやポータルサイトでは、不特定多数の人に用いられるように設計されているために、特定のキーワードで知識を探したとしてもその探し出した知識は同じような知識で固まってしまう。この環境に

において、ユーザが新しい知識を得るためには多大な労力が必要であり、知識探索コストが掛かってしまう。知識探索コストの上昇は知識探索行為と新たな知識獲得を阻害することとなり、結果として得られる情報量（情報利益）が減少してしまう。逆に、個人や特定分野に特化した場合、知識の方向性が固定化されることによって情報の多様性が失われ、イノベーションがおきにくくなる。従って、情報利益と情報の多様性、そしてイノベーションを考慮したシステムの構築が重要となる。

第2の問題は知識提供量が巨大になったことに関連する。変化する知識の例としては、株価情報やオークションでの商品価格などがある。ユーザがそのような知識を必要としているとき、汎用検索エンジンを用いる場合はキーワードを与えることによって知識を探し出す。しかし、このような知識は変化し続けているために、キーワードとインデックスの対応が追いつかず最新の情報を探し出すことは難しい。特に知識を構成する複数の属性が連動して変化する場合、必要な知識を獲得するために、複数の属性値を指定し、問い合わせする必要がある。この作業は検索コストを必要とする。従って、変化する知識に対応できるようなシステムを構築することが重要となる。

第3の問題は知識の記述・表現方法に関連する。従来、繰り返し利用されることを前提とした知識はある決まった形式で記述されてきた。その一方、一時的な出来事を記録した知識は、定型的・形式的な記述で記載するには十分な情報が無かったりその型に収まらなかったりするために、非定型で記述されることも少なくない。この一時的な知識を処理することは、人間にとっては容易であっても、形式的な情報しか扱うことのできない計算機にとって困難な事項であり、計算機によって意味的な内容によって認識しなければ大量の情報を処理することが難しい。よって、計算機で一時的な知識を整理・体系化し、再利用できることが重要となる。

そこで、社会的ネットワークと様々な行動履歴から、以上の3つの問題を解決する知的システムの構成法をテーマとし、研究を行ってきた。本論文は以下の構成からなる。

第1章では、本研究の動機、目的、意義について述べる。まず、本論文におけるコミュニティの種類と社会的ネットワークや履歴情報を用いることの意義、そしてその必要性を述べる。

第2章では、主に第1の問題である情報の多様性の維持とイノベーションについて注目し、考察する。情報利益を獲得し、情報の多様性を維持するには、コミュニティをまたがる弱い紐帯を経て届いた情報が有効ということがわかっている。そして、その弱い紐帯をもつユーザ（媒介者）は多様性が高くなり、得られる情報利益が大きくなる。しかしながら、どの媒介者が重要かを認識することは難しく、またネットワークの規模の変化を考慮して、低コストで定期的に算出できなければならない。

従って、低コストでの媒介者の重要度を認識する必要がある。その重要度を認識するために用いられる手法として、ネットワーク分析指標である媒介中心性を用いた方法がある。しかしながら、従来の媒介中心性を計算するには最短経路を求める必要があるため巨大な計算量が必要であり、低コストでの算出は難しく、要件を満たすことは難しい。

そこで第2章では、局所的なネットワークであるエゴネットワークによって媒介中心性を近似する手法を提案する。エゴネットワークを利用することによって、時間計算量や空間計算量を削減することができる。その特性を生かすため、エゴネットワークのリンクの接続性を用いて媒介中心性を代替する指標を導き出す。数値実験では、Watts-Strogatz モデル、Barabashi モデル、ランダムネットワークモデルといったモデルによる評価の他、実際のネットワークデータを用いて検証する。そして、その結果を考察しながら情報の多様性を考慮したシステム構成法について述べる。

第3章では、主に第2の問題である知識が変化する場合における必要知識の抽出について考察する。ここではその一例として TPO（Time、Place、Occasion）商品を対象にした電子商取引での推薦システ

ムの構成法について言及する。TPO 商品の特性は、その名前の通り、利用時期、利用する場所、そして利用時の機会に依存している。そして、TPO 商品の情報は随時生成・更新される。

推薦システムは、過去の評価から嗜好を予測し推薦を行うことでユーザの知識探索を効率化する。しかし、そのシステムにおいて TPO 商品を推薦する場合、ある時期に第 3 者の利用した商品の価値がユーザの利用時期などの条件によっては変化することを考慮しなくてはならない。例えば、その商品の 1 つであるホテルでは、宿泊の時期、利用する場所、予約のタイミングや利用場所でのイベントの有無などによって価値が変化し、その変化した価値は価格となって表れる。

