



HAL
open science

Projet CHOUCAS : Intégration de données hétérogènes et raisonnement spatial pour l'aide à la localisation des victimes en montagne

Ana-Maria Olteanu-Raimond, Paule-Annick Davoine, Mauro Gaio, Eric Gouarderes, Marie-Dominique van Damme, Marlène Villanova-Oliver, Mickaël Brasebin, Catherine Domingues, Cécile Duchêne, Olivier Favre, et al.

► To cite this version:

Ana-Maria Olteanu-Raimond, Paule-Annick Davoine, Mauro Gaio, Eric Gouarderes, Marie-Dominique van Damme, et al.. Projet CHOUCAS : Intégration de données hétérogènes et raisonnement spatial pour l'aide à la localisation des victimes en montagne. Spatial Analysis and GEomatics 2017, INSA de rouen, Nov 2017, Rouen, France. hal-01649156

HAL Id: hal-01649156

<https://hal.science/hal-01649156>

Submitted on 27 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Projet CHOUCAS : Intégration de données hétérogènes et raisonnement spatial pour l'aide à la localisation des victimes en montagne

Ana-Maria Olteanu-Raimond¹, Paule-Annick Davoine², Mauro Gaio³, Eric Gouardères³, Marie-Dominique Van Damme¹, Marlène Villanova-Oliver², Mickael Brasebin¹, Catherine Domingues¹, Cécile Duchêne¹, Olivier Favre⁴, Sébastien Mustière¹, Florent Devin³, Yannick Le Nir³, Ludovic Moncla³, Sylvain Bouveret², Philippe Genoud², Jérôme Gensel², Danielle Ziebelin²

1. LASTIG COGIT, IGN, ENSG, F-94160 Saint-Mandé, France

{ana-maria.raimond ;marie-dominique.vandamme ;mickael.brasebin ;
catherine.domingues; cecile.duchene ; sebastien.mustiere}@ign.fr

2. Univ. Grenoble Alpes, CNRS, Grenoble INP (Institute of Engineering Univ. Grenoble Alpes), LIG, F-38000 Grenoble, France

{paule-annick.davoine ;marlene.villanova-oliver ;sylvain.bouveret ;
philippe.genoud ;danielle.ziebelin ;gensel.jerome}@imag.fr

3. LIUPPA, Université de Pau et des Pays de l'Adour, 64013 Pau, France

{mauro.gaio; eric.gouarderes}@univ-pau.fr; {fd ;yl}@eisti.eu;
moncla.ludovic@gmail.com

4. PGHM de Grenoble, 38000 Grenoble, France

olivier.favre@ gendarmerie.interieur.gouv.fr

RESUME. L'objectif de l'article est de présenter le projet ANR CHOUCAS qui ambitionne de proposer des méthodes et des outils pour constituer, enrichir, interroger et visualiser des données spatiales issues de sources hétérogènes, et mener sur ces données un raisonnement spatial flou, avec comme contexte applicatif l'aide à la localisation de victimes en montagne.

ABSTRACT. The goal of the paper is to present the ANR project CHOUCAS. The aim of the project is to provide methods and tools for enriching, structuring, querying and visualising spatial data coming from heterogeneous sources and reasoning to help the rescue teams to speed up the phase of victims location and to better estimate a probable location area.

MOTS-CLES : intégration de données hétérogènes, gestion de connaissances, traitement automatique du langage, fusion d'informations imparfaites, géoréférencement indirect, géovisualisation, architecture SOA, secours à la personne, entité nommée spatiale.

2 SAGEO'2017

KEYWORDS: heterogeneous geographic data integration, knowledge modelling, natural language processing, imperfect data fusion, relative location, fuzzy modelling, geovisualisation, SOA architecture, emergency assistance, spatial named entity.

1. Introduction

Les principaux acteurs du secours en montagne sont les Pelotons de Gendarmerie de Haute Montagne (PGHM). La localisation de la victime pendant la phase d'alerte repose sur un dialogue entre le requérant (personne qui donne l'alerte, qui peut être la victime ou un tiers) et les secouristes. Il est crucial de réduire cette phase de localisation à un temps aussi court que possible, afin de faciliter l'assistance et augmenter les chances de survie de la victime.

