

6-3 実世界を反映したジオメディア収集と時空間分析技術

6-3 *Spatiotemporal Analysis of Geoweb Media*

金 京淑

Kyoung-Sook Kim

要旨

近年、地理空間ウェブ（Geoweb）サービスにより、ユーザがインターネットを通じて時空間地理コンテンツを容易に作成したり共有したりすることができる。そのため、地理空間情報と時間情報を含むジオウェブメディアを管理することが新しい課題になっており、それらのメディアによって実世界の事象に関する包括的な知識の獲得が始まっている。本稿では、場所、時間、および内容の3つの観点に沿った時空間データモデルにより、ジオウェブコンテンツを収集および管理するプラットフォーム（移動現象プラットフォーム）を紹介する。特に、提案した時空間データモデルは、台風の被害、洪水の広がり、そして病気や噂の広がりなど、ジオウェブによって記述された自然現象や社会現象を表現するために用いられる。また、新しい問い合わせ言語は、現象に関連するコンテンツを取得、集約、および要約するように定義されている。最後に我々は、提案した移動現象プラットフォームを用いて、Twitterなどのミニブログサイトのメッセージを管理し、そのメッセージ内容の傾向に関する時空間パターンを分析するアプリケーションシステムについて説明する。

Recently, Geospatial Web (Geoweb) services help people to easily create and share geospatiotemporal contents through Internet. As a result, it becomes an emerging issue to manage Geoweb media that contain geo-spatial and temporal information, and people start to obtain comprehensive knowledge about real-world events through those media. This paper introduces a novel platform for collecting and managing Geoweb contents along three aspects: location, time, and topic, called as moving phenomena platform. In particular, a new spatiotemporal data model is proposed for the representation of natural and social phenomena described by over the Geoweb, such as typhoon damages, flooding expansion, and even the spread of diseases or rumors. Also, a new query language is defined to retrieve, aggregate, and abstract contents relevant to phenomena. Finally, we present an application system on the basis of the moving phenomena platform that manages micro-blogging messages like Twitter and analyzes spatiotemporal patterns of topic trends on the messages.

[キーワード]

時空間情報の管理, 移動現象, 実世界の事象, 相関, ジオウェブメディア

Spatiotemporal information management, Moving phenomena, Real-world events, Correlations, Geoweb media

1 はじめに

“Web 2.0”により、ユーザはインターネットを介して情報を共有したり交換したりするようになった。また、Web 2.0 サービスと、スマートフォンなどのGPS付きのモバイルデバイスを組

み合わせるにより、地理情報を含むウェブコンテンツが強化されたため、地理的な位置情報は、インターネットを介して情報を作成したり情報にアクセスしたりするための1つの手段となっている。「地理空間ウェブ（略して、ジオウェブ）」という言葉は、旧式の地理情報システ

ム (GIS) の単純なウェブマッピングメカニズムから始まったが、現在は Google Maps、Yahoo Maps、OpenLayers などのマップマッシュアップおよびマッピング API などに機能が拡張されている [1]。文献 [2] の Andrew Turner は、「個人的な活動や、専門家でないユーザグループが利用するための複雑な地理手法およびツール」を表すために「ネオジオグラフィ」という言葉を使用している。旧式の GIS では、高精度かつ正確な地理情報を処理することが中心であったが、ネオジオグラフィでは、位置、事象、現象に関係のある個々の意見、活動、または経験について説明する位置情報ベースのコンテンツを共有することが重要である。Goodchild は、ユーザが生成したジオウェブ上の地理空間コンテンツの特別な場合としてボランティアな地理空間情報 (VGI) を定義し、世界を監視するセンサーとしての人の役割について議論している [3]。

ここ最近では、ドキュメント、写真、RSS (Really Simple Syndication) フィード、地理情報によって統合されたミニブログなどのジオウェブメディアが大量にあふれている。たとえば、スマートフォンのユーザが写真をアップロードすると、写真には自動または手動で、ユーザが GPS で取得した現在の位置または撮影場所の地理情報が付く。ジオ RSS は、サイト上の更新および新しい記事について、フィードの地理的な座標を含んでいる。特に、Twitter、Face book、YouTube などのソーシャルメディアは、実世界の事象のリアルタイムな情報共有手段として台頭してきたので、それらのジオウェブメディアは実世界の情報を反映する影としてみなされる [4]。そのため、我々は、絶えず生成される多量のジオウェブメディアを使用して、実世界に関連する包括的な知識を獲得しようとしている。

ジオウェブメディアを解析するには、時空間処理が、自然言語およびマルチメディア処理と共に、必須技術の 1 つである。つまり、位置情報と時間情報は、実世界で発生中または発生した事象の関連情報の作成、インデックス付け、および検索において、重要な役割を果たしている。2010 年のハイチ地震および 2011 年の東日本大震災に関して [5][6][9] で優れた例を参照することができる。本稿では、「移動現象プラットフォーム」

と呼ばれるジオウェブメディアを管理するシステムの研究・開発についてまとめている。文献 [7] に示す新しい時空間データモデルによって設計されたプラットフォームは、2つの主要なサブシステム、時空間管理の移動現象エンジンおよび視覚化ツールとしての Sticker で構成されている。移動現象プラットフォームの主な目的は、空間、時間、および内容に関する情報を含む異種のジオウェブメディアから認知実世界事象または現象を検出および追跡するための、時空間知識検出インフラストラクチャを提供することである。

本稿の残りの部分は、次のように構成されている。2では、我々の動機となった関連する研究について説明し、3では、我々のデータモデルと新しい問い合わせ言語について紹介する。4では、移動現象プラットフォームのシステムアーキテクチャと応用例、5では、今後の研究方向について結論を示す。

2 関連する研究

この章では、既存の時空間データモデルおよび関連するジオウェブテクノロジーを紹介し、本研究の理解の助けとする。

2.1 時空間データモデル

空間データベースは、計算機上で実世界を表現するために開発されてきた。初期の空間データベースには、山、建物、川などの静的な地理的特徴の表現が含まれていた。ところが、空間データベースの研究者は、実世界に存在している台風、洪水、交通などの動的な地理的特徴に注目してい

