

論文の内容の要旨

論文題目	実世界行動マイニングによる時系列行動ネットワークの構築 ～震災時の迅速な避難情報の共有に向けて～
学位 申請者	NGUYEN MINH THE

第1章：序論

計算機がユーザの状況に応じた行動を推薦することは、コンテキストウェアコンピューティングやユビキタスコンピューティング、特に避難やレスキューといった応用分野において重要な課題とされている。代表的な例として、震災時における避難行動の推薦などが挙げられる。

計算機が人間に行動を推薦するためには、まず推薦すべき行動の獲得が必要である。2011年3月11日の大震災当日、電話回線が遮断された中、避難行動や避難所の状況など震災に関する情報を交換するためTwitterが頻繁に利用された。しかし、つぶやきにはノイズ文字列が多く、他のテキストメディアより複雑かつ文法的に正しくない文が多い。また、緊急時であるため、迅速な処理が求められる。

また、人間の行動と行動間の関係を事前に形式化しておくことも必要である。そして、やはり緊急時であることを考慮すると時間情報が重要となり、各行動の時刻を明確に記述し、時系列的に整理することが必要である。

更に、Twitterなどのソーシャルメディアはユーザの自主性に依存しているため、つぶやかれなかった行動も数多く存在すると思われる。ユーザの行動に応じた情報の推薦を実現するため、これらの欠損行動を推測する必要がある。欠損行動を推測するにあたっては、人間の行動は必ず前後の行動と依存関係があること、また例えばそれがマイノリティであっても成功した行動はきちんと考慮すべきであることなど、いくつかの重要な課題がある。

そこで本研究では、まず日本語の文中に現れる行動を自動的に抽出する手法を提案する。次に、各時刻における行動と行動間の関係を表す時系列行動ネットワークをWebオントロジー言語（OWL）に従って設計する。最後に、行動の性質を考慮した行動ベース協調フィルタリング手法を提案し、行動ネットワークを補完する。

第2章：行動の自動抽出

震災等の緊急時には避難情報など行動に関する情報が特に重要であるため、本論文では行動属性の抽出に焦点をあてる。

本章では、条件付確率場と自己教師あり学習を利用して、日本語の文中に現れる行動属性（行動主、動作、対象、時間、場所）を自動的に抽出する手法を提案する。評価実験を行った結果、本手法は先行研究の課題を解決でき、一般のWebの文書だけでなく、ソーシャルメディアにおける文法的に正しくない文書にも対応できることが確認できた。本手法の主な長所は以下の通りである。

- ・ 訓練データを自動的に作成することで、ラベル付けのコストがかからない。
- ・ 行動の種類に依存せず、汎用性が高い手法である。
- ・ 行動の頻度に依存せず、マイノリティ行動も抽出できる。
- ・ 行動属性に限定することで、高速な処理が可能である。

しかし、現時点では、本手法は以下のような短所がある。

- ・ 抽出精度は文の長さや複雑さに依存する。
- ・ 文中に行動が明確に表される必要がある。

第3章：時系列行動ネットワークの構築

本章では、各時刻における行動と行動間の関係を表す時系列行動ネットワークをOWLに従って設計する。このネットワークでは、ノードは行動属性の概念（クラス）、リンクは行動属性間の関係（プロパティ）を表す。そして、上記の行動抽出手法を用いて時系列行動ネットワークを自動構築する手法を提案する。評価実験を行った結果、時系列行動ネットワークはOWLに従い、正しく設計・記述されていることが確かめられた。本手法の長所は以下の通りである。

- ・ 各時刻における行動と行動間の関係を表すことができる。
- ・ 緊急時にリアルタイムで行動の集合知を自動的に構築できる。

しかし、構築の精度は行動属性の自動抽出手法に依存するという短所が挙げられる。

第4章：時系列行動ネットワークの補完

本章では、行動の性質とユーザのゴール（避難、帰宅など）を考慮した行動ベース協調フィルタリング手法を提案し、行動ネットワークを補完する。つぶやかれなかった時間に複数の行動を行った場合でも、本手法は行動の実行時間を考慮して補完することができる。また、協調フィルタリングと言語処理を統合することで、マイノリティであるが価値がある行動に一定の重み付けを行う。評価実験で、提案手法はベースライン手法と比べて大幅に抽出精度が向上したことを示した。

第5章：震災時の応用

東日本大震災発生当日のTwitter投稿を対象に評価実験を行った結果、本研究の時系列行動ネットワークが以下の避難に関する情報を把握するのに有効であることが確かめられた。

- ・ 一時的に無料で提供してくれた飲み物、携帯バッテリーなど必需品
- ・ 電車、バスなど交通機関の運行状況
- ・ 一時的に開放してくれたトイレと避難所

また、時系列行動ネットワークでは、上記の情報を場所ごとに時系列的に整理するので、場所と時刻に応じて把握可能である。更に、時系列行動ネットワークを辿ることで、目的を達成するための行動の系列も把握可能である。

