

平成 30 年 6 月 18 日現在

機関番号：82626

研究種目：若手研究(B)

研究期間：2015～2017

課題番号：15K15995

研究課題名(和文) 集合移動パターン分析によるセマンティック軌跡データベースの研究

研究課題名(英文) Development of Semantic Trajectory Management with Collective Movement Patterns

研究代表者

金 京淑 (KIM, Kyoung-Sook)

国立研究開発法人産業技術総合研究所・情報・人間工学領域・研究チーム長

研究者番号：20738728

交付決定額(研究期間全体)：(直接経費) 2,900,000円

研究成果の概要(和文)：本研究ではヒトやモノなどの移動軌跡データの有用性を高めるため、ツイートを利用しソーシャルメディアデータから時空間分布パターンに基づく形態的な特徴を抽出し、その特徴の時間・空間的な変化による軌跡情報を検出する手法を確立した。また、時空間3次元セル基盤のクラスター構成技術、クラスターの時空間変化による傾向やそれらの間の位相関係のパターンによる意味情報を検出する等、より高度な知識を獲得できるセマンティック軌跡のフレームワークの要素技術の研究開発を行なった。さらに、データ管理のプロトタイプシステムを実装し、国際標準化団体OGCにおける移動体の位置情報の国際標準規格提案に本研究の成果の一部が利用された。

研究成果の概要(英文)：This research focused on the analysis and management of semantic trajectory of moving objects, especially a collective pattern of movements. In order to discover a useful pattern of the trajectory data of humans, mobile sensors, even social media, we analyzed the morphological features based on the spatio-temporal distribution of data and their topological relationships. In particular, we succeeded in extracting the movement information of topic trends in social media and tried to combine its semantic information with other movement data to construct semantic trajectories. In addition, we developed a prototype system for data management of semantic trajectories, and part of the results of this research was used for proposing international standards for moving object data in international standardization organization OGC.

研究分野：データベース

キーワード：移動軌跡データ データベース セマンティックコンピューティング 時空間データ 地理情報システム 国際標準化

1. 研究開始当初の背景

通信技術、測位を始めとするセンサ技術の発展に伴い、GPS やセンサ等から取得された大量の移動軌跡データ(ヒトやモノなど移動体の位置情報を含むデータ)を効率的に管理し、動きに関する問合せ処理の高速化や移動パターン分析に基づく将来の位置情報の予測など様々な研究方法が提案されてきた[1]。移動軌跡データは、人々の活動や移動状況に即したマイクロマーケティングやロボットを利用した災害時の効率的な避難誘導、細街路のネットワークを活用する超小型車両交通システムなどへの応用が期待されている。こうした応用を実現するには、蓄積された大規模な軌跡データから精度の高い知識、特にヒトやモノの「動き」に関する有用なパターンや規則を効率良く獲得する必要がある。

近年では、生の移動軌跡データだけではなく、移動体の動きに関わる特徴(速度、方向など)や地理空間情報(ランドマークや店舗、道路など)、移動手段情報を利用することで、移動軌跡データに意味情報を付与するセマンティック軌跡データ(Semantic Trajectory)の研究が盛んになりつつある[2]。SeMiTri[3]、SeTraStream[4]では、軌跡データから「移動」「停止」といった状態を検出し、地理空間情報を意味情報として付与するフレームワークを提案している。また、Baquara[5]では、Linked Data とオントロジー(概念間の関係の定義)を利用して自動的に意味情報を付加する手法を述べている。たとえば、図 1(a)の生の移動軌跡データに対して意味的なラベルを付与したセマンティック軌跡の例を図 1(b)に示す。

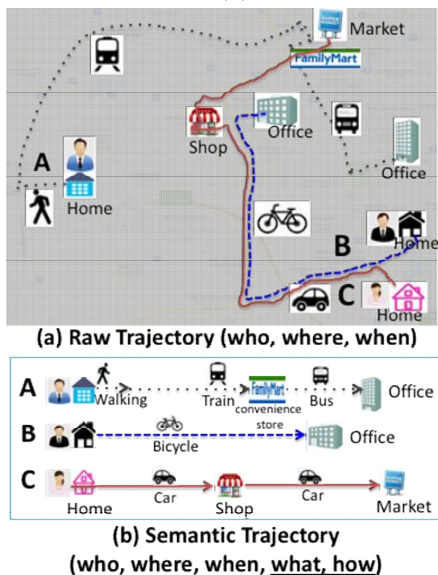


図 1. 移動行動によるセマンティック軌跡データ

一方、これら従来研究の多くは一種類の移動体の軌跡データを対象を限定することで、異なる種類のデータから動き情報や状況の変化を検出し、分析することが出来ない。しかしながら、実世界で起きる様々な現象をセマンティック軌跡として取り扱うためには、これまで研究対象とされてこなかった気象

