6-3 Spatiotemporal Analysis of Geoweb Media

Kyoung-Sook Kim

Recently, Geospatial Web (Geoweb) services help people to easily create and share geo-spatiotemporal 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 Geoweb media, such as typhoon damages, flooding expansion, and 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 tweets and analyzes spatiotemporal patterns of topic trends on the messages.

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

1 Introduction

The technologies of ‘Web 2.0’ have encouraged people to share and exchange information via the Internet. In addition, the coupling between Web 2.0 services and GPS-equipped mobile devices, especially smart phones, has augmented Web contents including geographic information, and geographical location becomes a means to create and access information over the Internet. Although the term ‘Geospatial Web (in short, Geoweb)’ started from the simple Web mapping mechanisms of traditional Geographic Information Systems (GIS) in early phase, map mash-ups and mapping API, such as Google Maps, Yahoo Maps, and OpenLayers, aggrandize Geoweb [1]. Andrew Turner in [2] uses the term of “Neogeography” to represent “complex geographical techniques and tools used for personal activities or for utilization by a non-expert group of users.” While traditional GIS have focused on handling geographic information itself with precision and accuracy, Neogeography is more important to share location-based contents that describe individual views, activities, or experiences related to places, events, even phenomena. Goodchild defines volunteered geographic information (VGI) as a special case of user-generated geospatial content on the Geoweb and discusses the role of people as sensors in monitoring the world [3].

In recent, there is a flood of Geoweb media such as documents, photos, videos, Really Simple Syndication (RSS) feeds, and (micro-) blogs integrated by geographic information. For example, when a user of smart phone uploads a picture, the picture automatically or manually is geo-tagged by the current location of the user acquired through GPS or the photographing place. GeoRSS represents geographic coordinates and features in feeds for updates and new articles on sites. In particular, as social media of Twitter, Facebook, and YouTube have emerged as real-time methods of information sharing of real-world events, those Geoweb media are regarded as information shadow of real world[4]. Consequently, we are trying to obtain comprehensive knowledge related to the real world using large amount of Geoweb media generated every moment.

In order to analyze Geoweb media, the spatiotemporal processing is one of essential technologies with natural language and multimedia processing. In other words, location and time play an important role in creating, indexing, and searching the related information about events happening/happened in the real world. We can easily refer to good examples in [5][6][9] about the Haiti earthquake in 2010 and the Great East Japan Earthquake in 2011. This paper summarizes the work of researching and developing systems to manage Geoweb media, called by 'Moving Phenomena Platform'. The platform, designed by a new spatiotemporal data model shown in [7], consists of two main sub-systems: Moving Phenomena Engine for the spatiotemporal data management and Sticker as a visualization tool. The main purpose of moving phenomena platform is to provide a spatiotemporal knowledge discovery infrastructure to detect and track cognitive real-world events or phenomena over the heterogeneous Geoweb media considering spatial, temporal, and thematic information.

The rest of the paper is organized as follows: Section **2** explains the related work that we have motivated, and Section **3** presents our data model and a new query language. The system architecture of moving phenomena platform and an example application are shown in Section **4**, and Section **5** concludes this paper with future directions of research and development.

2 Related work

This section introduces existing spatiotemporal data models and related Geoweb technologies to help the understanding of this work.

2.1 Spatiotemporal data models

Spatial databases have developed to represent the real world in computing spaces. The static representation of geographic features such as mountains, buildings, and rivers had been issued in early spatial databases. However, researchers of spatial databases were focused on the dynamic geographic features such as hurricanes, floods, and vehicles existing in the real world. From a high-level ontological perspective, we can make a distinction between continuants (objects) and occurrents (events) as shown in Fig. 1. While a continuant is an entity in the world that endures through time, such as people, airplanes, and volcanos, an occurrent represents an entity that happens or occurs and unfolds itself through time such as a human life, a flight, and an eruption. In [10], Grenon and Smith introduced SNAP ontology for continuants and SPAN ontology for occurrents as upper-level ontologies to represent the dynamic aspects of the world in spatiotemporal domains. According to this point of view, we can regard a flight as a collection of events of an airplane whose location changes over time.