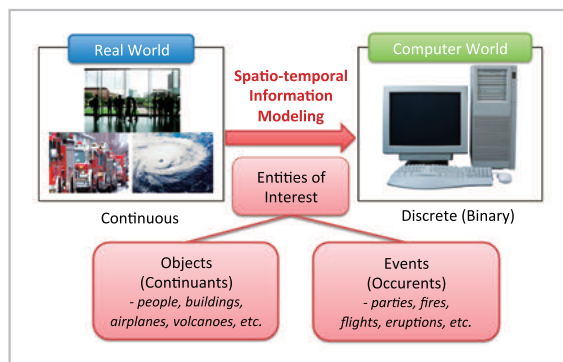


図 1 実世界において対象となる時空間エントリ

た。高次元の存在論の視点からは、図1に示すように、実物（オブジェクト）と出来事（事象）を区別することができる。実物は、人、飛行機、火山など、時間の経過と関係なく、一貫的な構造で継続される世界のエンティティのことであるが、出来事は、人生、フライト、噴火など、時間の経過に伴い、それ自身が発生して発展することで時空（space-time）の中に位置し、その変化に沿った形で構造化されるエンティティのことである。文献[10]では、GrenonとSmithが参照オントロジー（reference ontology）として実物にSNAPオントロジー、出来事にSPANオントロジーを導入し、時空間領域で世界の動的な観点を示した。この観点に従うと、時間によって位置が変化する飛行機の事象の集合としてフライトを捉えられる。

数多くの時空間データモデルが、文献[11]で評価される、動的な地理オブジェクトまたは地理現象の時間変化を表すために開発されてきた。それらは、移動型エンティティにおける離散的な変化だけでなく連続的な変化を表現しようとする。特に、移動オブジェクトのデータモデルでは、点、線、面積、または体積として表される空間オブジェクトの時間経過による連続的な変化を表示する。これらは概念上、時間の関数 $mobject: time \rightarrow spatial-object$ として定義される。これは、そのデータモデルでは、有効期間中ならいつでもオブジェクトの位置を推定できることを示している[12][13]。移動オブジェクトのモデルでは、地理現象の連続的な履歴変化を分析するため、基本データタイプと処理を提供している。ところが、ジオウェブメディアによって記述された特定の現象の操作に適用される際にいくつかの問題があった。1つ目に、それはオブジェクト指向のアプローチで設計されている点である。GISコミュニティ内では、オブジェクトは明示的な地理的境界を持つ離散的かつ独立的なエンティティに相当する。しかし、ジオウェブメディアによって記述される自然現象や社会現象の多くは、位置情報のドメイン領域を特定したりそれらの変化を追跡したりすることは非常に困難である。2つ目に、移動オブジェクトモデルでは、「特定の時間にオブジェクトはどこにあるか」という質問に答えるため、時間の経過を伴う位置情報のみに注目してい

る点である。しかし、人間は「特定の場所と時間に発生すること」を知りたい。そのため、地理空間、時間、および内容に関するジオウェブメディアの分析には別のアプローチが必要である。

2.2 ジオウェブメディアの調査

ジオマップおよびタイムラインを使用したユーザインターフェイスでは、ジオウェブメディアによって、ある位置と時間で、人々が注目していることや発生した事象についての考えを容易に取得できる。文献[14]で提案されているWeb-a-Whereシステムは、ページ上に表示されている位置の名前を使用して、ウェブページを位置情報と関連付ける。また、GPS搭載のデジタルカメラまたはウェブ上のユーザによって自動または手動で撮られる地理情報付きの写真の数が増加している[15]。文献[16]で説明されているGeoTrackerは、RSSフィードの時空間ナビゲーターである。これは、RSSフィードのテキストから位置マイニングを実行し、世界地図上にフィードを示す。さらに、タイムスライディングバー付きのフィードの時空間ナビゲーションに対応している。文献[17]で提案されているNewsStandは、マップインターフェイスを使用してニュース記事を表示するナビゲーターの良い例である。コンテンツが空間と時間の次元に沿って記述されるGeoTrackerと比べると、NewsStandは空間と内容を中心に扱っている。

また、ミニブログサービスはモバイルLBS（location-based services）に拡張されており、それらのメッセージの各エントリには、GPSまたはユーザプロフィールの位置情報フィールドに対応する地理的な位置情報が含まれている。Twitterのようなミニブログのメッセージは、ジオウェブメディアの大部分を占めている。文献[18]のTwitterStandは、ミニブログのメッセージに対するNewsStandの別バージョンとして地理情報付きのツイートをクラスタリングしてニュースを処理する。文献[6]のSakakiらは、スポーツイベント、事故、台風、地震などの事象の発生を検出するため、Twitterサービスのリアルタイムな性質を調べ、時間と地理の位置情報と関連付けられているツイートを使用している。近年、地理空間と時間のコンテンツを使用したソーシャルメ

ディアの視覚化およびナビゲーションは、ホットトピックであるモニタリングと同様に、一般的になっている。

3 移動型データ管理

ここでは、以前の研究 [7] に示されている移動現象モデルについて評価し、新しい SQL 型問い合わせ言語「移動現象問い合わせ言語 (Moving Phenomena Query Language: MPQL)」について説明する。詳細については文献 [9] で紹介されている。

3.1 データモデル

動物、人間、乗り物、台風などのオブジェクトや現象に関する各種のエントリの「動き」を分析することは、社会分野や科学分野の調査を行うために重要である。時空間エントリの「動き」のパターン、すなわち軌跡 (trajectory) を理解することは、現実の生活で意思決定を行うのに役立つ。しかし、病気の感染やうわさなどの特定の現象の変化 (増加、進行、減少) は、人間、乗り物、または動物の動きと比べると観測するのが難しい。我々は通常、観測可能な事象に基づいて位置を推測する。移動現象データモデルは、対象となる内容に関連するウェブコンテンツの時空間伝播を表現することが動機付けとなった。