そこで第 3 章では、ユーザが選択した情報を基に、統計的手法によって個人の嗜好を予測し、俯瞰型推薦、仰瞰型推薦、近傍型推薦という 3 つの視点から商品を推薦する方法を提案する。これら 3 種類の推薦方法を組み合わせることで、提案手法は現在の嗜好と過去の嗜好、そして両方を考慮した推薦を可能にする。また、記憶ベース推論を用いて俯瞰型推薦と近傍型推薦を改善する手法を提案する。それらの提案手法を 1 万人のユーザデータを用いた実験にて有効性を確認する。そして、それらの結果を交えながら、常に変化する情報にたいしての情報獲得に関するシステム構成法について述べる。

第 4 章では、主に第 3 の問題である一時的な知識の再利用法について考察する。この一例として、企業や組織におけるワークフローシステムについて言及する。企業内でのワークフローは、定型的なトランザクション型タスクと一時的なアドホック型タスクを扱う。第 3 の問題はこの後者のアドホック型タスクに相当し、アドホック型タスクの中には一時的であるが重要な知識が存在する。

しかし、アドホック型タスクに含まれる知識を保存し、再利用することは難しい。トランザクション型タスクはそのタスクについての情報が定型的に形式的情報で記述されるため、企業のルーチンワークとして利用される。一方、アドホック型タスクは、イベントによって発生する機会が多いため、トランザクション型タスクのようにタスク要素が定型的記述によって記述されることが少なく、意味的情報が重要視される。そのために、アドホック型タスクの構築は、属人的作業が必要なり、作業漏れといったヒューマンエラーを発生することが少なくない。また、個人・グループでの一時的な ToDo として処理されることが多く、再利用されにくいというえ、生産性が考慮されにくい。

そこで第 4 章ではこの問題を解決するために、アドホック型タスクをオントロジによって概念的に整理し・体系化し、再利用するための方式を提案する。ここでは、タスクでの対象物に関する『もの』のオントロジと、人物や環境に関する『コンテキスト』のオントロジを用いる。また、ビジネスプロセスのベストプラクティスから構成された MIT プロセスハンドブックを援用し、それを用いてアドホック型タスクの生産性を向上させることを考える。

第 5 章では、各章での内容を統括し将来への展望を示すことで、本論文のまとめとする。

## 審査結果の要旨

本論文は、知的システムを構成する上で生じる、知識の共有と獲得、利用についての問題点に言及し、特に、社会的ネットワークと履歴を用いた構成に関する新技術について論じたものである。具体的な知的システムの試作を通して得られた知見を考察しており、次のような成果を得ている。

- (1) 社会的ネットワークにおいて情報伝達を行う「媒介者」をグラフにおけるノードとして定義し、媒介度という指標を導入し、媒介度の高いノードを近似的に高速に検出するためのアルゴリズム

を示した。また、数値実験を通してこのアルゴリズムの有用性と弱点を示しており、大規模なピア・ツー・ピアにおける知的情報検索システムの基盤として利用することが期待される。

- (2) 場所や日時により価値が変わる商品を推薦する知的システムをとりあげ、暗黙知評価である購買履歴に着目した推薦アルゴリズムを示した。嗜好分布という概念を導入したことにより、日々変化する商品価値に追随できるところに新規性が見受けられる。このアルゴリズムを40万件の実データで評価し、推薦精度の向上を確認している。これにより、電子商取引で一般的になってきた知的推薦システムの適用領域を大きく広げるための一方策が示されたと評価できる。
- (3) アドホック型タスクを対象とし、非定型タスクの処理履歴に着目したワークフローシステムを実現するためのシステム構造を明らかにした。特に、事例の類似性と再利用の視点に基づく適合性を算出するアルゴリズムを「オントロジー」の導入により示した点に特徴がある。具体的な業務を例に知的システムの構築手順を示しており、従来定型処理のみを対象としていたワークフローシステムの適用領域を広げる方策が示されたと評価できる。

以上の研究成果は、知能情報工学分野における知的システムの構築方法論の発展に貢献するところ大である。また、申請者が自立して研究活動を行うに必要な能力と学識を有することを証したものである。学位論文審査委員会は、本論文の審査および最終試験の結果から、申請者に対して博士（工学）の学位を授与することを適当と認める。