A number of spatiotemporal data models have developed to represent dynamic geographic objects or temporal changes in geographic phenomena, which are reviewed in [11]. They are trying to capture not only the discrete changes but also the continuous changes in moving entities. In particular, moving-object data models represent spatial ob-

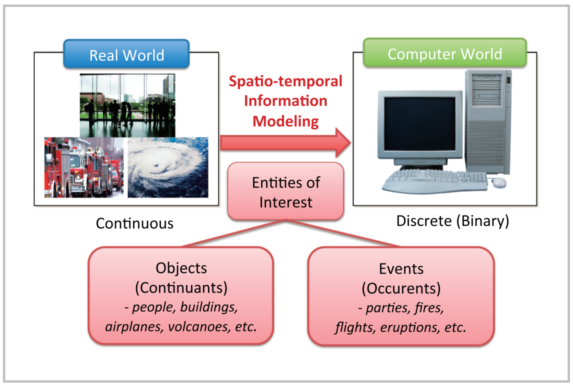


Fig.1 *Spatiotemporal entities of interest in the real world*

jects as points, lines, areas, or volumes that change continuously over time. They are conceptually defined as temporal function *mobject: time → spatial-object*; this means that the model supports to estimate the position of an object at any time during its lifetime [12][13]. The moving-object models have provided basic data types and operations to analyze the continuous historical changes of the geographic phenomena. However, they have some issues to be applied into handling a certain phenomena described by Geoweb media. First, it is designed by using an object-oriented approach. Within GIS communities, an object represents a discrete and independent entity having an explicit geographical boundary. In many cases of natural/social phenomena described by Geoweb media, however, it is very difficult to identify the domain area of their locations and track their changes. Second, the moving-object model focuses on only the location of objects over time to answer the question "Where is the object at a specific time?" However, we human want to know "What happens at a certain place and time?" Thus, we need another approach for the analysis of Geoweb media with respect to geo-space, time, and topic.

2.2 Exploring geoweb media

The user interfaces using a geo-map and time-line help a user to easily get an idea of what people are paying attention to and what is happening, at a location at a time through Geoweb media. The Web-a-Where system proposed in [14] associates Web pages with locations by using the names of the places appearing on the pages. In addition, there is an increase in the number of geo-tagged pictures that are automatically or manually generated by digital cameras with GPS or by users on the Web [15]. GeoTracker presented in [16] is a spatiotemporal navigator of RSS feeds. It performs location mining from the text of RSS feeds and presents them on the world map. Moreover, it supports a temporal navigation for feeds with time sliding bar. NewsStand proposed in [17] is also a good example of a navigator to display news articles using a map interface. When compared to GeoTracker, in which content is described along spatial and temporal dimensions, NewsStand focuses on the space and topic.

In addition, micro-blogging services are being extended into mobile location-based services (LBS) and each entry of their messages contains geographic location information corresponding to its posting places with GPS or the location field of user profiles. The micro-blogging messages like tweets comprise a big proportion of Geoweb media. TwitterStand in [18] is a news processing system by clustering geo-tagged tweets as another version of NewsStand for micro-blogging messages. In [6] Sakaki et al. investigate the real-time nature of Twitter service and use tweets associated with time and geographic location information to detect event occurrences such as sports events, accidents, typhoons, and earthquakes. In recent, visualization and navigation of social media using geo-spatial and temporal contexts are common as well as monitoring hot topics.

3 Movement-based data management

This Section reviews the moving phenomena model shown in our previous study [7] and describes a new SQL-like query language, 'Moving Phenomena Query Language (MPQL)' simply introduced by [9] in more detail.

3.1 Data models

The analysis of movement of various entities involving objects and phenomena such as animals, people, vehicles, and hurricanes is important for carrying out research in social and scientific domains. The understanding movement behavior of spatiotemporal entities helps us to make a decision in real life. However, changes (growing/progress/shrinking) of a certain phenomena such as disease infection or rumor are difficult to observe compared with the movement of people, vehicles, or animals. We usually infer their locations on the basis of observable events. The moving-phenomena data model was motivated for the creation of spatiotemporal shape by context information of a group of Web contents, especially for the representation of movements (spatiotemporal propagations) related to interesting topics.