モデルには、地理事象 (GeoEvent) と移動現象 (MovingPhenomenon) の 2 つの主要なデータタイプがある。地理事象は次のように定義される。

定義 1. 地理事象

地理事象は 4 つ組 (g, ti, v, o) で定義される。ここで、 g は幾何形状で $\forall g \in \{point, line, region\}$ 、 ti は時間間隔で $ti = [ts, te] (ts \leq te)$ 、 v は m 次の特徴空間 $(f_1, f_2, \dots, f_m) (1 \leq i \leq m)$ におけるベクトル (v_1, v_2, \dots, v_m) 、 o は観測者の定義である。

図 2 では、異種なジオウェブメディアの表現を統合するため、4W (ユーザ、内容、位置、時間) のコンテキストを持つ地理事象の概念を示す。すなわち、これは「特定の位置、特定の時間に、あるオブジェクト (ユーザまたはデバイス) によって何が観測されるのか」ということについて

て表している。認識レベルでは、地理事象には OGC (Open GIS Consortium) で定義される「Simple Features」の幾何形状オブジェクトがある [19]。地理的な位置情報および時間情報を含むジオウェブコンテンツ (KML フィードなど) は地理事象で表されることができる (図 3)。

2 番目は、地理事象の集合を表す移動現象である。文献 [20] で述べられているように、抽象表現は各要素の列挙よりも有益な情報を引き出すことがある。たとえば、図 4 (a) では、点に関する 2 つのクラスタがある。その図では、点のセット A と点のセット B の間の空間関係を認識するこ

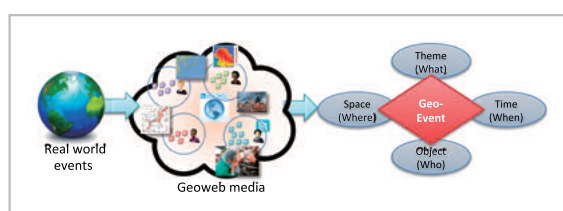


図 2 内容、時間、およびユーザの認知システムに基づく地理事象のコンテキスト情報

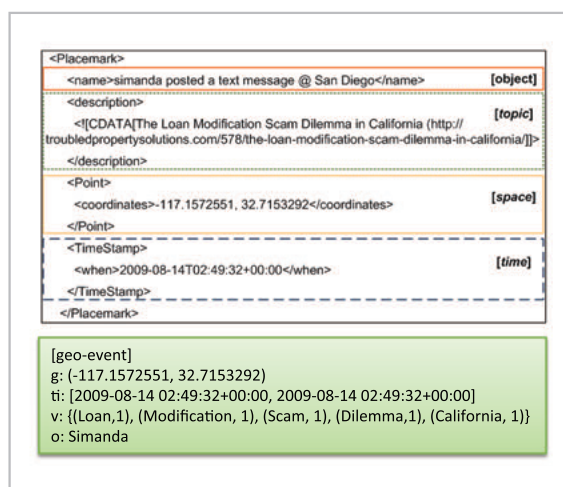


図 3 ジオウェブメディアからの地理事象の変換例

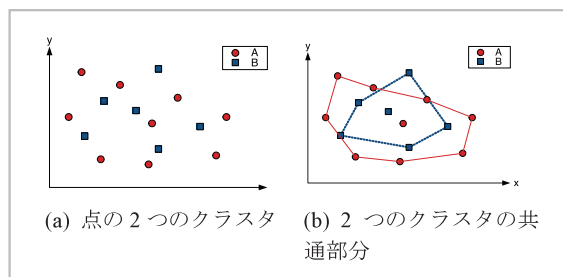


図 4 抽象化表現による新しい関係の発見

とはやや困難である。しかし、それらを2つの領域として一般化すると、図4(b)に示すように新しい位相関係が現れる。 $E = \{e_1, e_2, \dots, e_n\}$ を地理事象のデータセットとする。次に、Eの移動現象は定義2によって記述される。

定義2. 移動現象

移動現象は3つの成分 ($E, f_{\text{domain}}, f_{\text{interpolation}}$) で定義される。ここで、 f_{domain} は時空間領域を決定する関数 $g(t), \forall t \in [\min \{e.ts, e \in E\}, \max \{e.te, e \in E\}]$ 、 $f_{\text{interpolation}}$ は時空間の相関に基づく特徴範囲 E 内の未知の位置で特徴ベクトルを見積もるための補間関数である。

移動現象モデルは、病気やうわさの広がり、台風災害、地球温暖化など、特に離散している地理事象の集約サンプリングから連続的な現象を表すために設計された。時空間領域および補間の2つの連続的な関数によって、移動現象は異なる時空間境界(形状)と内容に関する特徴をもたらす。時空間領域の関数は、文献[8]で使用されている直方体、円錐、円筒、球、多面体などの時空間幾何形状のいずれかを持つ。補間関数は、'TIN'、'IDW'、'NEIGHBORS'、および'REGRESSION'のいずれかによって割り当てられる。特に、領域関数は次の章で示すように移動現象の視覚化形状によって直接マッピングされる。図5のクラスダイアグラムで、移動現象データモデルの主要部分を示す。