Actuellement, la localisation de la victime en montagne pendant la phase d'alerte s'appuie sur la recherche manuelle d'indices dans différentes sources de données et surtout sur le savoir-faire des secouristes et leur connaissance du terrain. Plusieurs éléments compliquent cette tâche de localisation. Le requérant, décrit la position de la victime et l'itinéraire suivi à l'aide de positions relatives imprécises par rapport à des objets repères eux-mêmes souvent ambigus. Les données nécessaires aux secouristes pour la localisation sont hétérogènes, multi-sources, réparties et insuffisamment structurées pour être interrogées efficacement. L'outillage logiciel permettant de combiner des sources de données géographiques hétérogènes sur lesquels il serait possible de mener des raisonnements déductifs est encore très peu développé. Le recours aux solutions de géolocalisation fondées sur des dispositifs mobiles de type smartphones n'est pas toujours possible : absence de couverture réseau, victime trop atteinte pour utiliser une application, etc.

Dans ce contexte, le projet ANR CHOUCAS a pour objectif de proposer des méthodes et des outils permettant 1) de constituer et enrichir des données géographiques issues de sources hétérogènes, et 2) d'élaborer des modèles de raisonnement pour favoriser le processus de décision dans le contexte de la localisation des victimes en montagne. Ce projet est une collaboration entre les équipes COGIT de l'IGN, STEAMER du Laboratoire d'Informatique de Grenoble et MOVIES du Laboratoire d'Informatique de l'Université de Pau et des Pays de l'Adour ainsi que le PGHM de Grenoble.

2. Objectifs scientifiques abordés dans le projet

CHOUCAS est un projet ancré dans la géomatique, définie comme pluridisciplinaire, et vise à répondre à quatre objectifs scientifiques.

–structuration des données issues de sources textuelles hétérogènes. Les descriptions spatiales à base de texte (DST) sont soumises à toutes sortes d'ambiguïtés qui empêchent un géocodage de qualité, et donc leur utilisation efficace. Même si des efforts sont engagés dans le géocodage qui nécessite d'annoter finement au préalable les DST dans les documents, la reconnaissance d'entités nommées se limite souvent à des DST intégrant un nom propre (Nouvel et al., 2016) ce qui peut conduire à un géocodage incomplet voire inapproprié. Nous soutenons que pour des tâches à grains fins, il est essentiel de considérer une plus grande variété de DST dans différentes situations verbales (Gaio et Moncla, 2017) en

intégrant en particulier des noms génériques et des dynamiques spatiales diverses issues de verbes de mouvement ('contourner le rocher situé à droite de la chapelle').

–raisonnement spatial qualitatif flou pour combiner différents éléments de localisation spatialisés. La victime est localisée à travers différents éléments de localisation exprimés sous la forme d'un couple d'objets de référence et d'un prédicat spatial ('en face des aiguilles de Chamonix'), temporel ('a marché 2h depuis le parking du Col des Montets') ou sensoriel ('voit le sommet de l'Aiguille du Midi'). Le premier verrou est de passer d'un géoréférencement relatif à un géoréférencement absolu. En partant de travaux sur la formalisation des relations spatiales entre objets géographiques (Bejaoui *et al.*, 2009, Mathet, 2000) et sur leur interprétation sémantique en fonction du contexte, nous proposerons des méthodes pour construire une géométrie pour chaque élément de localisation représentant une région de l'espace où la victime pourrait se trouver. Un deuxième verrou est la combinaison des régions ainsi obtenues pour inférer une localisation probable en tenant compte de l'imprécision des sources d'informations. Nous comptons recourir à une modélisation floue des régions, adaptant des modèles existants (Bejaoui *et al.*, 2009), et proposant un modèle de raisonnement spatial fondé sur les théories de l'incertain pour combiner plusieurs régions en gérant le conflit et la confiance des sources d'information associées.

–géovisualisation de données multidimensionnelles et imparfaites pour la prise de décision. La question centrale abordée ici est celle des moyens facilitant le recours à l'analyse géovisuelle comme support du raisonnement spatial dans un contexte de prise de décision (MacEachren *et al.* 2015). Un premier verrou est la définition de modèles et d'outils de restitution des informations et d'interaction capables d'accompagner de façon intuitive l'analyse géovisuelle fondée sur des données multidimensionnelles hétérogènes. Nous proposons de formaliser, modéliser et implémenter les stratégies convoquant les méthodes de visualisation les plus adaptées au contexte d'utilisation, et notamment aux questions que se pose le secouriste. Un second verrou concerne la visualisation du caractère imparfait des données. Au-delà des travaux existants (Berjawi *et al.* 2014), nous proposerons des méthodes prenant en compte l'imperfection dans le processus d'analyse géovisuelle par une sémiologie qui respecte le processus cognitif du secouriste (e.g. éviter la sur/sous-représentation de l'imperfection qui peut/peuvent être contreproductif/s).