In the model, there are two main data types: geo-events and moving phenomena. The geo-event is defined as followings:

Definition 1. Geo-Events

A geo-event is defined as a triple of (g, ti, v, o), where g is a geometry, $\forall g \in$ {point, line, region}, ti is a time interval, ti = [ts, te](ts $\leq$ te), v is a vector $(v_1, v_2, ..., v_m)$ in the *m*-dimension feature space $(f_1, f_2, . . . , f_m)(1 \leq i \leq m)$, and o is an observer identification.

Figure 2 shows the concept of geo-events with 4W (Who, What, Where, When) contexts to unify the expression of heterogeneous Geoweb media. Namely, it describes "what is observed at a specific location at a specific time by an object (person/device)?" For the realization level, a geo-event has a geometry object of "Simple Feature" defined by the OpenGIS Consortium (OGC) [17]. A geo-event represents a Geoweb content (e.g., a KML feed) containing geographic location and time information (see Fig. 3).

The second is moving phenomenon that represents an aggregation of geo-events. As mentioned in [18], the abstract representation sometimes derives useful information than enumeration of individual elements. In Figure 4(a), for example, there are two clusters of points. In the figure, it is a slightly difficult to recognize spatial relationships between point-set A and point-set B. However, if we generalize them as two regions, a new topological relationship appears, as shown in Fig. 4(b). Let $E = \{e_1, e_2, ..., e_n\}$ be a set of geo-events. Then a moving phenomenon of E is described by:

Definition 2. Moving Phenomena

A moving phenomenon is defined as a triple of $(E, f_{domain}, f_{interpolation})$, where f_{domain} is a function of a spatio-temporal domain g(t), $\forall t \in [\min\{e.ts, e \in E\}, \max\{e.te, e \in E\}]$, $f_{interpolation}$ is an interpolation function to estimate a feature vector at an unknown location within the feature range of E based on spatiotemporal correlations.

The moving-phenomena model was designed to represent especially continuous phe-

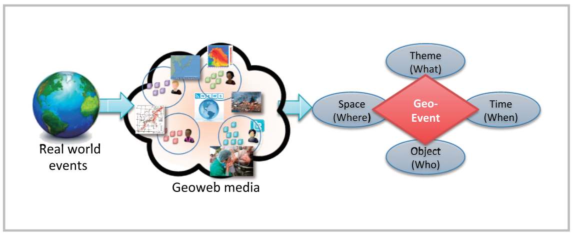


Fig.2 *Context information of geo-events on the basis of the cognitive systems of what, where, when, and who*

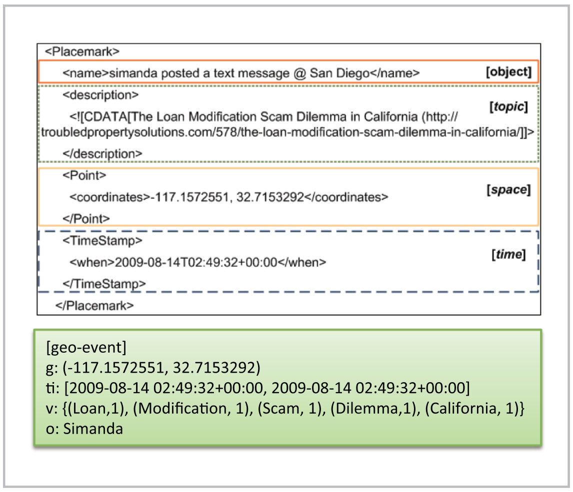
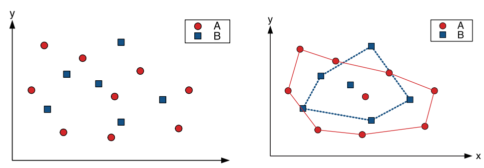


Fig.3 *An example of converting geo-events from Geoweb media*



(a) Two clusters of points (b) Intersection of two clusters

Fig.4 *Discovery of a new relationship by the abstracting representation* nomena aggregating sampling scattered geo-events, such as disease infection, rumor circulation, hurricane disaster, or global warming. According to two continuous functions for the spatiotemporal domain and interpolation, a moving phenomenon draws different spatio-temporal boundary (shape) and the thematic features inside area. The spatiotemporal domain function has one of spatiotemporal geometries such as box, cone, tube, sphere, and polyhedron used in [8]. The interpolation function is assigned by one of 'TIN', 'IDW', 'NEIGHBORS', and 'REGRESSION'. In particular, the domain function is directly mapped with the visualization shape of a moving phenomenon presented in the next chapter. The class diagram in Fig. 5 provides a main part of the moving phenomena model.