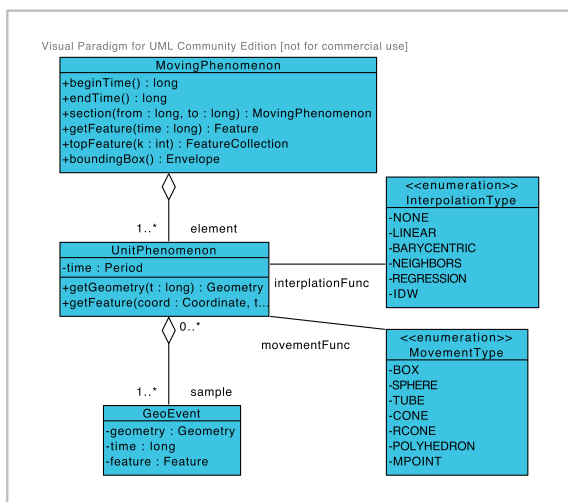


図5 移動現象データモデルのクラスダイアグラム

UnitPhenomenon クラスは、補間、移動関数、および地理事象のセットにより、MovingPhenomenon クラスと GeoEvent クラスの間で定義される。このデータモデルは、移動オブジェクトのデータモデルのスライスされた表現で動機付けされる。サンプルデータのセットが与えられたときに、サンプリングされていない位置で属性の値を見積もるため、いくつかの補間方法を事前に定義した。そして、移動関数のタイプは一定時間内の間、ユニット現象が時空の中にどのように移動したかまたは形状を持つかを決定する。最終的に、MovingPhenomenon のインスタンスは、地理事象のサンプル要素からなる UnitPhenomenon のインスタンスで構成される。

3.2 問い合わせ言語

データモデルに基づき、地理事象および移動現象を管理するため、MPQL と呼ばれる SQL に似た新しい問い合わせ言語を定義する。

● テーブル定義

MPQL の create ステートメントでは、'PERIOD'、'EVENT (地理事象の場合)'、'PHENOMENA (移動現象の場合)' などの新しいデータタイプを使用できる。drop ステートメントは一般的な SQL と同じである。以下に、表1で定義されるデータタイプのカラムを持つ新しいテーブルの作成例を示す。

```
CREATE TABLE geotweets (id BIGINT, message TEXT, occurs EVENT(TPOINT));
CREATE TABLE earthquakes (id BIGINT, magnitude FLOAT, occurs PHENOMENA(TUBE, NEIGHBORS));
```

● データ操作

manipulation ステートメントはテーブルに対するレコードの挿入、削除、アップデート、および取得に使用される。新しいデータタイプのインスタンスの表現には、表2に示すように、入力形式が必要である。event インスタンスを持つ列をテーブルに挿入するステートメントは以下のように与えられる。

```
INSERT INTO geotweets (occurs) VALUES (EVENT('TPOINT(2004-10-19T10:23:54, POINT(139.77 35.69))', 'hot spot'));
```

表1 MPQLの基本データタイプ

データクラス	データタイプ
Base	データベース対応タイプ
Temporal	PERIOD (データ/時間のタイプ)
Spatiotemporal	TPOINT/TLINESTRING/TPOLYGON/TGEOMETRYCOLLECTION/MPOINT/MLINESTRING/MPOLYGON/MGEOMETRYCOLLECTION/TUBE/SPHERE/CONE/RCONE/POLYHEDRON/MGEOMETRY
Thematic	FEATURE
SpatiotemporalおよびThematic	EVENT, PHENOMENA

表2 MPQLステートメントにおける新しいデータタイプのインスタンスの表現

タイプ	表現
period-instance	PERIOD (date-time, date-time) / [date-time, date-time]
date-time	yyyy-mm-ddThh:mm:ss
event-instance	EVENT ('wkt-extend', feature-instance)
wkt-extend	tgeometry-type (date-time, wkt) mgeometry-wkt
tgeometry-type	TPOINT/TLINESTRING/TPOLYGON/TGEOMETRYCOLLECTION
mgeometry-wkt	mgeometry-type ((date-time, wkt), (date-time, wkt), ...) / cylinder-type (period-instance, point-wkt, meter-distance) / tube-instance
mgeometry-type	MPOINT/MLINESTRING/MPOLYGON/MGEOMETRYCOLLECTION
cylinder-type	CONE/RCONE/SPHERE
tube-instance	TUBE ((date-time, point-wkt, meter-distance), (date-time, point-wkt, meter-distance), ...)
feature-instance	'text'/'text:valuetext:value ...'
phenomena-instance	PHENOMENA ('wkt-extend', feature-instance) / PHENOMENA (event-list, 'mgeometry-type', 'interpolation-type')
interpolation-type	LINEAR/BARYCENTRIC/NEIGHBORS/IDW/REGRESSION/NONE
event-list	{(event-instance), (event-instance), ...}
wkt	POINT (15 20) /*point-wkt*/ LINESTRING (0 0, 10 10, 20 25, 50 60) POLYGON ((0 0, 10 0, 10 10, 0 10, 0 0), (5 5, 7 5, 7 7, 5 7, 5 5)) MULTIPOINT (0 0, 20 20, 60 60) MULTILINESTRING ((10 10, 20 20), (15 15, 30 15)) MULTIPOLYGON (((0 0, 10 0, 10 10, 0 10, 0 0)), ((5 5, 7 5, 7 7, 5 7, 5 5))) GEOMETRYCOLLECTION (POINT(10 10), POINT(30 30), LINESTRING(15 15, 20 20))

ただし、PHENOMENAのインスタンスをテーブルのレコードに追加するには、次のような別の表現 'WITH' 句を使用する場合もある。

```
INSERT tab_name ([attr_name] [...])
VALUES ([attr_value] [...]) [WITH phenomena_col_name
EVENTS{event_list/event_select_statement};
```

したがって、3つの方法で phenomena インスタンスを持つレコードを挿入できる。

-Phenomena インスタンス: phenomena インスタンスを持つ列を直接挿入できる。

```
INSERT INTO earthquakes (magnitude, occurs)
VALUES
```

```
(3.2, PHENOMENA({EVENT...}, 'TUBE', 'NEIGHBORS');
```

```
INSERT INTO earthquakes (occurs)
VALUES (PHENOMENA('CONE ([2004-10-19T10:23:54, 2004-10-20T10:23:54], POINT(139.77 35.69), 1000)', 'magnitude: 3.2');
```

-事象リスト: WITH 句を使用して event インスタンスのセットから入力できる。このステートメントでは、WITH と EVENTS の間のカラム名は PHENOMENA 型によって定義されなければならない。

```
INSERT INTO earthquakes (magnitude)
VALUES (3.2) WITH occurs EVENTS {EVENT ('TPOINT(2004-10-19T10:23:54, POINT(10 20))', 'great earthquake'), EVENT ('TPOINT...', 'big damage')};
```