–intégration de sources hétérogènes spatialisables. La question centrale est l'exploitation de sources de données hétérogènes, voire conflictuelles. Nous proposons la conception d'un modèle de plate-forme générique de médiation orientée services pour permettre l'intégration transparente de sources de données internes et Web (Camp2Camp, DbPedia). Trois verrous sont identifiés : 1) construction d'un modèle d'abstraction orienté service des sources, 2) automatisation du processus de médiation des données et 3) structuration et interrogation des données. Pour aborder 1), on s'appuiera sur une caractérisation fonctionnelle et non fonctionnelle (qualité de service) des services (Itani *et al.*, 2015). Pour 2), on s'appuiera sur des techniques de raisonnement pour automatiser le processus de médiation par composition (Moncla *et al.*, 2016). Enfin pour le verrou 3) on étudiera la manipulation des données multidimensionnelles spatiales floues (Siqueira *et al.*, 2014).

Conclusion

Le projet ANR CHOUCAS vise à répondre initialement à un besoin exprimé par le PGHM, mais correspond à des problématiques sociétales abordées par d'autres acteurs de la sécurité civile tels que les SDIS, CRS ou SAMU. D'un point de vue scientifique, il doit aussi permettre des avancées sur l'accès à des ressources spatialisées hétérogènes, sur la structuration et l'enrichissement des données qui en sont issues, sur le référencement spatial direct à partir de référencement indirect sur le raisonnement spatial qualitatif, sur la modélisation et la manipulation des incertitudes, sur les aspects spatiaux dans le traitement automatique du langage et sur la géovisualisation de l'information dans un contexte de prise de décision. L'ensemble des développements qui seront réalisés sur la durée de projet (2017-2020) pour atteindre les objectifs scientifiques énumérés ci-dessus seront intégrés au sein d'un prototype construit autour d'une infrastructure de type SOA.

Bibliographie

- Bejaoui L., Pinet F., Bedard Y., Schneider M. (2009). Qualified topological relations between spatial objects with possible vague shape. *International Journal of Geographical Information Science*, vol.23, n° 7, p. 877-921.
- Berjawi B., Duchateau F., Chesneau E., Favetta F., Seccia G., Cuntly C., Miquel M., Laurini R. (2014). Uncertainty visualization of multi-providers cartographic integration. *Journal of Visual Languages and Computing*, vol. 25, p. 995-1002.
- Gaio M., Moncla, L. (2017) Extended Named Entity Recognition Using Finite-State Transducers: An Application To Place Names. *Actes du colloque 9th International Conference on Advanced Geographic Information Systems, 2017, Nice, France.*
- Itani J.K., Gouarderes E., Aniorde, P. (2015). A Framework for Managing Services in a Virtual Community Context, *Actes du colloque IEEE World Congress on Services 2015, New York City, USA.*
- MacEachren A-M., Robinson A., Hopper S., Gardner S., Murray R., Hetzler E. (2005). Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know, *Cartography and Geographic Information Science*, vol 32, n°3, p. 139-160.
- Mathet Y. (2000). New paradigms in space and motion: A model and an experiment. *Actes du colloque ECAI 2000 Workshop on Current Issues in Spatio-Temporal Reasoning*, Berlin, Allemagne.
- Moncla L., Gaio M., Noguerras-Iso J., Mustière S. (2016). Reconstruction of itineraries from annotated text with an informed spanning tree algorithm. *International Journal of Geographical Information Science*, vol. 30, n° 6, p. 1137-1160.
- Nouvel D., Ehrman M., Rosset S. (2016). *Named Entities for Computational Linguistics*, ISTE Editions and Wiley.
- Siqueira T. L. L., Ciferri C. D. A., Times V. C., Ciferri R. R. (2014). Modeling Vague Spatial Data Warehouses using the VSCube Conceptual Model, *Geoinformatica*, vol. 18, n° 2, p. 313-356.