

LIRIS

UMR 5205 CNRS

Thèses et HDR 2018

Laboratoire d'InfoRmatique
en Image et Systèmes d'information



INSA



— université
— LUMIÈRE
— LYON 2



HDR

Optimal Transport for Computer Graphics and Temporal Coherence of Image Processing Algorithms Nicolas Bonneel	7
Contributions to local and non-local surface analysis Julie Digne	11
Architectures orientées services : haute disponibilité, confiance, et sociabilité Noura Faci	15
Contributions to Pattern Mining in Augmented Graphs Marc Plantevit	23

Thèses

Integration Framework for Artifact-centric Processes in the Internet of Things Maroun Abi Assaf	31
Suivi visuel d'objets dans un réseau de caméras intelligentes : application aux systèmes de manutention automatisés Adel Benamara	33
Détection d'anomalies dans les séries temporelles : application aux masses de données sur les pneumatiques Seif Eddine Benkabou	37
Analyse d'images pour l'identification multi-organes d'espèces végétales Sarah Bertrand	39
Opérateur de Laplace-Beltrami discret sur les surfaces digitales Thomas Caissard	43
Modèles de programmation par contraintes pour le clustering conceptuel : application à un problème de configuration d'ERP Maxime Chabert	45

Ré-identification de personnes dans des images par apprentissage automatique Yiqiang Chen	47
Acceptation et usage d'un environnement numérique du travail dans l'enseignement primaire Elena Codreanu	51
Criticalité, identification et jeux de suppression de sommets dans les graphes Antoine Dailly	53
Migration et Enrichissement Sémantique d'Entités Culturelles Joffrey Decourselle	57
Fouille de grilles spatio-temporelles appliquée à la classification d'images et à l'analyse d'automates cellulaires Romain Deville	59
Mécanismes de négociation multilatérale pour la prise de décision collective Ndeye Arame Diago	61
Extraction et analyse du réseau acoustique d'oiseaux sociaux Marie Fernandez	63
Data-driven computational modelling for some of the implications of dopamine in the brain : From subcellular signalling to area networks Alexandre Foncelle	65
Représentation et échange de données tridimensionnelles géolocalisées de la ville Jérémy Gaillard	67
Towards Smart Services with Reusable and Adaptable Connected Objects: An Application to Wearable Non-Invasive Biomedical Sensors Arthur Gatouillat	69
Modèle biomécanique patient-spécifique pour la prédiction du mouvement respiratoire Matthieu Giroux	71
Hybrid 3D Mass Spring System for Soft Tissue Simulation Karolina Golec	73
Curare : curation et gestion de grandes collections de données sur le nuage Gavin Kemp	75

Détection non supervisée basée sur les relations spatiales, application à la détection d'objets et la reconnaissance de structure de documents d'entreprises Louisa Kessi	79
Traitement distribué de requêtes sur des flux variants Roland Kotto Kombi	81
Détection, localisation et typage de texte dans des images de documents hétérogènes par Réseaux de Neurones Profonds Bastien Moysset	83
Apprentissage ensembliste : étude comparative et améliorations par la sélection dynamique Anil Narassiguin	85
Prise en compte de la qualité des données lors de l'extraction et de la sélection d'évolutions dans les séries temporelles de champs de déplacements en imagerie satellitaire Hoang Viet Tuan Nguyen	87
Recommendation of Activity Sequences during Distributed Events Diana Nurbakova	91
Modélisation et utilisation de ressources et services Web et indexation de données dans un contexte d'incertitude Asma Omri	95
Analyse de formes et de textures : application à l'identification et à l'aide à la gradation de pièces de monnaie Xingyu Pan	99
Résolution de quelques problèmes d'optimisation dans les jeux et les graphes Claudia Gabriela Paris Sierra	101
Échantillonneurs basse discrétion anti-aliasés pour du rendu réaliste avec estimateurs de Monte Carlo Hélène Perrier	103
Practically Preserving and Evaluating Location Privacy Vincent Primault	105
Système de recommandation de ressources pédagogiques fondé sur les liens sociaux : formalisation et évaluation Mohammed Tadlaoui	107

Optimal Transport for Computer Graphics and Temporal Coherence of Image Processing Algorithms

Nicolas Bonneel

Jury

Filippo SANTAMBROGIO, PR, Lyon 1,
Julie DELON, PR, Paris Descartes,
Quentin MÉRIGOT, PR, Paris Sud,
Justin SOLOMON, Assistant Professor, MIT,
Nicolas COURY, PR, Bretagne Sud,

rapporteur
rapporteuse
rapporteur
examineur
examineur

Résumé

Ces dernières années, mes travaux ont porté sur deux sujets principaux : le transport optimal et le traitement de vidéos. Mon HDR détaille ces deux aspects.

Le transport optimal est un cadre en vogue pour manipuler des distributions de probabilité, des histogrammes ou plus généralement des fonctions. Il consiste à considérer une fonction comme un tas de sable se déplaçant dans l'espace en un coût minimum. Cela permet par exemple de définir une manière d'interpoler entre deux distributions de probabilité (ou plus), ou de définir une distance pertinente entre des histogrammes. Cette théorie a vu beaucoup d'applications aussi bien en informatique graphique qu'en apprentissage profond.

Ici, mes travaux se sont d'abord tournés vers le développement d'algorithmes efficaces pour résoudre des problèmes de transport optimal. Ensuite, j'ai défini et résolu des problèmes inverses utilisant la géométrie du transport optimal, comme le calcul de coordonnées barycentriques ou l'apprentissage de dictionnaires pour des histogrammes.

Le traitement vidéo a vu une croissance rapide due à la grande disponibilité d'appareils photo grand public, et d'applications faciles à utiliser permettant des traitements vidéo puissants, telles que SnapChat, Instagram ou Tik Tok. Cependant, re-adapter le long historique d'algorithmes de traitement d'images pour les faire fonctionner sur des vidéos est un travail titanesque. En effet, appliquer trivialement un traitement d'images sur chaque image d'une vidéo va souvent produire des artefacts temporels tels que du scintillement. Mes travaux ont cherché à adapter les algorithmes de traitement d'images les plus courants aux vidéos, de manière temporellement cohérente. J'ai d'abord adapté des filtres spécifiques (étalonnage de couleurs, et images intrinsèques), et ai finalement trouvé une manière d'adapter un grand nombre de traitements d'images aux vidéos sans même connaître leurs formulations.

Ma présentation d'HDR décrira mes contributions les plus représentatives.

Abstract

During the past few years, my research has focused on two main areas: optimal transport and video processing. My HDR details these two aspects.

Optimal transport is a trending framework for manipulating probability distributions, histograms, or more generally, functions. It consists in seeing a function as a pile of sand moving in space at a minimum cost. This allows, for instance, to define a way to interpolate between two (or more) probability distributions, or to define a meaningful distance between histograms. This theory has seen many applications from computer graphics to deep learning.

HDR

Soutenu le 09/11/2018

Établissement :
Université Claude Bernard Lyon1

Contact :
Nicolas.Bonneel@liris.cnrs.fr

Here, my work has first focused on building efficient algorithms to solve optimal transport problems. Then, I defined and solved inverse problems making use of the optimal transport geometry, such as computing barycentric coordinates or performing dictionary learning for histograms.

Video processing has seen a large growth due to the wide availability of consumer cameras, and user-friendly apps featuring powerful video processing capabilities such as SnapChat, Instagram or Tik Tok. However, re-adapting the long history of image processing algorithms to make them work on videos is a daunting task. Indeed, trivially applying image processing filters to all frames of a video most often results in temporal artifacts such as flickering. My work has focused on bringing common image processing algorithms to the realm of videos, in a temporally consistent fashion. I first worked on bringing specific filters to videos (color grading and intrinsic images), and ultimately found a way to bring many image processing filters to videos without even knowing their formulation.

My HDR presentation will describe my most representative contributions.

Publications

Revue internationale avec comité de lecture

- Morgan Schmitz, Matthieu Heitz, Nicolas Bonneel, Fred Maurice Ngolè Mboula, David Coeurjolly, Marco Cuturi, Gabriel Peyré & Jean-Luc Starck (2018). « Wasserstein Dictionary Learning: Optimal Transport-based unsupervised non-linear dictionary learning ». *SIAM Journal on Imaging Sciences*, vol. 11, n°1, pp. 643-678. doi : 10.1137/17M1140431. ArXiv : 1708.01955. HAL : hal-01717943.
- Nicolas Bonneel, David Coeurjolly, Pierre Gueth & Jacques-Olivier Lachaud (2018). « Mumford-Shah Mesh Processing using the Ambrosio-Tortorelli Functional ». *Computer Graphics Forum*, vol. 37, n°7, pp. 75-85. doi : 10.1111/cgf.13549. HAL : hal-01870901.
- Nicolas Bonneel, Balazs Kovacs, Sylvain Paris & Kavita Bala (2017). « Intrinsic Decompositions for Image Editing ». *Computer Graphics Forum*, vol. 36, n°2, pp. 593-609. HAL : hal-01483773.
- Nicolas Bonneel, James Tompkin, Deqing Sun, Oliver Wang, Kalyan Sunkavalli, Sylvain Paris & Hanspeter Pfister (2017). « Consistent Video Filtering for Camera Arrays ». *Computer Graphics Forum*, vol. 36, n°2, pp. 397-407. HAL : hal-01483753.
- Vincent Léon, Nicolas Bonneel, Guillaume Lavoué & Jean-Philippe Vandeborre (2016). « Continuous semantic description of 3D meshes ». *Computers and Graphics*, vol. 54, pp. 47-56. doi : 10.1016/j.cag.2015.07.018. HAL : hal-01196895.
- Nicolas Bonneel, Gabriel Peyré & Marco Cuturi (2016). « Wasserstein Barycentric Coordinates: Histogram Regression Using Optimal Transport ». *ACM Transactions on Graphics*, vol. 35, n°4, pp. 71:1-71:10. doi : 10.1145/2897824.2925918. HAL : hal-01303148.
- Nicolas Bonneel, Julien Rabin, Gabriel Peyré & Hanspeter Pfister (2015). « Sliced and Radon Wasserstein Barycenters of Measures ». *Journal of Mathematical Imaging and Vision*, vol. 1, n°51, pp. 22-45. doi : 10.1007/s10851-014-0506-3. HAL : hal-00881872.
- Nicolas Bonneel, James Tompkin, Kalyan Sunkavalli, Deqing Sun, Sylvain Paris & Hanspeter Pfister (2015). « Blind video temporal consistency ». *ACM Transactions on Graphics*, vol. 34, n°6, pp. 196:1-196:9. doi : 10.1145/2816795.2818107. HAL : hal-01264081.
- Nicolas Bonneel, Kalyan Sunkavalli, James Tompkin, Deqing Sun, Sylvain Paris & Hanspeter Pfister (2014). « Interactive intrinsic video editing ». *ACM Transactions on Graphics*, vol. 33, n°6, pp. 197:1-197:10. doi : 10.1145/2661229.2661253. HAL : hal-01264124.
- Fangyang Shen, Kalyan Sunkavalli, Nicolas Bonneel, Szymon Rusinkiewicz, Hanspeter Pfister & Xin Tong (2014). « Time-Lapse Photometric Stereo and Applications ». *Computer Graphics Forum*, vol. 33, n°7, pp. 359-367. doi : 10.1111/cgf.12504. HAL : hal-01264129.

Conférences internationales avec comité de lecture

- Beatrix-Emőke Fülöp-Balogh, Nicolas Bonneel & Julie Digne (2018). « Correcting Motion Distortions in Time-of-Flight Imaging ». *Motion, Interaction and Games 2018*, 8 novembre 2018, Limassol (Chypre), 7p. doi : 10.1145/3274247.3274512. HAL : hal-01890418.
- Vincent Léon, Vincent Itier, Nicolas Bonneel, Guillaume Lavoué & Jean-Philippe Vandeborre (2017). « Semantic correspondence across 3D models for example-based modeling ». *Eurographics Workshop on 3D Object Retrieval 2017 (3DOR 2017)*, 24 avril 2017, Lyon (France), 7 pages. HAL : hal-01535056.
- Morgan Schmitz, Matthieu Heitz, Nicolas Bonneel, Fred Ngolè, David Coeurjolly, Marco Cuturi, Gabriel Peyré & Jean-Luc Starck (2017). « Optimal transport-based dictionary learning and its application to Euclid-like Point Spread Function representation ». *Wavelets and Sparsity XVII*, 10 août 2017, San Diego (États-Unis), 103940H. doi : 10.1117/12.2270641. HAL : hal-01635342.
- Evgeny Levinkov, James Tompkin, Nicolas Bonneel, Steffen Kirchhoff, Bjoern Andres & Hanspeter Pfister (2016). « Interactive Multicut Video Segmentation ». *Pacific Graphics*, 14 octobre 2016, Okinawa (Japon), pp. 33-38. HAL : hal-01378800.
- M Keuper, E Levinkov, Nicolas Bonneel, Guillaume Lavoué, T Brox & B Andres (2015). « Efficient Decomposition of Image and Mesh Graphs by Lifted Multicuts ». *The IEEE International Conference on Computer Vision (ICCV)*, 16 décembre 2015, Santiago (Chili), 9 pages. HAL : hal-01264120.

Conférences nationales avec comité de lecture

- Vincent Léon, Nicolas Bonneel, Guillaume Lavoué & Jean-Philippe Vandeborre (2014). « Géodésiques sémantiques pour la description et la labélisation automatique de parties ». *Journées de l'Association Française d'Informatique Graphiques, Reims Image 2014*, 28 novembre 2014, Reims (France). HAL : hal-01171234.

Autres

- Nicolas Bonneel, David Coeurjolly, Pierre Gueth & Jacques-Olivier Lachaud (2018). « Mumford-Shah Mesh Processing using the Ambrosio-Tortorelli Functional ». HAL : hal-01815283.
- Nicolas Bonneel (2017). « Le transport optimal pour des applications en informatique graphique ». *Interstices*. HAL : hal-01533684.

Julie Digne

Jury

Bruno LÉVY, DR, INRIA Nancy Grand Est,
Niloy MITRA, PR, University College London,
Gabriel PEYRÉ, DR, CNRS/ENS Ulm,
Raphaëlle CHAINE, PR, Université Lyon 1,
Stefanie HAHMAN, PR, Université de Grenoble,
Ovsjanikov MAKS, PR, Ecole Polytechnique,

rapporteur
rapporteur
rapporteur
examinatrice
examinatrice
examineur

Résumé

Face à l'explosion des systèmes d'acquisition de surfaces, les algorithmes de traitement numérique de la géométrie évoluent rapidement pour s'adapter à la diversité des types de données et des qualités d'acquisition.

Tout d'abord, pour des données de qualité moindre, il est nécessaire d'améliorer la mesure par des processus de débruitage et de super-résolution. L'analyse de l'auto-similarité des surfaces permet de développer des approches d'amélioration de la mesure que ce soit pour des surfaces d'objet réel ou des formes généralisées (des formes dont la dimension intrinsèque peut varier). Au delà des données basse-résolution, la prise en compte de cette similarité permet également de compresser efficacement des données de surfaces, que l'on peut ensuite rééchantillonner pendant la décompression.

De plus, si les données géométriques sont en elles-mêmes un sujet d'étude, elles peuvent être accompagnées de données d'image ou d'autres mesures, qui permettent de compléter ou d'augmenter les données géométriques par une analyse jointe. Cette habilitation aborde ce problème de la fusion de données multi-capteur pour enrichir des nuages de points représentant des scènes urbaines par une collection de photos, qui, une fois correctement recalées permettent par exemple de coloriser le nuage de points.

Enfin, la dernière partie de cette habilitation s'intéresse à la reconstruction de maillages surfaciques. Pour certaines applications il est en effet intéressant de représenter une surface sous forme de maillage, et de remplacer ainsi localement plusieurs points par une unique facette plane. Pour cela, il est possible de quantifier la distance de transport optimal entre le nuage de points initial et un maillage. La reconstruction et l'optimisation du maillage peut être ainsi guidée par la minimisation de cette métrique.

Abstract

Following the recent progress of surface acquisition systems, geometry processing algorithms quickly evolve to deal with the variety of data types and acquisition quality. This habilitation manuscript details some recent approaches to tackle this challenge. First, for low-quality data, it is necessary to improve the measure by denoising or super-resolution algorithms. Self-similarity analysis yields efficient methods for improving the acquisition quality either for real object surfaces, or generalized shapes (shapes whose intrinsic dimension is not constant). Beyond low-resolution acquisition, taking this similarity into account also permits to compress point set surfaces, that can then be resampled during decompression. While geometric data are per se a research topic, additional image data or other type of measures can be acquired simultaneously, which allows to complete or augment the geometric information through a joint analysis. This manuscript addresses this multi-captor data problem to augment urban scenes point sets by using

HDR

Soutenu le 27/09/2018

Établissement :
Université Claude Bernard Lyon 1

Contact :
Julie.Digne@liris.cnrs.fr

a collection of pictures, which permits to colorize point clouds, once images are accurately registered. Finally, for specific purposes, it is interesting to represent surfaces as polygonal meshes potentially replacing several points by a single planar facet. To do so, this manuscript describes an Optimal Transportation metric between the initial point cloud and a mesh. The reconstruction and optimization of the mesh can then be driven by the minimization of this distance.

Publications

Revue internationale avec comité de lecture

- Julie Digne, Sébastien Valette & Raphaëlle Chaine (2018). « Sparse Geometric Representation Through Local Shape Probing ». *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, n°7, pp. 2238-2250. doi : 10.1109/TVCG.2017.2719024. HAL : hal-01547820.
- Azzouz Hamdi-Cherif, Julie Digne & Raphaëlle Chaine (2018). « Super-resolution of Point Set Surfaces using Local Similarities ». *Computer Graphics Forum*, vol. 37, n°1, pp. 60-70. doi : 10.1111/cgf.13216. HAL : hal-01521806.
- Johann Béarzi, Julie Digne & Raphaëlle Chaine (2018). « Wavejets: A Local Frequency Framework for Shape Details Amplification ». *Computer Graphics Forum*, vol. 37, n°2, pp. 13-24. doi : 10.1111/cgf.13338. HAL : hal-01722993.
- Oscar Argudo, Carlos Andujar, Antonio Chica, Eric Guérin, Julie Digne, Adrien Peytavie & Eric Galin (2017). « Coherent multi-layer landscape synthesis ». *Visual Computer*. doi : 10.1007/s00371-017-1393-6. HAL : hal-01520366.
- Eric Guérin, Julie Digne, Eric Galin, Adrien Peytavie, Christian Wolf, Bedrich Benes & Benoît Martinez (2017). « Interactive Example-Based Terrain Authoring with Conditional Generative Adversarial Networks ». *ACM Transactions on Graphics*. HAL : hal-01583706.
- Julie Digne & Carlo De Franchis (2017). « The Bilateral Filter for Point Clouds ». *Image Processing On Line*, vol. 7, pp. 278-287. doi : 10.5201/ipol.2017.179. HAL : hal-01636966.
- Maximilien Guislain, Julie Digne, Raphaëlle Chaine & Gilles Monnier (2017). « Fine scale image registration in large-scale urban LIDAR point sets ». *Computer Vision and Image Understanding*, vol. 157, pp. 90-102. doi : 10.1016/j.cviu.2016.12.004. HAL : hal-01468091.
- Eric Guérin, Julie Digne, Eric Galin & Adrien Peytavie (2016). « Sparse representation of terrains for procedural modeling ». *Computer Graphics Forum*, vol. 35, n°2, pp. 177-187. doi : 10.1111/cgf.12821. HAL : hal-01258986.
- Julie Digne (2015). « An Implementation and Parallelization of the Scale-Space Meshing Algorithm ». *Image Processing On Line*, vol. 5, pp. 282-295. doi : 10.5201/ipol.2015.102. HAL : hal-01238790.
- Julie Digne & Jean-Michel Morel (2014). « Numerical analysis of differential operators on raw point clouds ». *Numerische Mathematik*, vol. 127, n°2, pp. 255-289. doi : 10.1007/s00211-013-0584-y. HAL : hal-01135993.
- Julie Digne, Raphaëlle Chaine & Sébastien Valette (2014). « Self-similarity for accurate compression of point sampled surfaces ». *Computer Graphics Forum*, vol. 33, n°2, pp. 155-164. doi : 10.1111/cgf.12305. HAL : hal-00983003.
- Julie Digne (2014). « An analysis and implementation of a parallel ball pivoting algorithm ». *Image Processing On Line*, vol. 4, pp. 149-168. doi : 10.5201/ipol.2014.81. HAL : hal-01267724.

- Julie Digne, David Cohen-Steiner, Pierre Alliez, Fernando De Goes & Mathieu Desbrun (2013). « Feature-Preserving Surface Reconstruction and Simplification from Defect-Laden Point Sets ». *Journal of Mathematical Imaging and Vision*, pp. 1-14. doi : 10.1007/s10851-013-0414-y. HAL : hal-00827623.

Conférences internationales avec comité de lecture

- Beatrix-Emőke Fülöp-Balogh, Nicolas Bonneel & Julie Digne (2018). « Correcting Motion Distortions in Time-of-Flight Imaging ». *Motion, Interaction and Games 2018*, 8 novembre 2018, Limassol (Chypre), 7p. doi : 10.1145/3274247.3274512. HAL : hal-01890418.
- Maximilien Guislain, Julie Digne, Raphaëlle Chaine, Dimitri Kudelski & Pascal Lefebvre-Albaret (2016). « Detecting and Correcting Shadows in Urban Point Clouds and Image Collections ». *2016 International Conference on 3D Vision (3DV)*, Oct 2016, Stanford, United States, 28 octobre 2016, Stanford (États-Unis), 9p. HAL : hal-01393998.

Conférences nationales avec comité de lecture

- Johann Béarzi & Julie Digne (2016). « Surface derivatives computation using Fourier Transform ». *JFIG 2016*, 30 novembre 2016, Grenoble (France). HAL : hal-01735374.

Éditions scientifique d'Ouvrages (livres, chapitres, colloques, congrès, n° spéciaux)

- Julie Digne, Mariella Dimiccoli, Neus Sabater & Philippe Salembier (2015). « Neighborhood Filters and the Recovery of 3D Information ». *Handbook of Mathematical Methods in Imaging (2nd Edition)*, Otmar Scherzer, Springer, pp. 1645-1673. doi : 10.1007/978-1-4939-0790-8_27. HAL : hal-01238753.

Autres

- Maximilien Guislain, Julie Digne, Raphaëlle Chaine & Gilles Monnier (2016). « Recalage d'image dans des nuages de points de scènes urbaines ». *Actes des journées du Groupe de Travail en Modélisation Géométrique (GTMG) 2016*. HAL : hal-01320263.
- Julie Digne, Raphaëlle Chaine & Sébastien Valette (2014). « Auto-similarité pour la compression haute précision des surfaces échantillonnées ». *Actes des GTMG*, pp. 1-11. HAL : hal-01301060.
- Julie Digne, Pierre Alliez & David Cohen-Steiner (2013). « Transport optimal pour la reconstruction robuste de formes à partir de nuages de points ». *Actes des GTMG*, pp. 1-10. HAL : hal-01339238.

Jury

Jamal BENTAHAR, Full Professor, Concordia University, Canada	Rapporteur
Khalil DRIRA, DR, LAAS-CNRS, Toulouse	Rapporteur
Daniela GRIGORI, PR, Université Paris-Dauphine, Paris	Rapporteur
Djamal BENSILIMANE, PR, Université Claude Bernard, Lyon,	Examineur
François CHAROY, PR, Université de Lorraine, Nancy,	Examineur
Walid GAALOUL, PR, Institut Mines Télécom, Paris,	Examineur
Franck MORVAN, PR, Université Paul Sabatier, Toulouse,	Examineur
Marta RUKOZ, PR, Université Paris Ouest, Nanterre,	Examinatrice

Résumé

Mon agenda de recherche au cours des 10 dernières années a principalement porté sur la robustesse des applications orientées-services. Cela consiste à découvrir, sélectionner, et composer des services pouvant être sujets à des défaillances au moment de l'exécution, et susceptibles de faire échouer ces applications. Pour atténuer les risques de défaillance, nous nous sommes intéressés à la pertinence de la diversité pour exécuter les applications orientées-services avec succès. Plusieurs questions de recherche ont été identifiées telles que (i) comment définir et configurer la diversité; et (ii) comment sélectionner les services pour satisfaire les besoins en tolérance aux fautes. Une phase préliminaire a examiné comment la réplication pourrait aller de pair avec la diversité, de sorte que, par exemple, plusieurs services sémantiquement équivalents (mais pas nécessairement les mêmes d'un point de vue non fonctionnel) soient regroupés ensemble. La phase suivante consistait à déterminer comment améliorer la qualité de découverte de service en présence de fautes. L'idée était de s'appuyer sur les principes des réseaux sociaux pour sélectionner les services les plus appropriés, les dotant ainsi de certaines caractéristiques sociales comme qui collabore avec qui. Le travail effectué a permis de définir des réseaux sociaux de services, de les mettre en place, et d'en extraire des informations pertinentes de ces réseaux pour les besoins de la découverte. Assurer le bon usage de ces réseaux sociaux, une autre piste de recherche consistait à chercher des moyens de réglementer les actions des services (par exemple, établir et maintenir des réseaux de contacts) en utilisant la notion d'engagements. En fait, la robustesse au niveau "service" est nécessaire mais pas suffisant. Il est donc aussi important de l'inclure au niveau du processus de sélection lui-même, et donc, étendant notre processus de réflexion sur la question de robustesse. Les systèmes d'évaluation de la confiance des services sont principalement basés sur les expériences utilisateur lors de l'invocation de ces services. La question de recherche fut de comment évaluer la confiance d'un service en présence d'attaques telles que des expériences faussées. Le travail a abouti à la définition d'un modèle de crédibilité basé sur un clustering flou, un mécanisme de filtrage des utilisateurs ayant plusieurs identités et un modèle de confiance basé sur des bases de données probabilistes.

Il est bien connu que les architectures orientées-services et la gestion des processus métier vont de pair pour développer des applications d'entreprise. Nous avons jugé approprié d'explorer le tissage de principes sociaux, tels que la proximité, dans la conception et l'exécution des processus métier pour garantir la continuité de l'activité. La principale préoccupation est de comment assurer un alignement parfait des technologies Web 2.0 avec les stratégies de développement et les meilleures pratiques en entreprise. Le travail effectué a favorisé les relations sociales entre les employés d'une entreprise pour améliorer leur performance. En effet, il a été observé que

HDR

Soutenue le 14/12/2018

Établissement :
Université Claude Bernard Lyon 1

Contact :
Noura.Faci@liris.cnrs.fr

les relations informelles entre les personnes existent dans les entreprises aux niveaux stratégique et opérationnel. Les questions abordées sont de (i) comment concevoir des processus métier sur la base des principes sociaux et (ii) comment assurer une exécution efficace de ces processus lors de conflits de ressources. Une phase de recherche préliminaire était de développer des réseaux dédiés basés sur les relations sociales (par exemple, la supervision et le partenariat) entre les trois composants d'un processus (tâche, personne et machine) et d'analyser la valeur ajoutée de ces réseaux pour les entreprises. Ces réseaux capturent les différentes situations de collaboration entre tâches, entre personnes et entre machines. La phase suivante consistait à aborder la gestion des ressources dans les entreprises. L'idée était de capitaliser sur ces réseaux dédiés pour coordonner la production de ressources, leur consommation, et leur utilisation. Ce travail a permis: (i) de classer les ressources en fonction de leur nature et de leur type, (ii) identifier les conflits par catégorie de ressources et (iii) proposer des solutions pour ces conflits en utilisant les réseaux appropriés. Garantir la stabilité de ces réseaux devient alors une nécessité. À cette fin, l'approche proposée repose sur des engagements métier et sociaux pour réguler le fonctionnement au sein de ces réseaux.

Pour les 4 prochaines années, mon agenda de recherche abordera les nouveaux défis liés à l'Internet des objets, (r)évolution du Web, tels que la combinaison à la volée des objets et les objets cognitifs. Quelques résultats préliminaires ont déjà été partagé avec la communauté scientifique.

Mots-clés composition de services, tolérance aux fautes, diversité, systèmes de confiance, robustesse, gestion des processus métier, réseaux sociaux.

Abstract

My research agenda over the last 10 years has mainly revolved around the topic of robustness of service-oriented applications. This meant discovering, selecting, and composing services that could be faulty at run-time, and hence could make these applications fail. To mitigate the risks of failure, we examined diversity appropriateness for successful service-oriented applications. We raised a couple of questions that are (i) how to define and configure diversity; and (ii) how to select services so that fault tolerance requirements are met. A preliminary investigation examined how replication could work hand-in-hand with diversity so, that, for instance many replicate services (not necessarily the same from a non-functional perspective) could be grouped together. The next phase was to focus on how to improve the quality of service discovery in the presence of faults. The idea was to embrace the principles of social networks to select the most appropriate services, thus endowing them with some social characteristics like who collaborates with who. The work carried out allowed defining and setting up social networks of services and extracting relevant details from these networks for the needs of discovery. To ensure proper use of social networks, another avenue of research consisted of looking for ways that would regulate services' actions (e.g., establish and maintain networks of contacts) using the concept of commitments. In fact, robustness at the level of "service" is necessary but not sufficient. Thus, it is important to include the robustness in the selection process itself as well, and thus, extends our reflection process. We raised the question of how to evaluate a service's trust. Trust systems are mostly based on user/service experiences when requesting these services. The work resulted in the definition of a credibility model based on a fuzzy clustering, a mechanism for filtering users with multiple identities, and a trust model based on probabilistic databases.

It is well known that service-oriented architectures and business process management go hand in hand for developing enterprise applications. We deemed appropriate exploring the weaving of social principles like proximity into the design and execution of business processes so, that, business continuity is guaranteed. The main concern is how to ensure perfect alignment of Web 2.0 technologies with business development strategies and best practices. The work carried out fostered the social relations among an enterprise's employees to perform better. Indeed, it was noted that informal relations between people exist in enterprises at strategic, management, and operational levels. We address the questions of (i) how to design business processes based on social principles and (ii) how to ensure effective execution of these processes during resource conflicts. A preliminary research phase was to develop dedicated networks based on social relations (e.g., supervision and partnership) between a process's three components (task, person, and machine) and to analyze the added value of these networks to enterprises. These networks capture the different situations of collaboration between tasks, between persons, and between machines. The next phase was to address resource management in enterprises. The idea was to capitalize on dedicated networks to coordinate resource production, consumption, and use. The work allowed: (i) to categorize resources according to their nature and type, (ii) to identify conflicts per resource category, and (iii) to propose solutions to these conflicts by using the appropriate networks. Guaranteeing the stability of these networks is a necessity. To this end, the proposed approach relies on business and social commitments to regulate the functioning within these networks.

For the next 4-years, my research agenda will tackle new challenges related to Internet of Things, (r)evolution of the Web, such as on-the-fly combination of things and cognitive things. Some early findings have already been shared with the community.