-事象の取得: phenomena インスタンスは SELECT 句を使用してテーブルから事象を取得した後に挿入できる。この表現の場合、select カラムが EVENT タイプであることに注意する。

```
INSERT INTO earthquakes (magnitude)
VALUES (3.2) WITH occurs EVENTS {SELECT occurs FROM geotweets};
```

MPQL SELECT ステートメントでは、2つの固有の述語として CORRELATE と CLUSTERING が定義されている。CLUSTERING 述語は GROUP BY 句と共に使用される。1番目のパラメーターは事象タイプである必要があり、2番目のパラメーターはグループ数である。たとえば、次のステートメントを使用すると、時空間の距離

```

別に事象に関する3つのグループが作成される。
SELECT * FROM geotweets WHERE CORRELATES (occurs, EVENT(...), 'M = COSINE') < 0.5 OR CORRELATES (occurs, PERIOD(...), 'M = OVERLAP') = 1
GROUP BY CLUSTERING (occurs, 3);

```

述語 CORRELATE は、測定によって計算される各列の間の補正係数と問い合わせの引数を表現するため、0と1の間の値を返す。MPQL では3つの測定タイプとして、COSINE、OVERLAP、および EUCLIDEAN をあらかじめ定義する。ただし、ユーザは CREATE FUNCTION を使用して独自の測定を定義し、'M = name' によって書式設定される測定パラメーターとしてその名前を使用できる。

4 移動現象プラットフォーム

ここで、移動現象モデルに基づいてジオウェブメディアを管理するプラットフォームの実装について説明する。図6では、移動現象プラットフォームの概要を示す。プラットフォームは3つの主要なコンポーネントである、地理事象コレクター、移動現象エンジン、および Sticker で構成されている。各コンポーネントについては、4.1 で説明する。

4.1 地理事象コレクター

地理事象コレクターは、ウェブコンテンツから地理事象を生成するが、これらの地理事象は、コンバーターにコンテンツの位置情報と時間情報に関するメタデータが必要な場合は、時空間情報を追加するために地理時間情報が付加された後に生

成される。フレームワークを認識するため、ジオパージングおよびジオコーディング向けの Geonames service [21] を使用する単純な方法を用いて、地理時間情報を付ける。たとえば、各種サイトから Geonames service の RSS-to-GeoRSS コンバーターによって取得される GeoRSS フィードを収集できる。GeoRSS の場合、地理事象の生成のため、コンテンツで位置情報と時間情報のタグを使用する。しかし、地理情報または時空間情報のない特定のコンテンツに対しては、名前の付いたエンティティ認識に基づいて地理時間情報を付ける必要がある。名前付きのエンティティ認識は、自然言語処理によって位置、組織、時間などのカテゴリにテキスト要素を分類し、各要素に情報を付加するために使用される。前述のカテゴリ間では、位置および組織を示すエンティティは位置を推測するために使用され、時間のエンティティは地理事象の時間のスタンプのために使用される。名前付きエンティティから位置と時間の情報を判断するには、より洗練された方法が必要であるが、ここでは単純なヒューリスティック方法を使用している。次に、コレクターは、テキストタイプのジオウェブメディアの場合に、地理事象の内容に関する情報を表現するため、ペア（キーワード、重要度）の特徴ベクトルを計算する。情報検索（IR）システム [22] で使用されるよく知られた方法のように、ベクトル空間モデルのキーワードの重み付けにいくつかの種類の中で各キーワードの発生回数に基づく *tf-idf* 重み付けスキームを適用する。このスキームは、 $w(t, e) = tf_{t,e} \times ief_t$ によって与えられる。ここで、 $w(t, e)$ は地理事象ベクトル e におけるキーワード t の重み付けの値、 $tf_{t,e}$ は地理事象ベクトルにおける t の発生回数、 ief_t は事象の合計数を考慮することによってキーワード t の一般的な重要性を示す逆事象の頻度である。最後のパラメーターは式 $ief_t = \log N / ef_t$ によって計算される。ここで、 N は地理事象の合計数、 ef_t はキーワード t を含む地理事象の数である。さらに、各キーワードの重要度を計算する方法では、IR システムの検索効率を向上させることが非常に重要であるが、本研究では、高精度の IR システムについて説明することが目的ではなく、時空間次元を持つ地理事象を集約および分析するモデルを提示しようとしてい

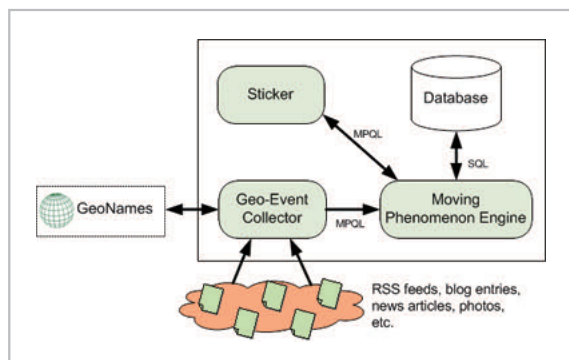


図6 移動現象プラットフォームの概要

る。最後に、変換された地理事象は、MPQL の insert ステートメントによって EVENT タイプのインスタンスとしてデータベースに格納される。

4.2 移動現象エンジン

移動現象エンジンは我々のプラットフォームでキーとなる役割を果たす。それは、3つのモジュール、MPQL Parser、GeoEvent Processor、および Moving Phenomenon Generator によって構成されている (図7参照)。

-MPQL Parser: MPQL の問い合わせの構文をチェックしてから、内部形式に翻訳する。

-GeoEvent Processor: トピックの移動を分析するための地理事象の取得およびフィルタリングを行う地理事象関連エンジンを備えている。事前に定義したユークリッド距離、重なり、コサイン類似度などで、地理事象間の空間、時間、および内容の相関係数を測る。これらの相関係数を組み合わせて演算することで地理情報付きの Twitter のメッセージ (ツイート) を容易に集約できる。たとえば、CORRELATE 関数に測定 COSINE のパラメータを持つと、2つの地理事象の相関値は次のように計算される。

