

論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Ontology-assisted Methods for the Detection and Clustering of Hierarchical Topics on the Social Web
著者(和文)	SLABBEKOORNKristian
Author(English)	Kristian Slabbekoorn
出典(和文)	学位:博士(学術), 学位授与機関:東京工業大学, 報告番号:甲第10256号, 授与年月日:2016年3月26日, 学位の種別:課程博士, 審査員:徳田 雄洋,佐伯 元司,徳永 健伸,権藤 克彦,西崎 真也
Citation(English)	Degree:Doctor (Academic), Conferring organization: Tokyo Institute of Technology, Report number:甲第10256号, Conferred date:2016/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

(博士課程)

論文審査の要旨及び審査員

報告番号	甲第		号	学位申請者氏名		Kristian Slabbekoorn	
論文審査 審査員		氏 名	職 名		氏 名	職 名	
	主査	徳田雄洋	教授	審査員	西崎真也	准教授	
	審査員	佐伯元司	教授				
		徳永健伸	教授				
		権藤克彦	教授				

論文審査の要旨（2000 字程度）

本論文は、"Ontology-assisted Methods for the Detection and Clustering of Hierarchical Topics on the Social Web"（オントロジーを用いたソーシャル Web 上の階層的トピックの検出とクラスタリング手法）と題し、英文 8 章からなっている。

第 1 章"Introduction"ではソーシャル Web の解析の重要性が近年高まっていること、従来のソーシャル Web 解析の主要な手法がトピックモデリング法、ドキュメントクラスタリング法、コミュニティ検出法であり、これらの手法はそれぞれ問題点を持っていると述べている。

第 2 章"Background and Related Work"では、準備として、ベクトル空間法について述べ、次に従来のソーシャル Web 解析の主要手法であるトピックモデリング法、ドキュメントクラスタリング法、コミュニティ検出法の概要とその問題点について述べている。

第 3 章"Entity Recognition and Ontological Expansion"では本論文の提案手法の共通基盤となる手法について述べている。基盤となる手法は、エンティティ認識、オントロジーへのマッピング、複数のオントロジーからのクラス抽出である。これらのクラスは利用者の関心の類似性を求める際にクラスベクトルを構成すると述べている。

第 4 章"Improving Relation-based User Recommendation on Twitter with Ontological Timeline Analysis"では、オントロジーを用いる解析の 1 番目の応用技術として Twitter 利用者の利用者関係に基づく推薦手法の改善手法を示している。すなわち従来手法で推薦された利用者に対して、トピックの広がりを見計らし、一定の範囲内におさまらない利用者はツイートにおけるトピックの一貫性がないものとして推薦リストから除外する手法について述べている。

第 5 章"Ontology-assisted Discovery of Hierarchical Topic Clusters on the Social Web"では、オントロジーを用いる解析技術の 2 番目の応用技術として Twitter 利用者ツイートのトピックのクラスタリング法を示している。従来手法は事前にトピックの個数を決めなければならなかった。本提案手法では、スコープ付きトピックからクラスベクトルを算出し、利用者間の類似性を求める CF/IUF 法を用いて、利用者の類似性グラフを作成し、これらのグラフから HCS クラスタリングアルゴリズムの適用により主要トピックを抽出する手法について述べている。

第 6 章"Implementation and Algorithmic Analysis"では、本論文のトピッククラスタリング手法を Web アプリケーションとして実現する方式について述べている。入力のは出発点となる一人の利用者あるいは 1 つのキーワードである。可視化と同時進行で、利用者をトピック類似性に基づいてクラスタリングすることができると述べている。

第 7 章"Evaluation"では、従来の利用者関係に基づく推薦法との比較実験、およびクラスタリングを行う LDA 法、k-means 法との比較実験について述べている。Twitter 利用者推薦については提案手法がツイートのトピックに一貫性のある利用者を的確に推薦できることを示している。

トピッククラスタリングについては LDA 法および k-means 法と比べて 27%程度までの改善を実現できたとしている。トピックスコープのパラメータを変えることにより適切なトピックを検出することができ、例えばアメリカンフットボールとサッカーに関する話題を正しく区別して分類することが可能となったと述べている。ニュースグループのトピック分類の場合には 1 つのアーティクルが短い場合は現状手法程度の性能であったと述べている。

第 8 章"Conclusions and Future Work"では、従来手法の限界を外部オントロジーのスコープ付き展開と利用者に関心を持つトピックを表すクラスベクトルの類似性計算により解決できたことを述べている。

以上より、本論文は、外部オントロジーへのスコープ付きマッピングとそれに基づくクラスベクトルによる類似性計算が、従来手法の制限を取り除くことに有効であることを示しており、学術上、工学上、寄与するところが大きい。よって本論文は博士（学術）として十分価値あるものと認める。

注意：「論文審査の要旨及び審査員」は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。