Keywords Service composition, fault tolerance, diversity, trust management systems, robustness, business process management, social networks.

Publications

Revue internationale avec comité de lecture

- Wehbi Benallal, Mahmoud Barhamgi, Djamel Benslimane, Zakaria Maamar, Noura Faci & Ameni Bellaaj (2018). « A Knowledge-based approach to manage configurable business processes ». *Concurrency and Computation: Practice and Experience*, vol. 2, n°6, pp. 22-44. doi : 10.1002/cpe.4920. HAL : hal-01857537.
- Thar Baker, Emir Ugljanin, Noura Faci, Mohamed Sellami, Zakaria Maamar & Ejub Kajan (2018). « Everything as a resource: foundations and illustration through Internet-of-things ». *Computers in Industry*, vol. 94, pp. 62-74. doi : 10.1016/j.compind.2017.10.001. HAL : hal-01636978.
- Raul Lara-Cabrera, Antonio Gonzalez Pardo, Karim Benouaret, Noura Faci, Djamel Benslimane & David Camacho (2017). « Measuring the Radicalisation Risk in Social Networks ». *IEEE Access*, vol. 5, pp. 10892-10900. doi : 10.1109/ACCESS.2017.2706018. HAL : hal-01636970.
- Zakaria Maamar, Noura Faci, Mohamed Sellami, Khoulood Boukadi, Fadwa Yahya, Ahmed Barnawi & Sherif Sakr (2017). « On business process monitoring using cross-flow coordination ». *Service Oriented Computing and Applications*, vol. 11, n°2, pp. 203-215. doi : 10.1007/s11761-017-0206-0. HAL : hal-01636954.

- Noura Faci, Zakaria Maamar, Vanilson Buregio, Emir Ugljanin & Djamal Benslimane (2017). « Web 2.0 applications in the workplace: How to ensure their proper use? ». *Computers in Industry*, vol. 88, pp. 1-11. doi : 10.1016/j.compind.2017.03.003. HAL : hal-01636958.
- Zohra Saoud, Noura Faci, Zakaria Maamar & Djamal Benslimane (2016). « A Fuzzy-based Credibility Model to Assess Web Services Trust under Uncertainty ». *Journal of Systems and Software*, vol. 122, n°4, pp. 496-506. doi : 10.4018/IJSSOE.2016100103. HAL : hal-01207317.
- Zakaria Maamar, Noura Faci, Sherif Sakr, Mohamed Boukhebouze & Ahmed Barnawi (2016). « Network-based social coordination of business processes ». *Information Systems*, vol. 58, pp. 56-74. doi : 10.1016/j.is.2016.02.005. HAL : hal-01636895.
- Vanilson Buregio, Ejub Kajan, Mohamed Sellami, Noura Faci, Zakaria Maamar & Djamal Benslimane (2016). « Revisiting Software Engineering in the Social Era ». In *International Journal of Systems and Service-Oriented Engineering*, vol. 6, n°4, pp. 36-46. doi : 10.4018/IJSSOE.2016100103. HAL : hal-01636905.
- Sanjay Modgil, Nir Oren, Noura Faci, Felipe Meneguzzi, Miles Simon & Michael Luck (2015). « Monitoring compliance with E-contracts and norms ». *Artificial Intelligence and Law*, vol. 23, n°2, pp. 161-196. doi : 10.1007/s10506-015-9167-9. HAL : hal-01212164.
- Noura Faci, Marinella Petrocchi, Gianpiero Costantino, Fabio Martinelli & Zakaria Maamar (2015). « A Quality Model for Social Networks Populated with Web Services ». *Service Oriented Computing and Applications*, vol. 9, n°2, pp. 139-155. HAL : hal-01207713.
- Zakaria Maamar, Noura Faci, Ejub Kajan, Sherif Sakr, Boukhebouze Mohamed & Ahmed Barnawi (2015). « How to Make Business Processes "Socialize"? ». *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, vol. 15, n°5, 15 pages. doi : 10.4108/eai.17-9-2015.150284. HAL : hal-01212094.
- Zakaria Maamar, Noura Faci, Khoulood Boukadi, Quan Z. Sheng & Lina Yao (2014). « Commitments to Regulate Social Web Services Operation ». *IEEE Transactions on Services Computing*, vol. 7, n°2, pp. 154-167. doi : 10.1109/TSC.2013.29. HAL : hal-01339229.
- Ejub Kajan, Noura Faci, Zakaria Maamar, Alfred Loo, Aldina Pljaskovic & Quan Z Sheng (2014). « The Network-based Business Process ». *IEEE Internet Computing*, vol. 18, n°2, pp. 63-69. doi : 10.1109/MIC.2014.35. HAL : hal-01207772.
- Felipe Meneguzzi, Sanjay Modgil, Nir Oren, Simon Miles, Michael Luck & Noura Faci (2012). « Applying Electronic Contracting to the Aerospace Aftercare Domain ». In *International Scientific Journal Engineering Applications of Artificial Intelligence (EAAI)*, vol. 25, pp. 1471-1487. doi : 10.1016/j.engappai.2012.06.004. HAL : hal-01353108.
- Zakaria Maamar, Noura Faci, Leandro Wives Krug, Youakim Badr, Pedro Santos & José Palazzo Moreira De Oliveira (2011). « Using Social Networks for Web Services Discovery ». *IEEE Internet Computing*. doi : 10.1109/MIC.2011.27. HAL : hal-01207288.
- Zakaria Maamar, Noura Faci, Soraya Kouadri Mostfaoui & Fahim Akhter (2011). « Towards a Framework for Weaving Social Networks into Mobile Commerce ». In *International Journal of Systems and Service-Oriented Engineering*, pp. 32-46. doi : 10.4018/jssoe.2011070103. HAL : hal-01354573.

- Zahia Guessoum, Jean-Pierre Briot, Nora Faci & Olivier Marin (2010). « Towards reliable multi-agent systems: An adaptive replication mechanism ». *Multiagent and Grid Systems - An International Journal of Cloud Computing*, vol. 6, n°1, pp. 1-24. doi : 10.3233/MGS-2010-0139. HAL : hal-01170014.
- Zakaria Maamar, Leandro Wives Krug, Youakim Badr, Said Elnaffar, Khoulood Boukadi & Noura Faci (2010). « LinkedWS: A Novel Web Services Discovery Model Based on the Metaphor of Social Networks ». *Simulation Modelling Practice and Theory*, pp. doi : 10.1016/j.simpat.2010.06.018. HAL : hal-00528092.

Revue nationale avec comité de lecture

- Zohra Saoud, Noura Faci, Zakaria Maamar & Djamal Benslimane (2015). « Un modèle de crédibilité basé sur le clustering flou pour une évaluation probabiliste de la confiance des ressources sur le Web ». *Revue des Sciences et Technologies de l'Information - Série ISI : Ingénierie des Systèmes d'Information*, vol. 20, n°6, pp. 79-98. doi : 10.3166/isi.20.6.79-98. HAL : hal-01256174.

Conférences internationales avec comité de lecture

- Abir Masmoudi, Mahmoud Barhamgi, Noura Faci, Zohra Saoud, Khalid Belhajjame, Djamal Benslimane & David Camacho (2018). « An Ontology-Based Approach for Mining Radicalization Indicators from Online Messages ». *2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA)*, 18 mai 2018, Krakow (Pologne), pp. 609-616. doi : 10.1109/AINA.2018.00094. HAL : hal-01878010.
- Zakaria Maamar, Noura Faci, Khoulood Boukadi, Emir Ugljanin, Mohamed Sellami, Thar Baker & Rafael Angarita (2018). « How to agentify the Internet-of-Things? ». *RCIS 2018 - IEEE 12th International Conference on Research Challenges in Information Science*, 31 mai 2018, Nantes (France), pp. 1-6. doi : 10.1109/RCIS.2018.8406665. HAL : hal-01865726.
- Zakaria Maamar, Noura Faci, Mohamed Sellami, Emir Ugljanin & Ejub Kajan (2018). « Everything-as-a-Thing for abstracting the Internet-of-Things ». *ICSOFT 2018: International Conference on Software Technologies*, 28 juillet 2018, Porto (Portugal), pp. 733-740. doi : 10.5220/0006869907670774. HAL : hal-01865730.
- Zakaria Maamar, Thar Baker, Noura Faci, Emir Ugljanin, Yacine Atif & Mohamed Sellami (2018). « Cognitive computing meets the Internet of Things ». *ICSOFT 2018: International Conference on Software Technologies*, 28 juillet 2018, Porto (Portugal), pp. 741-746. doi : 10.5220/0006877507410746. HAL : hal-01865729.
- Noura Faci, Zakaria Maamar, Thar Baker, Emir Ugljanin & Mohamed Sellami (2018). « In Situ mutation for active things in the IoT context ». *ICSOFT 2018: 13th International Conference on Software Technologies*, 28 juillet 2018, Porto (Portugal), pp. 725-732. doi : 10.5220/0006840607590766. HAL : hal-01865732.
- Zakaria Maamar, Mohamed Sellami, Noura Faci, Emir Ugljanin & Quan Z Sheng (2018). « Storytelling integration of the Internet of Things into business processes ». *BPM 2018: International Conference on Business Process Management*, 14 septembre 2018, Sidney (Australie), pp. 127-142. doi : 10.1007/978-3-319-98651-7_8. HAL : hal-01865733.

- Zakaria Maamar, Mohamed Sellami, Noura Faci & Sylvain Lefebvre (2017). « Detecting and Tackling Run-Time Obstacles in Social Business Processes ». International Conference on Advanced Information Networking and Applications (AINA), 27 mars 2017, Tapei (Taiwan), pp. 371-378. HAL : hal-01636945.
- Emir Ugljanin, Noura Faci, Mohamed Sellami & Zakaria Maamar (2016). « Tracking Users' Actions over Social Media: Application to Facebook ». Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), Demo paper, 13 juin 2016, Paris (France), pp. 255-256. HAL : hal-01636864.
- Emir Ugljanin, Noura Faci, Vanilson Buregio & Zakaria Maamar (2016). « How to Restrict Web 2.0 Applications Use in the Workplace? Example of Google+ Hangouts ». Enterprise Distributed Object Computing (EDOC) Workshops, 5 septembre 2016, Vienna (Autriche), pp. 1-4. HAL : hal-01636887.
- Zohra Saoud, Noura Faci, Zakaria Maamar & Djamel Benslimane (2015). « Sybil Tolerance and Probabilistic Databases to Compute Web Services Trust ». 19th East-European Conference on Advances in Databases and Information Systems (ADBIS 19), 11 septembre 2015, Poitiers (France), pp. 458-471. HAL : hal-01164279.
- Zohra Saoud, Noura Faci, Zakaria Maamar & Djamel Benslimane (2015). « Web Services Trust Assessment based on Probabilistic Databases ». The International Conference on NETWORKED SYSTEMS NETYS 2015, 15 mai 2015, Agadir (Maroc), pp. 397-410. HAL : hal-01160758.
- Zakaria Maamar, Vanilson Buregio, Noura Faci, Djamel Benslimane & Quan Z. Sheng (2015). « Controlling Web 2.0 Applications in the Workplace ». IEEE Enterprise Computing Conference (EDOC), 25 septembre 2015, Adelaide (Australie), pp. 191-200. HAL : hal-01212070.
- Zakaria Maamar, Vanilson Buregio, Noura Faci, Djamel Benslimane & Quan Sheng (2015). « Controlling Web 2.0 Applications in the Workplace ». 2015 IEEE 19th International Enterprise Distributed Object Computing Conference (EDOC), 25 septembre 2015, Adelaide (Australie), pp. 191-200. doi : 10.1109/EDOC.2015.35. HAL : hal-01857562.
- Zohra Saoud, Noura Faci, Zakaria Maamar & Djamel Benslimane (2014). « A Fuzzy Clustering-Based Credibility Model for Trust Assessment in a Service-Oriented Architecture ». 23rd IEEE International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises WETICE, 25 juin 2014, Parme (Italie), pp. 56-61. doi : 10.1109/WETICE.2014.35. HAL : hal-01132728.
- Zakaria Maamar, Sherif Sakr, Noura Faci, Boukhebouze Mohamed & Ahmed Barnawi (2014). « SUPER: Social-Based Business Process Management Framework ». International Conference in Service-Oriented Computing ICSSOC Workshops, 6 novembre 2014, Paris (France), pp. 413-417. HAL : hal-01591783.
- Zakaria Maamar, Noura Faci, S Kouadri Mostefaoui & Ejub Kajan (2013). « Network-based Conflict Resolution in Business Processes ». International Conference on e-Business Engineering (ICEBE), 11 septembre 2013, Coventry (Royaume-Uni), pp. 132-137. doi : 10.1109/ICEBE.2013.20. HAL : hal-01339230.
- Zakaria Maamar, Noura Faci, Alfred Loo & Parisa Ghodous (2012). « Towards a Quality of Social Network (QoSN) Model in the Context of Social Web Services ». International Conference on Exploring Services Science, Lecture Notes in Business Information Processing, 15 février 2012, Geneva (Suisse), pp. 297-310. doi : 10.1007/978-3-642-28227-0_22. HAL : hal-01352956.
- Zakaria Maamar, Noura Faci, Michael Luck & Salahdine Hachimi (2012). « Specifying and Implementing Social Web Services Operation using Commitments ». Symposium On Applied Computing (SAC), 26 mars 2012, Riva del Garda (Trento) (Italie), pp. 1955-1960. doi : 10.1145/2245276.2232099. HAL : hal-01352947.
- Hanane Abdeldjelil, Noura Faci, Zakaria Maamar & Djamel Benslimane (2012). « A Diversity-Based Approach For Managing Faults in Web Services ». IEEE International Conference on Advanced Information Networking and Applications (AINA), 26 mars 2012, Fukuoka (Japon), p. 81. doi : 10.1109/AINA.2012.85. HAL : hal-01352942.
- Noura Faci, Zakaria Maamar & Parisa Ghodous (2012). « Which Social Networks Should Web Services Sign-Up In? ». AAAI Spring Symposium on Intelligent Web Services Meet Social Computing, 26 mars 2012, Palo Alto (États-Unis), pp. 22-26. HAL : hal-01352959.
- Zakaria Maamar, Noura Faci, Quan Z. Sheng & Lina Yao (2012). « Towards a User-Centric Social Approach to Web Service Composition, Execution, and Monitoring ». International Conference on Web Information System Engineering (WISE), 28 novembre 2012, Paphos (Chypre), pp. 72-86. doi : 10.1007/978-3-642-35063-4_6. HAL : hal-01353109.
- Zakaria Maamar, Noura Faci, Leandro Krug Wives, Hamdi Yahyaoui & Hakim Hacid (2011). « Towards a Method for Engineering Social Web Services ». 4th Working Conference on Method Engineering (ME), 22 avril 2011, Paris (France), pp. 153-167. doi : 10.1007/978-3-642-19997-4_15. HAL : hal-01354411.
- Noura Faci, Hanane Abdeldjelil, Zakaria Maamar & Djamel Benslimane (2011). « Using Diversity to Design and Deploy Fault Tolerant Web Services ». International Conference on Collaboration Technologies and Infrastructures (WETICE), 27 juin 2011, Paris (France), pp. 73-78. doi : 10.1109/WETICE.2011.53. HAL : hal-01354423.
- Zakaria Maamar, Noura Faci, Youakim Badr, Leandro Krug Wives, Djamel Benslimane & José Palazzo M. De Oliveira (2011). « Towards a Framework for Weaving Social Networks Principles into Web Services Discovery ». International Conference on Web Intelligence, Mining and Semantics (WIMS'11), 25 mai 2011, Sogndal (Norvège), p. 51. doi : 10.1145/1988688.1988748. HAL : hal-01354384.
- Youakim Badr, Noura Faci, Zakaria Maamar & Frédérique Biennier (2011). « Multi-level Social Networking to Enable and Foster Collaborative Organization ». 12th Working Conference on Virtual Enterprises (PROVE), 19 octobre 2011, São Paulo (Brésil), pp. 3-10. doi : 10.1007/978-3-642-23330-2_1. HAL : hal-01354551.
- Noura Faci, Zakaria Maamar, Hanane Abdeldjelil & Djamel Benslimane (2011). « Vers un framework intégrant les principes des réseaux sociaux dans la découverte de services Web ». Conférence Internationale sur les NOuvelles Technologies de la REpartition (NOTERE'11), 9 mai 2011, Paris (France), inconnue. HAL : hal-01354422.
- Zakaria Maamar, Ghazi Alkhatib, Nora Faci & Youakim Badr (2010). « Towards a Social-Network Framework for Service Computing: A Research Action Plan ». The International Congress on Computer Applications and Computational Science (CACs 2010), 6 décembre 2010 (Singapour), pp. HAL : hal-00528498.

- Zakaria Maamar, Nora Faci, Youakim Badr, Djamal Benslimane, Leandro Wives Krug, Pedro Bispo Dos Santos & José Palazzo Moreira De Oliveira (2010). Weaving Social Networking Principles into Web Services Functioning, 3 décembre 2010, Ayia Napa (Chypre), pp. HAL : hal-00528477.
- Felipe Meneguzzi, Sanjay Modgil, Nir Oren, Simon Miles, Michael Luck, Noura Faci, Camden Holt & Malcom Smith (2009). « Monitoring and Explanation of Contract Execution: A Case Study in the Aerospace Domain ». AAMAS Conference Industry Track 2009, 10 mai 2009, Budapest (Hongrie), inconnue. HAL : hal-01437627.
- Sanjay Modgil, Noura Faci, Felipe Meneguzzi, Nir Oren, Simon Miles & Michael Luck (2009). « A Framework for Monitoring Agent-Based Normative Systems ». Eighth International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'09), 10 mai 2009, Budapest (Hongrie), pp. 153-160. HAL : hal-01437626.

Éditions scientifique d'Ouvrages (livres, chapitres, colloques, congrès, n° spéciaux)

- Zakaria Maamar, Youakim Badr, Noura Faci & Quan Z. Sheng (2014). « Realizing a Social Ecosystem of Web Services ». Advanced Web Services, pp. 455-473. doi : 10.1007/978-1-4614-7535-4. HAL : hal-01277054.
- Zakaria Maamar, Jamel Bentahar, Noura Faci & Philippe Thiran (2012). « Social Web Services Research Roadmap: Present & Future ». Distributed Computing Innovations for Business, Engineering and Science book, IGI Global, pp. 220-233. doi : 10.4018/978-1-4666-2533-4.ch011. HAL : hal-01353110.
- Salahdine Hachimi, Noura Faci & Zakaria Maamar (2012). « Impact Analysis of Web Services Substitution on Configurable Compositions ». Distributed Computing Innovations for Business, Engineering and Science book, IGI Global, pp. 300-313. doi : 10.4018/978-1-4666-2533-4.ch016. HAL : hal-01353111.
- Simon Miles, Nir Oren, Michael Luck, Sanjay Modgil, Felipe Meneguzzi, Noura Faci, Camden Holt & Gary Vickers (2010). « Electronic Business Contracts between Services ». Handbook of Research on P2P and Grid Systems for Service-Oriented Computing: Models, Methodologies and Applications, IGI Global, IGI Global, pp. 732-747. doi : 10.4018/978-1-61520-686-5.ch031. HAL : hal-01381429.

Autres

- Zohra Saoud, Noura Faci, Zakaria Maamar & Djamal Benslimane (2015). « Calcul de la confiance des services web dans un contexte d'utilisateurs multi-identités ». Conférence francophone. HAL : hal-01175783.

Jury

Toon CALDERS, PR, University of Antwerp
Florent MASSEGLIA, DR, INRIA Sophia Antipolis - Méditerranée
Dino PEDRESCHI, PR, University of Pisa
Angela BONIFATI, PR, Université Claude Bernard Lyon 1,
Bruno CRÉMILLEUX, PR, Université de Caen Normandie,
Eric GAUSSIER, PR, Université de Grenoble Alpes,
Céline ROBARDET, PR, INSA Lyon,
Alexandre TERMIER, PR, Université de Rennes 1,

Rapporteur
Rapporteur
Rapporteur
Examinatrice
Examineur
Examineur
Examinatrice
Examineur

Résumé

Dans cette Habilitation à Diriger des Recherches, je présente les principaux résultats auxquels j'ai contribué dans le domaine de la fouille de motifs dans les graphes augmentés. Les graphes sont un puissant outil mathématique permettant de modéliser de nombreux phénomènes réels où les entités sont décrites par des sommets et les relations entre elles via des arêtes. Ces graphes sont bien souvent augmentés par des informations décrivant plus précisément l'activité d'un sommet ou le contexte de l'interaction (l'arête). On parle alors de graphes attribués. Les graphes peuvent également être dynamiques, la structure et/ou les valeurs des attributs évoluant au cours du temps. La découverte de motifs dans de tels graphes permet de fournir à l'utilisateur des informations exploitables et d'enrichir ses connaissances.

Cette thèse se décompose en deux parties. Dans la première partie, je présente les différents domaines de motifs pour les graphes augmentés que j'ai proposés. Cela inclut la découverte de motifs de co-évolution dans des graphes attribués dynamiques, l'étude de liens entre la structure du graphes et les attributs des sommets et la découverte de sous-graphes attribués exceptionnels. J'introduis d'abord les motifs de co-évolution pour l'analyse de graphes dynamiques attribués ainsi que des mesures d'intérêt visant à évaluer un motif par rapport à chacune des dimensions du graphe (i.e., le temps, les attributs et les sommets). Des exemples de motifs découverts dans des données spatio-temporelles sont retournés. Je présente ensuite deux domaines de motifs pour analyser les liens entre la structure du graphe et les attributs propres aux sommets. Ces deux types de motifs sont illustrés au travers de l'analyse de réseaux de co-auteurs issus de DBLP. Cette première partie se termine par la découverte de sous-graphes attribués. Les méthodes proposées sont formalisées dans le cadre de la découverte de sous-groupes et de découverte de modèles exceptionnels.

Dans la seconde partie, j'étends ces domaines de motifs pour extraire des motifs plus intéressants en prenant en compte les connaissances du domaine, les retours de l'utilisateur ainsi que ses connaissances a priori. Je présente une méthode s'appuyant sur des modèles de mobilités issus de la physique statistique pour évaluer le caractère inattendu des trajectoires dans graphes de mobilité. Cela permet de prendre en compte des informations spatiales (i.e., distances entre les sommets, population d'un sommet) pour découvrir des sous-graphes réellement exceptionnels par rapport à ces informations et éliminer certaines trajectoires qui deviennent attendues dès lors que l'on connaît ces informations spatiales. Je présente ensuite une méthode qui vise à prendre en compte les retours de l'utilisateur dans une mesure de qualité biaisée lors de processus interactif de découvertes de motifs. Cette méthode est définie dans le contexte de l'analyse de medias sociaux, plus particulièrement la détection d'événements géo-localisés. S'appuyant

HDR

Soutenu le 14/12/2018

Établissement :
Université Claude Bernard Lyon 1

Contact :
Marc.Plantevit@liris.cnrs.fr

sur les retours de l'utilisateur, les termes ou les zones aimés sont favorisés grâce à une mesure de qualité biaisée. Enfin, je considère le problème de l'intérêt subjectif dans les graphes attribués afin de prendre en compte les apriori de l'utilisateur. Dans ce contexte, un intérêt particulier est donné à l'assimilation des motifs par l'utilisateur. Afin de faciliter cette assimilation, des descriptions alternatives – plus facilement interprétables – des sous-graphes exceptionnels sont construites.

Finalement, je conclus le manuscrit en discutant des perspectives de recherche.

Mots-clés: Graphes attribués, graphes dynamiques, fouille de motifs interactive, découverte de sous-groupes, modèles exceptionnels, intérêt subjectif.

Abstract

In this Habilitation à Diriger des Recherches thesis, I present the main results I have contributed to in the field of pattern mining in augmented graphs.

Graphs are a powerful mathematical abstraction that enables to depict many real world phenomena. Vertices describe entities and edges identify relations between entities. Such graphs are often augmented with additional pieces of information. For instance, the vertices or the edges are enriched with attributes describing them and are called vertex (respectively edge) attributed graphs. Graphs can also be dynamic, i.e., the structure and the values of vertex attributes may evolve through time. The discovery of patterns in such graphs may provide actionable insights and boost the user knowledge.

This manuscript is structured in two parts. In the first part, I discuss the different pattern domains for augmented graphs I contributed to define. This includes the discovery of co-evolution patterns in dynamic attributed graphs, the study of links between the graph structure and the vertex attributes and the discovery of exceptional attributed subgraphs. I first introduce the co-evolution patterns to analyze dynamic attributed graphs as well as interestingness measures to assess these patterns according to each dimension of the graphs (i.e., the dynamics, the vertex attributes organized within a hierarchy or not, and the vertices). Examples of co-evolution patterns in spatio-temporal data are reported. I then present two pattern domains to analyze the links between the vertex attributes and the graph structure. These two types of patterns are illustrated on co-authorship network built from DBLP. This part ends with the discovery of exceptional attributed subgraphs in edge or vertex attributed graphs. The proposed methods are rooted in Subgroup Discovery / Exceptional Model Mining.

In the second part, I discuss how to find pattern of higher interest by taking into account the domain knowledge, user feedback and user's prior knowledge through different contributions. I first present a method that borrows mobility models from statistical physics to assess the unexpectedness of some trajectories in mobility network. This allows the discovery of exceptional attributed subgraphs by taking into account spatial information (i.e., distance between vertices, importance of the vertices in term of population). I then present a method to take the user feedback into biased quality measures in interactive explorations. This method is defined in the context of social media analysis, especially geo-located event detection. Based on user feedback, liked tags and areas are fostered thanks to a biased quality measure. Eventually, I address the problem of subjective interestingness in attributed

graphs to take into account the user's prior knowledge. A particular interest is given to the assimilation of the patterns by the user. To ease this assimilation, alternative descriptions of exceptional attributed subgraphs are provided. Finally, I conclude this thesis by discussing some research perspectives.

Publications

Revue internationale avec comité de lecture

- Anes Bendimerad, Marc Plantevit & Céline Robardet (2018). « Mining exceptional closed patterns in attributed graphs ». *Knowledge and Information Systems (KAIS)*, vol. 56, n°1, pp. 1-25. doi : 10.1007/s10115-017-1109-2. HAL : hal-01625007.
- Mehdi Kaytoue, Marc Plantevit, Albrecht Zimmermann, Anes Bendimerad & Céline Robardet (2017). « Exceptional contextual subgraph mining ». *Machine Learning*, vol. 106, n°08, pp. 1171-1211. doi : 10.1007/s10994-016-5598-0. HAL : hal-01488732.
- Willy Ugarte, Patrice Boizumault, Bruno Crémilleux, Alban Lepailleur, Samir Loudni, Marc Plantevit, Chedy Raïssi & Arnaud Soulet (2017). « Skypattern mining: From pattern condensed representations to dynamic constraint satisfaction problems ». *Artificial Intelligence*, vol. 244, pp. 48-69. doi : 10.1016/j.artint.2015.04.003. HAL : hal-02048224.
- Marc Plantevit, Céline Robardet & Vasile-Marian Scuturici (2016). « Graph Dependency Construction Based on Interval-based Event Dependencies Detection in Data Streams ». *Intelligent Data Analysis*, vol. 202, pp. 223-256. doi : 10.3233/IDA-160803. HAL : hal-01130544.
- Mehdi Kaytoue, Yoann Pitarch, Marc Plantevit & Céline Robardet (2015). « What effects topological changes in dynamic graphs? ». *Social Network Analysis and Mining*, vol. 5, n°55, 55:1-55:17. doi : 10.1007/s13278-015-0294-9. HAL : hal-01221698.
- Peggy Cellier, Thierry Charnois, Marc Plantevit, Christophe Rigotti, Bruno Crémilleux, Olivier Gandrillon, Jiri Klema & Jean-Luc Manguin (2015). « Sequential pattern mining for discovering gene interactions and their contextual information from biomedical texts ». *Journal of Biomedical Semantics*, vol. 6, pp. 1-27. doi : 10.1186/s13326-015-0023-3. HAL : hal-01192959.
- Günce Orman, Vincent Labatut, Marc Plantevit & Jean-François Boulicaut (2015). « Interpreting communities based on the evolution of a dynamic attributed network ». *Social Network Analysis and Mining*, vol. 5, p. 20. doi : 10.1007/s13278-015-0262-4. ArXiv : 1506.04693. HAL : hal-01163778.
- Pierre-Nicolas Mougel, Christophe Rigotti, Marc Plantevit & Olivier Gandrillon (2014). « Finding maximal homogeneous clique sets ». *Knowledge and Information Systems (KAIS)*, vol. 39, n°3, pp. 1-30. doi : 10.1007/s10115-013-0625-y. HAL : hal-00827164.
- Thi Kim Ngan Nguyen, Loic Cerf, Marc Plantevit & Jean-François Boulicaut (2013). « Discovering Descriptive Rules in Relational Dynamic Graphs ». *Intelligent Data Analysis*, vol. 17, pp. 49-69. HAL : hal-01351698.
- Adriana Bechara Prado, Marc Plantevit, Céline Robardet & Jean-François Boulicaut (2013). « Mining Graph Topological Patterns: Finding Co-variations among Vertex Descriptors ». *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, pp. 2090-2104. doi : 10.1109/TKDE.2012.154. HAL : hal-01351727.

- Marc Plantevit, Anne Laurent, Dominique Laurent, Maguelonne Teisseire & Yeow Wei Choong (2010). « Mining Multi-Dimensional and Multi-Level Sequential Patterns ». *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 4, pp. 1-37. HAL : lirmm-00617320.
- Marc Plantevit, Anne Laurent, Dominique Laurent, Maguelonne Teisseire & Yeow Wei Choong (2010). « Mining multidimensional and multilevel sequential patterns ». *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 4, pp. 1-37. doi : 10.1145/1644873.1644877. HAL : hal-01381826.
- Thierry Charnois, Marc Plantevit, Christophe Rigotti & Bruno Crémilleux (2009). « Fouille de données séquentielles pour l'extraction d'information dans les textes ». *Traitement Automatique des Langues*, pp. 59-87. HAL : hal-01011618.

Revue nationale avec comité de lecture

- Yoann Pitarch, Anne Laurent, Marc Plantevit & Pascal Poncelet (2010). « Fenêtres sur Cube ». *Revue des Sciences et Technologies de l'Information - Série ISI : Ingénierie des Systèmes d'Information*, vol. 15, n°1, pp. 9-33. doi : 10.3166/isi.15.1.9-33. HAL : lirmm-00798834.