-Moving Phenomenon Generator: 地理事象のグループ化および抽象化の処理によって移動現象のインスタンスのセットを生成する。地理事象の同じセットが与えられるとしても、移動現象のタイプは抽象化および補間の関数によって違う時空間形状を持つ。たとえば、移動現象は移動関数と補関数としてそれぞれ POLYHEDRON および BARYCENTRIC タイプによって定義される場合、図8 (a) に示すように、3次元の位置-時間でドローネー三角分割法に基づいて抽象化される。次に、図8 (b) に示すように、BARYCENTRIC タイプは4面体の内側で重心補間法を使用して任意の位置にある特徴ベクトルを測定する。根本的に、提案された移動現象モデルでは、各種の地理事象に関する空間、時間、および内容の情報を組み合わせる処理が可能になる。

4.3 Sticker

Sticker (時空間情報クラスタリングおよび知

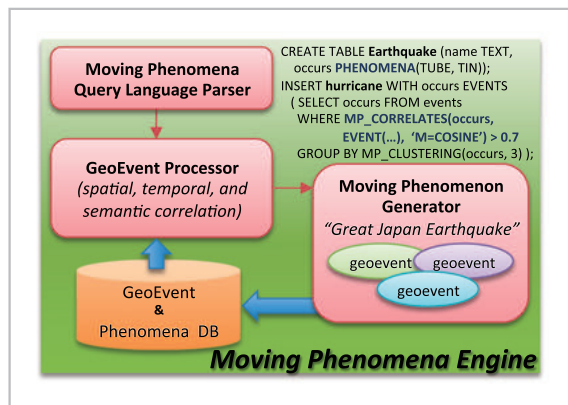


図7 移動現象エンジンのモジュール

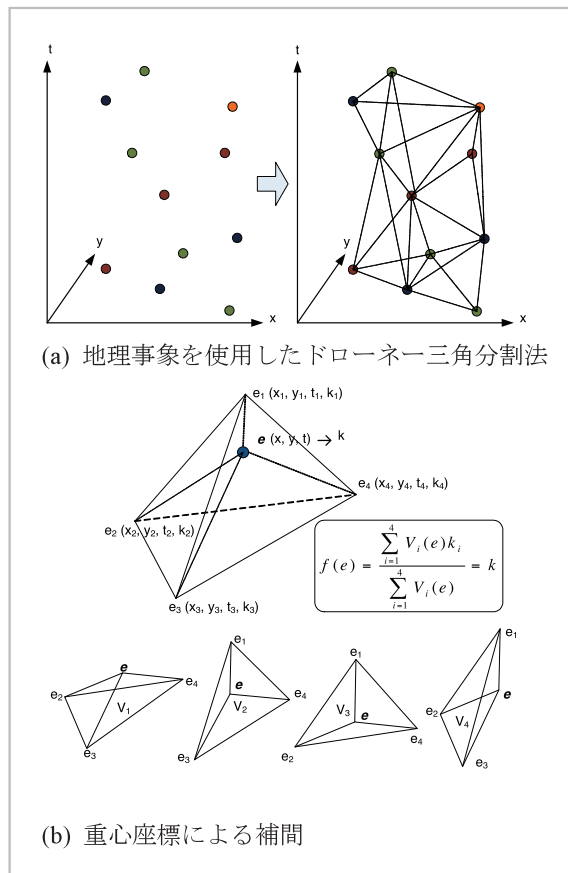


図8 移動関数および補関数による移動現象生成の例

識抽出器) は、データベース内で地理事象および移動現象をナビゲートする視覚化ツールの一種である。これは3つの主要なブラウザである、時空間3次元形状、タグクラウド、および特徴ストリームブラウザで構成されている。ユーザは3つのブラウザを組み合わせながら移動現象に関する内容、位置、および時間を確認したり、それら

の時空間または特徴の変化を比較したりできる。メインブラウザは、x軸とy軸が地理空間を表し、z軸（高さ）が時間を表すことにより、3次元（3D）の空間と時間の幾何形状を使用して移動/時空間の変化を示す。3D幾何形状を設計するため、我々は時間の幾何形状のコンセプトを動機付けした。時間の幾何形状は、図9（a）に示すように、3D座標（x, y, t）を持つ空間と時間の六面体、パス、およびプリズムに基づく時空間コンテキストにおいて個々の人間の活動と彼らのやり取りを理解するのに役立つ。空間および時間のパスとプリズムはそれぞれ、個人の移動履歴と、時間の経過を伴う空間内の特定の制約下でアクセス可能な移動の範囲を記述する。本研究では、いくつかの空間と時間の体積が移動現象および個々の地理事象の視覚化のために定義される。図9（b）では、空間と時間における3D幾何形状の基本タイプを示す。

5 デモンストレーション

ここでは、文献 [9] で紹介されている移動現象プラットフォーム上に開発された mTrend (moving trends) アプリケーションシステムについて説明する。これにより、ツイートから特定の話題の連続的な時空間傾向を分析しようとしている。ツイートおよびトピックの移動（時空間パターン）はそれぞれ、地理事象および移動現象のインスタンスによって表現される（図10）。

このデモンストレーションは、2011年の東日本大震災に関連するシナリオを提供する。図11（a）では、3月2日から3月26日までの地理的な位置によって地理情報が付加されたツイートの時空間分布を示し、図11（b）では、「地震」、「津波」、および「原発」に関連するトピックの時空間傾向を図示する。それらのツイートと時空間傾向は、移動現象エンジンによって管理され、Sticker コンポーネントによって視覚化される。まず第1に、仮想空間におけるトピック傾向の外観は、Stickerの時空間3D視覚化によると、地震→津波→原発事故のように、実世界の事象の順番に従っていることを観測できた。