データなどのセンサデータやソーシャルメディアなども対象とし、横断的なデータ収集や、時空間変更の検出、集約的分析が必要となる。たとえば、台風の被害については気象センサデータから雨雲の動きの情報を、ソーシャルメディアや官公庁等の発表する災害情報から避難状態や被災地域等の状況の変化をトラッキングできることが望ましい。また、従来研究では個々の移動体の軌跡に対する意味付けのみを扱っており、それらを集約することではじめて発見されるパターン(集合移動パターン)における意味付けについての研究は行われていない。これが可能になると、個々の移動軌跡データの集合から獲得される「複数の移動体が特定の場所に集合している」という集合移動パターンに対して、その場所の周辺の異種の地理空間データ(センサデータやソーシャルメディア等)から得られる「会議」や「ストライキ」などといった意味情報の付与が可能になる。

2. 研究の目的

本研究は、従来研究で着目されていない上記課題を解決することを目的として、様々な実世界データを収集・分析し、それらの集合パターンへの意味付けを行う新たなセマンティック軌跡を検出する技術と、生成されるデータを効率的に管理するためのデータベース技術の確立を目指す。その過程で、1)セマンティック軌跡データの生成フレームワークの要素技術となる、軌跡の抽出(Trajectory Extraction)、エピソードの検出(Episode Detection)、そして意味付け(Semantic Annotation)の方法、2)フレームワークから生成されるセマンティック軌跡データを RDF 形式で記述し、地理空間の問い合わせが処理できる GeoSPARQL[6]等を利用してデータを効率的に管理するデータベース技術を研究開発する。また、提案方式の有効性を明らかにするため、デジタル台風[7]や T-Drive[8]のデータのもとにセマンティック軌跡データベースを構築し、実証実験として構築されたセマンティック軌跡データと Twitter のソーシャルメディアとの連携及び可視化を実現するマッシュアップサービスを開発・提供することで、データの有用性の評価を行う。

3. 研究の方法

セマンティック軌跡データの生成フレームワークの要素技術として、軌跡の抽出、エピソードの検出の方法、大量なセマンティック軌跡データを管理する初期的なプロトタイプシステムを開発し、実験と改良を重ねる。A) 異種のセンサデータの統合と集合移動パターンの発見: ウェブ上に公開されている GPS ログデータ、気象データ、ソーシャルメディア(Twitter API で取得できるデータのみ)を中心にデータを収集する。さらに提案者の先行研究[9]を拡張して様々な実世界デ

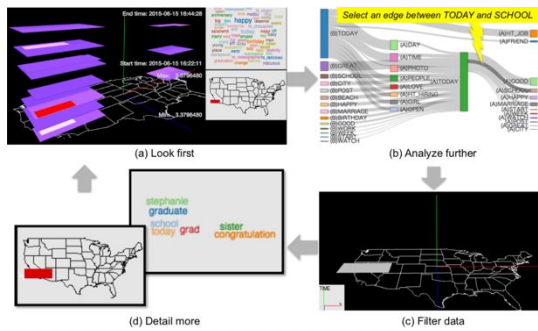


図 2. 可視化システムの意味付け処理画面

ータ（環境に設置されたセンサから得られるデータ、ソーシャルメディアなどから得られる人々の活動・状態データなど）に含まれる動き情報を網羅性高く抽出し、それらを組み合わせた集合移動パターンを分析する。

B) 集合移動パターンへの意味付け：軌跡データに応じてクラスタリングや統計的な分析を行い、分析結果の集合移動パターンに基づいて意味的なエピソードを定義する。エピソードに基づいて A) から検出された軌跡データを構造化された軌跡データに変換し、センサデータの特徴情報やソーシャルメディアを最大限活用して意味情報を付与する。