Conférences internationales avec comité de lecture

- Anes Bendimerad, Ahmad Mel, Jeffrey Lijffijt, Marc Plantevit, Céline Robardet & Tijl De Bie (2018). « Mining Subjectively Interesting Attributed Subgraphs ». *INTERNATIONAL WORKSHOP ON MINING AND LEARNING WITH GRAPHS*, held with SIGKDD 2018, 20 août 2018, Londres (Royaume-Uni). HAL : hal-02060190.
- Maëlle Moranges, Marc Plantevit, Arnaud Fournel, Moustafa Bensafi & Céline Robardet (2018). « Exceptional Attributed Subgraph Mining To Understand The Olfactory Percept ». *21st International Conference on Discovery Science*, 31 octobre 2018, Limassol (Chypre), pp. 276-291. HAL : hal-01878375.
- Adnene Belfodil, Sylvie Cazalens, Philippe Lamarre & Marc Plantevit (2017). « Flash points: Discovering exceptional pairwise behaviors in vote or rating data ». *ECML/PKDD*, 22 septembre 2017, Skopje (Macédoine), pp. 442-458. doi : 10.1007/978-3-319-71246-8_27. HAL : hal-01587041.
- Anes Bendimerad, Rémy Cazabet, Marc Plantevit & Céline Robardet (2017). « Contextual Subgraph Discovery With Mobility Models ». *COMPLEX NETWORKS 2017*, 29 novembre 2017, Lyon (France), pp. 477-489. HAL : hal-01625068.
- Guillaume Bosc, Marc Plantevit, Jean-François Boulicaut, Moustafa Bensafi & Mehdi Kaytoue (2016). « h(odor): Interactive Discovery of Hypotheses on the Structure-Odor Relationship in Neuroscience ». *ECML/PKDD 2016 (Demo)*, 23 septembre 2016, Riva del Garda (Italie), pp. 17-21. HAL : hal-01346679.
- Anes Bendimerad, Marc Plantevit & Céline Robardet (2016). « Unsupervised Exceptional Attributed Sub-graph Mining in Urban Data ». *IEEE International Conference on Data Mining (ICDM 2016)*, 15 décembre 2016, Barcelone (Espagne), pp. 21-30. HAL : hal-01430622.
- Guillaume Bosc, Jérôme Golebiowski, Moustafa Bensafi, Céline Robardet, Marc Plantevit, Jean-François Boulicaut & Mehdi Kaytoue (2016). « Local subgroup discovery for eliciting and understanding new structure-odor relationships ». *Discovery Science: 19th International Conference, DS 2016*, 21 octobre 2016, Bari (Italie), pp. 19-34. doi : 10.1007/978-3-319-46307-0_2. HAL : hal-01346660.

- Albrecht Zimmermann, Mehdi Kaytoue, Marc Plantevit, Céline Robardet & Jean-François Boulicaut (2015). « Profiling users of the Vélo 'v bike sharing system ». *2nd International Workshop on Mining Urban Data (MUD)*, 11 juillet 2015, Lille (France), pp. 63-64. HAL : hal-01193017.
- Houdyer Pierre, Albrecht Zimmermann, Mehdi Kaytoue, Marc Plantevit, Joseph Mitchell & Céline Robardet (2015). « Gazouille: Detecting and Illustrating Local Events from Geolocalized Social Media Streams ». *European Conference on Machine Learning and Knowledge Discovery in Databases*, Porto (Portugal), pp. 276-280. doi : 10.1007/978-3-319-23461-8_29. HAL : hal-01193030.
- Leon Gaillard, Guillaume Ruedin, Stéphanie Giroux-Julien, Marc Plantevit, Mehdi Kaytoue, Syamimi Saadon, Christophe Ménézo & Jean-François Boulicaut (2015). « Data-driven performance evaluation of ventilated photovoltaic double-skin facades in the built environment ». *6th International Building Physics Conference*, 17 juin 2015, Turin (Italie), pp. 447-452. doi : 10.1016/j.egypro.2015.11.694. HAL : hal-01272611.
- Marc Plantevit, Vasile-Marian Scuturici & Céline Robardet (2014). « Temporal Dependency Detection Between Interval-Based Event Sequences ». *New Frontiers in Mining Complex Patterns - Third International Workshop, NFMCP*, Held in Conjunction with ECML-PKDD, Nancy, France, Revised Selected Papers, 19 septembre 2014, Nancy (France), pp. 132-146. HAL : hal-01979495.
- Mehdi Kaytoue, Yoann Pitarch, Marc Plantevit & Céline Robardet (2014). « Triggering Patterns of Topology Changes in Dynamic Graphs ». *The 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 17 août 2014, Beijing, China (Chine), 55:1-55:17. doi : 10.1109/ASONAM.2014.6921577. HAL : hal-01301075.
- Günce Keziban Orman, Vincent Labatut, Marc Plantevit & Jean-François Boulicaut (2014). « A Method for Characterizing Communities in Dynamic Attributed Complex Networks ». *IEEE/ACM International Conference on Advances in Social Network Analysis and Mining (ASONAM)*, 20 août 2014, Pékin (Chine), pp. 481-484. doi : 10.1109/ASONAM.2014.6921629. ArXiv : 1406.6597. HAL : hal-01011913.
- Pierre Holat, Marc Plantevit, Chedy Raïssi, Nadi Tomeh, Thierry Charnois & Bruno Crémilleux (2014). « Sequence Classification Based on Delta-Free Sequential Pattern ». *IEEE International Conference on Data Mining*, 17 décembre 2014, Shenzhen (Chine). HAL : hal-01100929.
- Elise Desmier, Marc Plantevit, Céline Robardet & Jean-François Boulicaut (2014). « Granularity of co-Evolution Patterns in Dynamic Attributed Graphs ». *The Thirteenth International Symposium on Intelligent Data Analysis IDA 2014*, 30 octobre 2014, Leuven (Belgique), pp. 84-95. HAL : hal-01301086.
- Céline Robardet, Vasile-Marian Scuturici, Marc Plantevit & Antoine Fraboulet (2013). « When TEDDY meets Grizzly: Temporal Dependency Discovery for Triggering Road Deicing Operations (Demo) ». *KDD*, 11 août 2013, Chicago, IL (États-Unis), pp. 1490-1493. doi : 10.1145/2487575.2487706. HAL : hal-01339189.
- Elise Desmier, Marc Plantevit, Céline Robardet & Jean-François Boulicaut (2013). « Trend Mining in Dynamic Attributed Graphs ». *Machine Learning and Knowledge Discovery in Databases - European Conference, ECML PKDD 2013*, 23 septembre 2013, Prague (République Tchèque), pp. 654-669. doi : 10.1007/978-3-642-40988-2_42. HAL : hal-01339225.

- Elise Desmier, Marc Plantevit, Céline Robardet & Jean-François Boulicaut (2012). « Cohesive Co-Evolution Patterns in Dynamic Attributed Graphs ». *Discovery Science - 15th International Conference (DS 2012)*, 29 octobre 2012, Lyon (France), pp. 110-124. doi : 10.1007/978-3-642-33492-4_11. HAL : hal-01353051.
- Thi Kim Ngan Nguyen, Marc Plantevit & Jean-François Boulicaut (2012). « Mining Disjunctive Rules in Dynamic Graphs ». *2012 IEEE RIVF International Conference on Computing and Communication Technologies, Research, Innovation, and Vision for the Future (RIVF)*, 27 février 2012, Ho Chi Minh (Viêt Nam), pp. 74-79. doi : 10.1109/rivf.2012.6169829. HAL : hal-01352944.
- Arnaud Soulet, Chedy Raïssi, Marc Plantevit & Bruno Crémilleux (2011). « Mining Dominant Patterns in the Sky ». *The 11th IEEE International Conference on Data Mining - ICDM 2011*, 14 décembre 2011, Vancouver, B.C (Canada). HAL : inria-00623566.
- Arnaud Soulet, Bruno Crémilleux & Marc Plantevit (2011). « Summarizing Contrasts by Recursive Pattern Mining ». *Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference on*, 11 décembre 2011, vancouver (Canada), pp. 1155-1162. doi : 10.1109/ICDMW.2011.161. HAL : hal-01021980.
- Thi Kim Ngan Nguyen, Loïc Cerf, Marc Plantevit & Jean-François Boulicaut (2011). « Multidimensional Association Rules in Boolean Tensors ». *11th SIAM International Conference on Data Mining SDM'11*, 28 avril 2011, Phoenix, Arizona (États-Unis), pp. 570-581. doi : 10.1137/1.9781611972818.49. HAL : hal-01354377.
- Pierre-Nicolas Mougél, Marc Plantevit, Christophe Rigotti, Olivier Gandrillon & Jean-François Boulicaut (2010). « Constraint-Based Mining of Sets of Cliques Sharing Vertex Properties ». *Workshop on Analysis of Complex NETWORKS ACNE'10 co-located with ECML PKDD 2010*, 20 septembre 2010, Barcelona (Espagne), pp. 48-62. HAL : hal-01381539.
- Peggy Cellier, Thierry Charnois, Marc Plantevit & Bruno Crémilleux (2010). « Recursive Sequence Mining to Discover Named Entity Relations ». *9th International Symposium on Intelligent Data Analysis (IDA'10)*, 21 mai 2010, Tucson, Arizona, United States (États-Unis), pp. 30-41. HAL : hal-01016928.
- Thi Kim Ngan Nguyen, Loïc Cerf, Marc Plantevit & Jean-François Boulicaut (2010). « Discovering Inter-Dimensional Rules in Dynamic Graphs ». *Workshop on Dynamic Networks and Knowledge Discovery DyNaK'10 co-located with ECML PKDD 2010*, 24 septembre 2010, Barcelona, Spain (Espagne), pp. 1-12. HAL : hal-01381535.
- Peggy Cellier, Thierry Charnois & Marc Plantevit (2010). « Sequential Patterns to Discover and Characterise Biological Relations ». *11th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing'10)*, 27 mars 2010, Iasi, Romania (Roumanie), pp. 537-548. HAL : hal-01017207.

Conférences nationales avec comité de lecture

- Aimene Belfodil, Mehdi Kaytoue, Céline Robardet, Marc Plantevit & Julien Zarka (2016). « Une méthode de découverte de motifs contextualisés dans les traces de mobilité d'une personne ». *16ème Journées Francophones Extraction et Gestion des Connaissances*, 22 janvier 2016, Reims (France), pp. 63-68. HAL : hal-01265204.

- Guillaume Bosc, Mehdi Kaytoue, Marc Plantevit, Fabien De Marchi, Moustafa Bensafi & Jean-François Boulicaut (2015). « Vers la découverte de modèles exceptionnels locaux : des règles descriptives liant les molécules à leurs odeurs ». *15e Journées Internationales Francophones Extraction et Gestion des Connaissances (EGC 2015)*, 30 janvier 2015, Luxembourg (Luxembourg), pp. 305-316. HAL : hal-01346739.
- Léon Gaillard, Guillaume Ruedin, Stéphanie Giroux-Julien, Marc Plantevit, Mehdi Kaytoue, Syamimi Saadon, Christophe Ménézo & Jean-François Boulicaut (2015). « Évaluation des performances réelles de double-peaux PV ventilées en milieu urbain suivant une approche orientée données ». *2e journées nationales de l'énergie solaire - JNES 2015*, 3 juillet 2015, Perpignan (France). HAL : hal-01479040.
- Léon Gaillard, Guillaume Ruedin, Stéphanie Giroux-Julien, Marc Plantevit, Mehdi Kaytoue, Syamimi Saadon, Christophe Ménézo & Jean-François Boulicaut (2015). « Evaluation des performances réelles de double-peaux PV ventilées en milieu urbain suivant une approche orientée données, ». *XLII Colloque Inter-Universitaire Franco-Québécois*, 10 juin 2015, Sherbrooke (Canada), pp. 01-09. HAL : hal-01478901.
- Leon Gaillard, Guillaume Ruedin, Stéphanie Giroux-Julien, Marc Plantevit, Mehdi Kaytoue, Christophe Menezo & Jean-François Boulicaut (2015). « Fouille de données pour l'analyse du comportement complexe de systèmes photovoltaïque-thermiques en vrai grandeur et in situ intégrés aux bâtiments ». *Congrès français de thermique, la thermique de l'habitat et de la ville*, 29 mai 2015, La Rochelle (France). HAL : hal-01478948.
- Günce Keziban Orman, Vincent Labatut, Marc Plantevit & Jean-François Boulicaut (2014). « Une méthode pour caractériser les communautés des réseaux dynamiques à attributs ». *14ème Conférence Extraction et Gestion des Connaissances (EGC)*, 31 janvier 2014, Rennes (France), pp. 101-112. ArXiv : 1312.4676. HAL : hal-00918181.
- Elise Desmier, Marc Plantevit & Jean-François Boulicaut (2014). « Granularité des motifs de co-variation dans des graphes attribués dynamiques. ». *Extraction et Gestion des Connaissances EGC 2014*, 30 janvier 2014, Rennes (France), pp. 431-442. HAL : hal-01270736.
- Adriana Bechara Prado, Marc Plantevit, Céline Robardet & Jean-François Boulicaut (2012). « Extraction de co-variations entre des propriétés de sommets et leur position topologique dans un graphe attribué. ». *Actes Extraction et Gestion de Connaissances EGC'12*, 31 janvier 2012, Bordeaux (France), pp. 267-278. HAL : hal-01352957.
- Pierre-Nicolas Mougél, Marc Plantevit, Christophe Rigotti, Olivier Gandrillon & Jean-François Boulicaut (2011). « Extraction sous Contraintes d'Ensembles de Cliques Homogènes ». *11eme Conference Francophone sur l'Extraction et la Gestion des Connaissances (EGC'11)*, Brest (France), pp. 1-12. HAL : hal-00744808.
- Marc Plantevit, Chedy Raïssi & Bruno Crémilleux (2011). « Motifs séquentiels δ -libres ». *Extraction et gestion des connaissances (EGC'2011)*, 29 janvier 2011, Brest (France). HAL : hal-00653579.

Éditions scientifique d'Ouvrages (livres, chapitres, colloques, congrès, n° spéciaux)

- Jean-François Boulicaut, Marc Plantevit & Céline Robardet (2016). « Local Pattern Detection in Attributed Graphs ». Solving Large Scale Learning Tasks. Challenges and Algorithms - Essays Dedicated to Katharina Morik on the Occasion of Her 60th Birthday, pp. 168-183. doi : 10.1007/978-3-319-41706-6_8. HAL : hal-02016506.

Autres

- Julien Salotti, Marc Plantevit, Céline Robardet & Jean-François Boulicaut (2012). « Supporting the Discovery of Relevant Topological Patterns in Attributed Graphs ». Demo Session of the IEEE International Conference on Data Mining (IEEE ICDM 12), 10 décembre 2012, Brussels (Belgique). Poster. doi : 10.1109/ICDMW.2012.38. HAL : hal-01353106.
- Pierre-Nicolas Mougél, Marc Plantevit, Christophe Rigotti, Olivier Gandrillon & Jean-François Boulicaut (2010). « A Data Mining Approach to Highlight Relations Between Functional Modules ». IPG (Integrative Post-Genomics), 25 novembre 2010, Lyon (France). Poster. HAL : hal-01381584.

Jury

Christine VERDIER, PR, Université de Grenoble Alpes
Anne LAURENT, PR, Université de Montpellier
Frédérique LAFOREST, PR, Université Jean Monet
Corine CAUVET, PR, Université Aix-Marseille
Youakim BADR, MC, INSA-Lyon
Youssef AMGHAR, PR, INSA-Lyon
Kablan BARBAR, PR, Université Libanaise,

Rapporteuse
Rapporteuse
Examinatrice
Examinatrice
Directeur de thèse
Co-directeur de thèse
Co-directeur de thèse

Résumé

La démocratisation des objets communicants fixes ou mobiles pose de nombreux défis concernant leur intégration dans des processus métiers afin de développer des services intelligents. Dans le contexte de l'Internet des objets, les objets connectés sont des entités hétérogènes et dynamiques qui englobent des fonctionnalités et propriétés cyber-physiques et interagissent via différents protocoles de communication. Pour pallier aux défis d'interopérabilité et d'intégration, il est primordial d'avoir une vue unifiée et logique des différents objets connectés afin de définir un ensemble de langages, outils et architectures permettant leur intégration et manipulation à grande échelle. L'artéfact métier a récemment émergé comme un modèle d'objet (métier) autonome qui encapsule ses données, un ensemble de services, et manipulant ses données ainsi qu'un cycle de vie à base d'états. Le cycle de vie désigne le comportement de l'objet et son évolution à travers ses différents états pour atteindre son objectif métier. La modélisation des objets connectés sous forme d'artéfact métier étendu nous permet de construire un paradigme intuitif pour exprimer facilement des processus d'intégration d'objets connectés dirigés par leurs données. Face aux changements contextuels et à la réutilisation des objets connectés dans différentes applications, les processus dirigés par les données, (appelés aussi « artéfacts » au sens large) restent relativement invariants vu que leurs structures de données ne changent pas. Or, les processus centrés sur les services requièrent souvent des changements dans leurs flux d'exécution. Cette thèse propose un cadre d'intégration de processus centré sur les artéfacts et leur application aux objets connectés. Pour cela, nous avons construit une vue logique unifiée et globale d'artéfact permettant de spécifier, définir et interroger un très grand nombre d'artéfacts distribués, ayant des fonctionnalités similaires (maisons intelligentes ou voitures connectées, ...). Le cadre d'intégration comprend une méthode de modélisation conceptuelle des processus centrés artéfacts, des algorithmes d'appariement inter-artéfacts et une algèbre de définition et de manipulation d'artéfacts. Le langage déclaratif, appelé AQL (Artifact Query Language) permet en particulier d'interroger des flux continus d'artéfacts. Il s'appuie sur une syntaxe de type SQL pour réduire les efforts d'apprentissage. Nous avons également développé un prototype pour valider nos contributions et mener des expériences dans le contexte de l'Internet des objets.

Abstract

The emergence of fixed or mobile communicating objects poses many challenges regarding their integration into business processes in order to develop smart services. In the context of the Internet of Things, connected devices are heterogeneous and dynamic entities that encompass cyber-physical features and properties and interact through different communication protocols. To overcome the challenges related to interoperability and integration, it is essential to build a unified and logical view of different

Thèse

Soutenue le 09/07/2018

Établissement :
INSA Lyon

Contact :
Youakim.Badr@iris.cnrs.fr

Suivi visuel d'objets dans un réseau de caméras intelligentes : application aux systèmes de manutention automatisés

Adel Benamara

connected devices in order to define a set of languages, tools and architectures allowing their integrations and manipulations at a large scale. Business artifact has recently emerged as an autonomous (business) object model that encapsulates attribute-value pairs, a set of services manipulating its attribute data, and a state-based lifecycle. The lifecycle represents the behavior of the object and its evolution through its different states in order to achieve its business objective. Modeling connected devices and smart objects as an extended business artifact allows us to build an intuitive paradigm to easily express integration data-driven processes of connected objects. In order to handle contextual changes and reusability of connected devices in different applications, data-driven processes (or artifact processes in the broad sense) remain relatively invariant as their data structures do not change. However, service-centric or activity-based processes often require changes in their execution flows. This thesis proposes a framework for integrating artifact-centric processes and their application to connected devices. To this end, we introduce a logical and unified view of a «global» artifact allowing the specification, definition and interrogation of a very large number of distributed artifacts, with similar functionalities (smart homes or connected cars, ...). The framework includes a conceptual modeling method for artifact-centric processes, inter-artifact mapping algorithms, and artifact definition and manipulation algebra. A declarative language, called AQL (Artifact Query Language) aims in particular to query continuous streams of artifacts. The AQL relies on a syntax similar to the SQL in relational databases in order to reduce its learning curve. We have also developed a prototype to validate our contributions and conducted experimentations in the context of the Internet of Things.

Publications

Revue internationale avec comité de lecture

- Maroun Abi Assaf, Youakim Badr, Hicham El Khoury & Kablan Barbar (2018). « Generating Database Schemas from Business Artifact Models ». *International Journal of Information Technology and Computer Science(IJITCS)*, vol. 10, n°2, pp. 10-17. doi : 10.5815/ijitcs.2018.02.02. HAL : hal-01709001.

Conférences internationales avec comité de lecture

- Maroun Abi Assaf, Youakim Badr & Amghar Youssef (2017). « A Continuous Query Language for Stream-based Artifacts ». 28th International Conference on Database and Expert Systems Applications (DEXA), 31 août 2017, Lyon (France), pp 80-89. HAL : hal-01575959.
- Maroun Abi Assaf, Youakim Badr, Kablan Barbar & Youssef Amghar (2016). « AQL: A Declarative Artifact Query Language ». *Advances in Databases and Information Systems - 20th East European Conference, ADBIS 2016*, 28 août 2016, Prague (République Tchèque), pp. 119-133. doi : 10.1007/978-3-319-44039-2_9. HAL : hal-01407850.
- Maroun Abi Assaf (2016). « Towards an Integration System for Artifact-centric Processes ». *SIGMOD PhD Symposium 2016*, 26 juin 2016, San Francisco (États-Unis). doi : 10.1145/2926693.2929904. HAL : hal-01407818.
- Maroun Abi Assaf, Youakim Badr, Kablan Barbar & Youssef Amghar (2015). « On the Integration of Artifact Lifecycles ». *The 7th International Conference on Management of Computational and Collective Intelligence in Digital EcoSystems: MEDES'15*, 29 octobre 2015, Caraguatubá, Sao Paulo (Brésil), pp. 59-63. doi : 10.1145/2857218.2857228. HAL : hal-01212857.

Jury

Jenny BENOIS-PINNEAU, PR, Université de Bordeaux,
Thierry CHATEAU, PR, Université de Clermont Auvergne,
Catherine ACHARD, MC, HDR, Université UPMC Paris 6,
François BRÉMOND, DR, INRIA Sophia Antipolis,
Serge MIGUET, PR, Université Lumière Lyon 2,
Mihaela SCUTURICI, MC, Université Lumière Lyon 2,
David ZAK, Directeur, Fives CortX,
Matthieu MILER, CR, Fives Intralogistics,

Rapporteur
Rapporteur
Examinatrice
Examineur
Directeur de thèse
Co-directrice
Invité
Invité

Résumé

Cette thèse se place dans le cadre de l'intralogistique (ou logistique interne), un domaine qui s'intéresse au traitement et à l'optimisation des flux physiques au sein des entrepôts, centres de distribution et usines. Les systèmes de manutention automatisés sont au cœur de la logistique interne de plusieurs industries comme le commerce en ligne, la messagerie postale, la grande distribution, l'industrie manufacturière, le transport aéroportuaire, etc. Ces équipements composés de lignes de convoyage haute cadence permettent un transport sûr et fiable d'un volume considérable de biens et de marchandises tout en réduisant les coûts.

L'automatisation de l'acheminement des flux physiques par les systèmes de manutention repose sur l'identification et le suivi en temps réel des charges transportées. Dans cette thèse, nous explorons une solution de suivi qui emploie un réseau de caméras intelligentes à champs recouvrants. L'objectif final est de fournir l'information de suivi sur les charges transportées pour le pilotage d'un système de manutention.

Le suivi d'objets est un problème fondamental de la vision par ordinateur qui a de nombreuses applications comme la vidéosurveillance, la robotique, les voitures autonomes, etc. Nous avons intégré plusieurs briques de base traditionnellement appliquées aux scènes de surveillance automobile ou de surveillance des activités humaines pour constituer une chaîne de suivi de référence. Cette chaîne d'analyse vidéo étalon nous a permis de caractériser des hypothèses propres au convoyage d'objets. Nous proposons d'incorporer cette connaissance métier dans la chaîne de suivi pour en améliorer les performances. Nous avons notamment pris en compte, dans l'étape de segmentation des images, le fait que les objets doivent pouvoir s'arrêter sans pour autant être intégrés aux modèles d'arrière-plan. Nous avons également exploité la régularité des trajectoires des objets convoyés dans les installations, ce qui nous a permis d'améliorer les modèles prédictifs de la position et de la vitesse des objets, dans les étapes de suivi. Enfin, pour réidentifier les objets dans les situations ambiguës, comme par exemple celle des objets proches les uns des autres, nous avons intégré des contraintes de stricte monotonie dans l'ordre des colis sur le convoyeur, contraintes qui n'existent pas dans les scènes généralistes.

Nous nous sommes par ailleurs attelés à un problème pratique d'optimisation des performances sur l'architecture multi-cœurs couplée aux caméras intelligentes. Dans ce cadre, nous avons mis en place un apprentissage dynamique de la zone de l'image contenant le convoyeur. Cela nous a permis de limiter la mise à jour du modèle de fond à cette seule zone d'intérêt. Nous avons, par la suite, proposé une stratégie de parallélisation qui partitionne de manière adaptative cette région d'intérêt de l'image, afin d'équilibrer au

Thèse

Soutenue le 19/12/2018

Établissement :

Université Lumière Lyon2

Contact :

Serge.Miguet@iris.cnrs.fr

mieux la charge de travail entre les différents cœurs de l'architecture des caméras intelligentes.

Nous avons également traité la problématique du suivi sur plusieurs caméras. Nous avons proposé une approche basée sur un système de composition d'événements. Cette approche nous a permis de fusionner les données de suivi local pour former les trajectoires globales des colis, tout en intégrant des informations issues du processus métier, par exemple la saisie de l'information de destination par des opérateurs sur un terminal avant la dépose des colis.

Nous avons validé cette approche sur un système de manutention mis en place dans un centre de tri postal de grande envergure. Le réseau de caméras déployé est composé de 32 caméras qui assurent le suivi de plus de 400.000 colis/jour sur des lignes de dépose. Le taux d'erreur du suivi obtenu est inférieur à 1 colis sur 1000 (0,1%).

Mots clés : Intra-logistique, systèmes de manutention automatisée, suivi d'objets, soustraction de fond, caméras intelligentes, réseau de caméras.

Abstract

Intra-logistics (or internal logistics) focuses on the management and optimization of internal production and distribution processes within warehouses, distribution centers, and factories. Automated handling systems play a crucial role in the internal logistics of several industries such as e-commerce, postal messaging, retail, manufacturing, airport transport, etc. These systems are composed by multiple high-speed conveyor lines that provide safe and reliable transportation of a large volume of goods and merchandises while reducing costs.

The automation of the conveying process relies on the identification and the real-time tracking of the transported objects. In this thesis, we designed a tracking system that employs a network of smart cameras with an overlapping field of view. The goal is to provide a real-time tracking information about conveyed objects to control the automated handling system.

Multiple object tracking is a fundamental problem of computer vision that has many applications such as video surveillance, robotics, autonomous cars, etc. We integrated several building blocks traditionally applied to traffic surveillance or human activities monitoring to constitute a tracking pipeline. We used this baseline tracking pipeline to characterize contextual scene information proper to the conveying scenario. We integrated this contextual information to the tracking pipeline to enhance its performance. In particular, we took into account the state of moving objects that become stationary in the background subtraction step to prevent their absorption to the background model. We have also exploited the regularity of objects' trajectory to enhance the motion model associated with the tracked objects. Finally, we integrated the precedence ordering constraint among the conveyed object to reidentify them when they are close to each other.

We have also tackled practical problems related to the optimization of the execution of the proposed tracking pipeline in the multi-core architectures of smart cameras. In particular, we proposed a dynamic learning process that extracts the region of the image that corresponds to the conveyor lines. We reduced the number of the processed pixels by restricting the processing to this region of interest. We also proposed a parallelization strategy that adaptively partitions this region of interest of the image, in order to balance

the workload between the different cores of the smart cameras.

Finally, we proposed a multiple cameras tracking algorithm based on event composition. This approach fuses the local tracking generated by the smart cameras to form global object trajectories. It also includes information from third-party systems such as the object destination entered by human operators.

We deployed and validated the proposed approach for the control of a sorting system in a postal distribution warehouse. A network of 32 cameras tracks more than 400.000 parcel/day on the injections lines feeding the sorting system. The tracking error rate is less than 1 parcel in 1000 (0.1%).

Keywords: Intra-logistics, automated handling systems, multiple object tracking, background subtraction, smart cameras, cameras network.

Détection d'anomalies dans les séries temporelles : application aux masses de données sur les pneumatiques

Seif Eddine
Benkabou

Jury

Younes BENNANI, PR, Université Paris 13
Yann GUERMEUR, PR, LORIA-Nancy
Salima BENBERNOU, PR, Université Paris 5
Lydia BOUDJELOUD-ASSALA, PR, Université de Lorraine
Marc SEBBAN, PR, Université de Saint-Etienne
Khalid BENABDESLEM, PR, Université Lyon1
Bruno CANITIA, PR, LIZEO ONLINE MEDIA GROUP,

Rapporteur
Rapporteur
Examinatrice
Examinatrice
Examinateur
Directeur de thèse
Invité

Résumé

La détection d'anomalies est une tâche cruciale qui a suscité l'intérêt de plusieurs travaux de recherche dans les communautés d'apprentissage automatique et fouille de données. La complexité de cette tâche dépend de la nature des données, de la disponibilité de leur étiquetage et du cadre applicatif dont elles s'inscrivent. Dans le cadre de cette thèse, nous nous intéressons à cette problématique pour les données complexes et particulièrement pour les séries temporelles uni et multi-variées. Le terme «anomalie» peut désigner une observation qui s'écarte des autres observations au point d'éveiller des soupçons. De façon plus générale, la problématique sous-jacente (aussi appelée détection de nouveautés ou détection des valeurs aberrantes) vise à identifier, dans un ensemble de données, celles qui diffèrent significativement des autres, qui ne se conforment pas à un «comportement attendu» (à définir ou à apprendre automatiquement), et qui indiquent un processus de génération différent. Les motifs «anormaux» ainsi détectés se traduisent souvent par de l'information critique. Nous nous focalisons plus précisément sur deux aspects particuliers de la détection d'anomalies à partir de séries temporelles dans un mode non-supervisé. Le premier est global et consiste à ressortir des séries relativement anormales par rapport une base entière. Le second est dit contextuel et vise à détecter localement, les points anormaux par rapport à la structure de la série étudiée. Pour ce faire, nous proposons des approches d'optimisation à base de clustering pondéré et de déformation temporelle pour la détection globale ; et des mécanismes à base de modélisation matricielle pour la détection contextuelle. Enfin, nous présentons une série d'études empiriques sur des données publiques pour valider les approches proposées et les comparer avec d'autres approches connues dans la littérature. De plus, une validation expérimentale est fournie sur un problème réel, concernant la détection de séries de prix aberrants sur les pneumatiques, pour répondre aux besoins exprimés par le partenaire industriel de cette thèse

Abstract

Anomaly detection is a crucial task that has attracted the interest of several research studies in machine learning and data mining communities. The complexity of this task depends on the nature of the data, the availability of their labeling and the application framework on which they depend. As part of this thesis, we address this problem for complex data and particularly for uni and multivariate time series. The term «anomaly» can refer to an observation that deviates from other observations so as to arouse suspicion that it was generated by a different generation process. More generally, the underlying problem (also called novelty detection or outlier detection) aims to identify, in a set of data, those which differ significantly from others, which do not conform to an «expected behavior» (which could be defined or learned), and which indicate a different mechanism. The «abnormal» patterns

Thèse

Soutenu le 21/03/2018

Établissement :
Université Claude Bernard Lyon 1

Contact :
Karim.Benabdeslem@iris.cnrs.fr

Sarah Bertrand

thus detected often result in critical information. We focus specifically on two particular aspects of anomaly detection from time series in an unsupervised fashion. The first is global and consists in detecting abnormal time series compared to an entire database, whereas the second one is called contextual and aims to detect locally, the abnormal points with respect to the global structure of the relevant time series. To this end, we propose an optimization approaches based on weighted clustering and the warping time for global detection ; and matrix-based modeling for the contextual detection. Finally, we present several empirical studies on public data to validate the proposed approaches and compare them with other known approaches in the literature. In addition, an experimental validation is provided on a real problem, concerning the detection of outlier price time series on the tyre data, to meet the needs expressed by, LIZEO, the industrial partner of this thesis

Publications

Revue internationale avec comité de lecture

- Seif-Eddine Benkabou, Khalid Benabdeslem & Canitia Bruno (2017). « Unsupervised Outlier Detection for Time Series by Entropy and Dynamic Time Warping ». Knowledge and Information Systems (KAIS). doi : 10.1007/s10115-017-1067-8. HAL : hal-01528252.