Stickerには時空間3D視覚化以外にも2つ以上のビューアー、Tag-Compare および Stream-

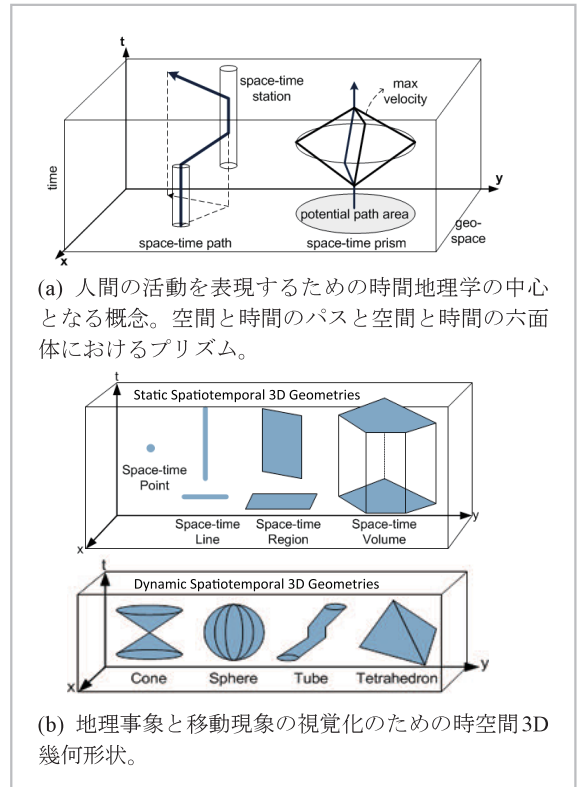


図9 空間と時間の六面体内の時空間3D幾何形状

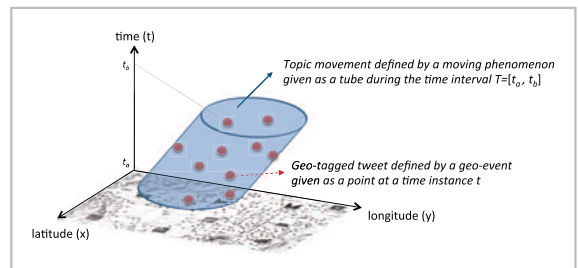


図10 時空間領域における地理情報付きのツイートから導かれる管状で示されるトピックの移動

Compareがある。図12では、各時空間傾向内に含まれている top-k の特徴（この場合、キーワードになる）を与える Tag-Compare を示し、特徴が共通か異なるかを比較する。このデモンストレーションでは、「日本」、「地震」、および「情報」が主要な共通のキーワードであり、各傾向には、関連するキーワードや位置の名前が含まれている。たとえば、「地震」の傾向には「太平洋」、「東北地方」、「被害」が、「津波」の傾向には「神奈川」、「江の島」、「新幹線」が含まれている。もう一方は、図13で示すように、各特徴の頻度を

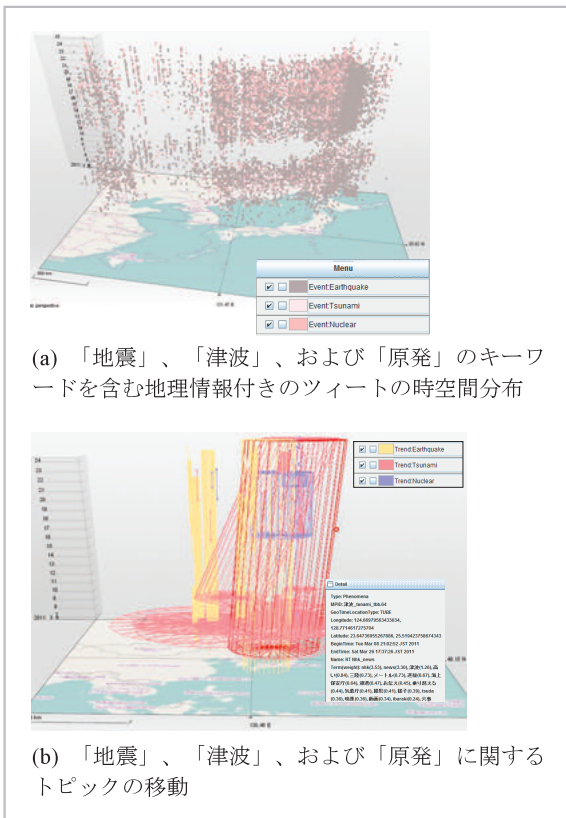


図 11 mTrend のスナップショット



図 12 トピック移動間でのタグクラウドの比較

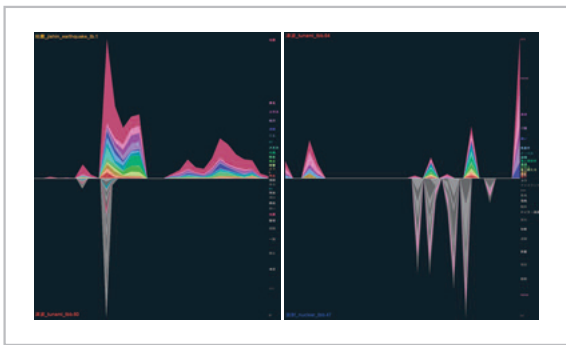


図 13 時間の経過を伴うストリームの頻度の比較

時系列で記述する Stream-Compare である。これは、時間の経過を伴うトピックの変化の傾向および差の有効期間を理解するのに役立つ。たとえば、「津波」の傾向には、早期の「地震」と同様のトピックが含まれているが、時間が経過すると、それには「原発」の傾向と同じキーワードが含まれている。移動現象プラットフォームでは、ジオウェブメディアを管理したり、新しい時空間パターンおよび知識を検出したりするのに便利な関数を提供している。