C) 多次元セマンティック軌跡データモデル設計：生成されたセマンティック軌跡データの時空間情報と任意の粒度での意味情報を格納し、共有するデータ構造を設計する。

D) データベース管理：既存の意味情報の拡張や推論機能を使用するため、RDF スキーマ基盤のデータベースを採用し、地理空間におけるトポロジーに基づいた検索が可能なセマンティックデータベースを元にフレキシブルな意味粒度を支援するオペレータとデータ管理機能を実装する。

4. 研究成果

A) 異種のセンサデータの統合と集合移動パターンの発見：ツイートを利用しソーシャルメディアデータから時空間分布パターンに基づく形態的な特徴を抽出し、その特徴の時間・空間的な変化による軌跡情報を検出する手法を確立した。従来の文書クラスタリングによるトピック抽出手法や空間・時間クラスタリングによるイベント情報抽出手法等と比べ、提案手法は時空間クラスタの位置情報の変化（移動）とクラスタ間の位相関係のパターンを利用することで、俯瞰的な分析だけでは見逃してしまう地域性の高い事象情報を分析できる。多様な時空間データ分布パターン分析技術に関する成果を GIS 分野のトップジャーナル(IF: 1.66)で誌上発表した。

B) 集合移動パターンへの意味付け：大規模かつ複雑なデータに潜在する時空間的なパターンを視覚的に発見し、意味情報を融合するための可視化分析技術を提案した。従来技術では困難であった時間・位置情報に基づいた特徴的なトピックの抽出を容易にする対話的なインタフェースである。時空間 3 次元

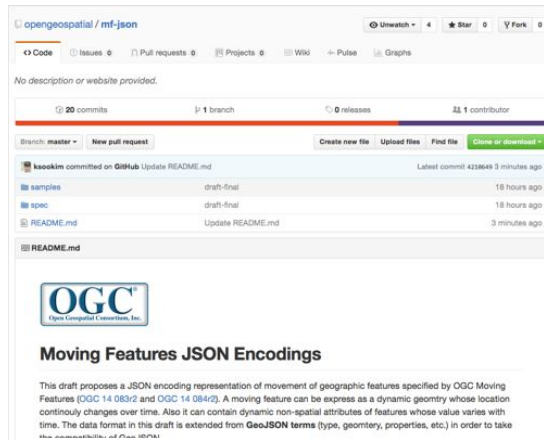


図 3. OGC 国際標準ベストプラクティス

セル基盤のクラスター構成からクラスターの時空間変化による傾向やそれらの間の位相関係のパターンによる意味情報を簡単に付けることができる（図 2）。本可視化システムは論文にまとめ、国際会議にて発表した。

C) 多次元セマンティック軌跡データモデル設計：異種のセマンティック軌跡データを活用するため、JSON 形式のデータ交換フォーマットと REST に忠実な実装を行ったデータサービスの API 仕様を地理空間情報の国際標準化団体 Open Geospatial Consortium(OGC)の国際標準ベストプラクティスとして発表した。既存の XML より簡潔に構造化されたデータを記述ことができ、人間が理解しやすくなり、特に、従来の線形補間の奇跡データ構造と比べ、様々な動きパターンを表現できる。また、移動体の位置情報とともに多次元的な意味情報を一緒に表現することで、軌跡データの有効性を向上した。

D) データベース管理：フレームワークから生成される軌跡データに意味付けを行い、大量なセマンティック軌跡データを管理するデータベース技術について研究開発を行った。研究開始当初では GeoSPARQL の RDF スキーマ基盤のデータベースの採用を検討したが、大量の移動軌跡データを扱うにはスケラビリティの問題があるため、意味的なエピソードの意味情報および時空間的なトポロジー関係をフレキシブルに組み合わせ検索できるデータベースの機能を Cassandra という NoSQL 基盤に実装した。また、移動軌跡データの収集と利用のポジティブフィードバックを推進するため、移動体の位置情報へのデータアクセス仕様開発に本研究の一部分の成果を利用し、OGC 国際標準規格として採択された。

E) 応用：本研究の成果を活用し、個人線量計など放射線に関わるセンサ情報の利活用するシステムを実装した。システムでは、1) Bluetooth 搭載型個人線量計と連動し、刻々と変化する個人の位置情報（移動軌跡データ）と線量計データを簡便に収集するモバイルアプリ、2) モバイルアプリで収集されるセンサ情報をリアルタイムに蓄積・分析する

クラウドベース NoSQL データ管理基盤、3) 蓄積された線量計データと原子力規制庁が公開する放射線モニタリングデータを組み合わせ、今後被ばくすると見込まれる放射線量を推定し、それらを解析・検索・可視化する基盤技術、を研究開発した。本応用システムは論文にまとめ、国際会議にて発表した。