Conférences internationales avec comité de lecture

- Vivien Kraus, Seif-Eddine Benkabou, Khalid Benabdeslem & Frederic Cherqui (2018). « An improved Laplacian semi-supervised regression ». 30th International Conference on Tools with Artificial Intelligence, 7 novembre 2018, Volos (Grèce). HAL : hal-01954836.
- Seif-Eddine Benkabou, Khalid Benabdeslem & Bruno Canitia (2017). « L2-Type Regularization-based Unsupervised Anomaly Detection from Temporal Data ». IEEE IJCNN, 14 mai 2017, Alaska (États-Unis). HAL : hal-01458412.
- Seif-Eddine Benkabou, Khalid Benabdeslem & Canitia Bruno (2017). « Local-to-Global unsupervised anomaly detection from temporal data ». PAKDD, 23 mai 2017, Jeju (Corée du Sud). HAL : hal-01458411.
- Seif-Eddine Benkabou, Khalid Benabdeslem & Canitia Bruno (2016). « Entropy-based clustering for anomaly detection from time-series data ». ICML Workshop on Anomaly detection, New York (États-Unis). HAL : hal-01340978.

Conférences nationales avec comité de lecture

- Seif-Eddine Benkabou, Khalid Benabdeslem & Canitia Bruno (2017). « Régularisation Ridge pour la détection non-supervisée à partir de séries temporelles ». SFC: Société Francophone de Classification, 30 juin 2017, Lyon (France). HAL : hal-01581660.
- Seif-Eddine Benkabou, Khalid Benabdeslem & Canitia Bruno (2016). « Une approche embedded pour la détection de nouveautés à partir de séries temporelles ». Société Francophone de Classification (SFC), Marrakech (Maroc). HAL : hal-01340981.
- Seif-Eddine Benkabou, Khalid Benabdeslem & Canitia Bruno (2015). « Une approche à deux niveaux séquentiels pour la détection de nouveautés à partir de séries temporelles ». 22ème Rencontres de la Société Francophone de Classification, 9 septembre 2015, Nantes (France), pp. 129-132. HAL : hal-01198439.

Jury

Laure TOUGNE, PR, Université Lumière Lyon 2
Guillaume CERUTTI, Docteur, IRNA
Anne VERROUST-BLONDET, CR, INRIA
Pascal DESBARATS, PR, Université de Bordeaux
Christophe DUCOTTET, PR, Université Jean Monnet
Jasmine BURGUET, CR, INRA,
David ROUSSEAU, PR, Université d'Angers,

Directrice de thèse
Co-directeur de thèse
Rapportrice
Rapporteur
Rapporteur
Examinatrice
Examineur

Résumé

Cette thèse s'inscrit dans le cadre de l'ANR ReVeRIES dont l'objectif est d'utiliser les technologies mobiles pour aider l'homme à mieux connaître son environnement et notamment les végétaux qui l'entourent. Plus précisément, le projet ReVeRIES s'appuie sur une application mobile, nommée Folia développée dans le cadre du projet ANR ReVeS, capable de reconnaître les espèces d'arbres et arbustes à partir de photos de leurs feuilles. Ce prototype se différencie des autres outils car il est capable de simuler le comportement du botaniste. Dans le contexte du projet ReVeRIES, nous nous proposons d'aller beaucoup plus loin en développant de nouveaux aspects : la reconnaissance multimodale d'espèces, l'apprentissage par le jeu et les sciences citoyennes. L'objet de cette thèse porte sur le premier de ces trois aspects, à savoir l'analyse d'images d'organes de végétaux en vue de l'identification. Plus précisément, nous considérons les principaux arbres et arbustes, endémiques ou exotiques, que l'on trouve en France métropolitaine. L'objectif de cette thèse est d'étendre l'algorithme de reconnaissance en prenant en compte d'autres organes que la feuille. Cette multi-modalité est en effet essentielle si nous souhaitons que l'utilisateur apprenne et s'entraîne aux différentes méthodes de reconnaissance, pour lesquelles les botanistes utilisent la variété des organes (i.e. les feuilles, les fleurs, les fruits et les écorces). La méthode utilisée par Folia pour la reconnaissance des feuilles étant dédiée, car simulant le botaniste, ne peut s'appliquer directement aux autres organes. Ainsi, de nouveaux verrous se posent, tant au niveau du traitement des images qu'au niveau de la fusion de données. Une première partie de la thèse a été consacrée à la mise en place de méthodes de traitement d'images pour l'identification des espèces végétales. C'est l'identification des espèces d'arbres à partir d'images d'écorces qui a été étudiée en premier. Les descripteurs développés prennent en compte la structure de l'écorce en s'inspirant des critères utilisés par les botanistes. Les fruits et les fleurs ont nécessité une étape de segmentation avant leur description. Une nouvelle méthode de segmentation réalisable sur smartphone a été développée pour fonctionner sur la grande variabilité des fleurs et des fruits. Enfin, des descripteurs ont été extraits sur les fruits et les fleurs après l'étape de segmentation. Nous avons décidé de ne pas faire de séparation entre les fleurs et les fruits car nous avons montré qu'un utilisateur novice en botanique ne sait pas toujours faire la différence entre ces deux organes sur des arbres dits « d'ornement » (non fruitiers). Pour les fruits et les fleurs, la prédiction n'est pas seulement faite sur les espèces mais aussi sur les genres et les familles, groupes botaniques traduisant d'une similarité entre ces organes. Une deuxième partie de la thèse traite de la combinaison des descripteurs des différents organes que sont les feuilles, les écorces, les fruits et les fleurs. En plus des méthodes de combinaison basiques, nous proposons de prendre en compte la confusion entre les espèces, ainsi que les prédictions d'appartenance aux taxons botaniques supérieurs à l'espèce. Enfin, un chapitre d'ouverture est consacré au traitement de ces images par des réseaux de neurones à convolutions. En effet, le Deep-Learning est

Thèse

Soutenue le 10/12/2018

Établissement :

Université Lumière Lyon2

Contact :

Laure.Tougne@liris.cnrs.fr

de plus en plus utilisé en traitement d'images, notamment appliqué aux organes végétaux. Nous proposons dans ce contexte de visualiser les filtres de convolution extrayant de l'information, afin de faire le lien entre les informations extraites par ces réseaux et les éléments botaniques.

Abstract

This thesis is part of the ANR ReVeRIES, which aims to use mobile technologies to help people better understand their environment and in particular the plants that surround them. More precisely, the ReVeRIES project is based on a mobile application called Folia developed as part of the ANR ReVeS project and capable of recognising tree and shrub species based on photos of their leaves. This prototype differs from other tools in that it is able to simulate the behavior of the botanist. In the context of the ReVeRIES project, we propose to go much further by developing new aspects: multimodal species recognition, learning through play and citizen science. The purpose of this thesis is to focus on the first of these three aspects, namely the analysis of images of plant organs for identification. More precisely, we consider the main trees and shrubs, endemic or exotic, found in metropolitan France. The objective of this thesis is to extend the recognition algorithm by taking into account other organs in addition to the leaf. This multi-modality is indeed essential if we want the user to learn and practice the different methods of recognition for which botanists use the variety of organs (i.e. leaves, flowers, fruits and bark). The method used by Folia for leaf recognition being dedicated, because simulating the work of a botanist on the leaf, cannot be applied directly to other organs. Thus, new challenges are emerging, both in terms of image processing and data fusion. The first part of the thesis was devoted to the implementation of image processing methods for the identification of plant species. The identification of tree species from bark images was the first to be studied. The descriptors developed take into account the structure of the bark inspired from the criteria used by botanists. Fruits and flowers required a segmentation step before their description. A new segmentation method that can be used on smartphones has been developed to work in spite of the high variability of flowers and fruits. Finally, descriptors were extracted on fruits and flowers after the segmentation step. We decided not to separate flowers and fruits because we showed that a user new to botany does not always know the difference between these two organs on so-called «ornamental» trees (not fruit trees). For fruits and flowers, prediction is not only made on their species but also on their genus and family, botanical groups reflecting a similarity between these organs. The second part of the thesis deals with the combination of descriptors of the different organs: leaves, bark, fruits and flowers. In addition to basic combination methods, we propose to consider the confusion between species, as well as predictions of affiliations in botanical taxa higher than the species. Finally, an opening chapter is devoted to the processing of these images by convolutional neural networks. Indeed, Deep Learning is increasingly used in image processing, particularly for plant organs. In this context, we propose to visualize the learned convolution filters extracting information, in order to make the link between the information extracted by these networks and botanical elements.

Publications

Revue internationale avec comité de lecture

- Sarah Bertrand, Rihab Ben Ameer, Guillaume Cerutti, Didier Coquin, Lionel Valet & Laure Tougne (2018). « Bark and Leaf Fusion Systems to Improve Automatic Tree Species Recognition ». *Ecological Informatics*. doi : 10.1016/j.ecoinf.2018.05.007. HAL : hal-01811039.

Conférences internationales avec comité de lecture

- Sarah Bertrand, Guillaume Cerutti & Laure Tougne (2017). « Segmentation Algorithm on Smartphone Dual Camera: Application to Plant Organs in the Wild ». 10th International Conference on Machine Vision, 15 novembre 2017, Vienna (Autriche). HAL : hal-01675655.
- Sarah Bertrand, Guillaume Cerutti & Laure Tougne (2017). « Bark Recognition to Improve Leaf-based Classification in Didactic Tree Species Identification ». VISAPP 2017 - 12th International Conference on Computer Vision Theory and Applications, 1 mars 2017, Porto (Portugal). HAL : hal-01486591.

Autres

- Sarah Bertrand, Guillaume Cerutti & Laure Tougne (2018). « Visualization of Leaf Botanical Features Extracted from AlexNet Convolutional Layers ». IAMPS - International Workshop on Image Analysis Methods for the Plant Sciences 2018, 23 janvier 2018, Nottingham (Royaume-Uni). Poster. HAL : hal-01691924.

Thomas Caissard

Jury

Isabelle DEBLED-RENNESSON, PR, Université de Lorraine, rapporteure
Nicolas PASSAT, PR, Université de Reims Champagne-Ardenne, rapporteur
Raphaëlle CHAINE, PR, Université Lyon 1, examinatrice
David COHEN-STEINER, CR, INRIA Sophia Antipolis, examinateur
David COEURJOLLY, DR, CNRS, directeur de thèse
Jacques-Olivier LACHAUD, PR, Univ. Savoie Mont Blanc, co-directeur de thèse
Tristan ROUSSILLON, MC, INSA Lyon, co-directeur de thèse

Résumé

La question centrale de cette thèse est le développement d'un opérateur Laplace-Beltrami discret sur des surfaces numériques. Ces surfaces proviennent de la théorie de la géométrie discrète, c'est-à-dire de la géométrie qui se concentre sur des sous-ensembles d'entiers relatifs. Nous nous situons ici dans un cadre théorique où les surfaces digitales sont le résultat d'un processus d'approximation, ou de discrétisation, d'une surface lisse sous-jacente. Cette méthode permet à la fois de prouver des théorèmes de convergence de quantités discrètes vers des quantités continues, mais aussi, par des analyses numériques, de confirmer expérimentalement ces résultats. Pour la discrétisation de l'opérateur, nous rencontrons deux problèmes : d'une part, notre surface n'est qu'une approximation de la surface continue sous-jacente, et d'autre part, l'estimation triviale des grandeurs géométriques sur la surface numérique ne nous donne généralement pas une bonne estimation de cette grandeur. Nous avons déjà des réponses au deuxième problème : ces dernières années, de nombreux articles se sont concentrés sur le développement de méthodes d'approximation de certaines grandeurs géométriques sur des surfaces digitales (comme les normales ou la courbure), méthodes que nous allons décrire dans cette thèse. Ces nouvelles techniques d'approximation nous permettent d'injecter des informations de mesure dans les éléments de notre surface. Nous utilisons donc l'estimation des normales pour répondre au premier problème, ce qui nous permet en fait d'approximer avec précision le plan tangent en un point de la surface et, par une méthode d'intégration, de surmonter les problèmes topologiques liés à la surface discrète. Nous présentons un résultat théorique de convergence du nouvel opérateur discrétisé, puis nous illustrons ses propriétés par une analyse numérique. Nous effectuons une comparaison détaillée du nouvel opérateur avec ceux de la littérature adaptée sur les surfaces digitales, ce qui permet, au moins pour la convergence, de montrer que seul notre opérateur possède cette propriété. Nous illustrons également l'opérateur via certaines de ces applications telles que sa décomposition spectrale ou le flux de courbure moyenne.

Abstract

The central issue of this thesis is the development of a discrete Laplace-Beltrami operator on digital surfaces. These surfaces come from the theory of discrete geometry, i.e. geometry that focuses on subsets of relative integers. We place ourselves here in a theoretical framework where digital surfaces are the result of an approximation, or discretization process, of an underlying smooth surface. This method makes it possible both to prove theorems of convergence of discrete quantities towards continuous quantities, but also, through numerical analyses, to experimentally confirm these results. For the discretization of the operator, we face two problems: on the one hand, our surface is only an approximation of the underlying continuous surface,

Thèse

Soutenue le 13/12/2018

Établissement :
Université Claude Bernard Lyon 1

Contact :
David.Coeurjolly@liris.cnrs.fr

and on the other hand, the trivial estimation of geometric quantities on the digital surface does not generally give us a good estimate of this quantity. We already have answers to the second problem: in recent years, many articles have focused on developing methods to approximate certain geometric quantities on digital surfaces (such as normals or curvature), methods that we will describe in this thesis. These new approximation techniques allow us to inject measurement information into the elements of our surface. We therefore use the estimation of normals to answer the first problem, which in fact allows us to accurately approximate the tangent plane at a point on the surface and, through an integration method, to overcome topological problems related to the discrete surface. We present a theoretical convergence result of the discretized new operator, then we illustrate its properties using a numerical analysis of it. We carry out a detailed comparison of the new operator with those in the literature adapted on digital surfaces, which allows, at least for convergence, to show that only our operator has this property. We also illustrate the operator via some of these applications such as its spectral decomposition or the mean curvature flow.

Publications

Conférences internationales avec comité de lecture

- Thomas Caissard, David Coeurjolly, Jacques-Olivier Lachaud & Tristan Roussillon (2017). « Heat kernel Laplace-Beltrami operator on digital surfaces ». 20th International Conference on Discrete Geometry for Computer Imagery, 21 septembre 2017, Vienna (Autriche), pp. 241-253. HAL : hal-01575544.

Conférences nationales avec comité de lecture

- Thomas Caissard, David Coeurjolly, Tristan Roussillon & Jacques-Olivier Lachaud (2016). « Laplace-Beltrami operator on Digital Curves ». JFIG, 2 décembre 2016, Grenoble (France). HAL : hal-01497255.

Autres

- Thomas Caissard, David Coeurjolly, Jacques-Olivier Lachaud & Tristan Roussillon (2016). « Heat kernel Laplace-Beltrami operator on digital surfaces ». HAL : hal-01498293.

Jury

Thi-Bich-Hanh DAO, MC, Université d'Orléans,	Rapporteur
Christian BESSIERE, DR, CNRS,	Rapporteur
Élisa FROMONT, PR, Universités de Rennes 1,	Examinatrice
Valérie BOTTA-GENOULAZ, PR, INSA-LYON,	Examinatrice
Christian SCHULTE, PR, KTH Royal Institute of Technology,	Examinateur
Christine SOLNON, PR, INSA-LYON,	Directrice de thèse
Amélie CORDIER, MC, Hoomano,	Co-directrice
Pierre-Antoine CHAMPIN, MC, Univ. Claude Bernard Lyon 1,	Co-directeur
Justinian OPRESCU, Docteur, Infologic,	Invité

Résumé

Les ERP (Enterprise Resource Planning) sont incontournables dans les systèmes d'information des sociétés industrielles: ils jouent un rôle crucial pour automatiser et suivre leurs processus afin d'améliorer leur compétitivité. Un ERP est un logiciel générique qui est utilisé par plusieurs sociétés industrielles ayant des besoins et des processus différents. C'est pourquoi de nombreux paramètres permettent d'adapter le fonctionnement du système aux besoins d'une société. Le déploiement d'un ERP, qui vise à paramétrer le système en fonction des besoins collectés, est donc une tâche complexe qui requiert une profonde expertise du système mais aussi du métier de l'entreprise industrielle. Infologic est une société qui développe et installe son propre ERP appelé Copilote. La difficulté liée au déploiement de Copilote dans une société industrielle est un réel frein pour la croissance d'Infologic et réduire la complexité du paramétrage de Copilote est un enjeu vital pour Infologic. C'est pourquoi nous avons étudié le processus de déploiement de Copilote et particulièrement la phase de paramétrage du système. Nous proposons une approche visant à extraire, depuis l'ensemble des paramétrages existants, un catalogue de paramétrages correspondant à des besoins fonctionnels précédemment rencontrés afin de les réutiliser lors des prochains déploiements de Copilote. Nous proposons d'utiliser la programmation par contraintes pour cela, afin de pouvoir facilement personnaliser les solutions calculées en ajoutant des contraintes et des critères d'optimisation variés. Nous introduisons de nouveaux modèles à base de contraintes pour résoudre des problèmes de clustering conceptuel, ainsi qu'une contrainte globale pour le problème de couverture exacte avec plusieurs algorithmes de propagation. Nous montrons qu'elle permet de modéliser facilement des problèmes de clustering conceptuel, et de les résoudre plus efficacement que les approches déclaratives de l'état de l'art.

Abstract

Enterprise Resource Planning (ERP) systems are essential for industrial companies to automatize and monitor their business processes in order to boost their competitiveness. ERP systems are generic software designed to serve a large variety of companies with different business processes. Therefore, they have many configuration options to support various business processes used in different companies. The implementation process of an ERP system consists in assigning values to ERP parameters according to the company requirements: It determines the exact operations and processes supported by the system in the specific company. Infologic is a French company that develops and integrates their own ERP system called Copilote. It has thousands of parameters that are used to adapt it as precisely as possible to customer requirements. However, this flexibility makes the

Soutenue le 18/12/2018

Établissement :
INSA Lyon

Contact :
Christine.Solnon@liris.cnrs.fr

implementation of Copilote a time consuming task that requires a deep knowledge of its functionalities and parameters. Reducing the complexity of the implementation of Copilote is a critical issue for Infologic who needs to integrate efficiently new system integrators to meet the demand of new customers. In this thesis, we study the implementation process of Copilote in order to understand the main issues encountered by Infologic. We propose a new approach for extracting a catalog of configuration parts from existing configurations of Copilote, and each configuration part is associated with the business requirement it fulfills in order to reuse it for next implementations of Copilote. To this aim, we propose to use constraint programming (CP) to easily integrate feedbacks of experts by means of new constraints or criteria. We introduce new CP models to solve conceptual clustering problems and a new global constraint for the exact cover problem with several propagation algorithms. We show it allows to model easily conceptual clustering problems and to solve it more efficiently than existing declarative approaches.

Jury

Jenny BENOIS-PINEAU, PR, Université de Bordeaux,	Rapporteur
François BREMOND, DR, INRIA Sophia Antipolis,	Rapporteur
Nicolas THOME, PR, CNAM,	Examinateur
Cathrine ACHARD, MC, Sorbonne Université,	Examinatrice
Jean-Yves DUFOR, Docteur Ingénieur, Thales ThereSIS Lab,	Examinateur
Atila BASKURT, PR, INSA-LYON,	Directeur de thèse
Stefan DUFFNER, MC, INSA-LYON,	Co-directeur de thèse

Résumé

La vidéo surveillance est d'une grande valeur pour la sécurité publique. En tant que l'un des plus importantes applications de vidéo surveillance, la ré-identification de personnes est définie comme le problème de l'identification d'individus dans des images captées par différentes caméras de surveillance à champs non-recouvrants. Cependant, cette tâche est difficile à cause d'une série de défis liés à l'apparence de la personne, tels que les variations de poses, de point de vue et de l'éclairage etc. Pour régler ces différents problèmes, dans cette thèse, nous proposons plusieurs approches basées sur l'apprentissage profond de sorte d'améliorer de différentes manières la performance de ré-identification. Dans la première approche, nous utilisons les attributs des piétons tels que genre, accessoires et vêtements. Nous proposons un système basé sur un réseau de neurones à convolution(CNN) qui est composé de deux branches : une pour la classification d'identité et l'autre pour la reconnaissance d'attributs. Nous fusionnons ensuite ces deux branches pour la ré-identification. Deuxièmement, nous proposons un CNN prenant en compte différentes orientations du corps humain. Le système fait une estimation de l'orientation et, de plus, combine les caractéristiques de différentes orientations extraites pour être plus robuste au changement de point de vue. Comme troisième contribution de cette thèse, nous proposons une nouvelle fonction de coût basée sur une liste d'exemples. Elle introduit une pondération basée sur le désordre du classement et permet d'optimiser directement les mesures d'évaluation. Enfin, pour un groupe de personnes, nous proposons d'extraire une représentation de caractéristiques visuelles invariante à la position d'un individu dans une image de group. Cette prise en compte de contexte de groupe réduit ainsi l'ambiguïté de ré-identification. Pour chacune de ces quatre contributions, nous avons effectué de nombreuses expériences sur les différentes bases de données publiques pour montrer l'efficacité des approches proposées.

Abstract

Video surveillance systems are of a great value for public safety. As one of the most important surveillance applications, person re-identification is defined as the problem of identifying people across images that have been captured by different surveillance cameras without overlapping fields of view. With the increasing need for automated video analysis, this task is increasingly receiving attention. However, this problem is challenging due to the large variations of lighting, pose, viewpoint and background. To tackle these different difficulties, in this thesis, we propose several deep learning based approaches to obtain a better person re-identification performance in different ways. In the first proposed approach, we use pedestrian attributes to enhance the person re-identification. The attributes are defined as semantic mid-level descriptions of persons, such as gender, accessories, clothing etc. They could be helpful to extract characteristics that are invariant to the pose

Thèse

Soutenue le 12/10/2018

Établissement :
INSA Lyon

Contact :
Atilla.Baskurt@liris.cnrs.fr

and viewpoint variations thanks to the descriptor being on a higher semantic level. In order to make use of the attributes, we propose a CNN-based person re-identification framework composed of an identity classification branch and of an attribute recognition branch. At a later stage, these two cues are combined to perform person re-identification. Secondly, among the challenges, one of the most difficult is the variation under different viewpoint. The same person shows very different appearances from different points of view. To deal with this issue, we consider that the images under various orientations are from different domains. We propose an orientation-specific CNN. This framework performs body orientation regression in a gating branch, and in another branch learns separate orientation-specific layers as local experts. The combined orientation-specific CNN feature representations are used for the person re-identification task. Thirdly, learning a similarity metric for person images is a crucial aspect of person re-identification. As the third contribution, we propose a novel listwise loss function taking into account the order in the ranking of gallery images with respect to different probe images. Further, an evaluation gain-based weighting is introduced in the loss function to optimize directly the evaluation measures of person re-identification. At the end, in a large gallery set, many people could have similar clothing. In this case, using only the appearance of single person leads to strong ambiguities. In realistic settings, people often walk in groups rather than alone. As the last contribution, we propose to learn a deep feature representation with displacement invariance for group context and introduce a method to combine the group context and single-person appearance. For all the four contributions of this thesis, we carry out extensive experiments on popular benchmarks and datasets to demonstrate the effectiveness of the proposed systems.

Publications

Revue internationale avec comité de lecture

- Yiqiang Chen, Stefan Duffner, Andrei Stoian, Jean-Yves Dufour & Atilla Baskurt (2018). « Deep and Low-level Feature based Attribute Learning for Person Re-identification ». *Image and Vision Computing*, vol. 79, pp. 25-34. doi : 10.1016/j.imavis.2018.09.001. HAL : hal-01895367.

Conférences internationales avec comité de lecture

- Yiqiang Chen, Stefan Duffner, Andrei Stoian, Jean-Yves Dufour & Atilla Baskurt (2018). « Person Re-identification using group context ». *Advanced Concepts for Intelligent Vision systems*, 24 septembre 2018, Poitiers (France). HAL : hal-01895373.
- Yiqiang Chen, Stefan Duffner, Andrei Stoian, Jean-Yves Dufour & Atilla Baskurt (2018). « Person Re-identification with a Body Orientation-Specific Convolutional Neural Network ». *Advanced Concepts for Intelligent Vision systems*, 24 septembre 2018, Poitiers (France). HAL : hal-01895374.
- Yiqiang Chen, Stefan Duffner, Andrei Stoian, Jean-Yves Dufour & Atilla Baskurt (2018). « SIMILARITY LEARNING WITH LISTWISE RANKING FOR PERSON RE-IDENTIFICATION ». *International conference on image processing*, 8 octobre 2018, Athenes (Grèce). HAL : hal-01895355.
- Yiqiang Chen, Stefan Duffner, Andrei Stoian, Jean-Yves Dufour & Atilla Baskurt (2018). « Pedestrian attribute recognition with part-based CNN and combined feature representations ». *VISAPP2018*, 27 janvier 2018, Funchal (Portugal). HAL : hal-01625470.

- Yiqiang Chen, Stefan Duffner, Andrei Stoian, Jean-Yves Dufour & Atilla Baskurt (2017). « Triplet CNN and pedestrian attribute recognition for improved person re-identification ». *14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS 2017)*, 1 septembre 2017, Lecce (Italie). doi : 10.1109/AVSS.2017.8078542. HAL : hal-01625479.

Jury

André TRICOT, PR Université Toulouse 2	rapporteur
Cécile VAN DE LEEMPUT, PR Université Libre de Bruxelles	rapporteur
Olivier VIGNEAU, Expert	examineur
Christine MICHEL, MC INSA Lyon	co-encadrante
Jean-François CERISIER, PR université de Poitiers	président
Marc-Eric BOBILLIER-CHAUMONT, PR Université Lumière Lyon 2	directeur

Résumé

Ce travail de thèse a pour objectif d'évaluer les facteurs intervenant dans l'acceptation d'un Espace Numérique de Travail (ONE) dans le contexte des écoles primaires en France. Dans ce cadre, nous avons mis en place un modèle d'acceptation triangulaire qui associe trois types de facteurs d'acceptation : facteurs pratiques (liés à la conception de l'outil), facteurs de perception (liés aux attitudes des utilisateurs) et facteurs situationnels liés aux caractéristiques du système d'activité scolaire. Afin d'évaluer ces facteurs, nous avons implémenté trois opérations de recueil de données. Dans la première opération (Axe 1), nous avons évalué, à partir de statistiques d'usage (N=266), les différences en termes d'utilisation entre les écoles ayant participé à la conception de l'outil et les écoles non participantes. Dans la seconde opération (Axe 2), nous avons évalués les attitudes des enseignants (N=8) et des parents (N=158) à l'aide de questionnaires d'intention d'usage et d'entretiens semi-directifs. Enfin, la troisième opération (Axe 3) est constituée d'entretiens approfondis à l'aide de méthodes d'explicitation et des analyses ethnographiques de publications réalisées sur ONE par les enseignants, les élèves et les parents d'élèves. Les résultats montrent que l'acceptation des ENT dans les écoles primaires est un processus lent, qui se construit sur plusieurs années. Nous constatons un rôle déterminant de la coopération entre les concepteurs et les enseignants dans l'acceptation future de ONE. Les enseignants et les parents sont globalement favorables à l'utilisation de ONE et ont des perceptions positives concernant sa facilité d'usage et le rôle de cet outil dans la communication entre les écoles et les familles et dans le suivi des élèves. Ils craignent cependant que l'outil allonge le temps passé devant les écrans par les élèves. Les facteurs situationnels montrent certaines difficultés pour intégrer ONE dans les pratiques enseignantes : problèmes de gestion de la classe, manque d'infrastructure, allongement du temps des séances et un accompagnement insuffisant sur le plan pédagogique. En revanche, les enseignants apprécient le rôle important de l'outil, d'une part, dans la formation des compétences numériques et la responsabilisation digitale des élèves et, d'autre part, dans la valorisation des connaissances des élèves et la construction d'une expérience d'apprentissage plus autonome.

Abstract

The objective on this PhD thesis is to assess the factors involved in the acceptance of a VLE (Virtual Learning Environment), called One, in a context of primary schools in France. Within this framework, we developed a triangulated acceptancy model which associates three types of determinants : practical factors (related to the technology conception), perceptual factors (associated to users' attitudes toward One), and situated factors related to the school activity system's specifications. In order to assess these factors, we implemented three phases of data collection. In the first stage (Topic 1), we assessed, from usage statistics, the differences between schools that were

Thèse

Soutenue le 20/06/2018

Établissement :

Contact :

Christine.Michel@liris.cnrs.fr

involved in the platform's design and those who were not. In the second stage (Topic 2), we evaluated the teachers and parents' attitudes toward the VLE with specific one-to-one interviews and surveys. Finally, the third stage (Topic 3) was based on qualitative interviews and ethnographic analysis of activities developed by teacher, students and parents on the platform. The results attest the fact that the acceptance of VLE in primary schools is a slow process which is built step by step and achieved over several years. The collaboration between designers and teachers plays a crucial role in the future acceptance of One. The teachers and parents are globally in favour of platform's adoption and they show positive views on its ease of use, the facilitator role played in the communication between families and schools and concerning a better monitoring of pupils. They fear however that One could increase the length of childrens' exposure to screens. The situated factors highlight some difficulties to integrate technologies in teachers' everyday practices : organisational issues, class management, lack of infrastructure, increase in classes' length, insufficient pedagogical support for teachers. Nonetheless, the teachers appreciate the contribution of the platform in the development of children's digital skills on the one hand, and in a better exploitation and recognition of children's knowledge and a more autonomous learning experience on the other hand

Publications

Revue nationale avec comité de lecture

- Elena Codreanu, Christine Michel, Marc-Eric Bobillier-Chaumont & Olivier Vigneau (2017). « L'acceptation et l'appropriation des ENT (Espaces Numériques de Travail) par les enseignants du primaire ». vol. 24, n°1, 39p. doi : 10.23709/sticf.24.1.1. ArXiv : 1806.01539. HAL : hal-01512377.