The class of UnitPhenomenon is defined between the classes of MovingPhenomenon and GeoEvent with interpolation, movement function, and a set of geo-events. This data model is motivated from the sliced representation of moving-object data models. We consider several interpolation methods in order to estimate the value of the attributes at an unsampled location when a set of sample data is given. The movement/shape of a unit phenomenon depends on the type of movement functions, and this movement/shape remains constant for a certain period of time. Eventually, an instance of MovingPhenomenon consists of instances of UnitPhenomenon, which comprises sample elements of geo-events.

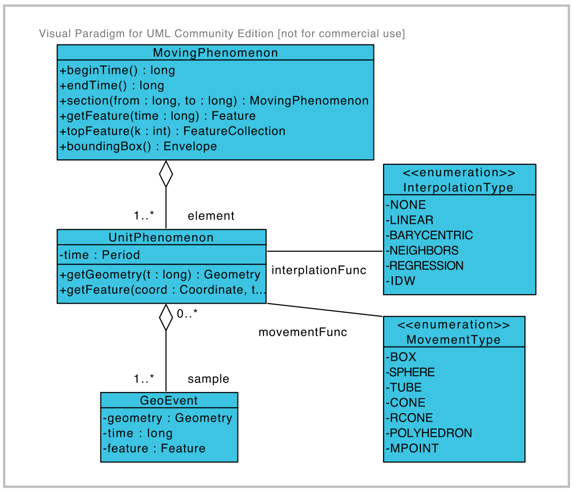


Fig.5 Class diagram of moving phenomena model

3.2 Query languages

Based on the data model, we define a new SQL-like query language for managing geo-events and moving phenomena, called MPQL (Moving Phenomena Query Language).

● Table Definition

In the create statement of MPQL, you can use new data types such as 'PERIOD', 'EVENT (for geo-events)' and 'PHENOMENA (for moving phenomena)'. The drop statement is the same as common SQL. The followings present examples of the creation of a new table with columns of data types defined in Table 1.

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

● Data Manipulation

The manipulation statement is used to insert, delete, update, and retrieve records into a table. The representation of instances of new data types requires input formats, as shown in Table 2. The statement that inserts a row with an event instance into a table is given like:

Table 1 Basic data types of MPQL

Data Classes	Data Types
Base	Database supported types
Temporal	PERIOD (Data/Time types)
Spatiotemporal	TPOINT/TLINESTRING/TPOLYGON/TGEOMETRY-COLLECTION/MPOINT/MLINESTING/MPOLYGON/MGEOMETRYCOLLECTION/TUBE/SPHERE/CONE/RCONE/POLYHEDRON/MGEOMETRY
Thematic	FEATURE
Spatiotemporal & Thematic	EVENT, PHENOMENA

Table 2 *The representation of instances of new data types on MPQL statements*

Types	Representation
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:value text: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))

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

However, the INSERT statement of PHENOMENA has another expressions ‘WITH’ clause as follows:

INSERT *tab_name* ([*attr_name*] [,...]) VALUES ([*attr_value*] [,...]) **[WITH *phenomena_col_name* EVENTS{*event_ list/ event_select_statement*}];**

Thus, you can use three expressions to insert a record with a phenomena instance.

- Phenomena instance: You can directly insert a row with a phenomena instance.

```
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)', 'magni-
tude:3.2');
```

- Event list: It is possible to populate a table with a set of event instances using WITH clause. In this statement, the column name between WITH and EVENTS should be defined by PHENOMENA.

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

- Retrieval of events: A phenomena instance can be inserted after retrieving events from a table using the SELECT clause. In the case of this expression, note that the select column is the EVENT type.

