
Visualisation d'événements enrichis par des images satellites et des données ouvertes : application aux données sismiques

Helbert Arenas, Nathalie Aussenac-Gilles,
Catherine Comparot, Cassia Trojahn

Institut de Recherche en Informatique de Toulouse, Toulouse, France
helbert.arenas,aussenac,catherine.comparot,cassia.trojahn@irit.fr

RÉSUMÉ. Cet article présente un mécanisme d'enrichissement d'événements à partir de sources de données ouvertes, géolocalisées et datées, en exploitant les relations spatiales et temporelles. Nous validons cette approche sur des données sismiques, des enregistrements de métadonnées d'images satellites et des données météorologiques. On peut ainsi voir les zones impactées par un séisme et connaître les conditions météorologiques avant, pendant et après l'événement. Nous présentons un outil de visualisation qui permet de sélectionner une zone d'intérêt et de visualiser les données enrichies correspondantes.

ABSTRACT. This paper presents a mechanism for enriching events with geolocalized and temporal open data thanks to spatial and temporal relations. We validate the approach by integrating seismic data, metadata records of satellite imagery and weather information. Together they enable to see impacted areas and to know weather conditions before, during and after a seismic event. We developed a visualisation tools allowing the user to select an area of interest and to visualise the corresponding enriched data.

MOTS-CLÉS : intégration de données, données sismiques, images satellites, données météorologiques

KEYWORDS: data integration, seismic data, satellite images, weather data

1. Introduction

Les technologies actuelles nous permettent d'accéder quasi-instantanément à un grand éventail de données issues de dispositifs de détection (des capteurs). Cette richesse d'information offre des possibilités nouvelles mais soulève aussi des défis. Par exemple, comment identifier les faits pertinents au sein de grandes quantités de données qui sont constamment mises à jour par des dispositifs? Dans cet objectif, nous nous intéressons aux événements et aux données caractérisés par une dimension temporelle et spatiale, et montrons qu'une modélisation explicite des relations temporelles et spatiales entre ces données en facilite l'intégration.

Nous illustrons notre propos en nous focalisant sur les événements sismiques. De nos jours, un réseau sophistiqué de capteurs permet de relever des événements sismiques pratiquement en temps réel. L'institut géologique des États-Unis (USGS), à travers son programme d'études des risques de tremblements de terre, permet d'accéder à cette information par le biais d'un service Web. Cependant, le contexte des données fournies est pauvre, ce qui limite les possibilités d'analyse, alors que de nombreuses sources disponibles sur le web pourraient apporter de l'information utile par exemple pour intervenir dans les zones concernées. Ainsi, les images satellites de la zone impactée par un important tremblement de terre pourraient fournir des informations sur les conditions avant et après l'événement, ce qui aiderait les opérations de secours. Les données météo pourraient également fournir des indications sur le contexte de la catastrophe et les possibilités d'intervention pour ces opérations. L'ensemble aiderait le processus de prise de décision.

Bien que les informations à intégrer soient de nature hétérogène, elles ont en commun de disposer de propriétés spatio-temporelles. Il suffit d'intégrer ces différentes données ou tout au moins d'unifier leur accès en exploitant ces propriétés. Ce type de démarche a été identifié comme indispensable très tôt dans le domaine des SIG (Flowerdew, 1991). Le contexte peut alors être automatiquement déterminé à partir d'une mise en relation des propriétés spatio-temporelles puis par l'analyse de ces relations, à l'instar des travaux menés par (Peuquet, Duan, 1995).

Dans la lignée des travaux de (Jensen *et al.*, 2004) et de (Noy, 2004), nous proposons donc un modèle qui relie les données en s'appuyant sur les propriétés spatio-temporelles des événements (sismiques) et des données provenant d'autres sources (méta-données des images satellite, informations météorologiques, et zones d'intérêt de l'utilisateur). Nous avons développé un logiciel de visualisation de ces données exploitant ce modèle. Dans ce qui suit, nous décrivons les sources de données, les mécanismes de collecte de données et proposons un modèle mettant en relation les données à partir des relations entre leurs propriétés spatio-temporelles. Nous présentons enfin le logiciel de visualisation d'événements sismiques permettant d'accéder à ces données.