Conférences internationales avec comité de lecture

- Elena Codreanu, Christine Michel, Marc-Eric Bobillier-Chaumont & Olivier Vigneau (2016). « A Triangulated Model to Assess Adoption of Virtual Learning Environments in Primary Schools ». CSEDU 2016 - Proceedings of the 8th International Conference on Computer Supported Education, 23 avril 2016, Rome (Italie), pp. 287-293. HAL : hal-01274646.
- Elena Codreanu, Christine Michel, Marc-Eric Bobillier-Chaumont & Olivier Vigneau (2016). « The Acceptance of VLEs (Virtual Learning Environments) by Primary School Teachers ». CSEDU 2016 - Proceedings of the 8th International Conference on Computer Supported Education, 23 avril 2016, Rome (Italie), pp. 299-307. HAL : hal-01267042.

Éditions scientifique d'Ouvrages (livres, chapitres, colloques, congrès, n° spéciaux)

- Elena Codreanu, Christine Michel, Marc-Eric Bobillier-Chaumont & Olivier Vigneau (2017). « Assessing the Adoption of Virtual Learning Environments in Primary Schools: An Activity Oriented Study of Teacher's Acceptance ». Computers Supported Education, Costagliola, G, Uhomoihi, J, Zvacek, S, McLaren, B.M., Springer International Publishing, pp. 513-531. HAL : hal-01562342.

Jury

Frédéric HAVET, DR, CNRS, Sophia Antipolis,	rapporteur
Mickaël MONTASSIER, PR, Université de Montpellier,	rapporteur
Cristina BAZGAN, PR, Université Paris Dauphine,	examinatrice
Elzbieta SIDOROWICZ, MC, University of Zielona Góra,	examinatrice
Hamza SI KADDOUR, MC, Université Lyon 1,	examineur
Hamamache KHEDDOUCI, PR, Université Lyon 1,	directeur de thèse
Éric DUCHÊNE, MC, Université Lyon 1,	co-directeur de thèse
Aline PARREAU, CR, CNRS, Université Lyon 1,	co-directrice de thèse

Résumé

Dans cette thèse, nous étudions des problématiques de graphes et de jeux combinatoires. Il existe de nombreux liens entre ces deux domaines : ainsi, les jeux sont un bon moyen de modéliser une opposition dans un problème d'optimisation, et dans l'autre sens plusieurs jeux classiques sont définis sur les graphes. Nous allons étudier deux problèmes de graphes et adapter des jeux combinatoires classiques pour y jouer sur des graphes. Dans un premier temps, nous étudions un problème de criticalité. Un graphe qui vérifie une certaine propriété, mais tel qu'une simple modification (ajout ou suppression d'arête ou de sommet) la lui fait perdre est appelé critique pour cette propriété. Nous nous intéressons au problème des graphes critiques pour la propriété «avoir un diamètre égal à 2», appelés graphes D2C. La conjecture de Murty-Simon donne une borne supérieure sur le nombre d'arêtes d'un graphe D2C en fonction de son nombre de sommets. Or, des recherches récentes laissent supposer que cette borne peut être améliorée pour les graphes D2C non-bipartis. Nous démontrons donc une borne amoindrie pour une sous-famille de graphes D2C. Dans un deuxième temps, nous considérons un problème d'identification, laquelle consiste à assigner une étiquette à toutes les arêtes ou à tous les sommets d'un graphe, cette assignation devant engendrer une étiquette différente pour chaque sommet. Nous définissons une coloration d'arêtes par des ensembles d'entiers induisant une identification des sommets, et démontrons que cette coloration nécessite au plus un nombre logarithmique d'entiers par rapport à l'ordre du graphe pour l'identifier. Ce résultat est mis en comparaison avec d'autres types de colorations identifiantes, qui nécessitent dans le pire des cas un nombre linéaire d'entiers pour identifier tous les sommets. Dans un troisième temps, nous étudions des jeux de suppression de sommets, qui sont des jeux dans lesquels deux joueurs suppriment d'un graphe des sommets en respectant certaines règles prédéfinies, le premier joueur incapable de jouer perdant la partie. Nous proposons un cadre global pour l'étude de nombreux jeux de suppression de sommets dans les graphes, qui inclut plusieurs jeux classiques comme Arc-Kayles et permet une généralisation des jeux de soustraction et des jeux octaux sur les graphes. Dans leur définition classique, ces jeux ont généralement des comportements réguliers : tous les jeux de soustraction finis sont ultimement périodiques et il est conjecturé que c'est également le cas des jeux octaux. Nous étudions plus spécifiquement les jeux de soustraction connexes CSG(S), dans lesquels les joueurs peuvent supprimer k sommets induisant un sous-graphe connexe sans déconnecter le graphe si $k \in S$ (avec S fini). Nous démontrons que tous ces jeux sont ultimement périodiques, dans le sens ou pour un graphe et un sommet donnés, un chemin attaché à ce sommet peut être réduit à partir d'un certain rang sans modifier la valeur de Grundy du graphe pour le jeu. Nous trouvons également des résultats de périodicité pure, en particulier sur les étoiles subdivisées : pour certains ensembles S, les chemins des étoiles peuvent être réduits à leur longueur modulo une

Thèse

Soutenue le 27/09/2018

Établissement :
Université Claude Bernard Lyon 1

Contact :
Hamamache.Kheddouci@iris.cnrs.fr

certaine période sans changer l'issue du jeu. Enfin, nous définissons une variante pondérée de Arc-Kayles, appelée Weighted Arc-Kayles (ou WAK), dans laquelle les joueurs doivent sélectionner une arête pour réduire le poids de ses extrémités, les sommets ayant un poids nul étant supprimés du graphe. Nous montrons une réduction entre WAK et Arc-Kayles, puis que les valeurs de Grundy de WAK sont non-bornées, ce qui répond à une question ouverte sur Arc-Kayles. Nous montrons également que les valeurs de Grundy de WAK sont ultimement périodiques lorsque tous les poids du graphe sauf un sont fixes

Abstract

In this thesis, we study both graphs and combinatorial games. There are several links between those two domains : games are useful for modeling an opponent in optimization problems on graphs, and in the other direction several classical games are played on graphs. We will study two graph problems and adapt some classical combinatorial games to be played on graphs. In a first chapter, we study a criticality problem. A graph that verifies some property, and such that any modification (vertex or edge addition or deletion) breaks the property is called critical for this property. We focus on the critical graphs for the property «having diameter 2», called D2C graphs. The Murty-Simon conjecture gives an upper bound on the number of edges in a D2C graph with a given number of vertices. However, recent research suggests that this bound can be improved for non-bipartite D2C graphs. We show the validity of this approach by proving a smaller upper bound for a subfamily of non-bipartite D2C graphs. In a second chapter, we consider an identification problem. Identification consists in assigning some data to every edge or vertex of a graph, such that this assignment induces a label to every vertex with the added condition that two distinct vertices must have a different label. We define an edge-coloring using sets of integers inducing an identification of the vertices, and prove that this coloring requires at most a logarithmic number of integers (with respect to the order of the graph) in order to successfully identify the vertices. This result is compared with other identifying colorings, for which the number of colors required to successfully identify the vertices can be linear with respect to the order of the graph. In order to show the link between graphs and games, we adapt a well-known family of games to be played on graphs. We propose a general framework for the study of many vertex deletion games (which are games in which the players delete vertices from a graph under predefined rules) such as Arc-Kayles. This framework is a generalization of subtraction and octal games on graphs. In their classical definition, those games exhibit a high regularity : all finite subtraction games are ultimately periodic, and Guy conjectured that this is also true for all finite octal games. We specifically study the connected subtraction games CSG(S) (with S being a finite set). In those games, the players can remove k vertices from a graph if and only if they induce a connected subgraph, the graph remains connected after their deletion, and $k \in S$. We prove that those games are all ultimately periodic, in the sense that for a given graph and vertex, a path attached to this vertex can be reduced (after a certain preperiod) without changing the Grundy value of the graph for the game. We also prove pure periodicity results, mostly on subdivided stars : for some sets S, the paths of a subdivided star can be reduced to their length modulo a certain period without changing the outcome of the game. Finally, we define a weighted version of Arc-Kayles, called Weighted Arc-Kayles (WAK for short). In this game, the players select an edge and reduce the weight of its endpoints. Vertices with weight 0 are removed from the graph. We show a reduction between WAK and Arc-Kayles, then we prove that the Grundy values of WAK are unbounded, which answers an open question

on Arc-Kayles. We also prove that the Grundy values of WAK are ultimately periodic if we fix all but one of the weights in the graph

Publications

Revue internationale avec comité de lecture

- Laurent Beaudou, Pierre Coupechoux, Antoine Dailly, Sylvain Gravier, Julien Moncel, Aline Parreau & Eric Sopena (2018). « Octal Games on Graphs: The game 0.33 on subdivided stars and bistars ». *Theoretical Computer Science*, vol. 746, pp. 19-35. doi : 10.1016/j.tcs.2018.06.018. ArXiv : 1612.05772. HAL : hal-01418153.
- Nicolas Bousquet, Antoine Dailly, Eric Duchene, Hamamache Kheddouci & Aline Parreau (2017). « A Vizing-like theorem for union vertex-distinguishing edge coloring ». *Discrete Applied Mathematics*, vol. 232, pp. 88-98. doi : 10.1016/j.dam.2017.07.002. ArXiv : 1605.02588. HAL : hal-01313088.

Conférences internationales avec comité de lecture

- Laurent Beaudou, Pierre Coupechoux, Antoine Dailly, Sylvain Gravier, Julien Moncel, Aline Parreau & Eric Sopena (2017). « Octal Games on Graphs ». *Combinatorial Game Theory Colloquium II, CGTC 2*, 27 janvier 2017, Lisbonne (Portugal). HAL : hal-02264751.

Autres

- Antoine Dailly, Florent Foucaud & Adriana Hansberg (2018). « Strengthening the Murty-Simon conjecture on diameter 2 critical graphs ». ArXiv : 1812.08420. HAL : hal-01959683.
- Laurent Beaudou, Pierre Coupechoux, Antoine Dailly, Sylvain Gravier, Julien Moncel, Aline Parreau & Eric Sopena (2018). « Octal Games on Graphs: The game 0.33 on subdivided stars and bistars. Full proof of Theorem 22 ». HAL : hal-01807116.
- Antoine Dailly, Eric Duchene, Urban Larsson & Gabrielle Paris (2018). « Partition games are pure breaking games ». ArXiv : 1803.02621. HAL : hal-01723190.

Jury

Amel BOUZEGHOUB, PR, Télécom SudParis, Rapporteur
Abdelkader HAMEURLAIN, PR, Université Toulouse 3, Rapporteur
Karine ZEITOUNI, PR, Université de Versailles Saint-Quentin, Examinatrice
Mohand-Saïd HACID, PR, Université Lyon 1, Directeur de thèse
Nicolas LUMINEAU, MC, Université Lyon 1, Co-directeur de thèse

Résumé

De nombreux efforts ont été faits ces dernières années pour faciliter la gestion et la représentation des entités culturelles. Toutefois, il existe encore un grand nombre de systèmes souvent isolés et encore utilisés dans les institutions culturelles reposant sur des modèles non sémantiques qui rendent difficile la validation et l'enrichissement des données. Cette thèse a pour but de proposer de nouvelles solutions pour améliorer la représentation et l'enrichissement sémantique de données culturelles en utilisant les principes du Web Sémantique. Pour ce faire, la recherche est focalisée d'une part sur l'adoption de modèles plus sémantiques comme selon les principes de FRBR qui permet de représenter des familles bibliographiques complexes en utilisant un modèle entités associations avec différents niveaux d'abstraction. Toutefois, la qualité d'une telle transformation est cruciale et c'est pourquoi des améliorations doivent être faites au niveau de la configuration et de l'évaluation d'un tel processus. En parallèle, la thèse cherche à profiter de ces nouveaux modèles sémantiques pour faciliter l'interconnexion des données avec des sources externes comme celles du Linked Open Data ou des sources moins structurées (Sites Web, Flux). Cela doit permettre de générer des bases de connaissances thématiques plus en accord avec les besoins des utilisateurs. Cependant, l'agrégation d'informations depuis des sources hétérogènes implique des étapes d'alignement à la fois au niveau du schéma et au niveau des entités

Abstract

Many efforts have been done these last two decades to facilitate the management and representation of cultural heritage data. However, many systems used in cultural institutions are still based on flat models and are generally isolated which prevents any reuse or validation of information. This Ph.D. aims at proposing new solutions for enhancing the representation and enrichment of cultural entities using the Semantic Web technologies. This work consists in two major steps to reach this objective. On the one hand, the research is focused on the metadata migration process to transform the schema of existing knowledge catalogs to new semantic models. This study is based on a real-world case study using the concepts from the Functional Requirements for Bibliographic Records (FRBR) which allows to generate graph-based knowledge bases. Yet, the quality of such a migration is the cornerstone for a successful adoption. Thus, several challenges related to the tuning and the evaluation of such a process must be faced. On the other hand, the research aims at taking advantage of these semantic models to facilitate the linkage of information with external and structured sources (e.g., Linked Open Data) and extracting additional information from other sources (e.g., microblogging) to build a new generation of thematic knowledge bases according to the user needs. However, in this case, the aggregation of information from heterogeneous sources requires additional steps to match and merge both correspondences at schema and instance level

Thèse

Soutenue le 28/09/2018

Établissement :
Université Claude Bernard Lyon 1

Contact :
Mohand-Said.Hacid@liris.cnrs.fr

Publications

Reuves internationales avec comité de lecture

- Joffrey Decourselle (2016). « Towards a Pattern-based Semantic Enrichment of Bibliographic Entities ». IEEE TCDL. HAL : hal-01404651.

Conférences internationales avec comité de lecture

- Joffrey Decourselle & Frédéric Riondet (2017). « Adopting Semantic Technologies in Public Health Documentation ». International Semantic Web Conference [Industry Track], 25 octobre 2017, Vienne (Autriche). HAL : hal-01584080.
- Joffrey Decourselle, Fabien Duchateau, Trond Aalberg, Naimdjon Takhirov & Nicolas Lumineau (2016). « BIB-R: a Benchmark for the Interpretation of Bibliographic Records ». Theory and Practice of Digital Libraries (TPDL), 9 septembre 2016, Hannover (Allemagne), pp. 163-174. doi : 10.1007/978-3-319-43997-6_13. HAL : hal-01324529.
- Joffrey Decourselle, Fabien Duchateau & Nicolas Lumineau (2015). « A Survey of FRBRization Techniques ». Theory and Practice of Digital Libraries (TPDL), 14 septembre 2015, Poznań (Pologne), pp. 185-196. doi : 10.1007/978-3-319-24592-8_14. HAL : hal-01198487.
- Joffrey Decourselle, Audun Vennesland, Trond Aalberg, Fabien Duchateau & Nicolas Lumineau (2015). « A Novel Vision for Navigation and Enrichment in Cultural Heritage Collections ». Semantic Web For Cultural Heritage (SW4CH), 8 septembre 2015, Poitiers (France), pp. 488-497. doi : 10.1007/978-3-319-23201-0_49. HAL : hal-01194308.

Autres

- Joffrey Decourselle (2016). « Case-oriented Semantic Enrichment of Bibliographic Entities ». Theory and Practice of Digital Libraries, 9 septembre 2016, Hannover (Allemagne). Poster. HAL : hal-01346830.
- Joffrey Decourselle, Fabien Duchateau, Trond Aalberg, Naimdjon Takhirov & Nicolas Lumineau (2016). « Open Datasets for Evaluating the Interpretation of Bibliographic Records ». Joint Conference on Digital Libraries, 23 juin 2016, Newark (États-Unis). Poster. doi : 10.1145/2910896.2925457. HAL : hal-01302830.
- Joffrey Decourselle, Fabien Duchateau & Ronald Ganier (2016). « Syrtis: New Perspectives for Semantic Web Adoption ». BOBCATSSS, 29 janvier 2016, Lyon (France). Poster. HAL : hal-01258556.

Jury

Bruno CRÉMILLIEU, PR, Université de Caen Normandie,	Rapporteur
Jean-Yves RAMEL, PR, Polytech' Tours	Rapporteur
Jean-Christophe JANODET, PR, Université d'Evry	Examineur
Celine ROUVEIROL, PR, Université Paris XIII	Examinatrice
Christine SOLNON, PR, INSA Lyon	Directrice de thèse
Elisa FROMONT, PR, Université Rennes 1	Co-directrice de thèse
Baptiste JEUDY, MC, Université de Saint Étienne	Co-directeur de thèse

Résumé

Durant cette thèse, nous abordons le problème de la fouille exhaustive de motifs pour un cas particulier de graphes : les grilles. Ces grilles peuvent être utilisées pour modéliser des objets ayant une structure régulière. Ces structures sont naturellement présentes dans de nombreux jeux de plateaux (les dames, les échecs ou le go par exemple) ou encore dans les modélisations d'écosystèmes utilisant des automates cellulaires. On les retrouve également à un plus bas niveau dans les images, qui sont des grilles 2D de pixels ou encore les vidéos, qui sont des grilles spatio-temporelles 2D+t de pixels. Au cours de cette thèse, nous avons proposé un nouvel algorithme de fouille de motifs fréquents dédié aux grilles spatio-temporelles, GriMA. L'usage des grilles régulières permet à notre algorithme de réduire la complexité des tests d'isomorphismes. Ces tests sont souvent utilisés par les algorithmes génériques de fouilles de graphes mais ayant une complexité importante, cela limite leur usage sur des données réelles. Deux applications ont été proposées pour évaluer notre algorithme : la classification d'images pour la fouille de grilles 2D et la prédiction d'automates cellulaires pour la fouille de grilles 2D+t.

Abstract

During this thesis, we consider the exhaustive graph mining problem for a special kind of graphs : the grids. These grids can be used to model objects that present a regular structure. These structures are naturally present in multiple board games (checkers, chess or go for instance) or in ecosystems models using cellular automata. It is also possible to find this structure in a lower level in images, which are 2D grids of pixels, or even in videos, which are 2D+t spatio-temporal grids of pixels. In this thesis, we proposed a new algorithm to find frequent patterns dedicated to spatio-temporal grids, GriMA. Use of regular grids allow our algorithm to reduce the complexity of the isomorphisms test. These tests are often use by generic graph mining algorithm but because of their complexity, they are rarely used on real data. Two applications were proposed to evaluate our algorithm: image classification for 2D grids mining and prediction of cellular automata for 2D+t grids mining.

Publications

Conférences internationales avec comité de lecture

- Romain Deville, Elisa Fromont, Baptiste Jeudy & Christine Solnon (2017). « Mining Frequent Patterns in 2D+t Grid Graphs for Cellular Automata Analysis ». Graph-Based Representations in Pattern Recognition: 11th IAPR-TC-15 International Workshop, GBRPR 2017, 18 mai 2017, Anacapri (Italie), pp. 177-186. doi : 10.1007/978-3-319-58961-9. HAL : hal-01494623.

Thèse

Soutenue le 30/05/2018

Établissement :
INSA Lyon

Contact :
Christine.Solnon@liris.cnrs.fr

- Romain Deville, Elisa Fromont, Baptiste Jeudy & Christine Solnon (2016). « GriMa: a Grid Mining Algorithm for Bag-of-Grid-Based Classification ». The joint IAPR International Workshops on Structural and Syntactic Pattern Recognition (SSPR 2016) and Statistical Techniques in Pattern Recognition (SPR 2016), 2 décembre 2016, Merida (Mexique), pp. 132-142. HAL : hal-01376019.

Conférences nationales avec comité de lecture

- Romain Deville, Elisa Fromont, Baptiste Jeudy & Christine Solnon (2016). « Recherche efficace de motifs fréquents dans des grilles ». RFIA 2016, 1 juillet 2016, Clermont Ferrand (France). HAL : hal-01325043.

Jury

Leila MERGHEM-BOULAHIA, PR, Univ. de Technologie Troyes	Rapporteur
Nicolas MAUDET, PR, Université Pierre et Marie Curie-Paris 6	Rapporteur
Bo AN, PR, Nanyang Technological University (Singapour),	Examinateur
Élise BONZON, MC, Université Paris Descartes - Paris 5	Examinatrice
Hamamache KHEDDOUCI, PR, Univ. Claude Bernard Lyon1	Examinateur
Samir AKNINE, PR, Univ. Claude Bernard de Lyon 1	Directeur de thèse
Mbaye SÈNE, MC, Univ. Cheikh Anta Diop de Dakar (Sénégal)	Co-directeur

Résumé

La prise de décision collective est un processus dans lequel un groupe d'individus, ayant des intérêts différents, se réunit pour trouver une solution collective à un problème. Ce processus est inhérent aux activités de toute organisation politique, économique ou sociale. Le développement de l'Intelligence Artificielle notamment les Systèmes Multi-Agents a permis la modélisation et l'automatisation des processus de prise de décision afin de mieux comprendre et d'analyser leur fonctionnement. Une décision collective peut être prise par vote ou par négociation. Dans le cadre de cette thèse, nous abordons les mécanismes de négociation multilatérale pour la prise de décision collective basés sur des approches heuristiques. Les agents construisent la solution à leur problème à travers leurs interactions à la différence des modèles basés sur la théorie des jeux dont l'espace des solutions est supposé connu par tous les agents. Le problème des négociations heuristiques réside dans les mécanismes de raisonnement des agents dont la complexité augmente lorsque le nombre d'agents et d'attributs à négocier devient important. L'objectif de cette thèse est ainsi de proposer des mécanismes de négociation décentralisés (sans médiateur) et distribués en mettant en exergue l'aspect organisationnel des agents. Notre approche s'inspire du concept diviser pour régner et permet aux agents de négocier de façon incrémentale. Le but est de faciliter la recherche d'accords et de limiter la complexité du raisonnement des agents. Les travaux de cette thèse ont abouti à trois contributions abordant la négociation multi-agents sous différents angles tels que l'organisation des agents, le protocole d'interaction et les stratégies de concession et de choix de solutions équitables et justes. Pour valider nos propositions, nous avons implémenté sous JavaJade les mécanismes de négociation proposés. Les critères de performance que nous avons évalués sont, notamment, la convergence, le temps de négociation et la qualité de la solution. Nous avons comparé nos modèles avec ceux existants et les résultats obtenus montrent leur efficacité pour l'obtention des accords entre les agents

Abstract

Collective decision making is a process in which many participants with different interests interact in order to build a solution to their problem. It is inherent to many organizations and companies. Nowadays, the advances in Artificial Intelligence, notably, Multi-Agents Systems enabled the automation of decision-making processes in order to analyze and to better understand how these mechanisms work. A collective decision may be made by using a voting system or by using negotiation. In this thesis, we focus on multilateral negotiation for collective decision making by proposing negotiation models. The proposed models based on heuristic approach. The agents interact with them in order to build a solution to their problem. This context is different from models based on game theory where the set of possible solutions are

Ndeye Arame Diago

Thèse

Soutenue le 17/09/2018

Établissement :
Université Claude Bernard Lyon 1

Contact :
Samir.Aknine@liris.cnrs.fr

Marie Fernandez

supposed to be known by all agents. So heuristic negotiation issue is that agents' reasoning may be very complex. This complexity grows where the number of agents and issues to be negotiated are important. The goal of this research work consists of devising negotiation mechanisms where agents' interaction are fully decentralized. We focus on organization aspect of the multi-agent system by using divide and conquer approach in order to reduce the negotiation complexity and hence to facilitate research of agreements. Our works tackle negotiation under different contexts which lead us to bring three contributions which focus on agents' organization, interaction protocols, negotiation object, concession strategies and effective and fair solution concept. The proposed mechanisms are implemented in JavaJade. We analyze the convergence of the negotiation, negotiation time and quality of the solution. Our models are compared with a centralized approach where all of the agents are gathered around one group to negotiate. Our empirical analyses show that our propositions allow the agents to reach collectives agreements

Jury

Hugues BERRY,
Clémentine VIGNAL,
Guillaume BESLON,
François-Xavier DECHAUME-MONCHARMONT,
Valérie DUFOUR,
Sébastien DEREGNAUCOURT,
Isabelle GEORGE,

rapporteur
rapporteur

Résumé

Posséder des données fiables, à jour et précises sur les populations d'oiseaux peut se révéler central aux décisions de politique environnementale. La bioacoustique est un outil de suivi non invasif de populations animales et avantageux lorsque les méthodes d'observation ou les captures sont difficiles. De plus, il a été montré chez de nombreuses espèces que l'étude de la communication acoustique peut largement contribuer à comprendre la dynamique des interactions sociales au sein d'un groupe. Cependant, l'étude des interactions vocales peut se révéler difficile, notamment lorsque l'on souhaite s'intéresser à une échelle fine des échanges. C'est pourquoi la bioacoustique n'a que peu été utilisée pour la caractérisation de la structure sociale de populations. L'objectif de ce projet de thèse était le développement de techniques d'extraction de vocalisations individuelles au sein d'un groupe, ainsi que la modélisation de leur dynamique fine. Après avoir été développée, testée et validée, notre méthode a permis d'étudier le réseau acoustique chez une espèce d'oiseau social, le diamant mandarin, et d'explorer le lien entre réseau acoustique et réseau social. A travers plusieurs études, nous avons montré que la dynamique vocale d'un groupe dépend à la fois de la composition de ce groupe (sa taille, la présence de couples ou de juvéniles) et du contexte environnemental (sans perturbation, puis avec séparation visuelle ou présence d'un danger). Ainsi, avec le développement de méthodes d'extraction de réseau acoustique, ce projet contribue à la fois à la recherche fondamentale et appliquée dans ce domaine : en recherche fondamentale car l'étude de la dynamique des interactions vocales permet de mieux comprendre le réseau social, et en recherche appliquée pour le suivi de population.!

Abstract

Bird populations represent a significant proportion of urban and rural biodiversity. For this purpose, the acquisition of reliable, updated and precise data on bird population can be a central factor for environmental decisions. The current classical techniques are difficult regarding human resources (banding, tracking, counting) and often invasive. Bioacoustics is a non-invasive tool for animal populations monitoring (density, migration paths...). Moreover, it has been shown in many species that the study of vocal exchanges can largely help to understand the social interactions occurring in a group. However, studying vocal exchanges can be difficult, especially when we want to assess fine scale interactions. For this reason bioacoustics have rarely been used to characterize groups' social structure. The aim of this project was to develop techniques for the extraction of individual vocalizations in a group, and the modelling of their dynamics at a fine scale. After we developed, tested and validated our method, we used it to extract the acoustic network in a bird social species, the zebra finch, and investigate the link between acoustic and social network. Throughout different studies we showed that

Thèse

Soutenue le 09/04/2018

Établissement :
INSA Lyon

Contact :
Huges.Berry@liris.cnrs.fr

Data-driven computational modelling for some of the implications of dopamine in the brain : From subcellular signalling to area networks

Alexandre Foncelle

the group composition, more particularly its size, the presence of couples or the presence of juveniles can shape parts of the vocal dynamics. We also found that the environmental context (without any perturbation, then a context of separation for a couple, or predation in a group) can impact the vocal interactions dynamics. Thus, this project make contribution to both fundamental and applied research: in fundamental research by contributing to the study of vocal interactions dynamics to better understand the social network, and in applied research by contributing to define new standards for population monitoring.

Jury

Krasimira TSANEVA-ATANASOVA, PR, University of Exeter
Philippe FAURE, PR, Université Pierre et Marie Curie
Hédi SOULA, PR, Université Pierre et Marie Curie
Guillaume BESLON, PR, INSA Lyon
Hugues BERRY, PR, INRIA

Rapporteure
Rapporteur
Examineur
Examineur
Directeur de thèse

Résumé

Dans le cerveau, il est difficile de mettre au point des expériences avec un niveau de contrôle approprié à cause du haut niveau de connectivité. Pour traiter ce problème, les modèles mathématiques sont utilisés pour représenter le cerveau d'une façon plus compréhensible. En effet, les modèles mathématiques peuvent être plus pratiques que les expériences pour tester des hypothèses et chercher à extraire l'essence même du principe étudié, en le simplifiant. De plus, la modélisation computationnelle forme une branche spécifique de la modélisation mathématique, permettant de résoudre de gros calculs numériques. Dans cette thèse, j'ai utilisé la modélisation computationnelle à travers différentes approches pour étudier certaines régions cérébrales. Nous avons collaboré avec des neurobiologistes en appliquant nos modèles à des données expérimentales pour contribuer à mieux comprendre l'action de la dopamine, un neuromodulateur. J'ai étudié la diversité de l'action de la dopamine à trois échelles: la région cérébrale, le niveau cellulaire et le niveau moléculaire. La dopamine a un gros impact sur le cerveau et elle est principalement connue pour son implication dans le système de récompense. En effet, c'est une molécule associée à la prédiction de récompense et de punition. Peu de régions produisent de la dopamine et ces régions sont altérées par la maladie de Parkinson ou perturbées par la dépression. Pour la maladie de Parkinson, j'ai conçu un modèle de type taux de décharge pour reproduire l'activité neuronale des ganglions de la base. Ce modèle montre des réponses neuronales significativement différentes, entre la condition témoin et la condition parkinsonienne. Par ailleurs, avec un modèle de type Hodgkin-Huxley prenant en compte la dynamique de l'ion potassium, j'ai pu appuyer l'hypothèse que la région cérébrale appelée l'habenula, lorsqu'elle est hyperactive, induirait la dépression. Cette dépression serait due à un déséquilibre de la concentration en potassium à cause d'une dysfonction de l'astrocyte (surexpression des canaux Kir4.1). Enfin, la dopamine est aussi impliquée dans la plasticité synaptique, un phénomène à la base de la mémoire. Je l'ai étudié avec un troisième modèle, prenant en compte plusieurs résultats expérimentaux relatifs à la plasticité en fonction du timing des potentiels d'action et de sa modulation.