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

In an MPQL SELECT statement, a query can be represented as same as SQL statement except two native predicates: CORRELATE and CLUSTERIN G. The CLUSTERING can use with GROUP BY together. The first parameter should be the event type and the second is the number of groups. For example, the following statement makes three groups of events by their spatiotemporal distance.

SELECT * FROM geotweets WHERE CORRELATES(occurs, EVENT(...), 'M=COSINE') < 0.5 OR CORRELATES (occurs, PERIOD(...), 'M=OVERLAP') = 1 GROUP BY CLUSTERING(occurs, 3);

The predicate CORRELATE returns a real value to represent the coefficient of correlation between each row and the query argument calculated by the measurement. We define three measurement types: COSINE, OVERLAP, and EUCLIDEAN. However, a user can define his/her own measurement using CREATE FUNCTION and use its name as a measurement parameter formatted by 'M=name'.

4 Moving phenomena platform

Here, we explain the implementation of the platform to manage Geoweb media on the basis of the moving-phenomena model. Figure 6 shows the overall of moving phenomena platform. The platform consists of three main components: geo-event collector, moving phenomenon engine, and Sticker. Each component is discussed in the following sub-chapter.

4.1 Geo-event collector

The geo-event collector generates geo-events from Web contents; however, these geo-events are generated after the geotime-tagging to add spatiotemporal information, if the convertor requires metadata about the location and time of the contents. In order to realize the framework, we perform geotime-tagging by employing a simple method that uses Geonames service [21] for geo-parsing and geo-coding. For example, we can collect GeoRSS feeds obtained via the RSS-to-GeoRSS convertor of Geonames service from various sites. In the case of GeoRSS, we use tags of location and time on the contents for the generation of geo-events. However, we need to carry out geotime-tagging on the basis of named entity recognition for certain contents that do not contain geographic or temporal tags. The named entity recognition is used to classify text elements into categories such as location, organization, and times by natural language processing and to put tags to each element. Among the above-mentioned categories, the entities that indicate location and organization are used to derive positions, and time entities are used for the time stamping of geo-events. We need more sophisticated methods in order to determine the location and time information from the named entities; however, we eliminate the requirement of such methods by using a heuristic and simple method and by concentrating upon the problem to capture the movements of phenomena, rather than geotime-tagging. Next, the collector encodes the feature vectors of pairs (term, weight) to represent thematic information of geo-events in case of text-typed Geoweb media. The weight of a term denotes the importance of a term in the content of a geo-event. Even though there are several variants of term-weighting for the vector space model, such as a well-known method used in information retrieval (IR) systems [22], we apply the *tf-idf* weighting scheme that is based on the number of occurrences of each term. This scheme is given by $w(t,e) = tf_{t,e} \times ief_t$, where $w(t,e)$ is the weight value of a term t in a geo-event vector e, $tf_{t,e}$ is the number of occurrences of t in the geo-event vector, and ief_t is the inverse event frequency that indicates the general importance of the term t by considering the total number of events. The last parameter is computed by the expression $ief_t = logN/ef_t$, where N is the total number of geo-events and ef_t is the number of geo-events that contain term t. Further,

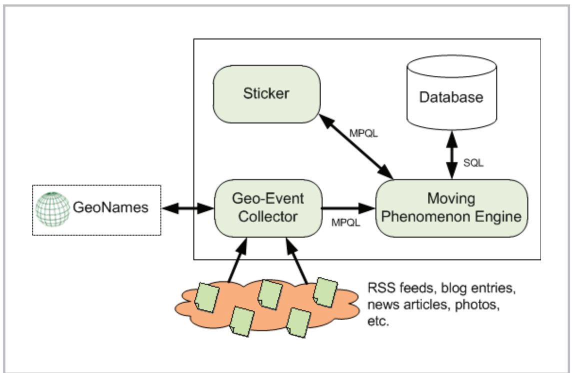


Fig.6 Overview of moving phenomena platform

it is very important to assign a weight to each term in each content to improve the search effectiveness of an IR system. However, this study does not aim to describe high-precision IR system, but rather seeks to present a model to aggregate and analyze the geo-events with spatiotemporal dimensions. Finally, the converted geo-events are stored in a database via MPQL insert statement as the instances of type EVENT.