2. Processus d'intégration et sources de données

2.1. Les données

Données sismiques : A travers le programme des risques liés aux tremblements de terre, la *United States Geological Survey* (USGS) (USGS, 2016) fournit des données sur les événements sismiques du monde entier, en temps réel et actualisées toutes les 5 minutes. Les données sont disponibles dans plusieurs formats (ATOM, KML, GeoJSON) via un service Web. Différents critères de recherche peuvent être utilisés, tels que ceux basés sur les attributs temporels, spatiaux, ou alphanumériques (contributeur, ampleur, profondeur, etc.) d'un séisme. Toutes les 30 minutes, nous récupérons la liste des événements sismiques détectés au cours de la dernière heure au format GeoJSON. Les données de ce fichier sont ensuite stockées dans une base de données. La figure 1 présente une vue simplifiée des données via la table *SeismData*.

Données météo : Le *NOAA's Global Historical Climatological Network* (NOAA, 2016) est un référentiel des enregistrements climatiques constitué à partir de plus de 20 sources de données, y compris les stations météorologiques du monde entier. Des fichiers contenant les données propres à chaque station météorologique sont disponibles via FTP. Il existe des enregistrements vieux de plus de 175 ans, ce qui représente une quantité considérable de données. Nous avons développé une approche pour identifier uniquement les stations météorologiques les plus pertinentes et télécharger les fichiers correspondants. Dans la figure 1, Ces données sont représentées par la table *WeatherReports*.

Métadonnées d'images satellites : Le service Web RESTo (*REstful Semantic search Tool for Geospatial*) (Gasperi, 2014), déployé par le CNES, offre un accès aux métadonnées des images satellites Sentinel 1 et Sentinel 2, au format JSON et ATOM. Ce service met en oeuvre une extension du service OGC13-026 OpenSearch pour l'observation de la terre. L'accès à ces métadonnées peut être paramétré avec des critères temporels, spatiaux, liés à la couverture terrestre, au type de capteur, au niveau de traitement, etc. Dans la figure 1, ces données sont stockées dans la table *SatelliteImagery*. RESTo est mis à jour en continu, mais nous ne les collectons que toutes les heures.

2.2. Mise à jour de liens

Un grand nombre d'événements sismiques se produisent dans le monde, à tout moment. Nous nous concentrons uniquement sur les événements sismiques d'une magnitude supérieure à 5 ML. À ce stade, l'événement sismique est ressenti et peut causer des dégâts matériels. Dans notre modèle, le contexte d'un événement sismique est établi à partir d'analyses temporelles et spatiales. Du point de vue temporel, nous comparons deux valeurs temporelles (deux dates), déterminant si elles sont égales ou si l'une précède l'autre, la différence entre les deux valeurs pouvant aussi servir à définir un intervalle de temps. Si la

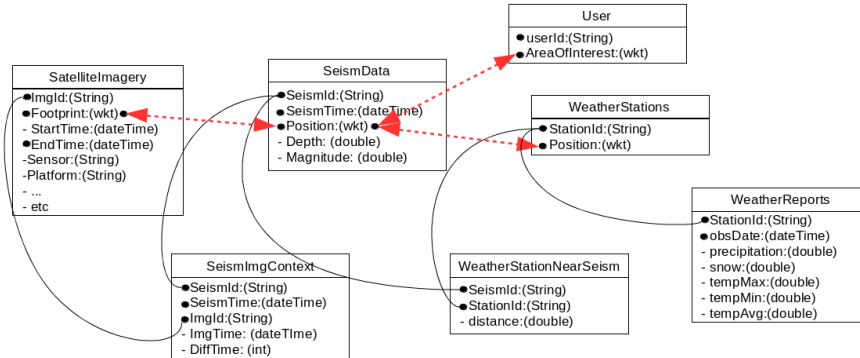


FIGURE 1. Diagramme UML du modèle mettant en relation les données via leurs propriétés temporelles et spatiales (relations spatiales en pointillés).

différence est inférieure à un seuil donné, alors les deux dates sont liées temporellement par la relation *isNear*. Pour l'analyse spatiale, les relations entre deux géométries proviennent d'un ensemble de relations possibles (Strobl, 2008). Par exemple, nous pouvons considérer qu'un point est lié à un polygone si ce point est situé *within* le polygone. Nous pouvons aussi établir la relation *isNear* entre deux géométries si la distance entre ces géométries est inférieure à un seuil préalablement défini par l'utilisateur.