Abstract

In the brain, the high connectivity level makes it difficult to set up experiments with an appropriate level of control. To address that issue, mathematical models are used to represent the brain in a more comprehensive way. Easier than experiments to test hypotheses, mathematical models can extend them closer to reality and aim to extract the studied principle essence, by simplifying it. Computational modelling is a specific branch of mathematical modelling allowing to solve large numerical calculations. In this thesis, I used computational modelling to study brain parts through different approaches, all in collaboration with neurobiologists and applied to experimental data. A common framework is given by the goal of contributing to a picture of the action of the neuromodulator dopamine. I studied the diversity of

Thèse

Soutenue le 05/04/2018

Établissement :
INSA Lyon

Contact :
Huges.Berry@liris.cnrs.fr

dopamine's action at three different scales: the brain region, the cellular level and the molecular level. Dopamine has a large impact on the brain and it is mainly known for its rewarding dimension, it is, indeed, the molecule associated with reward prediction and punishment. Few regions in the brain produce dopamine and these regions are impaired in Parkinson's disease or disrupted in major depressive disorders. Concerning Parkinson's disease, I designed a firing-rate model to fit experimental basal ganglia neural activity, which disclosed significant changes of the neural response between control and Parkinsonian condition. Furthermore, with a Hodgkin-Huxley model accounting for the dynamics of the potassium ion, I could support the hypothesis that the brain region called lateral habenula hyper-activates and induces major depressive disorders because of unbalanced potassium concentration due to astrocyte dysfunction (Kir4.1 channels overexpression). Dopamine is also involved in synaptic plasticity, a phenomenon at the basis of memory that I explored with a third model accounting for several experimental results pertaining to spike-timing-dependent plasticity and its modulation.

Jury

Sidonie CHRISTOPHE, CR, IGN	Rapporteuse
Jean-Pierre JESSEL, PR, Université de Toulouse	Rapporteur
Paule-Annick DAVOINE, PR, Université de Grenoble	Examinatrice
Hugo LEDOUX, PR, Université de Delft	Invité
Vincent PICAVET, Docteur, Oslandia	Invité
Adrien PEYTAVIE, MC, Université de Lyon	Co-encadrant de thèse
Gilles GESQUIERE, PR, Université Lyon 2	Directeur de thèse

Résumé

Le perfectionnement des modes d'acquisition 3D (relevés laser, photographiques, etc.) a conduit à la multiplication des données 3D géolocalisées disponibles. De plus en plus de villes mettent leur modèle numérique 3D à disposition en libre accès. Pour garantir l'interopérabilité des différentes sources de données, des travaux ont été effectués sur la standardisation des protocoles d'échange et des formats de fichier. En outre, grâce aux nouveaux standards du Web et à l'augmentation de la puissance des machines, il est devenu possible ces dernières années d'intégrer des contenus riches, comme des applications 3D, directement dans une page web. Ces deux facteurs rendent aujourd'hui possible la diffusion et l'exploitation des données tridimensionnelles de la ville dans un navigateur web. Ma thèse, dotée d'un financement de type CIFRE avec la société Oslandia, s'intéresse à la représentation tridimensionnelle de la ville sur le Web. Plus précisément, il s'agit de récupérer et de visualiser, à partir d'un client léger, de grandes quantités de données de la ville sur un ou plusieurs serveurs distants. Ces données sont hétérogènes : il peut s'agir de la représentations 3D des bâtiments (maillages) et du terrain (carte de hauteur), mais aussi d'informations sémantiques telles que des taux de pollution (volumes), la localisation de stations de vélos (points) et le nombre de vélos disponibles, etc. Durant ma thèse, j'ai exploré différentes manières d'organiser ces données dans des structures génériques afin de permettre une transmission progressive de fortes volumétries de données 3D. La prise en compte de l'aspect multi-échelle de la ville est un élément clef de la conception de ces structures. L'adaptation de la visualisation des données à l'utilisateur est un autre grand axe de ma thèse. Du fait du grand nombre de cas d'utilisations existants pour la ville numérique, les besoins de l'utilisateur varient grandement : des zones d'intérêts se dégagent, les données doivent être représentées d'une manière spécifique... J'explore différentes manières de satisfaire ces besoins, soit par la priorisation de données par rapport à d'autres lors de leur chargement, soit par la génération de scènes personnalisés selon les préférences indiquées par l'utilisateur.

Abstract

Advances in 3D data acquisition techniques (laser scanning, photography, etc.) has led to a sharp increase in the quantity of available 3D geolocated data. More and more cities provide the scanned data on open access platforms. To ensure the intercompatibility of different data sources, standards have been developed for exchange protocols and file formats. Moreover, thanks to new web standards and the increase in processing power of personal devices, it is now possible to integrate rich content, such as 3D applications, directly in a web page. These two elements make it possible to share and exploit 3D city data into a web browser. The subject of my thesis, co-financed by the Oslandia company, is the 3D representation of city data on the Web. More

Thèse

Soutenue le 22/05/2018

Établissement :
Université Claude Bernard Lyon 1

Contact :
Gilles.Gesquiere@liris.cnrs.fr

precisely, the goal is to retrieve and visualize a great quantity of city data from one or several distant servers in a thin client. This data is heterogeneous: it can be 3D representations of buildings (meshes) or terrain (height maps), but also semantic information such as pollution levels (volume data), the position of bike stations (points) and their availability, etc. During my thesis, I explored various ways of organizing this data in generic structures in order to allow the progressive transmission of high volumes of 3D data. Taking into account the multiscale nature of the city is a key element in the design of these structures. Adapting the visualization of the data to the user is another important objective of my thesis. Because of the high number of uses of 3D city models, the user's needs vary greatly: some specific areas are of higher interest, data has to be represented in a certain way... I explore different methods to satisfy these needs, either by prioritizing some data over others during the loading stage, or by generating personalized scenes based on a set of preferences defined by the user.

Publications

Revue internationale avec comité de lecture

- Jérémy Gaillard, Adrien Peytavie & Gilles Gesquière (2018). « Visualisation and personalisation of multi-representations city models ». *International Journal of Digital Earth*, pp. 1-18. HAL : hal-01946770.

Conférences internationales avec comité de lecture

- Jérémy Gaillard, Adrien Peytavie & Gilles Gesquière (2016). « A Data Structure for Progressive Visualisation and Edition of Vectorial Geospatial Data ». *3D Geoinfo*, 21 octobre 2016, Athènes (Grèce), pp. 201-209. doi : 10.5194/isprs-annals-IV-2-W1-201-2016. HAL : hal-01420117.
- Jérémy Gaillard, Alexandre Vienne, Rémi Baume, Frédéric Pedrinis, Adrien Peytavie & Gilles Gesquière (2015). « Urban data visualisation in a web browser ». *Web3D 2015*, 21 juin 2015, Heraklion (Grèce), pp. 81-88. doi : 10.1145/2775292.2775302. HAL : hal-01196834.

Jury

Atika RIVENQ, PR, UPHF,
Ye-Qiong SONG, PR, Université de Lorraine,
Jacques MALENFANT, PR, Sorbonne Universités,
Zahia GUESSOUM, MC, Sorbonne Université,
Youakim BADR, MC, INSA-Lyon,
Bertrand MASSOT, MC, INSA-Lyon,

Rapporteur
Rapporteur
Examinateur
Examinatrice
Directeur de thèse
Co-encadrant de thèse

Résumé

La prolifération des objets communicants fixes et mobiles soulève la question de leur intégration dans les environnements quotidiens, par exemple dans le cadre de la e-santé ou de la domotique. Les principaux défis soulevés relèvent de l'interconnexion et de la gestion de la masse de donnée produite par ces objets connectés. Notre objectif est de développer une infrastructure Machine to Machine interopérable, ouverte, et dans le respect des standards récemment spécifiés. Cette infrastructure permettra de réaliser des services à partir d'objets connectés. Afin de développer celle-ci, il est nécessaire de détailler un modèle abstrait d'objets connectés indépendant de matériels et de logiciels sous-jacents, de gérer leur masse d'événement mais également de comprendre et modéliser leur caractéristiques cyber-physiques. Pour mener à bien la réalisation de services intelligents à partir d'objets connectés, les trois axes de recherche suivant seront développés : la définition d'un modèle d'objet connecté à base de propriétés cyber-physiques ainsi qu'une algèbre de composition de services ; de pouvoir assurer une auto-adaptation intelligente dans un contexte évolutif permettant aux objets de raisonner sur eux même, la gestion du flux d'événements qui impose la définition d'un modèle d'événement facilitant l'interopérabilité des objets ; et enfin la définition d'un langage déclaratif pour détecter les événements significatifs dans la masse d'événements. La validation de ces contributions s'effectuera par le biais du développement et de l'expérimentation à grandeur nature d'un service de diagnostic médical continu basé sur la collecte de données médicales en masse par des réseaux non-intrusifs de capteurs biomédicaux portables sur le corps humain.

Abstract

The rapid growth of fixed and mobile communicating objects raises the issue of their integration in everyday environment, e.g. in a e-health or home-automation context. The main challenges of these objects are the interoperability and the handling of the massive amount of data that they generate. Our objective is to develop an open, interoperable machine to machine framework maintaining the recently elaborated standards. This framework will allow the realization of services based upon connected devices. To insure the efficient development of that framework, an abstract model of connected objects needs to be specified, along with the elaboration of a formal algebra describing the composition of the connected devices, the handling of the important amount of data they will generate and finally it is of major importance to understand and model their cyber-physical properties. In order to complete the development successfully, three research directions will be followed: the specification of a connected device model based on its cyber-physical properties along with a service composition algebra; being able to insure a smart self-adaptation in a changing context allowing the connected devices to reason about themselves, the handling of the events flow, which requires the definition of an event model to make the objects

Thèse

Soutenue le 20/12/2018

Établissement :
INSA Lyon

Contact :
Youakim.Badr@iris.cnrs.fr

interoperable; and finally the definition of a declarative language allowing the detection of significant events in the massive generated data. The testing of these contribution will be realized through the development and the large scale experimentation of a remote monitoring and diagnostics service based on the mass collection of medical data with non invasive embedded biomedical sensors network.

Publications

Revue internationale avec comité de lecture

- Arthur Gatouillat, Youakim Badr, Bertrand Massot & Ervin Sejdić (2018). « Internet of Medical Things: A Review of Recent Contributions Dealing with Cyber-Physical Systems in Medicine ». IEEE internet of things journal, vol. 5, n°5, pp. 3810-3822. doi : 10.1109/JIOT.2018.2849014. HAL : hal-01836236.
- Arthur Gatouillat, Youakim Badr & Bertrand Massot (2018). « Smart and Safe Self-Adaption of Connected Devices Based on Discrete Controllers ». IET Software, vol. 13, n°1, pp. 49-59. doi : 10.1049/iet-sen.2018.5029. HAL : hal-01854053.
- Arthur Gatouillat, Antoine Dumortier, Subashan Perera, Youakim Badr, Claudine Gehin & Ervin Sejdić (2017). « Analysis of the pen pressure and grip force signal during basic drawing tasks: The timing and speed changes impact drawing characteristics ». Computers in Biology and Medicine, vol. 87, pp. 124-131. doi : 10.1016/j.combiomed.2017.05.020. HAL : hal-01585955.

Conférences internationales avec comité de lecture

- Arthur Gatouillat, Bertrand Massot, Youakim Badr, Ervin Sejdić & Claudine Gehin (2018). « Evaluation of a real-time low-power cardiorespiratory sensor for the IoT ». 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 21 juillet 2018, Honolulu (États-Unis), pp. 5382-5385. doi : 10.1109/EMBC.2018.8513550. HAL : hal-01768798.
- Arthur Gatouillat, Bertrand Massot, Youakim Badr, Ervin Sejdić & Claudine Gehin (2018). « Building IoT-Enabled Wearable Medical Devices: An Application to a Wearable, Multiparametric, Cardiorespiratory Sensor ». Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2018), 21 janvier 2018, Funchal (Portugal), pp. 109-118. doi : 10.5220/0006729101090118. HAL : hal-01677241.
- Arthur Gatouillat, Youakim Badr & Bertrand Massot (2018). « Hybrid Controller Synthesis for the IoT ». The 33rd ACM/SIGAPP Symposium On Applied Computing (ACM SAC), 13 avril 2018, Pau (France), pp. 778-785. doi : 10.1145/3167132.3167219. HAL : hal-01644356.
- Arthur Gatouillat & Youakim Badr (2017). « Verifiable and Resource-Aware Component Model for IoT Devices ». 9th International Conference on Management of Digital EcoSystems (MEDES), 10 novembre 2017, Bangkok (Thaïlande), pp. 235-242. doi : 10.1145/3167020.3167056. HAL : hal-01619270.
- Arthur Gatouillat, Youakim Badr & Bertrand Massot (2017). « QoS-Driven Self-Adaptation for Critical IoT-Based Systems ». Workshop on Adaptive Service-oriented and Cloud Applications (ASOCA), 13 novembre 2017, Malaga (Espagne), pp. 93-105. doi : 10.1007/978-3-319-91764-1_8. HAL : hal-01619297.

Jury

Marc DANIEL, PR, Ecole Polytech Marseille	Rapporteur
Stefanie HAHMANN, PR, Université Grenoble INP-Ensimag	Rapporteur
Laurence CHÈZE, PR, Université Lyon 1	Examinatrice
Emmanuel PROMAYON, PR, Université Grenoble Alpes	Examineur
Behzad SHARIAT, PR, Université Lyon 1	Directeur de thèse
Hamid LADJAL, MC, Université Lyon 1	Co-directeur de thèse

Résumé

La Radio/Hadron-thérapie consiste à déposer une dose létale de rayonnement dans la tumeur tout en réduisant l'impact de cette dose sur les tissus sains. Les mouvements internes, en particulier ceux engendrés par la respiration modifient la forme, la position et la densité des organes, source d'erreur et d'incertitude sur la position du dépôt de dose. Lorsque la tumeur se trouve sur un organe en mouvement, la difficulté majeure est de cibler la tumeur pendant le traitement. Cette incertitude sur la position rend indispensable la mise en place d'une stratégie permettant la prédiction du mouvement tumoral. Ceci permet en effet de guider le faisceau de rayons ionisants de sorte qu'il suive les mouvements tumoraux. De plus, le traitement par hadronthérapie nécessite également l'accès à une description précise de la densité de l'ensemble des organes traversés par le faisceau, car la position du dépôt maximal de l'énergie véhiculée par les ions (le pic de Bragg) en dépend. Malheureusement, le mouvement respiratoire est complexe et sa prédiction n'est pas une tâche simple - en particulier, la respiration est commandée par l'action indépendante des muscles de la cage thoracique et du diaphragme. Les techniques actuelles basées sur l'imagerie, telles que le Cone-Beam ou le recalage déformable d'images, tentent de prédire la position des tumeurs pulmonaires. Ces méthodes font l'hypothèse d'un mouvement reproductible de l'appareil respiratoire dans le temps. D'autres techniques basées sur l'emploi de deux caméras à rayons X (cyberknife, tracking mis au point par l'équipe du Centre carbone d'Heidelberg [HIT]) peuvent permettre la pré-diction de la position des tumeurs, quand leur segmentation et leur contourage automatique en temps réel est possible. Cependant, ces méthodes sont, si ce n'est risquées, invasives, et elles ne permettent pas de calculer l'évolution des organes environnants, une information indispensable pour déterminer la position du pic de Bragg. Ainsi déduire le mouvement de la tumeur à partir de seules séries d'images médicales apparaît comme insuffisant. Une solution peut alors résider dans le développement d'un modèle biomécanique patient-spécifique du système respiratoire intégrant la variabilité du mouvement respiratoire. Pour que ce modèle soit précis, il doit comprendre la modélisation de la cage thoracique, du diaphragme et des poumons. Il est tout aussi important que ce modèle puisse être piloté par des paramètres mesurés en externe (capteurs 3D, spiromètre, etc.) an de préserver un caractère non-invasif et de corrélérer le mouvement externe du thorax et de l'abdomen, ainsi que le flux d'air échangé avec les mouvements internes. Les changements de propriétés mécaniques des milieux traversés par le faisceau doivent également être modélisés an de satisfaire les besoins de l'hadronthérapie.

Abstract

The 4D computational patient specific of the respiratory system could be potentially used in various medical contexts; for diagnosis, treatment planning, laparoscopic, dose computation or the registration between online

Thèse

Soutenue le 17/10/2018

Établissement :
Université Claude Bernard Lyon 1

Contact :
Behzad.Shariat@iris.cnrs.fr

imaging systems such as positron emission tomography (PET), computed-tomography (CT) as well as high delity and precise computer-based training simulators. The main novelty of this PhD project lies in the context of radiation therapy; we have developed a patient-specific biomechanical model of the respiratory system enabling the correlation of the internal organs motion with respiratory surrogate signal(s) during the treatment. This permits to take into account the respiratory motion variabilities. The deformation of the different structures is controlled and driven by simulated rib cage (mimic the external intercostal muscles) and diaphragm actions. For the diaphragm, we have applied the radial direction of muscle forces, and simple homogeneous dirichlet boundary condition is applied to the lower part of the diaphragm, which is attached to the rib cage. For each rib a rigid transformation is calculated automatically by finite helical axis method (rigid translation and rotation) and used to define displacement boundary conditions. The resulting widening of the thoracic cavity forces the lungs to expand due to an applied negative pressure in the pleural cavity. Other novelty of the PhD project, that the amplitude of the lung pressure and diaphragm force are patient-specific, and determined at different respiratory states by an optimization framework based on inverse FE analysis methodology, by minimizing the volume lungs errors, between the respiratory volume (calculated from CT scan images at each state) and the simulated volume (calculated by biomechanical simulation). All other structures are linked to each other, but feature different deformation behavior due to the assigned material properties. Our results are quite realistic compared to the 4D CT scan images and the proposed physically-based FE model is able to predict correctly the respiratory motion

Jury

Maud MARCHAL, PR, INSA Rennes, IRISA
Yohan PAYAN, PR, TIMC-IMAG
Laurence CHÈZE, PR, UCBL, LBMC
Sébastien LAPORTE, PR, ENSAM, IBHGC
Philippe MESEURE, PR, XLIM
Guillaume DAMIAND, PR, CNRS, LIRIS
Florence ZARA, MC, UCBL
Stéphane NICOLLE, PR, UCBL, LBMC
Jean-François PALIERNE, PR, Physique Lab.

Rapporteure
Rapporteur
Examinatrice
Examineur
Examineur
Directeur de thèse
Co-directrice de thèse
Co-directeur de thèse
Invité

Résumé

La nécessité de simulations de tissus mous, tels que les organes internes, se pose avec le progrès des domaines scientifiques et médicaux. Le but de ma thèse est de développer un nouveau modèle générique, topologique et physique, pour simuler les organes humains. Un tel modèle doit être facile à utiliser, doit pouvoir effectuer des simulations en temps réel avec un niveau de précision permettant l'utilisation à des fins médicales. Cette thèse explore de nouvelles méthodes de simulation et propose des améliorations pour la modélisation de corps déformables. Les méthodes proposées visent à pouvoir effectuer des simulations rapides, robustes et fournissant des résultats physiquement précis. L'intérêt principal de nos solutions réside dans la simulation de tissus mous élastiques à petites et grandes déformations à des fins médicales. Nous montrons que pour les méthodes existantes, la précision pour simuler librement des corps déformables ne va pas de pair avec la performance en temps de calcul. De plus, pour atteindre l'objectif de simulation rapide, de nombreuses approches déplacent certains calculs dans une étape de pré-traitement, ce qui entraîne l'impossibilité d'effectuer des opérations de modification topologiques au cours de la simulation comme la découpe ou le raffinement. Dans cette thèse, le cadre utilisé pour les simulations s'appelle TopoSim. Il est conçu pour simuler des matériaux à l'aide de systèmes masses-ressorts (MSS) avec des paramètres d'entrée spécifiques. En utilisant un MSS, qui est connu pour sa simplicité et sa capacité à effectuer des simulations temps réel, nous présentons plusieurs améliorations basées physiques pour contrôler les fonctionnalités globales du MSS qui jouent un rôle clé dans la simulation de tissus réels. La première partie de ce travail de thèse vise à reproduire une expérience réelle de simulation physique qui a étudié le comportement du tissu porcine à l'aide d'un rhéomètre rotatif. Son objectif était de modéliser un corps visco-élastique non linéaire. À partir de l'ensemble des données acquises, les auteurs de l'expérience ont dérivé une loi de comportement visco-élastique qui a ensuite été utilisée afin de la comparer avec nos résultats de simulation. Nous définissons une formulation des forces visco-élastiques non linéaires inspirée de la loi de comportement physique. La force elle-même introduit une non linéarité dans le système car elle dépend fortement de l'amplitude de l'allongement du ressort et de trois paramètres spécifiques à chaque type de tissu. La seconde partie de la thèse présente notre travail sur les forces de correction de volume permettant de modéliser correctement les changements volumétriques dans un MSS. Ces forces assurent un comportement isotrope des solides élastiques et un comportement correct du volume quel que soit la valeur du coefficient de Poisson utilisé. La méthode nécessite de résoudre deux problèmes: l'instabilité provoquant des plis et les contraintes de Cauchy. Nos solutions à ces limitations impliquent deux étapes. La première consiste à utiliser trois types de ressorts dans un maillage entièrement hexaédrique: les arêtes, les

Karolina Golec

Thèse

Soutenue le 19/01/2018

Établissement :
Université Claude Bernard Lyon 1

Contact :
Guillaume.Damiand@liris.cnrs.fr

faces diagonales et les diagonales internes. Les raideurs des ressorts dans le système ont été formulées pour obéir aux lois mécaniques de base. La deuxième étape consiste à ajouter des forces de correction linéaires calculées en fonction du changement de volume et des paramètres mécaniques du tissu simulé, à savoir le coefficient de Poisson et le module de Young [etc...]

Abstract

The need for simulations of soft tissues, like internal organs, arises with the progress of the scientific and medical environments. The goal of my PhD is to develop a novel generic topological and physical model to simulate human organs. Such a model shall be easy to use, perform the simulations in the real time and which accuracy will allow usage for the medical purposes. This thesis explores novel simulation methods and improvement approaches for modeling deformable bodies. The methods aim at fast and robust simulations with physically accurate results. The main interest lies in simulating elastic soft tissues at small and large strains for medical purposes. We show however, that in the existing methods the accuracy to freely simulate deformable bodies and the real-time performance do not go hand in hand. Additionally, to reach the goal of simulating fast, many of the approaches move the necessary calculations to pre-computational part of the simulation, which results in inability to perform topological operations like cutting or refining. The framework used for simulations in this thesis is designed to simulate materials using Mass Spring Systems (MSS) with particular input parameters. Using Mass-Spring System, which is known for its simplicity and ability to perform fast simulations, we present several physically-based improvements to control global features of MSS which play the key role in simulation of real bodies

Publications

Conférences internationales avec comité de lecture

- Karolina Golec, Florence Zara, Stéphane Nicolle, Jean-François Paliarne & Guillaume Damiand (2016). « New Mass Spring System formulation to model the behavior of soft tissues ». 22nd European Society of Biomechanics Congress (ESB 2016), 13 juillet 2016, Lyon (France). HAL : hal-01320714.
- Karolina Golec, Matthieu Coquet, Florence Zara & Guillaume Damiand (2015). « Improvement of a Topological-Physical Model to manage different physical simulations ». 23rd International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision 2015, 12 juin 2015, Plzen (République Tchèque), pp. 25-34. HAL : hal-01152120.

Autres

- Karolina Golec, Jean-François Paliarne, Florence Zara, S Nicolle & Guillaume Damiand (2017). « Hybrid 3D Mass-Spring System to simulate isotropic materials ». Work In Progress Session, VRIPHYS 2017: 13th Workshop on Virtual Reality Interaction and Physical Simulation, 23-24 April 2017, Lyon, France. HAL : hal-01504413.

Jury

Ladjel BELLATRECHE, PR, ENSMA Poitiers,	Rapporteur
Ernesto EXPOSITO, PR, Univ. de Pau et des Pays de l'Adour,	Rapporteur
Laurent D'ORAZIO, PR, Univ. de Rennes,	Examineur
Abderrafiaa KOUKAM, PR, UTBM,	Examineur
Salima HASSAS, PR, Univ. Claude Bernard Lyon 1,	Examinatrice
Parisa GHODOUS, PR, Univ. Claude Bernard Lyon 1,	Directrice de thèse
Genoveva VARGAS-SOLAR, CR, CNRS,	Co-directrice
Catarina FERREIRA DA SILVA, MC, Univ. Claude Bernard Lyon 1	Co-directrice

Résumé

L'émergence de nouvelles plateformes décentralisées pour la création de données, tel que les plateformes mobiles, les capteurs et l'augmentation de la disponibilité d'open data sur le Web, s'ajoute à l'augmentation du nombre de sources de données disponibles et apporte des données massives sans précédent à être explorées. La notion de curation de données qui a émergé se réfère à la maintenance des collections de données, à la préparation et à l'intégration d'ensembles de données (data set), les combinant avec une plateforme analytique. La tâche de curation inclut l'extraction de métadonnées implicites et explicites ; faire la correspondance et l'enrichissement des métadonnées sémantiques afin d'améliorer la qualité des données. La prochaine génération de moteurs de gestion de données devrait promouvoir des techniques avec une nouvelle philosophie pour faire face au déluge des données. Ils devraient aider les utilisateurs à comprendre le contenu des collections de données et à apporter une direction pour explorer les données. Un scientifique peut explorer les collections de données pas à pas, puis s'arrêter quand le contenu et la qualité atteignent des niveaux satisfaisants. Notre travail adopte cette philosophie et la principale contribution est une approche de curation des données et un environnement d'exploration que nous avons appelé CURARE. CURARE est un système à base de services pour curer et explorer des données volumineuses sur les aspects variété et variabilité. CURARE implémente un modèle de collection de données, que nous proposons, visant représenter le contenu structurel des collections des données et les métadonnées statistiques. Le modèle de collection de données est organisé sous le concept de vue et celle-ci est une structure de données qui pourvoit une perspective agrégée du contenu des collections des données et de ses parutions (releases) associées. CURARE pourvoit des outils pour explorer (interroger) des métadonnées et pour extraire des vues en utilisant des méthodes analytiques. Exploiter les données massives requière un nombre considérable de décisions de la part de l'analyste des données pour trouver quelle est la meilleure façon pour stocker, partager et traiter les collections de données afin d'en obtenir le maximum de bénéfice et de connaissances à partir de ces données. Au lieu d'explorer manuellement les collections des données, CURARE fournit de outils intégrés à un environnement pour assister les analystes des données à trouver quelle est la meilleure collection qui peut être utilisée pour accomplir un objectif analytique donné. Nous avons implémenté CURARE et expliqué comment le déployer selon un modèle d'informatique dans les nuages (cloud computing) utilisant des services de science des données sur lesquels les services CURARE sont branchés. Nous avons conçu des expériences pour mesurer les coûts de la construction des vues à partir des ensembles des données du Grand Lyon et de Twitter, afin de pourvoir un aperçu de l'intérêt de notre approche et notre environnement de curation de données

Thèse

Soutenue le 26/09/2018

Établissement :
Université Claude Bernard Lyon 1

Contact :
Parisa.Ghodous@liris.cnrs.fr

Abstract

The emergence of new platforms for decentralized data creation, such as sensor and mobile platforms and the increasing availability of open data on the Web, is adding to the increase in the number of data sources inside organizations and brings an unprecedented Big Data to be explored. The notion of data curation has emerged to refer to the maintenance of data collections and the preparation and integration of datasets, combining them to perform analytics. Curation tasks include extracting explicit and implicit meta-data; semantic metadata matching and enrichment to add quality to the data. Next generation data management engines should promote techniques with a new philosophy to cope with the deluge of data. They should aid the user in understanding the data collections' content and provide guidance to explore data. A scientist can stepwise explore into data collections and stop when the content and quality reach a satisfaction point. Our work adopts this philosophy and the main contribution is a data collections' curation approach and exploration environment named CURARE. CURARE is a service-based system for curating and exploring Big Data. CURARE implements a data collection model that we propose, used for representing their content in terms of structural and statistical meta-data organised under the concept of view. A view is a data structure that provides an aggregated perspective of the content of a data collection and its several associated releases. CURARE provides tools focused on computing and extracting views using data analytics methods and also functions for exploring (querying) meta-data. Exploiting Big Data requires a substantial number of decisions to be performed by data analysts to determine which is the best way to store, share and process data collections to get the maximum benefit and knowledge from them. Instead of manually exploring data collections, CURARE provides tools integrated in an environment for assisting data analysts determining which are the best collections that can be used for achieving an analytics objective. We implemented CURARE and explained how to deploy it on the cloud using data science services on top of which CURARE services are plugged. We have conducted experiments to measure the cost of computing views based on datasets of Grand Lyon and Twitter to provide insight about the interest of our data curation approach and environment

Publications

Revue internationale avec comité de lecture

- Hind Benfenatki, Catarina Ferreira Da Silva, Gavin Kemp, Benharkat Aïcha-Nabila, Parisa Ghodous & Zakaria Maamar (2017). « MADONA: a method for automated provisioning of cloud-based component-oriented business Applications ». *Service Oriented Computing and Applications*, vol. 11, n°1, pp. 87-100. doi : 10.1007/s11761-016-0199-0. HAL : hal-01374876.
- Gavin Kemp, Genoveva Vargas-Solar, Catarina Ferreira Da Silva, Parisa Ghodous, Christine Collet & Pedropablo López Amaya (2016). « Cloud big data application for transport ». *International Journal of Agile Systems and Management*, vol. 9, n°3, pp. 232-250. doi : 10.1504/IJASM.2016.079940. HAL : hal-01374885.
- Gavin Kemp, Pedropablo López Amaya, Catarina Ferreira Da Silva, Genoveva Vargas-Solar, Parisa Ghodous & Christine Collet (2016). « Big Data Collections and Services For Building Intelligent Transport Applications ». *International Journal of Electronic Business Management*, vol. 14, p. 11. HAL : hal-01374880.

Conférences internationales avec comité de lecture

- Gavin Kemp, Genoveva Vargas-Solar, Catarina Ferreira Da Silva & Parisa Ghodous (2015). « Aggregating and Managing Big Realtime Data in the Cloud - Application to Intelligent Transport for Smart Cities ». *Proceedings of the 1st International Conference on Vehicle Technology and Intelligent Transport Systems*, 22 mai 2015, Lisbon (Portugal), pp. 107-112. doi : 10.5220/0005491001070112. HAL : hal-01213310.
- Gavin Kemp, Genoveva Vargas-Solar, Catarina Ferreira Da Silva, Parisa Ghodous, Christine Collet & Pedropablo Lopez (2015). « Towards Cloud big data services for intelligent transport systems ». *Proc. of 22th ISPE International Conference on Concurrent Engineering, Advances in Transdisciplinary Engineering*, Vol. 2: *Transdisciplinary Lifecycle Analysis of Systems*, 23 juillet 2015, Delft (Pays-Bas), pp. 377-385. HAL : hal-01213302.

Éditions scientifique d'Ouvrages (livres, chapitres, colloques, congrès, n° spéciaux)

- Gavin Kemp, Genoveva Vargas-Solar, Catarina Ferreira Da Silva, Parisa Ghodous & Christine Collet (2015). « Service Oriented Big Data Management for Transport ». *Smart Cities, Green Technologies, and Intelligent Transport Systems / series Communications in Computer and Information Science*, pp. 267-281. doi : 10.1007/978-3-319-27753-0_15. HAL : hal-01372400.