4.2 Moving phenomena engine

The moving phenomenon engine plays a key role in our platform. It consists of three modules: MPQL Parser, GeoEvent Processor, and Moving Phenomena Generator (see Fig. 7).

- MPQL Parser: It checks the syntax of the MPQL query and then translates it into an internal form.
- GeoEvent Processor: It has a geo-event correlation engine for retrieving and filtering geo-events before generating topic movements. For calculating the correlation between geo-events, we predefined primitive correlation functions with respect to spatial, temporal, and semantic dimensions, such as Euclidean distance, overlap, cosine similarity, and so on. A user can easily aggregate geo-tagged Twitter messages (tweets) using those correlation functions or the user-defined functions composed of the primitive ones. For example, the function CORRELATE has the parameter of measurement COSINE, the correlation value of two geo-events is calculated as:

- Moving Phenomena Generator: It generates a set of instances of moving phenomena by the processes of grouping and abstracting the geo-events. Even though the same set of geo-events is given, the type of moving phenomena varies depending on its abstraction and interpolation function. For example, when a moving phenomena is defined by POLYHEDRON and BARYCENTRIC type as the movement and interpolation function, respectively, it is generated by the Delaunay triangulation in a three-dimensional space-time, as shown in Fig. 8(a). Then it estimates its feature vector at an arbitrary location using a barycentric interpolation method inside a tetrahedron, as shown in Fig. 8(b). Fundamentally, the proposed moving phenomena are combination of spatial, temporal, and thematic information of various geo-events, and our model enables the handling of such information.

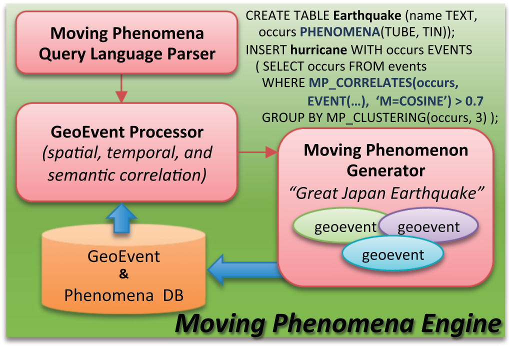
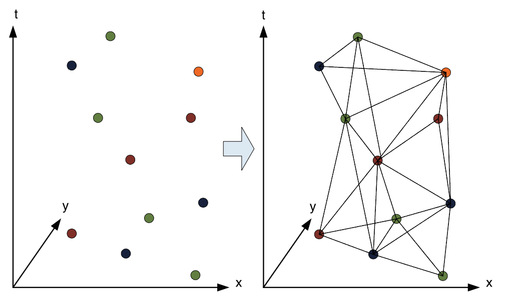


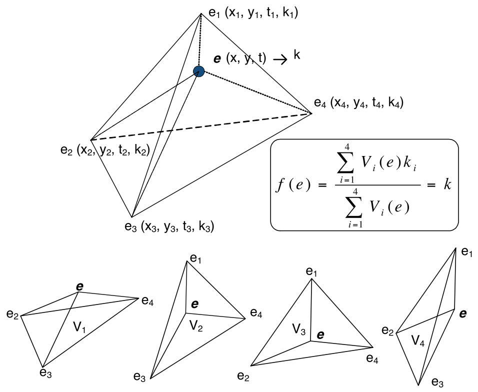
Fig.7 *Modules of moving phenomena engine*

4.3 Sticker

The Sticker (SpatioTemporal Information Clustering and Knowledge ExtRactor) is a visualization tool to navigate geo-events and moving phenomena in the database. It consists of three main browsers to explore spatiotemporal three-dimensional shapes, tag clouds, and feature streams. It lets users see what is/was issued in where and when and compare feature changes among those spatiotemporal movements by using the combination of three browsers. The main browser is a three-dimensional viewer whose x-axis and y-axis represent geographic space and z-axis (height) represents time to show the movements/spatiotemporal changes using three dimensional (3D) space-time geometries. In order to design 3D geometries, we motivated the concept of time geography. The time geography helps understand individual human activities and their interactions in a spatiotemporal context based on space-time cube, path, and prism with 3D (x, y, t) coordinates, as shown in Fig.