6 結論

ジオウェブ環境では、地理と時間の情報が付加されたコンテンツが多量に生成されるが、これらのコンテンツは事実、イベント、活動、状況など、実世界で発生する各種の現象について説明している。本研究では、空間、時間、および内容に関する情報を組み合わせてジオウェブでメディアを収集および管理する移動現象プラットフォームを開発した。特に、プラットフォームは、地理事象および移動現象で構成される移動現象データモデルに基づいて設計された。モデルでは、地理的な位置情報や内容に関する情報について特に連続的な時空間変化を記述する。さらに、プラットフォームは、空間、時間、および内容に関するコンテキストによって計算される補正の見積もりによって地理事象および移動現象を取得するために MPQL を提供し、ユーザとのインタラクションのために各種の視覚化ナビゲーターを備えている。デモンストレーションでは、移動現象プラットフォームのアプリケーションとして mTrend を導入した。これは、地理情報付きのツイートを集約し、実世界の事象に関連するソーシャルメディアの時空間傾向を検出する。我々は、2011年に発生した東日本大震災を用いて時間経過に伴うトピックの傾向および移動（パターン）を示した。

今後の研究では、実世界で発生している各種の現象を表すため、プラットフォームを拡張する予定である。また、地理事象および移動現象の間の社会的な関係に基づいて新しい相関測定を追加する予定である。最後に、同様の傾向の移動を分析し、それらの間の新しい時空間関係を追跡する。

参考文献

- 1 The geoweb: Spatially enabling the next-generation web, ESRI White Paper, 2006.
- 2 A. J. Turner, "Introduction to neogeography," O'REILLY Short Cut, 2006.
- 3 M. F. Goodchild, "Citizens as sensors: The world of volunteered geography," *Geo-Journal*, 69(4): 211–221, 2007.
- 4 T. O'Reilly and J. Battelle, "Web squared: Web 2.0 five years on," In Proc. of the 6th Annual Web 2.0 Summit, O'Reilly Media, Inc. and TechWeb, 2008.
- 5 S. Vieweg, A. L. Hughes, K. Starbird, and L. Palen, "Microblogging during two natural hazards events: what twitter may contribute to situational awareness," In Proc. of the 28th international conference on Human factors in computing systems (CHI), pp. 1079–1088, 2010.
- 6 T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes Twitter users: real-time event detection by social sensors," In Proc. of the 19th International Conference on World Wide Web, pp. 851–860, 2010.
- 7 K.-S. Kim, K. Zettsu, Y. Kidawara, and Y. Kiyoki, "Moving Phenomenon: Aggregation and Analysis of Geo-time-Tagged Contents on the Web," In Proc. of the 9th International Symposium on Web and Wireless Geographical Information Systems, pp. 7–24, 2009.
- 8 K.-S. Kim, K. Zettsu, Y. Kidawara, and Y. Kiyoki, "StickViz: A New Visualization Tool for Phenomenon-Based k-Neighbors Searches in Geosocial Networking Services," In Proc. of the 2010 12th International Asia-Pacific Web Conference (APWEB). pp. 22–28, 2010.
- 9 K.-S. Kim, R. Lee, and K. Zettsu, "mTrend: Discovery of topic movements on geo-microblogging messages," In Proc. of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (GIS), pp. 529–532, 2011.
- 10 P. Grenon and B. Smith, "SNAP and SPAN: Towards dynamic spatial ontology," *Spatial Cognition & Computation: An Interdisciplinary Journal*, 4(1): 69–104, 2004.
- 11 N. Pelekis, B. Theodoulidis, I. Kopanakis, and Y. Theodoridis, "Literature review of spatio-temporal database models," *The Knowledge Engineering Review*, 19(3): 235–274, 2004.
- 12 L. Forlizzi, R. H. Guñing, E. Nardelli, and M. Schneider, "A data model and data structures for moving objects databases," In Proc. of the 2000 ACM SIGMOD international conference on Management of data (SIGMOD), pp. 319–330, 2000.
- 13 R. H. Guñing, M. H. Böhlen, M. Erwig, C. S. Jensen, N. A. Lorentzos, M. Schneider, and M. Vazirgiannis, "A foundation for representing and querying moving objects," *ACM Transactions on Database Systems*, 25(1): 1–42, 2000.
- 14 E. Amitay, N. Har'El, R. Sivan, and A. Soffer, "Web-a-where: geotagging web content," In Proc. of the 27th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR), pp. 273–280, 2004.
- 15 C. Torniai, S. Battle, and S. Cayzer, "Sharing, discovering and browsing geotagged pictures on the world wide web," *The Geospatial Web*, pp. 159–170, 2007.
- 16 Y.-F. Chen, G. D. Fabbriozio, D. Gibbon, R. Jana, and S. Jora, "Geotracker: Geospatial and temporal rss navigation," In Proc. of the International World Wide Web Conference (WWW), pp. 41–50, 2007.
- 17 B. E. Teitler, M. D. Lieberman, D. Panozzo, J. Sankaranarayanan, H. Samet, and J. Sperling, "Newsstand: a new view on news," In Proc. of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems (ACM-GIS), pp. 1–10, 2008.
- 18 J. Sankaranarayanan, H. Samet, B. E. Teitler, M. D. Lieberman, and J. Sperling, "TwitterStand: News in Tweets," In Proc. of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 42–51, 2009.

- 19 OpenGIS Implementation Standard for Geographic information - Simple feature access - Part 1: Common architecture. <http://www.opengeospatial.org/standards/sfa/>
- 20 A. Galton and M. Duckham, "What is the region occupied by a set of points?" In GIScience," Lecture Notes in Computer Science, 4197, pp. 81–98, 2006.
- 21 GeoNames. <http://www.geonames.org>
- 22 A. Singhal, "Modern information retrieval: A brief overview," IEEE Data Engineering Bulletin, 24(4): 35–43, 2001.

(平成 24 年 6 月 14 日 採録)



金 京淑 (Kyoung-Sook Kim)

ユニバーサルコミュニケーション研究所
情報利活用基盤研究室研究員
博士 (理学)
時空間データベース、データマイニング、
地理情報システム、ユビキタスコンピューティング
ksookim@nict.go.jp