Détection non supervisée basée sur les relations spatiales, application à la détection d'objets et la reconnaissance de structure de documents d'entreprises

Louisa Kessi

Jury

L. WENDLING, PR, Université Paris Descartes
L. LIKFORMAN-SULEM, MC, Telecom ParisTech
R. INGOLD, PR, Université de Fribourg
A. ELMOATAZ, PR, Université de Caen
C. GARCIA, PR, INSA Lyon
F. LE BOURGEOIS, MC, INSA Lyon

Rapporteur
Rapporteur
Examineur
Examineur
Directeur de thèse
Co-encadrant

Résumé

Cette thèse a pour objectif de développer un système de reconnaissance de structures logique des documents d'entreprises sans modèle. Il s'agit de reconnaître la fonction logique de blocs de textes qui sont importants à localiser et à identifier. Ce problème est identique à celui de la détection d'objets dans une scène naturelle puisqu'il faut à la fois reconnaître les objets et les localiser dans une image. A la différence de la reconnaissance d'objets, les documents d'entreprises doivent être interprétés sans aucune information a priori sur leurs modèles de structures. La seule solution consiste à développer une approche non supervisée basée principalement sur les relations spatiales et sur les informations textuelles et images. Les documents d'entreprises possèdent des contenus et des formes très hétérogènes car chaque entreprise et chaque administration créent son propre formulaire ou ses propres modèles de factures. Nous faisons l'hypothèse que toute structure logique de document est constituée de morceaux de micro-structures déjà observées dans d'autres documents. Cette démarche est identique en détection d'objets dans les images naturelles. Tout modèle particulier d'objet dans une scène est composé de morceaux d'éléments déjà vu sur d'autres exemples d'objets de même classe et qui sont reliés entre eux par des relations spatiales déjà observées. Notre modèle est donc basé sur une reconnaissance partie par partie et sur l'accumulation d'évidences dans l'espace paramétrique et spatial. Notre solution a été testée sur des applications de détection d'objets dans les scènes naturelles et de reconnaissance de structure logique de documents d'entreprises. Les bonnes performances obtenues valident les hypothèses initiales. Ces travaux contiennent aussi de nouvelles méthodes de traitement et d'analyse d'image couleurs de documents et d'images naturelles.

Publications

Conférences internationales avec comité de lecture

- Louisa Kessi, Frank Le Bourgeois & Christophe Garcia (2016). « An Efficient New PDE-based Characters Reconstruction After Graphics Removal ». 15th International Conference on Frontiers in Handwriting Recognition (ICFHR-2016), 26 octobre 2016, Shenzhen (Chine). HAL : hal-01345831.
- Louisa Kessi, Frank Le Bourgeois & Christophe Garcia (2015). « An Efficient Image Registration Method based on Modified NonLocal - Means : Application to Color Business Document Images ». 10th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 14 mars 2015, Berlin (Allemagne), pp. 166-173. doi : 10.5220/0005315301660173. HAL : hal-01272993.

Thèse

Soutenue le 13/09/2018

Établissement :
INSA Lyon

Contact :
Christophe.Garcia@liris.cnrs.fr

- Louisa Kessi, Frank Le Bourgeois, Christophe Garcia & Jean Duong (2015). « ACoDPS : Robust and Unsupervised Automatic Color Document Processing System ». VISAPP 2015, 10th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 14 mars 2015, Berlin (Allemagne). doi : 10.5220/0005315801740185. HAL : hal-01272989.

Éditions scientifique d'Ouvrages (livres, chapitres, colloques, congrès, n° spéciaux)

- Louisa Kessi, Frank Le Bourgeois & Christophe Garcia (2016). « ACoDSS: Robust Unsupervised Automatic Color Segmentation System for Noisy Heterogeneous Document Images ». European Project Space on Computer Vision, Graphics, Optics and Photonics Berlin, Germany March, 2015 Filipa Duarte (Ed.), SCITERESS, Sponsored and Organized by INSTICC Published by SCITEPRESS. HAL : hal-01272998.

Roland Kotto Kombi

Jury

Bernd AMANN, PR, Sorbonne Université
Franck MORVAN, PR, Université Paul Sabatier
Hala SKAF-MOLLI, PR, Université de Nantes
Bruno DEFUDE, PR, Télécom Paris Sud
Philippe LAMARRE, PR, INSA Lyon
Nicolas LUMINEAU, MC, Université Lyon 1

Rapporteur
Rapporteur
Examinatrice
Examineur
Directeur de thèse
Co-encadrant de thèse

Résumé

Le traitement de flux de données est au cœur des problématiques actuelles liées au Big Data. Face à de grandes quantités de données (Volume) accessibles de manière éphémère (Vélocité), des solutions spécifiques tels que les systèmes de gestion de flux de données (SGFD) ont été développés. Ces SGFD reçoivent des flux et des requêtes continues pour générer de nouveaux résultats aussi longtemps que des données arrivent en entrée. Dans le contexte de cette thèse, qui s'est réalisée dans le cadre du projet ANR Socioplug (ANR-13-INFR-0003), nous considérons une plateforme collaborative de traitement de flux de données à débit variant en termes de volume et de distribution de valeurs. Chaque utilisateur peut soumettre des requêtes continues et contribue aux ressources de traitement de la plateforme. Cependant, chaque unité de traitement traitant les requêtes dispose de ressources limitées ce qui peut engendrer la congestion du système en fonction des variations des flux en entrée. Le problème est alors de savoir comment adapter dynamiquement les ressources utilisées par chaque requête continue par rapport aux besoins de traitement. Cela soulève plusieurs défis : i) comment détecter un besoin de reconfiguration ? ii) quand reconfigurer le système pour éviter sa congestion ? Durant ces travaux de thèse, nous nous sommes intéressés à la gestion automatique de la parallélisation des opérateurs composant une requête continue. Nous proposons une approche originale basée sur une estimation des besoins de traitement dans un futur proche. Ainsi, nous pouvons adapter le niveau de parallélisme des opérateurs de manière proactive afin d'ajuster les ressources utilisées aux besoins des traitements. Nous montrons qu'il est possible d'éviter la congestion du système mais également de réduire significativement la consommation de ressources à performance équivalente. Ces différents travaux ont été implémentés et validés dans un SGFD largement utilisé avec différents jeux de tests reproductibles.

Abstract

In a Big Data context, stream processing has become a very active research domain. In order to manage ephemeral data (Velocity) arriving at important rates (Volume), some specific solutions, denoted data stream management systems (DSMSs), have been developed. DSMSs take as inputs some queries, called continuous queries, defined on a set of data streams. A continuous query generates new results as long as new data arrive in input. In many application domains, data streams have input rates and distribution of values which change over time. These variations may impact significantly processing requirements for each continuous query. This thesis takes place in the ANR project Socioplug (ANR-13-INFR-0003). In this context, we consider a collaborative platform for stream processing. Each user can submit multiple continuous queries and contributes to the execution support of the platform. However, as each processing unit supporting treatments has limited resources in terms of CPU and memory, a significant increase in input rate may cause the congestion of the system. The problem is then how to adjust dynamically

Thèse

Soutenue le 29/06/2018

Établissement :
INSA Lyon

Contact :
Philippe.Lamarre@liris.cnrs.fr

resource usage to processing requirements for each continuous query ? It raises several challenges : i) how to detect a need of reconfiguration ? ii) when reconfiguring the system to avoid its congestion at runtime ? In this work, we are interested by the different processing steps involved in the treatment of a continuous query over a distributed infrastructure. From this global analysis, we extract mechanisms enabling dynamic adaptation of resource usage for each continuous query. We focus on automatic parallelization, or auto-parallelization, of operators composing the execution plan of a continuous query. We suggest an original approach based on the monitoring of operators and an estimation of processing requirements in near future. Thus, we can increase (scale-out), or decrease (scale-in) the parallelism degree of operators in a proactive many such as resource usage fits to processing requirements dynamically. Compared to a static configuration defined by an expert, we show that it is possible to avoid the congestion of the system in many cases or to delay itin most critical cases. Moreover, we show that resource usage can be reduced significantly while delivering equivalent throughput and result quality. We suggest also to combine this approach with complementary mechanisms for dynamic adaptation of continuous queries at runtime. These different approaches have been implemented within a widely used DSMS and have been tested over multiple and reproducible micro-benchmarks.

Publications

Revue nationale avec comité de lecture

- Cécile Favre, Chloé Artaud, Clément Duffau, Ophélie Fraisier & Roland Kotto-Kombi (2017). « Forum Jeunes Chercheurs à Inforsid 2016 ». Revue des Sciences et Technologies de l'Information - Série ISI : Ingénierie des Systèmes d'Information, vol. 22, n°2, pp. 121-147. HAL : hal-01914115.
- Cécile Favre, Chloé Artaud, Clément Duffau, Ophélie Fraisier & Roland Kotto-Kombi (2017). « Forum Jeunes Chercheurs à Inforsid 2016 ». Ingénierie des Systèmes d'Information, vol. 22, n°2, pp. 121-147. HAL : hal-01827542.

Conférences internationales avec comité de lecture

- Roland Kotto-Kombi, Nicolas Lumineau & Philippe Lamarre (2017). « A Preventive Auto-Parallelization Approach for Elastic Stream Processing ». 37th (IEEE) International Conference on Distributed Computing Systems, 8 juin 2017, Atlanta (États-Unis), pp. 1532-1542. doi : 10.1109/ICDCS.2017.253. HAL : hal-01585096.

Conférences nationales avec comité de lecture

- Roland Kotto-Kombi, Nicolas Lumineau & Philippe Lamarre (2017). « Approche préventive pour une gestion élastique du traitement parallèle et distribué de flux de données ». EGC 2017, 27 janvier 2017, Grenoble (France), pp. 57-68. HAL : hal-01460709.

Autres

- Roland Kotto-Kombi, Nicolas Lumineau, Philippe Lamarre & Yves Caniou (2015). « Parallel and Distributed Stream Processing: Systems Classification and Specific Issues ». HAL : hal-01215287.

Jury

Elisa FROMONT, PR, Université de Rennes 1	Rapporteur
Laurent HEUTTE, PR, Université de Rouen	Rapporteur
Matthieu CORD, PR, Sorbonne Université	Examineur
Camille COUPRIE, Docteur, Facebook AI Research	Examineur
Christian WOLF, MC, INSA-Lyon	Co-directeur de thèse
Christopher KERMORVANT, Docteur, TEKLIA	Co-directeur de thèse

Résumé

Lire automatiquement le texte présent dans les documents permet de rendre accessible les informations qu'ils contiennent. Pour réaliser la transcription de pages complètes, la localisation des lignes de texte est une étape cruciale. Les méthodes traditionnelles de détection de lignes, basées sur des approches de traitement d'images, peinent à généraliser à des jeux de données hétérogènes. Pour cela, nous proposons dans cette thèse une approche par réseaux de neurones profonds. Nous avons d'abord proposé une approche de segmentation mono-dimensionnelle des paragraphes de texte en lignes à l'aide d'une technique inspirée des modèles de reconnaissance, où une classification temporelle connexionniste (CTC) est utilisée pour aligner implicitement les séquences. Ensuite, nous proposons un réseau qui prédit directement les coordonnées des boîtes englobant les lignes de texte. L'ajout d'un terme de confiance à ces boîtes hypothèses permet de localiser un nombre variable d'objets. Nous proposons une prédiction locale des objets afin de partager les paramètres entre les localisations et, ainsi, de multiplier les exemples d'objets vus par chaque prédicteur de boîte lors de l'entraînement. Cela permet de compenser la taille restreinte des jeux de données utilisés. Pour récupérer les informations contextuelles permettant de prendre en compte la structure du document, nous ajoutons, entre les couches convolutionnelles, des couches récurrentes LSTM multi-dimensionnelles. Nous proposons trois stratégies de reconnaissance pleine page qui permettent de tenir compte du besoin important de précision au niveau des positions et nous montrons, sur la base hétérogène Maudor, la performance de notre approche pour des documents multilingues pouvant être manuscrits et imprimés. Nous nous comparons favorablement à des méthodes issues de l'état de l'art. La visualisation des concepts appris par nos neurones permet de souligner la capacité des couches récurrentes à apporter l'information contextuelle.

Abstract

Being able to automatically read the texts written in documents, both printed and handwritten, makes it possible to access the information they convey. In order to realize full page text transcription, the detection and localization of the text lines is a crucial step. Traditional methods tend to use image processing based approaches, but they hardly generalize to very heterogeneous datasets. In this thesis, we propose to use a deep neural network based approach. We first propose a mono-dimensional segmentation of text paragraphs into lines that uses a technique inspired by the text recognition models. The connexionist temporal classification (CTC) method is used to implicitly align the sequences. Then, we propose a neural network that directly predicts the coordinates of the boxes bounding the text lines. Adding a confidence prediction to these hypothesis boxes enables to locate a varying number of objects. We propose to predict the objects locally in order to share the network parameters between the locations and to increase the number of different objects that each single box predictor sees during training. This

Thèse

Soutenue le 28/05/2018

Établissement :
INSA Lyon

Contact :
Christian.Wolf@Liris.cnrs.fr

compensates the rather small size of the available datasets. In order to recover the contextual information that carries knowledge on the document layout, we add multi-dimensional LSTM recurrent layers between the convolutional layers of our networks. We propose three full page text recognition strategies that tackle the need of high preciseness of the text line position predictions. We show on the heterogeneous Maudor dataset how our methods perform on documents that can be printed or handwritten, in French, English or Arabic and we favorably compare to other state of the art methods. Visualizing the concepts learned by our neurons enables to underline the ability of the recurrent layers to convey the contextual information.

Publications

Revue internationale avec comité de lecture

- Bastien Moysset, Christopher Kermorvant & Christian Wolf (2018). « Learning to detect, localize and recognize many text objects in document images from few examples ». International Journal on Document Analysis and Recognition. doi : 10.1007/s10032-018-0305-2. HAL : hal-01800636.

Conférences internationales avec comité de lecture

- Bastien Moysset, Christopher Kermorvant & Christian Wolf (2017). « Full-Page Text Recognition: Learning Where to Start and When to Stop ». International Conference on Document Analysis and Recognition, 15 novembre 2017, Kyoto (Japon). HAL : hal-01563124.
- Bastien Moysset, Jérôme Louradour, Christopher Kermorvant & Christian Wolf (2016). « Learning text-line localization with shared and local regression neural networks ». International Conference on Frontiers in Handwriting Recognition, 26 octobre 2016, Shenzhen (Chine). HAL : hal-01345713.
- Bastien Moysset, Pierre Adam, Christian Wolf & Jérôme Louradour (2015). « Space Displacement Localization Neural Networks to locate origin points of handwritten text lines in historical documents ». ICDAR 2015 Workshop on Historical Document Imaging and Processing, 22 août 2015, Nancy (France). HAL : hal-01178342.
- Bastien Moysset, Christopher Kermorvant, Christian Wolf & Jérôme Louradour (2015). « Paragraph text segmentation into lines with Recurrent Neural Networks ». International Conference on Document Analysis and Recognition, 26 août 2015, Tunisia (Tunisie). HAL : hal-01151760.

Autres

- Bastien Moysset, Christopher Kermorvant & Christian Wolf (2017). « Pose-conditioned Spatio-Temporal Attention for Human Action Recognition ». ArXiv : 1703.10106. HAL : hal-01593572.

Jury

Marianne CLAUSEL, PR, Université de Lorraine,
Christophe GONZALES, PR, Université Paris 6,
Hanane AZZAG, MC, Université Paris 13,
Véronique EGLIN, PR, INSA Lyon,
Jesse READ, MC, École Polytechnique,
Marc SEBBAN, PR, Université de St-Etienne,
Alexandre AUSSEM, PR, Université Lyon 1,
Haytham ELGHAZEL, MC, Polytech Lyon,

Rapporteuse
Rapporteur
Examinatrice
Examinatrice
Examinateur
Examinateur
Directeur de thèse
Co-Directeur de thèse

Résumé

Les méthodes ensemblistes constituent un sujet de recherche très populaire au cours de la dernière décennie. Leur succès découle en grande partie de leurs solutions attrayantes pour résoudre différents problèmes d'apprentissage intéressants parmi lesquels l'amélioration de l'exactitude d'une prédiction, la sélection de variables, l'apprentissage de métrique, le passage à l'échelle d'algorithmes inductifs, l'apprentissage de multiples jeux de données physiques distribués, l'apprentissage de flux de données soumis à une dérive conceptuelle, etc... Dans cette thèse nous allons dans un premier temps présenter une comparaison empirique approfondie de 19 algorithmes ensemblistes d'apprentissage supervisé proposé dans la littérature sur différents jeux de données de référence. Non seulement nous allons comparer leurs performances selon des métriques standards de performances (Exactitude, AUC, RMS) mais également nous analyserons leur diagrammes kappa-erreur, la calibration et les propriétés biais-variance. Nous allons aborder ensuite la problématique d'amélioration des ensembles de modèles par la sélection dynamique d'ensembles (dynamic ensemble selection, DES). La sélection dynamique est un sous-domaine de l'apprentissage ensembliste où pour une donnée d'entrée x , le meilleur sous-ensemble en terme de taux de réussite est sélectionné dynamiquement. L'idée derrière les approches DES est que différents modèles ont différentes zones de compétence dans l'espace des instances. La plupart des méthodes proposées estime l'importance individuelle de chaque classifieur faible au sein d'une zone de compétence habituellement déterminée par les plus proches voisins dans un espace euclidien. Nous proposons et étudions dans cette thèse deux nouvelles approches DES. La première nommée ST-DES est conçue pour les ensembles de modèles à base d'arbres de décision. Cette méthode sélectionne via une métrique supervisée interne à l'arbre, idée motivée par le problème de la malédiction de la dimensionnalité : pour les jeux de données avec un grand nombre de variables, les métriques usuelles telle la distance euclidienne sont moins pertinentes. La seconde approche, PCC-DES, formule la problématique DES en une tâche d'apprentissage multi-label avec une fonction coût spécifique. Ici chaque label correspond à un classifieur et une base multi-label d'entraînement est constituée sur l'habilité de chaque classifieur de classer chaque instance du jeu de données d'origine. Cela nous permet d'exploiter des récentes avancées dans le domaine de l'apprentissage multi-label. PCC-DES peut être utilisé pour les approches ensemblistes homogènes et également hétérogènes. Son avantage est de prendre en compte explicitement les corrélations entre les prédictions des classifieurs. Ces algorithmes sont testés sur un éventail de jeux de données de référence et les résultats démontrent leur efficacité faces aux dernières alternatives de l'état de l'art

Thèse

Soutenue le 04/05/2018

Établissement :
Université Claude Bernard Lyon 1

Contact :
Alexandre.Aussem@liris.cnrs.fr

Prise en compte de la qualité des données lors de l'extraction et de la sélection d'évolutions dans les séries temporelles de champs de déplacements en imagerie satellitaire

Abstract

Ensemble methods has been a very popular research topic during the last decade. Their success arises largely from the fact that they offer an appealing solution to several interesting learning problems, such as improving prediction accuracy, feature selection, metric learning, scaling inductive algorithms to large databases, learning from multiple physically distributed data sets, learning from concept-drifting data streams etc. In this thesis, we first present an extensive empirical comparison between nineteen prototypical supervised ensemble learning algorithms, that have been proposed in the literature, on various benchmark data sets. We not only compare their performance in terms of standard performance metrics (Accuracy, AUC, RMS) but we also analyze their kappa-error diagrams, calibration and bias-variance properties. We then address the problem of improving the performances of ensemble learning approaches with dynamic ensemble selection (DES). Dynamic pruning is the problem of finding given an input x , a subset of models among the ensemble that achieves the best possible prediction accuracy. The idea behind DES approaches is that different models have different areas of expertise in the instance space. Most methods proposed for this purpose estimate the individual relevance of the base classifiers within a local region of competence usually given by the nearest neighbours in the euclidean space. We propose and discuss two novel DES approaches. The first, called ST-DES, is designed for decision tree based ensemble models. This method prunes the trees using an internal supervised tree-based metric; it is motivated by the fact that in high dimensional data sets, usual metrics like euclidean distance suffer from the curse of dimensionality. The second approach, called PCC-DES, formulates the DES problem as a multi-label learning task with a specific loss function. Labels correspond to the base classifiers and multi-label training examples are formed based on the ability of each classifier to correctly classify each original training example. This allows us to take advantage of recent advances in the area of multi-label learning. PCC-DES works on homogeneous and heterogeneous ensembles as well. Its advantage is to explicitly capture the dependencies between the classifiers predictions. These algorithms are tested on a variety of benchmark data sets and the results demonstrate their effectiveness against competitive state-of-the-art alternatives

Jury

Florence TUPIN, PR, Télécom ParisTech,	Présidente
Élisa FROMONT, PR, Université de Rennes 1,	Rapportrice
Bruno CRÉMILLEUX, PR, Université de Caen Normandie,	Rapporteur
Dino IENCO, PR, IRSTEA,	Examineur
Nicolas MÉGER, MC, Université Savoie Mont Blanc,	Directeur de thèse
Christophe RIGOTTI, MC, INSA Lyon,	Co-Directeur de thèse
Emmanuel TROUVÉ, PR, Univ. Savoie Mont Blanc,	Co-Directeur de thèse

Résumé

Ce travail de thèse traite de la découverte de connaissances à partir de Séries Temporelles de Champs de Déplacements (STCD) obtenues par imagerie satellitaire. De telles séries occupent aujourd'hui une place centrale dans l'étude et la surveillance de phénomènes naturels tels que les tremblements de terre, les éruptions volcaniques ou bien encore le déplacement des glaciers. En effet, ces séries sont riches d'informations à la fois spatiales et temporelles et peuvent aujourd'hui être produites régulièrement à moindre coût grâce à des programmes spatiaux tels que le programme européen Copernicus et ses satellites phares Sentinel. Nos propositions s'appuient sur l'extraction de motifs Séquentiels Fréquents Groupés (SFG). Ces motifs, à l'origine définis pour l'extraction de connaissances à partir des Séries Temporelles d'Images Satellitaires (STIS), ont montré leur potentiel dans de premiers travaux visant à dépouiller une STCD. Néanmoins, ils ne permettent pas d'utiliser les indices de confiance intrinsèques aux STCD et la méthode de swap randomisation employée pour sélectionner les motifs les plus prometteurs ne tient pas compte de leurs complémentarités spatio-temporelles, chaque motif étant évalué individuellement. Notre contribution est ainsi double. Une première proposition vise tout d'abord à associer une mesure de fiabilité à chaque motif en utilisant les indices de confiance. Cette mesure permet de sélectionner les motifs portés par des données qui sont en moyenne suffisamment fiables. Nous proposons un algorithme correspondant pour réaliser les extractions sous contrainte de fiabilité. Celui-ci s'appuie notamment sur une recherche efficace des occurrences les plus fiables par programmation dynamique et sur un élagage de l'espace de recherche grâce à une stratégie de push partiel, ce qui permet de considérer des STCD conséquentes. Cette nouvelle méthode a été implémentée sur la base du prototype existant SITS-P2miner, développé au sein du LISTIC et du LIRIS pour extraire et classer des motifs SFG. Une deuxième contribution visant à sélectionner les motifs les plus prometteurs est également présentée. Celle-ci, basée sur un critère informationnel, permet de prendre en compte à la fois les indices de confiance et la façon dont les motifs se complètent spatialement et temporellement. Pour ce faire, les indices de confiance sont interprétés comme des probabilités, et les STCD comme des bases de données probabilistes dont les distributions ne sont que partielles. Le gain informationnel associé à un motif est alors défini en fonction de la capacité de ses occurrences à compléter/affiner les distributions caractérisant les données. Sur cette base, une heuristique est proposée afin de sélectionner des motifs informatifs et complémentaires. Cette méthode permet de fournir un ensemble de motifs faiblement redondants et donc plus faciles à interpréter que ceux fournis par swap randomisation. Elle a été implémentée au sein d'un prototype dédié. Les deux propositions sont évaluées à la fois quantitativement et qualitativement en utilisant une STCD de référence couvrant des glaciers du Groenland construite à partir de données optiques Landsat. Une autre STCD que nous avons construite à partir de données radar TerraSAR-X couvrant le massif du Mont-Blanc est également

Hoang Viet Tuan
Nguyen

Thèse

Soutenue le 10/10/2018

Établissement :
INSA Lyon

Contact :
Christophe.Rigotti@liris.cnrs.fr

utilisée. Outre le fait d'être construites à partir de données et de techniques de télédétection différentes, ces séries se différencient drastiquement en termes d'indices de confiance, la série couvrant le massif du Mont-Blanc se situant à des niveaux de confiance très faibles. Pour les deux STCD, les méthodes proposées ont été mises en œuvre dans des conditions standards au niveau consommation de ressources (temps, espace), et les connaissances des experts sur les zones étudiées ont été confirmées et complétées.

Abstract

This PhD thesis deals with knowledge discovery from Displacement Field Time Series (DFTS) obtained by satellite imagery. Such series now occupy a central place in the study and monitoring of natural phenomena such as earthquakes, volcanic eruptions and glacier displacements. These series are indeed rich in both spatial and temporal information and can now be produced regularly at a lower cost thanks to spatial programs such as the European Copernicus program and its famous Sentinel satellites. Our proposals are based on the extraction of grouped frequent sequential patterns. These patterns, originally defined for the extraction of knowledge from Satellite Image Time Series (SITS), have shown their potential in early work to analyze a DFTS. Nevertheless, they cannot use the confidence indices coming along with DFTS and the swap method used to select the most promising patterns does not take into account their spatio-temporal complementarities, each pattern being evaluated individually. Our contribution is thus double. A first proposal aims to associate a measure of reliability with each pattern by using the confidence indices. This measure allows to select patterns having occurrences in the data that are on average sufficiently reliable. We propose a corresponding constraint-based extraction algorithm. It relies on an efficient search of the most reliable occurrences by dynamic programming and on a pruning of the search space provided by a partial push strategy. This new method has been implemented on the basis of the existing prototype SITS-P2miner, developed by the LISTIC and LIRIS laboratories to extract and rank grouped frequent sequential patterns. A second contribution for the selection of the most promising patterns is also made. This one, based on an informational criterion, makes it possible to take into account at the same time the confidence indices and the way the patterns complement each other spatially and temporally. For this aim, the confidence indices are interpreted as probabilities, and the DFTS are seen as probabilistic databases whose distributions are only partial. The informational gain associated with a pattern is then defined according to the ability of its occurrences to complete/refine the distributions characterizing the data. On this basis, a heuristic is proposed to select informative and complementary patterns. This method provides a set of weakly redundant patterns and therefore easier to interpret than those provided by swap randomization. It has been implemented in a dedicated prototype. Both proposals are evaluated quantitatively and qualitatively using a reference DFTS covering Greenland glaciers constructed from Landsat optical data. Another DFTS that we built from TerraSAR-X radar data covering the Mont-Blanc massif is also used. In addition to being constructed from different data and remote sensing techniques, these series differ drastically in terms of confidence indices, the series covering the Mont-Blanc massif being at very low levels of confidence. In both cases, the proposed methods operate under standard conditions of resource consumption (time, space), and experts' knowledge of the studied areas is confirmed and completed.

Publications

Revue internationale avec comité de lecture

- Tuan Nguyen, Nicolas Méger, Christophe Rigotti, Catherine Pothier, Emmanuel Trouvé, Noel Gourmelen & Jean-Louis Mugnier (2018). « A pattern-based method for handling confidence measures while mining satellite displacement field time series. Application to Greenland ice sheet and Alpine glaciers ». *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, n°11, pp. 4390-4402. doi : 10.1109/JSTARS.2018.2874499. HAL : hal-01912708.

Conférences internationales avec comité de lecture

- Tuan Nguyen, Nicolas Méger, Christophe Rigotti, Catherine Pothier, Emmanuel Trouvé & Jean-Louis Mugnier (2018). « Finding Complementary and Reliable Patterns in Displacement Field Time Series of Alpine Glaciers ». *IGARSS 2018 - IEEE International Geoscience and Remote Sensing Symposium*, 27 juillet 2018, Valencia (Espagne), pp. 4213-4216. HAL : hal-01868749.
- Tuan Nguyen, Nicolas Méger, Christophe Rigotti, Catherine Pothier, Emmanuel Trouvé & Noel Gourmelen (2017). « Handling coherence measures of displacement field time series: Application to Greenland ice sheet glaciers ». *9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp)*, 29 septembre 2017, Brugge (Belgique), pp. 1-4. doi : 10.1109/Multi-Temp.2017.8035228. HAL : hal-01591461.
- Tuan Nguyen, Nicolas Méger, Christophe Rigotti, Catherine Pothier & Rémi Andréoli (2016). « SITS-P2miner: Pattern-Based Mining of Satellite Image Time Series ». *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD) Demo*, 23 septembre 2016, Riva del Garda (Italie), pp. 63-66. doi : 10.1007/978-3-319-46131-1_14. HAL : hal-01367996.

Diana Nurbakova

Jury

Patrice BELLOT, PR, Université Aix Marseille - LIS,
Anne BOYER, PR, Université de Lorraine - LORIA,
Josiane MOTHE, PR, ESPE Toulouse Midi Pyrénées - IRIT,
Ilya MARKOV, Docteur, Université d'Amsterdam,
Sylvie CALABRETTO, PR, INSA Lyon - LIRIS,
Léa LAPORTE, MC, INSA Lyon - LIRIS,

Rapporteur
Rapporteuse
Examinatrice
Examineur
Directrice de thèse
Co-directrice

Résumé

Les événements distribués, se déroulant sur plusieurs jours et/ou sur plusieurs lieux, tels que les conventions, festivals ou croisières, sont de plus en plus populaires ces dernières années et attirant des milliers de participants. Les programmes de ces événements sont généralement très denses, avec un grand nombre d'activités se déroulant en parallèle. Ainsi, choisir les activités à entreprendre est devenu un véritable défi pour les participants. Les systèmes de recommandation peuvent constituer une solution privilégiée dans ce genre d'environnement. De nombreux travaux en recommandation se sont concentrés sur la recommandation personnalisée d'objets spatiaux (points d'intérêts immuables dans le temps ou événements éphémères) indépendants les uns des autres. Récemment, la communauté scientifique s'est intéressée à la recommandation de séquences de points d'intérêts, exploitant des motifs comportementaux des utilisateurs et incorporant des contraintes spatio-temporelles pour recommander un itinéraire de points d'intérêts. Néanmoins, très peu de travaux se sont intéressés à la problématique de la recommandation de séquence d'activités, problème plus difficile du fait du caractère éphémère des objets à recommander. Dans cette thèse, nous proposons tout d'abord une formalisation du problème de la recommandation de séquences d'activités. Dans ce cadre, nous proposons et discutons une classification des types d'influences pouvant avoir un impact sur l'estimation de l'intérêt des utilisateurs dans les activités. Ensuite, nous proposons ANASTASIA, une approche de recommandation personnalisée de séquences d'activités lors des événements distribués. Notre approche est basée sur trois composants clés : (1) l'estimation de l'intérêt d'un utilisateur pour une activité, prenant en compte différentes influences, (2) l'intégration de motifs comportementaux d'utilisateurs basés sur leurs historiques d'activités et (3) la construction d'un planning ou séquence d'activités prenant en compte les contraintes spatio-temporelles de l'utilisateur et des activités. Nous explorons ainsi des méthodes issues de l'apprentissage de séquences et de l'optimisation discrète pour résoudre le problème. Enfin, nous démontrons le manque de jeu de données librement accessibles pour l'évaluation des algorithmes de recommandation d'événements et de séquences d'événements. Nous pallions à ce problème en proposant deux jeux de données, librement accessibles, que nous avons construits au cours de la thèse: Fantasy_db et DEvIR. Fantasy_db comporte des données de participation à des événements lors d'une croisière, recueillies lors d'une étude utilisateur, tandis que DEvIR réunit des données de participation au Comic Con de San Diego, convention majeure dans le domaine.