<引用文献>

- [1] Moving Objects Databases. Ralf Hartmut Güting, Markus Schneider: Morgan Kaufmann 2005, ISBN 0-12-088799-1.
- [2] Semantic trajectories modeling and analysis. Christine Parent et al.: ACM Computing Surveys 45(4), pp.42:1-42:32, 2013.
- [3] SeMiTri: a framework for semantic annotation of heterogeneous trajectories. Zhixian Yan et al.: In EDBT, pp.259-270, 2011.
- [4] SeTraStream: semantic-aware trajectory construction over streaming movement data. Zhixian Yan et al.: In SSTD, pp.367-385, 2011.
- [5] Baquara: A Holistic Ontological Framework for Movement Analysis Using Linked Data. Renato Fileto et al.: LNCS 8217, pp.342-355, 2013.
- [6] GeoSPARQL, <http://www.opengeospatial.org/standards/geosparql>
- [7] Digital Typhoon, <http://agora.ex.nii.ac.jp/digital-typhoon/index.html.en>
- [8] T-Drive, <https://www.microsoft.com/en-us/research/project/t-drive-driving-directions-based-on-taxi-traces>
- [9] Moving Phenomenon: Aggregation and Analysis of Geotime-Tagged Contents on the Web, Kyoung-Sook Kim et al.: LNCS 5886, pp.7-24, 2009.

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 2 件)

Jun Lee, Hirotaka Ogawa, Yongjin Kwon, Kyoung-Sook Kim(C.A.): Spatial Footprints of Human Perceptual Experience in Geo-Social Media. ISPRS Int. J. Geo-Information 7(2): 71 (2018), 査読あり, <https://doi.org/10.3390/ijgi7020071>.

Kyoung-Sook Kim, Isao Kojima, Hirotaka Ogawa: Discovery of local topics by using latent spatio-temporal relationships in geo-social media. International Journal of Geographical Information Science 30(9): 1899-1922 (2016), 査読あり, <http://dx.doi.org/10.1080/13658816.2016.1146956>.

[学会発表](計 8 件)

Jun Lee, Kyoung-Sook Kim(C.A.), Ryong Lee, Sang-Hwan Lee: Visual insight of spatiotemporal IoT-generated contents. AVI 2018: 70:1-70:3.

Kyoung-Sook Kim, Dongmin Kim, Hyemi Jeong, Hirotaka Ogawa: Stinum: A Holistic Visual Analysis of Moving Objects with Open Source Software. SIGSPATIAL/GIS 2017: 85:1-85:4.

Teerawat Kumrai, Kyoung-Sook Kim(C.A.), Mianxiong Dong, Hirotaka Ogawa: Optimal KD-Partitioning for the Local Outlier Detection in Geo-Social Points. ISNN (1) 2017: 104-112.

Jun Lee, Kyoung-Sook Kim(C.A.), Yongjin Kwon, Hirotaka Ogawa: Understanding human perceptual experience in unstructured data on the web. WI 2017: 491-498.

Kyoung-Sook Kim, Melissa Bica, Isao Kojima, Hirotaka Ogawa: RendezView: Look at Meanings of an Encounter Region over Local Social Flocks. IWGS 2015: 87-96.

Kyoung-Sook Kim, Isao Kojima, Ryoichi Suzuki, Wataru Naito, Hirotaka Ogawa: RALFIE: a life-logging system for reducing potential radiation exposures. EM-GIS 2015: 18:1-18:4.

[図書](計 0 件)

[産業財産権]

出願状況(計 0 件)

取得状況(計 0 件)

[その他]

OGC Best Practice: OGC Moving Features Encoding Extension - JSON, <http://docs.opengeospatial.org/bp/16-14Or1/16-14Or1.html>

業界横断で位置情報の迅速かつ高度な活用を促進するデータアクセス仕様「Moving Features Access」が OGC 国際標準として採択, http://www.aist.go.jp/aist_j/press_release/pr2017/pr20170313/pr20170313.htm

6. 研究組織

(1) 研究代表者

金 京淑 (KIM, Kyoung-Sook)

産業技術総合研究所・人工知能研究センター・研究チーム長

研究者番号: 20738728

(2) 研究分担者

なし

(3) 連携研究者

なし

(4) 研究協力者

なし