Images satellites comme contexte des événements sismiques : Grâce aux relations précédentes, nous pouvons identifier un ensemble d'images satellites fournissant un contexte à un événement sismique important. Ces images sont sélectionnées en priorité selon la relation topologique entre la géométrie de l'événement sismique, fournie par l'épicentre, et l'empreinte de l'image satellite. Dans notre modèle, un script réalise cette tâche pour les événements sismiques de magnitude supérieure à 5ML. Dans la figure 1, cette comparaison spatiale est représentée par un lien rouge pointillé entre les tables *SatelliteImagery* et *SeismData*. Les relations identifiées sont stockées dans la table *SeismImgContext*. Les opérations spatiales, bien que très rapides, sont coûteuses en terme de traitement, il est donc plus efficace pour le long terme de les stocker au lieu de les calculer à la volée.

Données météo comme contexte des événements sismiques : Nous devons identifier pour chaque événement important, les stations météorologiques situées à proximité. Pour cela, au lieu de calculer la distance entre toutes les stations météorologiques et l'épicentre sismique, il s'avère plus efficace de sélectionner un sous-ensemble de stations candidates dans une zone délimitée autour de l'épicentre sismique. Nous calculons alors la distance uniquement pour

ces stations. Finalement, nous enregistrons dans la table *WeatherStationsNearSeism* les stations pour lesquelles la relation *isNear* existe; relation représentée sur la figure 1 par un lien rouge entre les tables *SeismData* et *WeatherStations*. Un script interrogeant les séismes importants nouvellement détectés, relie ces points sismiques aux stations météo les proches d'eux. Afin de réduire le trafic de données, nous ne demandons des données météorologiques que pour les stations météorologiques identifiées dans la table *WeatherStationNearSeism*.

3. Outil de visualisation

L'application de visualisation utilise une architecture client-serveur. Des scripts PHP enregistrent et mettent à jour une base de données MySQL toutes les 3 minutes. Côté client, la visualisation repose sur une carte Web développée avec la bibliothèque OpenLayers. Le composant de visualisation utilise JS pour interroger le serveur. La carte Web peut être consultée sur <http://geo-space.info/projects/dataIntegration/>.

La figure 2 présente un exemple de visualisation des données contextuelles. Un bouton avec le libellé *Display List of Important Seismic Events* permet d'afficher un panneau sur le côté droit de l'écran avec la liste des événements ayant une magnitude supérieure à 5, et qui se sont produits au cours des 15 derniers jours. En sélectionnant un événement sismique, la zone de l'épicentre est affichée. Cette zone est représentée par un point rouge sur la carte. Le panneau de gauche représente la liste des images satellites qui chevauchent l'épicentre. Pour chaque image, l'application affiche l'identifiant de l'image, la collection (Sentinel 1 ou Sentinel 2), l'écart de temps (en jours) entre l'image et l'événement sismique, la couverture nuageuse de l'image, la date où l'image a été prise et un lien pour voir l'image. La carte montre également ces images sous forme de polygones bleus avec un remplissage vert. Les stations météorologiques à proximité sont matérialisées sous forme de points de couleur magenta. Les stations météorologiques plus éloignées du seuil sont indiquées par des points bleus plus petits. En cliquant sur une station météorologique, le panneau de gauche montre les mesures relevées par la station météorologique sélectionnée, ainsi qu'un graphique montrant l'évolution temporelle de variables météorologiques disponibles pour cette station.

