

(続紙 1)

京都大学	博士 (情報学)	氏名	倉島 健
論文題目	ソーシャルメディアデータからの体験マイニングに関する研究		
<p>(論文内容の要旨)</p> <p>本論文では、ソーシャルメディアに存在する、人々の都市での体験に関する記述を構造化情報として蓄積し、さらに、大量の体験データに埋もれている有用な傾向を発見し活用するまでを扱う体験マイニング手法について提案を行っている。</p> <p>近年、ブログに代表されるソーシャルメディアの普及が著しい。現在のところ、ソーシャルメディアは、都市を実際に訪れた人の体験や、実体験に基づく主観的な記述を頻度高く含むという、他メディアとは異なる重要な特徴を持っている。しかし、自然言語形式で書かれた非構造データであり、その意味や属性を掴みづらく、また、日々、膨大な量のコンテンツが生成されているといった理由から、これらの情報は十分に利活用されていない現状がある。これまで、ソーシャルメディアを解析対象として「情報が新鮮である」観点に着目した話題分析技術や、特定の商品・サービスの「良し悪し」に着目した評判分析技術は盛んに行われてきたが、個々人のうちで直接的に得られた都市における体験情報を扱うことを指向し、「時間的・空間的な背景に紐づけられた都市における人間行動」を軸に分析をした研究はほとんど存在しない。</p> <p>そこで、本論文では、ソーシャルメディアに存在する、人々の体験を構造化情報として蓄積し、そこから有用な傾向を獲得して個人や企業の意志決定に利活用するまでの処理プロセスを支援するための体験マイニング技術に関する提案を行っている。具体的には、本論文では「ブログ文書からの体験情報抽出と構造化」、「人々の体験に関する有用な相関ルール (association rule, アソシエーションルール) の選択手法」、「写真共有サイトのジオタグ情報を利用したトラベルルート推薦」に関する研究を行い、提案した手法の評価を行っている。</p> <p>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。</p> <p>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。</p> <p>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、本研究との位置づけを議論している。</p> <p>第3章では、自然言語で記述された非構造データであるブログ文書から個人の体験情報を、時間、空間、行動属性から成る構造化情報として抽出する手法を提案し、これについて論じている。フィルモアの格文法解析に加え、時間的・空間的要因によって出現が規定されることを手がかりに、人間の行動内容を示す表現を適切に選択し、</p>			

抽出する特徴がある。実際のブログ文書からの体験情報抽出精度を評価する実験により、提案法の有効性を示している。また、提案手法に基づき、ユーザが体験情報集合を柔軟に検索・要約することのできる体験ブログマップ (Blog Map of Experiences) を実現している。

第4章では、構造化した人々の体験情報集合から有用な知識を抽出する手法を提案している。具体的には、アソシエーションルール抽出手法で得られた大量のルールの中から「ある特定の状況（時空条件）において特徴的に出現する行動の発見」、「人々に認知されている行動と都市で実際に人々がしている行動の差異発見」につながるものをそれぞれ選択する手法を述べている。評価実験では、時空間条件によって、また、出現するメディアの種類・性質によって行動の出現傾向がどの程度変化するかを評価する提案手法により、効率よく有用な知識を獲得できることを示している。

第5章では、蓄積した人々の体験情報を未来の意志決定に利活用する試みとして、旅行者のトラベルルートを自動拡張する手法を提案している。評価実験では、過去の場所訪問履歴をもとにユーザが好む場所の潜在的な特徴を推定すること、また、ユーザが置かれた状況（現在地、空き時間）を考慮して訪問先を絞り込むことで、提案法が旅行者の訪問先を高精度に予測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。

注) 論文内容の要旨と論文審査の結果の要旨は1頁を38字×36行で作成し、合わせて、3,000字を標準とすること。

論文内容の要旨を英語で記入する場合は、400～1,100 wordsで作成し
審査結果の要旨は日本語500～2,000字程度で作成すること。

(論文審査の結果の要旨)

本論文は、ソーシャルメディアに存在する、人々の都市での体験に関する記述を構造化情報として蓄積し、さらに、大量の体験データに埋もれている有用な傾向を発見し活用するまでを扱う体験マイニング手法を提案したものである。

ブログに代表されるソーシャルメディアは、都市を実際に訪れた人の体験や、実体験に基づく主観的な記述を頻度高く含むという、他メディアとは異なる重要な特徴を持っている。しかし、自然言語形式で書かれた非構造データであり、その意味や属性を掴みづらく、また、日々、膨大な量のコンテンツが生成され、発信されているといった理由から、これらの情報は十分に利活用されていない現状がある。

学位申請者は、ソーシャルメディアに存在する体験情報を人々が利活用する機会を生み出すための体験マイニング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. 自然言語で記述された非構造データであるブログ文書から個人の体験情報を、時間、空間、行動属性から成る構造化情報として抽出する手法を提案した。フィルモアの格文法解析に加え、時間的・空間的要因によって出現が規定されることを手がかりに、人間の行動内容を示す表現を抽出する。実際のブログ文書からの体験情報抽出精度を評価する実験により、提案法の有効性を示している。
2. 構造化した人々の体験情報集合から有用な知識を抽出する手法を提案した。具体的には、アソシエーションルール抽出手法で得られた大量のルールの中から「ある特定の状況（時空条件）において特徴的に出現する行動の発見」、「人々に認知されている行動と都市で実際に人々がしている行動の差異発見」につながるものをそれぞれ選択する手法を述べている。評価実験では、時空間条件によって、また、出現するメディアの種類によって行動の出現傾向がどの程度変化するかを評価する提案手法により、効率よく有用な知識を取得できることを示している。
3. 蓄積した人々の体験情報を未来の意志決定に利活用する試みとして、旅行者のトラベルルートを自動拡張する手法を提案した。評価実験では、過去の場所訪問履歴をもとにユーザが好む場所の潜在的な特徴を推定すること、また、ユーザが置かれた状況（現在地、空き時間）を考慮して訪問先を絞り込むことで、提案法が旅行者の訪問先を高精度に予測できることを示している。

以上、本論文は、ソーシャルメディア上の人々の体験情報を構造化し、構造化データから有用な知識を発見し活用するための新手法を提案しており、さらに、提案手法の有効性を示すために詳細な実験を行うなど、学術上、實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られる。また、平成26年8月8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。

注) 論文審査の結果の要旨の結句には、学位論文の審査についての認定を明記すること。更に、試問の結果の要旨（例えば「平成 年 月 日論文内容とそれに関連した口頭試問を行った結果合格と認めた。」）を付け加えること。

Webでの即日公開を希望しない場合は、以下に公開可能とする日付を記入すること。
要旨公開可能日： 年 月 日以降