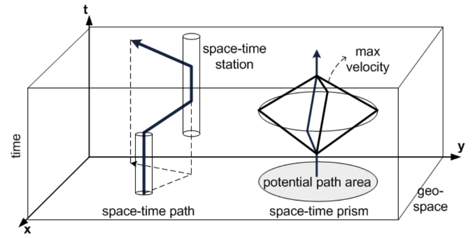


(a) Delaunay triangulation with geo-events

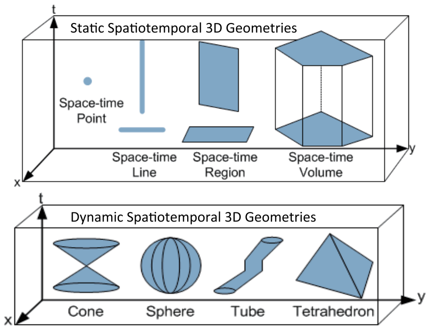


(b) Interpolation by barycentric coordinates

Fig.8 *Example of moving phenomenon generation with movement and interpolation functions*



(a) The central cencepts in time geography to represent human activities: the space-time path and prism in the space-time cube



(b) Spatiotemporal 3D geometries for the visualization of geo-events and moving phenomena

Fig.9 *Spatiotemporal 3D geometries in space-time cube*

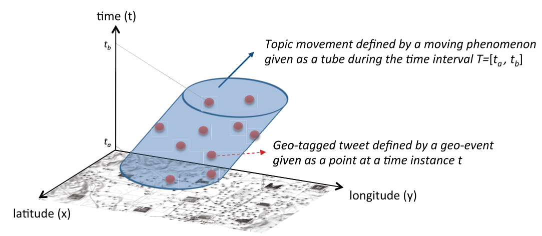


Fig.10 *Example of a topic movement given as a tube shape derived from geo-tagged tweets in a spatiotemporal domain*

9(a). The space-time path and prism describe the historical movement of an individual and the possible accessible extent of the movement under certain constraints in space over time, respectively. In this study, several space-time volumes not only prism for visualizing moving phenomena defined by several individual events. Figure 9(b) shows basic types of 3D geometries in the space-time.

5 Demonstration

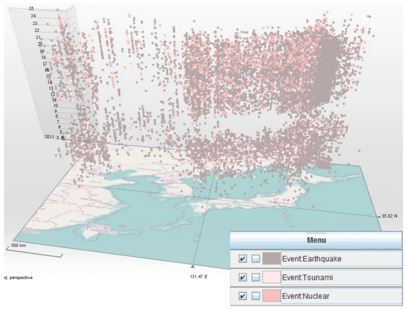
This Section shows an application system developed by the moving phenomena platform, mTrend (moving trends) introduced in [9]. It was trying to analyze the continuous spatiotemporal trend of tweets. A tweet and a topic movement (spatiotemporal pattern) is represented by an instance of geo-event and moving phenomena, respectively (see Fig. 10).

This demonstration presents a scenario related to the Great East Japan Earthquake 2011. Figure 11(a) shows the spatiotemporal distribution of tweets tagged by geographic positions from March 2 to 26, and their spatiotemporal trends of topics related to ‘earthquake’, ‘tsunami’, and ‘nuclear’ are illustrated by Fig. 11(b). Those tweets and their spatiotemporal trends are managed by moving phenomena engine and visualized by Sticker components. First of all, we were able to observe that the appearance of topic trends in virtual space follows the order of events in the real world, such as earthquake → tsunami → nuclear accident, via the spatiotemporal 3D visualization of Sticker.