4. Conclusions et perspectives

Cet article a présenté une approche d'enrichissement de données basé sur la notion d'événements, en utilisant un ensemble de relations spatio-temporelles entre ces événements et d'autres sources de données. Nous l'avons appliqué à des événements sismiques, associées aux méta-données d'images satellite et à des données météorologiques. Ceci permet de localiser des séismes, de retrouver des images satellites de la zone affectée ainsi que des données météorologiques des périodes autour de la date de l'événement. Un outil de visualisation des données

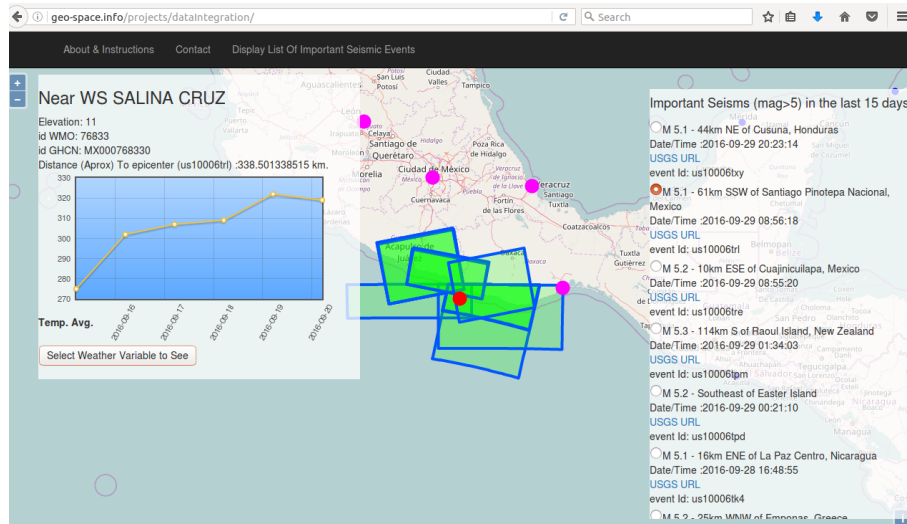


FIGURE 2. Carte sur l'interface web de l'application.

ainsi enrichies a été développé. Nous faisons actuellement évoluer ce logiciel en exploitant un modèle sémantique réutilisant des ontologies comme support à l'intégration, ce qui permet de définir ensuite des mécanismes de raisonnement sur les relations spatio-temporelles. Nous avons déjà mise en oeuvre ce modèle pour des données météorologiques et des images satellite (Arenas *et al.*, 2016)

Bibliographie

- Arenas H., Aussenac-Gilles N., Comparot C., Trojahn C. (2016). Semantic Integration of Geospatial Data from Earth Observations. In *Knowledge Engineering and Knowledge Management: 20th International Conference, EKAW 2016, poster and demo session*, p. 1- 4.
- Flowerdew R. (1991). Spatial data integration principles. In *Geographical information systems: principles and applications*, vol. 1, p. 375 – 387. Longman: Harlow.
- Gasperi J. (2014). Semantic Search Within Earth Observation Products Database Based on Automatic Tagging of Image Content. In *Proceedings of the Conference on Big Data from Space*, p. 4-6.
- Jensen J., Saalfeld A., Broome F., Cowen D., Price K., Ramsey D. *et al.* (2004). Spatial data acquisition and integration. In R. McMaste, E. Usery (Eds.), *A research agenda for geographic information science*, p. 17 – 60. Taylor and Francis.
- NOAA. (2016). *NOAA Global Historical Climatology Network*.

- Noy N. F. (2004). Semantic integration: A survey of ontology-based approaches. *SIGMOD Record*, vol. 33, p. 65-70.
- Peuquet D., Duan N. (1995). An event-based spatiotemporal data model (estdm) for temporal analysis of geographical data. *Int. Journal of Geographical Information Systems*, vol. 9, p. 7-24.
- Strobl C. (2008). Dimensionally Extended Nine Intersection Model (DE-9IM). In S. Shashi, H. Xiong (Eds.), *Encyclopedia of GIS*, p. 240–245. Springer.
- USGS. (2016). *Earthquake Hazards Program*, Last visited in September 2016.