第6章：結論

本論文は、震災時等においてユーザ状況に応じて必要な情報を提供するシステムの構築を目的として、①ソーシャルメディア（ブログ、Twitter）からユーザ行動を自動抽出する手法、②OWLを用いた時系列行動ネットワークの設計・構築、および③欠損行動を行動ベース協調フィルタリングにより推測し、行動ネットワークを補完する手法を提案している。

東日本大震災発生時のつぶやきを対象に評価実験を行った結果、提案手法が避難情報を迅速に提供するのに有効であることが確かめられた。なお、本研究成果は震災時だけでなく、様々な自然災害（台風、大雪）、人的災害（スト、テロ）時の行動把握、更に観光案内、マーケティングなど多くの分野に適用可能である。

今後の課題として、抽出した行動の信頼性を検討すべきである。また、本研究の時系列行動ネットワークを用いて、震災や観光案内などにおける、ユーザの状況に応じた行動を推薦するサービスを検討していきたい。

論文審査の結果の要旨

学位申請者氏名 NGUYEN MINH THE

審査委員主査 大須賀 昭彦

委員 岡本 敏雄

委員 川村 隆浩

委員 田野 俊一

委員 加藤 聰彦

本論文は、震災時等においてユーザ状況に応じて必要な情報を提供するシステムの構築を目的として、①ソーシャルメディア（ブログ、Twitter）からユーザ行動を自動抽出する手法、②Web オントロジー言語（OWL）を用いた時系列行動ネットワークの設計・構築、および③欠損行動（ユーザによって投稿されなかった行動）を行動ベース協調フィルタリングにより推測し、行動ネットワークを補完する手法を提案した。特に、協調フィルタリングと言語処理を統合することで、マイノリティであるが価値のある行動に一定の重み付けを行う点に特徴がある。そして、東日本大震災発生時のつぶやきを対象に評価実験を行った結果、提案手法が避難情報を迅速に提供するのに有効であることが確かめられた。なお、本研究成果は震災時だけでなく、様々な自然災害（台風、大雪）、人的災害（ストライキ、テロ）時の行動把握、更に観光案内、マーケティングなど多くの分野に適用可能であると考えられる。

第2章：行動の自動抽出

本章では、条件付確率場と自己教師あり学習を利用して、日本語の文中に現れる行動属性（行動主、動作、対象、場所、時間）を自動的に抽出する手法を提案した。評価実験を行った結果、本手法は先行研究の課題を解決でき、一般のWebの文書だけでなく、ソーシャルメディアにおける文法的に正しくない文書にも対応できることが確認できた。緊急時にリアルタイムで行動を人手で抽出するのは困難であるため、提案手法は重要な役割を果たすと考えられる。

提案手法では、既存の学習モデル（条件付確率場）を採用し、以下の工夫点を加えている。

- 訓練データを自動的に作成するため、ラベル付けのコストがかからない。
- 自己教師あり学習の適用と訓練データの追加を行うことにより、F-measureはベースライン手法と比べて大幅に向上した。
- 行動の頻度ではなく、文の構成と単語の品詞に基づいて行動属性のラベルを推測するため、低頻度であるマイノリティ行動も抽出できる。
- 行動のドメインに依存しないヒューリスティックを利用することで、震災だけでなく他の分野の行動も抽出できる。
- 行動属性に限定することで、高速な処理が可能である。

第3章：時系列行動ネットワークの構築

本章では、行動を時系列的に整理する時系列行動ネットワークをOWLに基づいて設計・構築した。評価実験を行った結果、時系列行動ネットワークはOWLに従い、正しく設計・記述されていることが確かめられた。時系列行動ネットワークは時間の流れで動的に変化する集合知であり、行動のインスタンスまで表す点に特徴がある。

第4章：時系列行動ネットワークの補完

本章では、行動の性質とユーザのゴール（避難、帰宅など）を考慮した行動ベース協調フィルタリング手法を提案し、行動ネットワークを補完する。つぶやかれなかった時間に複数の行動を行った場合でも、本手法は行動の実行時間を考慮して補完することができる。特に、協調フィルタリングと言語処理を統合することで、マイノリティであるが価値のある行動に一定の重み付けを行う点に特徴がある。評価実験を行った結果、提案手法はベースライン手法と比べて大幅に抽出精度が向上したことが確かめられた。

第5章：震災時の応用

東日本大震災発生当日のTwitter投稿を対象に評価実験を行った結果、本研究の時系列行動ネットワークが以下の避難に関する情報を把握するのに有効であることが確かめられた。

- ・ 一時的に無料で提供してくれた飲み物、携帯バッテリーなど必需品
- ・ 電車、バスなど交通機関の運行状況
- ・ 一時的に開放してくれたトイレと避難所

また、時系列行動ネットワークでは、上記の情報を場所ごとに時系列的に整理するので、場所と時刻に応じて把握可能である。更に、時系列行動ネットワークを辿ることで、目的を達成するための行動の系列も把握可能である。

以上述べたように、本研究成果はオリジナリティに富み、有用性と実用性の双方においても高い価値を持つものである。よって、本論文は博士(工学)の学位論文として十分な価値を有するものと認める。