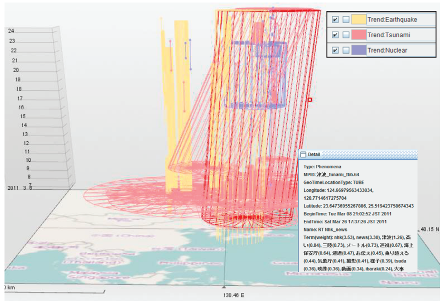
Besides spatiotemporal 3D visualization, there are two more viewers: Tag-Compare and Stream-Compare. Figure 12 shows Tag-Compare that presents the top-k features (in this case, it would be keywords) contained inside each spatiotemporal trend and compare which feature is common or different. In this demonstration, we found that ‘Japan’, ‘earthquake’, and ‘information’ are main common keywords and each trend contains its related keywords/place names, e.g., ‘earthquake’ and ‘tsunami’ trend includes ‘the Pacific Ocean’, ‘Tohoku distinct’, and ‘damage’ and ‘Kanagawa’, ‘Enoshima’, and ‘Shinkansen’. The other is Stream-Compare to draw time series of each feature frequency as shown by Fig. 13. It also helps to understand the life cycle of a trend and the difference of changes of topics over time. For example, ‘tsunami’ trend contains similar topics to ‘earthquake’ in the early time, but as time goes by it includes the same keywords as ‘nuclear’ trend. The moving phenomena platform provides useful functions to manage Geoweb media and to discover a new spatiotemporal pattern and knowledge.

6 Conclusion

The Geoweb environment produces a large amount of geotime-tagged contents that describe various phenomena occurring in the world, such as facts, events, activities, or situations. In this study, we developed moving phenomena platform to collect and manage those media on the Geoweb with combination of spatial, temporal, and thematic information. In particular, the platform was designed by moving phenomena data models consisting of geo-events and moving phenomena. The model describes especially continuous spatiotemporal changes not only geographic location but also semantic information. Furthermore, the platform provided MPQL to retrieve geo-events and moving phenomena with correlation measurements calculated by spatial, tem-



(a) Spatiotemporal distribution of geo-tweets containing keyword ‘earthquake’ , ‘tsunami’ , and ‘nuclear’



(b) The topic movements with respect to ‘earthquake’ , ‘tsunami’ , and ‘nuclear’

Fig.11 *Snapshots of mTrend*

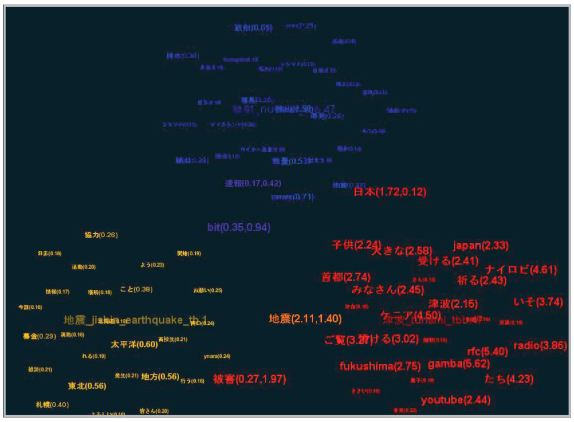


Fig.12 *The comparison of tag clouds among topic movements*

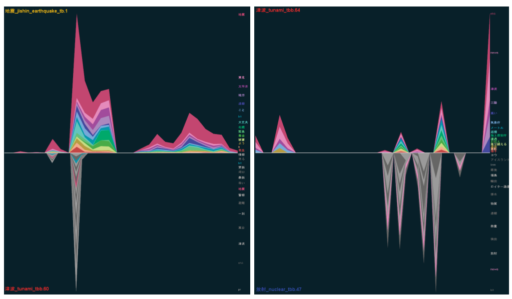


Fig.13 *The comparison of stream frequencies over time*

poral, and semantic contexts, and it has various visual navigators for user interactions. In the demonstration, we introduced mTrend as an application of moving phenomena platform, which aggregates geo-tagged tweets and discovers spatiotemporal trends of social media related to real-world events. We presented topic trends and their movements (patterns) over time using the Twitter messages before or after the Great East Japan Earthquake occurs in 2011.

For the future work, we will extend our platform in order to represent various phenomena occurring in the real world. Also, we will add new correlation measurements based on social relationships among geo-events and moving phenomena. Finally, the analyzing of the similar trend movements and identifying new spatiotemporal relationships between them will be followed.

References

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 Geotime-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 SIGSPTIAL 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. Guẗing, 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. Guẗing, 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. Fabbrizio, 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.

(Accepted June 14, 2012)

***Kyoung-Sook Kim**, Ph.D.*

Researcher, Information Services Platform Laboratory, Universal Communication Research Institute

Spatiotemporal Data Management, GIS, Ubiquitous Computing