Abstract

Multi-day events such as conventions, festivals, cruise trips, to which we refer to as distributed events, have become very popular in recent years, attracting hundreds or thousands of participants. Their programs are usually very dense,

Thèse

Soutenue le 13/12/2018

Établissement :
INSA Lyon

Contact :
Sylvie.Calabretto@liris.cnrs.fr

making it challenging for the attendees to make a decision which events to join. Recommender systems appear as a common solution in such an environment. While many existing solutions deal with personalized recommendation of single items, recent research focuses on the recommendation of consecutive items that exploits user's behavioral patterns and relations between entities, and handles geographical and temporal constraints. In this thesis, we first formulate the problem of recommendation of activity sequences, classify and discuss the types of influence that have an impact on the estimation of the user's interest in items. Second, we propose an approach (ANASTASIA) to solve this problem, which aims at providing an integrated support for users to create a personalized itinerary of activities. ANASTASIA brings together three components, namely: (1) estimation of the user's interest in single items, (2) use of sequential influence on activity performance, and (3) building of an itinerary that takes into account spatio-temporal constraints. Thus, the proposed solution makes use of the methods based on sequence learning and discrete optimization. Moreover, stating the lack of publicly available datasets that could be used for the evaluation of event and itinerary recommendation algorithms, we have created two datasets, namely: (1) event attendance on board of a cruise (Fantasy_db) based on a conducted user study, and (2) event attendance at a major comic book convention (DEVIR). This allows to perform evaluation of recommendation methods, and contributes to the reproducibility of results.

Publications

Conférences internationales avec comité de lecture

- Diana Nurbakova (2018). « Recommendation of Activity Sequences during Distributed Events ». 26th Conference on User Modeling, Adaptation and Personalization (UMAP 2018), 11 juillet 2018, Singapore (Singapour). doi : 10.1145/3209219.3213592. HAL : hal-01875718.
- Diana Nurbakova, Jie Yang, Iván Cantador, María Cortés-Cediel & Alessandro Bozzon (2017). « Recommender systems for citizens: the CitRec'17 workshop manifesto ». the International Workshop on Citizens for Recommender Systems, CitRec@RecSys 2017, 31 août 2017, Como (Italie). doi : 10.1145/3127325.3177871. HAL : hal-01723007.
- Diana Nurbakova, Léa Laporte, Sylvie Calabretto & Jérôme Gensel (2017). « Itinerary Recommendation for Cruises: User Study ». RecTour 2017: 2nd Workshop on Recommenders in Tourism, 27 août 2017, Como (Italie), pp. 31-34. HAL : hal-01577228.
- Diana Nurbakova, Léa Laporte, Sylvie Calabretto & Jérôme Gensel (2017). « Recommendation of Short-Term Activity Sequences During Distributed Events ». International Conference on Computational Science (ICCS 2017), 14 juin 2017, Zürich (Suisse), pp. 2069-2078. doi : 10.1016/j.procs.2017.05.154. HAL : hal-01568796.
- Diana Nurbakova, Léa Laporte, Sylvie Calabretto & Jérôme Gensel (2017). « Users Psychological Profiles for Leisure Activity Recommendation: User Study ». Proceedings of CitRec 2017: Workshop on Recommender Systems for Citizens, 31 août 2017, Como (Italie). doi : 10.1145/3127325.3127328. HAL : hal-01574704.

Conférences nationales avec comité de lecture

- Diana Nurbakova, Léa Laporte, Sylvie Calabretto & Jérôme Gensel (2018). « Recommendation de séquences d'activités lors d'évènements distribués ».

Conférence en Recherche d'Informations et Applications (CORIA) 2018, 15th French Information Retrieval Conference,, 16 mai 2018, Rennes (France). HAL : hal-01936776.

- Diana Nurbakova, Léa Laporte, Sylvie Calabretto & Jérôme Gensel (2016). « ANASTASIA : recommandation de séquences d'activités spatio-temporelles ». Rencontres Jeunes Chercheurs en Recherche d'Information (RJCRI CORIA-CIFED), 11 juin 2016, Toulouse (France), p. 10. HAL : hal-01328062.

Autres

- Diana Nurbakova, Léa Laporte, Sylvie Calabretto & Jérôme Gensel (2016). « ANASTASIA: recommendation of spatio-temporal activities sequences ». Rencontres Jeunes Chercheurs en Recherche d'Information (RJCRI CORIA-CIFED), 11 mars 2016, Toulouse (France). Poster. HAL : hal-01328145.
- Diana Nurbakova, Sylvie Calabretto, Léa Laporte & Jérôme Gensel (2016). « STAS : Recommendation of spatio-temporal activity sequences ». Journée Scientifique de l'Arc 6, 24 novembre 2016, Lyon (France). Poster. HAL : hal-01575728.
- Diana Nurbakova, Sylvie Calabretto, Léa Laporte & Jérôme Gensel (2015). « Recommendation de séquences d'activités en contexte mobile et dynamique ». Journée scientifique de l'ARC 6, 20 novembre 2015, Grenoble (France). Poster. HAL : hal-01328134.

Jury

Ouajdi KORBAA, Université de Sousse	Examineur
Hajer BAAZAOU, Université de la Mannouba	Examinatrice
Nadjet KAMEL, Université de Sétif	Examinatrice
Allel HADJALI, ENSMA Poitiers	Rapporteur
Ikram AAMOUS, Université de Sfax	Rapporteur
Djamal BENSLIMANE, Univ. Claude Bernard Lyon 1	Directeur de thèse
Mohamed Nazih OMRI, Université de Sousse,	Directeur de thèse
Karim BENOURET, Univ. Claude Bernard Lyon 1	Co-directeur de thèse

Résumé

Il est communément admis que la production de données connaît, depuis plusieurs années, un développement spectaculaire en raison de la multiplication des nouvelles technologies telles que les réseaux sociaux, les nouveaux appareils mobiles, les compteurs intelligents, les capteurs et le cloud computing. De fait, cette explosion de données devrait se poursuivre et même accélérer. S'interroger sur la façon dont on devrait traiter cette masse de qui devient de plus en plus variée, complexe et moins structurée, est alors essentiel. DaaS (Data As A Service) peut être définie comme l'approvisionnement, la gestion et la fourniture de données présentées dans un format immédiatement consommable aux utilisateurs professionnels des organisations en tant que service. Les données retournées par ces services se caractérisent généralement par l'incertitude et l'hétérogénéité. Nombreux sont les approches qui traitent les données selon le cycle de vie du service Web qui repose sur 6 phases à savoir la création, la sélection, la découverte, la modélisation, l'invocation et la composition des services, dans le but de résoudre le problème de volume de données, de son hétérogénéité ou de sa vitesse d'évolution. En revanche, il y a très peu d'approches qui s'intéressent à la qualité de données et au traitement de son incertitude dans le Web. Nous nous sommes naturellement intéressés, dans cette thèse, à la question des services Web dans un contexte de systèmes distribués et hétérogènes. La principale contribution à apporter dans le cadre de ce travail de recherche est d'étudier la composition de services et/ou de ressources Web et l'indexation de données dans un contexte incertain. Dans un premier temps, au travers des apports de la littérature, le cadre théorique relatif aux spécificités du concept de service DaaS incertain, est présenté en adoptant la théorie possibiliste. Le problème de la composition de services Web et l'impact de l'incertitude, qui peut être associée à la sortie d'un service, sur les processus de sélection et de composition des services sont explicités. Pour ce faire, nous avons proposé une approche possibiliste afin de modéliser l'incertitude des données renvoyées par des services incertains. Plus précisément, nous avons étendu les normes de description de service Web (par exemple, WSDL) pour représenter les degrés d'incertitude des sorties. Nous avons également étendu le processus d'invocation de service pour prendre en compte l'incertitude des données d'entrée. Cette extension est basée sur la théorie des mondes possibles utilisée dans les bases de données possibilistes. Nous avons également mis en avant un ensemble d'opérateurs de composition, sensibles aux valeurs d'incertitude, dans le but d'orchestrer des services de données incertains. Dans un deuxième temps, nous avons étudié l'impact de l'incertitude sur la représentation et la manipulation des ressources Web. Nous avons défini le concept de ressource Web incertaine et proposé des mécanismes de composition de ressources. Pour ce faire, un modèle de description de l'incertitude à travers le concept de ressource Web incertaine a été présenté. Celui-ci est basé sur un modèle probabiliste où

Thèse

Soutenue le 29/08/2018

Établissement :
Université Claude Bernard Lyon 1

Contact :
Djamal.Benslimane@liris.cnrs.fr

chaque ressource peut avoir plusieurs représentations possibles, avec une certaine probabilité. Enfin, et dans un dernier temps, nous avons proposé des méthodes d'indexation documentaire des données de type Big Data. Au commencement, nous avons adopté une approche d'indexation syntaxique de données incertaines, ensuite, nous avons suivi une méthode d'indexation sémantique incertaine. Enfin, et pour booster cette démarche, nous avons proposé une méthode hybride d'indexation dans un contexte incertain.

Abstract

It is commonly accepted that data production has several years, a spectacular development due to the multiplication of new technologies such as social networks, new mobile devices, smart meters, sensors and cloud computing. In fact, this explosion of data should continue and even accelerate. Wondering how we should deal with this mass of which becomes more and more varied, complex and less structured, is so essential. DaaS (Data As A Service) can be defined as the supply, the management and provision of data presented in an immediately consumable format business users of organizations as a service. The returned data These services are generally characterized by uncertainty and heterogeneity. There are many approaches that process data according to the life cycle of the service Web based on 6 phases namely creation, selection, discovery, modeling, invocation and composition of services, in order to solve the volume problem data, its heterogeneity or its speed of evolution. On the other hand, there is very little approaches that address data quality and uncertainty processing in the Web. Naturally, in this thesis we have been interested in the question of Web in a context of distributed and heterogeneous systems. The main contribution to bring in this research work is to study the composition of services and / or web resources and indexing data in an uncertain context. As a first step, through the contributions of the literature, the theoretical framework on the specificities of the DaaS concept of uncertain service, is presented by adopting the possibilistic theory. The problem of web services composition and the impact of uncertainty, which can be associated with the exit of a service, on the processes of selection and composition of the services are explained. To do this, we proposed an approach Possibilist to model the uncertainty of data returned by uncertain services. Specifically, we have extended Web service description standards (for example, example, WSDL) to represent the uncertainty levels of the outputs. We also have extended the service invocation process to take into account the uncertainty of input data. This extension is based on the theory of possible worlds used in the possibilistic databases. We also put forward a set of composition operators, sensitive to uncertainty values, in order to orchestrate uncertain data services. In a second step, we studied the impact of uncertainty on the representation and manipulating web resources. We defined the resource concept Web uncertain and proposed mechanisms of composition of resources. To do this, a pattern of uncertainty description through the concept of uncertain web resource has been presented. It is based on a probabilistic model where each resource can have several possible representations, with a certain probability. Finally, and finally, we have proposed methods of documentary indexingBig data type data. Initially, we adopted a approach of syntactic indexing of uncertain data, then we followed a uncertain semantic indexing method. Finally, and to boost this approach, we proposed a hybrid method of indexing in an uncertain context.

Publications

Revue internationale avec comité de lecture

- Asma Omri, Karim Benouaret, Djamel Benslimane & Mohamed Nazih Omri (2018). « Towards an Understanding of Cloud Services under Uncertainty: A Possibilistic Approach ». *International Journal of Approximate Reasoning*, vol. 98, pp. 146-162. doi : 10.1016/j.ijar.2018.04.004. HAL : hal-01857477.

Conférences internationales avec comité de lecture

- Asma Omri, Karim Benouaret, Mohammed Nazih Omri & Djamel Benslimane (2017). « Toward a New Model of Indexing Big Uncertain Data ». *International Conference on Management of Digital EcoSystems (MEDES)*, 7 novembre 2017, Bangkok (Thaïlande), pp. 93-98. doi : 10.1145/3167020.3167034. HAL : hal-01857547.
- Asma Omri, Karim Benouaret, Mohamed Nazih Omri & Djamel Benslimane (2016). « Querying Data Services in an Uncertain Environment: A Possibilistic-Based Approach ». *International Conference on Signal-Image Technology Internet-Based Systems (SITIS 2016)*, 1 décembre 2016, Naples (Italie), pp. 246-251. doi : 10.1109/SITIS.2016.47. HAL : hal-01857558.

Jury

Laure TOUGNE, PR, Université Lumière Lyon 2
Véronique EGLIN, PR, INSA Lyon
Antoine VACAVANT, MC, Université Clément-Ferrant 1
Alain CROUZIL, MC, Université Toulouse 3
Pascal DESARATS, PR, Université Bordeaux 1
Sylvie TREUILLET, MC, Université d'Orléans

Directrice de thèse
Examinatrice
Examinateur
Examinateur
Rapporteur
Rapporteur

Résumé

Objets de collection depuis les temps anciens, de nos jours les pièces de monnaie constituent un marché de plus en plus important. L'évaluation par des experts de l'état de conservation des pièces de monnaie, que nous nommons gradation, joue un rôle important pour déterminer leur valeur sur le marché. Dans le but de grader des pièces de monnaie de manière efficace et objective, la société GENI collabore avec le laboratoire LIRIS, afin d'automatiser le processus de gradation à partir de photos de pièces de monnaie. L'objectif principal de cette thèse est de fournir une aide à la gradation des pièces de monnaie à partir des photos de qualité. Le projet est composé de quatre phases : segmentation des monnaies, identification du type monétaire, détection et reconnaissance du millésime et gradation des monnaies. Dans la première phase, la pièce de monnaie est segmentée de sa photo de manière précise à l'aide d'un modèle paramétrique déformable. Ce dernier permet également d'extraire des caractéristiques de la pièce de monnaie telles que sa taille, son nombre de coins, de pans, etc. Lors de la deuxième phase, nous cherchons dans une base de données le type monétaire de référence correspondant à la pièce de monnaie requête à l'aide de scores de similarité basés sur des graphes. Le premier score se base sur des caractéristiques locales des contours en relief, et le second, qui est semi-global, permet de mettre en évidence des différences de motifs. La troisième phase concerne la reconnaissance du millésime. Il s'agit d'un sujet difficile car les caractères, dans ce contexte, ont un premier plan de couleur très similaire à l'arrière-plan. Après avoir localisé la zone du millésime et l'avoir découpée en imagerie de chiffres, nous proposons une méthode de reconnaissance de chiffres à l'aide de caractéristiques « topologiques ». Enfin, concernant la gradation des monnaies, nous proposons une méthode se basant sur une quantification des « éléments inattendus » comme les rayures et les taches. La pièce de monnaie est d'abord recalée sur une monnaie de référence, puis, nous détectons les « éléments inattendus » significatifs sur des zones d'intérêt. Enfin, concernant les « éléments inattendus » tenus difficiles à repérer individuellement, nous détectons les zones granuleuses à l'aide du Deep Learning. Le résultat obtenu par cette méthode, proche de ce que l'expert réalise « à la main », servira d'aide aux numismates.

Abstract

Coins have been collected and studied since ancient times. Today, coin collection has become a hobby for anyone who wants to participate in. Coin grading is a way of determining the physical condition of a coin to provide an indicator for its market value. For grading coins on a large scale and in a relatively objective way, the GENI company cooperates with the laboratory LIRIS to automate the process by using coin photos. The main objective of this thesis is to grade coins from well-conditioned photos. The project is composed of four steps : coin segmentation from raw photos, monetary type identification, coin date detection and recognition and, coin grading.

Thèse

Soutenue le 05/06/2018

Établissement :
Université Lumière Lyon2

Contact :
Laure.Tougne@liris.cnrs.fr

Claudia Gabriela
Paris Sierra

The first step is to extract coins from raw photos with a high precision. With a deformable geometric model, we segment precisely round coins, many-sided coins, wavyedged coins and holed coins by recognizing their shapes. The second step consists of coin recognition or monetary type identification. We match the query coin to the most identical type reference by using two similarity scores. The first similarity score is based on local features of relief contours. The second similarity score is a semi-global measure that highlights the difference between relief patterns. The third step is to detect and recognize coin dates. However, the fact that such characters have the same color as the background makes traditional optical character recognition methods difficult to apply. After extracting the date zone and cropping it into digit images, we propose a learning-free method to recognize those digits by analyzing their "topological" features. In the last step, the grading process is carried out by quantification of «unexpected elements» such as scratches and dirty marks. The coin to grade is registered to a reference coin. Then, large "unexpected elements" are detected in some regions of interest. However, some "micro-scratches" are difficult to extract individually but all together make a «grainy» surface. To deal with it, we use Deep Learning techniques to classify those grainy zones containing such "micro-scratches". The result of our system, which is close to the manual expert one, is considered as a useful help for numismatists.

Publications

Reuves internationales avec comité de lecture

- Xingyu Pan & Laure Tougne (2017). « A New Database of Digits Extracted from Coins with Hard-to-Segment Foreground for OCR Evaluation ». *Frontiers in information and communication technologies*. doi : 10.3389/fict.2017.00009. HAL : hal-01518293.

Conférences internationales avec comité de lecture

- Xingyu Pan & Laure Tougne (2018). « Image Anaysis and Deep Learning for Aiding Professional Coin Grading ». *International Conference on Image Video Processing and Artificial Intelligence*, 17 août 2018, Shanghai (Chine). doi : 10.1117/12.2500142. HAL : hal-01795304.
- Xingyu Pan & Laure Tougne (2016). « Topology-Based Character Recognition Method for Coin Date Detection ». *18th International Conference on Image Analysis and Processing*, 25 octobre 2016, Paris (France). HAL : hal-01371496.
- Xingyu Pan, Kitti Puritat & Laure Tougne (2014). « A New Coin Segmentation and Graph-Based Identification Method for Numismatic Application ». *10th International Symposium on Visual Computing*, 10 décembre 2014, Las Vegas (États-Unis). doi : 10.1007/978-3-319-14364-4_18. HAL : hal-01371488.

Jury

Sylvain GRAVIER, DR, Institut Fourier, University of Grenoble,	Rapporteur
Bostjan BRESAR, PR, University of Maribor,	Rapporteur
Myriam PREISSMAN, DR, G-SCOP, Grenoble INP,	Examinatrice
Aline PARREAU, CR, University of Lyon,	Examinatrice
Angela BONIFATI, PR, University of Lyon,	Examinatrice
Olivier TOGNI, PR, University of Burgundy,	Examinateur
Éric DUCHÊNE, MC, University of Lyon	directeur de thèse
Paul DORBEC, MC, University of Bordeaux,	co-Directeur de thèse

Résumé

J'ai étudié trois problèmes d'optimisation dans les graphes et les jeux combinatoires. Tout d'abord, les codes identifiants dans les graphes où les sommets font faces à des failles: les codes cherchent à repérer les failles pour les réparer. On s'est intéressé aux codes identifiants dans les graphes circulants en utilisant des plongements de ces graphes dans des grilles infinies. Ensuite, j'ai étudié le jeu de marquage de sommets et le jeu de coloration d'arêtes: ici deux joueurs se font face, le premier cherche à construire une coloration correcte (ou un marquage correct) et le deuxième cherche à l'en empêcher. Pour le jeu de marquage on s'est intéressé aux changements de stratégie gagnante lorsqu'on modifie le graphe. Pour le jeu de coloration d'arêtes on a donné une stratégie gagnante pour le premier joueur pourvu que le graphe considéré admette une certaine décomposition sur les arêtes. On améliore notamment des résultats sur les graphes planaires. Enfin j'ai étudié les jeux à tas purement de casse: deux joueurs à tour de rôle prennent un tas et le cassent en un certain nombre de tas non vides. On s'intéresse aux stratégies gagnantes lorsque les joueurs jouent sur un unique tas contenant n jetons. Ces jeux de pure casse semblent, à l'oeil nu, être réguliers. On a montré que c'est effectivement le cas pour certains et on a donné un test qui permet de déterminer la régularité cas par cas. Un seul cas ne semble pas correspondre à cette régularité: son comportement reste un mystère. En conclusion, je me suis intéressé à trois problèmes bilatéraux qui utilisent différentes méthodes et qui remplissent des propos différents dans le domaine de la combinatoire

Abstract

I studied three optimization problems on graphs and combinatorial games. First, identifying codes were studied : vertices couteract faults. Identifying codes help locate the fault to repair it. We focused on circulant graphs by embedding them on infinite grids. Then, the marking and the coloring games were studied : two player games were one player wants to build something (a proper coloration or a proper marking) and the other wants to prevent the first player from doing so. For the marking game we studied the evolution of the strategy when modifying the graph. For the coloring game we defined a new edge-wise decomposition of graphs and we defined a new strategy on this decomposition that improves known results on planar graphs. In the end, I studied pure breaking games : two players take turns to break a heap of tokens in a given number of non-empty heaps. We focused on winning strategies for the game starting with a unique heap on n tokens. These games seem, on first sight, to be all regular : we showed this is the case for some of them and we gave a test to study one game at a time. Only one of these games does not seem to be regular, its behavior remains a mystery. To sum up, I studied three bilateral problems that use different methods and have different purposes in combinatorics

Thèse

Soutenue le 09/10/2018

Établissement :
Université Claude Bernard Lyon 1

Contact :
Eric.Duchene@liris.cnrs.fr

Publications

Revue internationale avec comité de lecture

- Clément Charpentier, Brice Effantin & Gabrielle Paris (2018). « On the game coloring index of F+-decomposable graphs ». *Discrete Applied Mathematics*, vol. 236, pp. 73-83. HAL : hal-01883462.

Autres

- Antoine Dailly, Eric Duchene, Urban Larsson & Gabrielle Paris (2018). « Partition games are pure breaking games ». ArXiv : 1803.02621. HAL : hal-01723190.

Jury

Celine LOSCOS, PR, Université de Reims
Nicolas HOLZSCHUCH, PR, INRIA Grenoble
Christiane LEMIEUX, PR, Université de Waterloo
Mathias PAULIN, PR, Université de Toulouse
David COEURJOLLY, PR, Université de Lyon
Victor OSTROMOUKHOV, PR, Université de Lyon

Rapporteuse
Rapporteur
Examinatrice
Examineur
Co-Directeur de thèse
Directeur de thèse

Résumé

Lorsque l'on affiche un objet 3D sur un écran d'ordinateur, on transforme cet objet en une image, c.-à.-d en un ensemble de pixels colorés. On appelle Rendu la discipline qui consiste à trouver la couleur à associer à ces pixels. Calculer la couleur d'un pixel revient à intégrer la quantité de lumière arrivant de toutes les directions que la surface renvoie dans la direction du plan image, le tout pondéré par une fonction binaire déterminant si un point est visible ou non. Malheureusement, l'ordinateur ne sait pas calculer des intégrales on a donc deux méthodes possibles : Trouver une expression analytique qui permet de supprimer l'intégrale de l'équation (approche basée statistique). Approximer numériquement l'équation en tirant des échantillons aléatoires dans le domaine d'intégration et en en déduisant la valeur de l'intégrale via des méthodes dites de Monte Carlo. Nous nous sommes ici intéressés à l'intégration numérique et à la théorie de l'échantillonnage. L'échantillonnage est au cœur des problématiques d'intégration numérique. En informatique graphique, il est capital qu'un échantillonneur génère des points uniformément dans le domaine d'échantillonnage pour garantir que l'intégration ne sera pas biaisée. Il faut également que le groupe de points généré ne présente aucune régularité structurelle visible, au risque de voir apparaître des artefacts dit d'aliasage dans l'image résultante. De plus, les groupes de points générés doivent minimiser la variance lors de l'intégration pour converger au plus vite vers le résultat. Il existe de nombreux types d'échantillonneurs que nous classeront ici grossièrement en 2 grandes familles : Les échantillonneurs bruit bleu, qui ont une faible la variance lors de l'intégration tout en générant de groupes de points non structurés. Le défaut de ces échantillonneurs est qu'ils sont extrêmement lents pour générer les points. Les échantillonneurs basse discrédance, qui minimisent la variance lors de l'intégration, génèrent des points extrêmement vite, mais qui présentent une forte structure, générant énormément d'aliasage. Notre travail a été de développer des échantillonneurs hybrides, combinant à la fois bruit bleu et basse discrédance

Abstract

When you display a 3D object on a computer screen, we transform this 3D scene into a 2D image, which is a set of organized colored pixels. We call Rendering all the process that aims at finding the correct color to give those pixels. This is done by integrating all the light rays coming for every directions that the object's surface reflects back to the pixel, the whole being ponderated by a visibility function. Unfortunately, a computer can not compute an integrand. We therefore have two possibilities to solve this issue: We find an analytical expression to remove the integrand (statistic based strategy). Numerically approximate the equation by taking random samples in the integration domain and approximating the integrand value using Monte Carlo methods. Here we focused on numerical integration and sampling theory. Sampling is a fundamental part of numerical integration. A good sampler should generate

Thèse

Soutenu le 07/03/2018

Établissement :
Université Claude Bernard Lyon 1

Contact :
Victor.Ostromoukhov@liris.cnrs.fr

Vincent Primault

points that cover the domain uniformly to prevent bias in the integration and, when used in Computer Graphics, the point set should not present any visible structure, otherwise this structure will appear as artifacts in the resulting image. Furthermore, a stochastic sampler should minimize the variance in integration to converge to a correct approximation using as few samples as possible. There exists many different samplers that we will regroup into two families: Blue Noise samplers, that have a low integration variance while generating unstructured point sets. The issue with those samplers is that they are often slow to generate a pointset. Low Discrepancy samplers, that minimize the variance in integration and are able to generate and enrich a point set very quickly. However, they present a lot of structural artifacts when used in Rendering. Our work aimed at developing hybrid samplers, that are both Blue Noise and Low Discrepancy

Publications

Revue internationale avec comité de lecture

- Hélène Perrier, David Coeurjolly, Feng Xie, Matt Pharr, Pat Hanrahan & Victor Ostromoukhov (2018). « Sequences with Low-Discrepancy Blue-Noise 2-D Projections ». Computer Graphics Forum, vol. 37, n°2, pp. 339-353. HAL : hal-01717945.
- Abdalla Ahmed, Hélène Perrier, David Coeurjolly, Victor Ostromoukhov, Jianwei Guo, Dong-Ming Yan, Hui Huang & Oliver Deussen (2016). « Low-Discrepancy Blue Noise Sampling ». ACM Transactions on Graphics, vol. 35, n°6, 247:1-247:13. doi : 10.1145/2980179.2980218. HAL : hal-01372542.

Conférences internationales avec comité de lecture

- Hélène Perrier, Jérémy Levallois, David Coeurjolly, Jean-Philippe Farrugia, Jean-Claude Iehl & Jacques-Olivier Lachaud (2016). « Interactive Curvature Tensor Visualization on Digital Surfaces ». DGCI2016, 20 janvier 2016, Nantes (France), pp. 282-294. doi : 10.1007/978-3-319-32360-2_22. HAL : hal-01262663.

Jury

Richard CHBEIR, PR IUT de Bayonne et du Pays Basque
Benjamin NGUYEN, PR INSA Centre Val de Loire
Licia CAPRA, PR University College London
Kévin HUGUENIN, PR UNIL
Lionel BRUNIE, PR INSA Lyon
Sonia BEN MOKHTAR, PR CNRS
Cédric LAURADOUX, CR, INRIA,

rapporteur
rapporteur
examinatrice
examineur
directeur de thèse
examinatrice
examineur

Résumé

Depuis quelques dizaines d'années, l'utilisation de téléphones contenant un capteur GPS a fortement augmenté. Cependant, tous ces usages ne sont pas sans menace pour la vie privée des utilisateurs. En effet, les données de mobilité qu'ils envoient à ces services peuvent être utilisées pour inférer des informations sensibles telles que leur domicile ou leur lieu de travail. C'est à ce moment qu'entrent en action les mécanismes de protection, visant à redonner aux utilisateurs le contrôle sur leur vie privée. Nous commençons par répertorier les mécanismes de protection existants et les métriques utilisées pour les évaluer. Cette première analyse met en avant une information particulièrement sensible : les points d'intérêt. Ces derniers représentent tous les lieux où les utilisateurs passent la majeure partie de leur temps. Cela nous conduit à proposer un nouveau mécanisme de protection, PROMESSE, dont le but principal est de cacher ces points d'intérêt. Les mécanismes de protection sont en général configurés par des paramètres, qui ont un grand impact sur leur efficacité. Nous proposons ALP, une solution destinée à aider les utilisateurs à configurer leurs mécanismes de protection à partir d'objectifs qu'ils ont spécifiés. Enfin, nous présentons Accio, un logiciel regroupant la majeure partie du travail de cette thèse. Il permet de lancer facilement des expériences destinées à étudier des mécanismes de protection, tout en renforçant leur reproductibilité.

Abstract

In the past decades, the usage of GPS-enabled smartphones has dramatically risen. However, all these usages do not come without privacy threats. Indeed, location data that users are sending to these services can be used to infer sensitive knowledge about them, such as where they live or where they work. This is where protection mechanisms come into play, whose goal is to put users back in control of their privacy. We start by surveying existing protection mechanisms and metrics used to evaluate them. This first analysis highlights a particularly sensitive information, namely the points of interest. These are all the places where users use to spend most of their time. This leads us towards building a new protection mechanism, PROMESSE, whose main goal is to hide these points of interest. Protection mechanisms tend to be configured by parameters, which highly impact their effectiveness in terms of privacy and utility. We propose ALP, a solution to help users to configure their protection mechanisms from a set of objectives they specified. Finally, we introduce Accio, which is a software encompassing most of our work. Its goal is to allow to easily launch location privacy experiments and enforce their reproducibility.

Thèse

Soutenue le 01/03/2018

Établissement :
INSA Lyon

Contact :
Lionel.Brunie@liris.cnrs.fr

Publications

Revue internationale avec comité de lecture

- Sophie Cerf, Sara Bouchenak, Bogdan Robu, Nicolas Marchand, Vincent Primault, Sonia Ben Mokhtar, Antoine Boutet & Lydia Y. Chen (2018). « Automatic Privacy and Utility Preservation of Mobility Data: A Nonlinear Model-Based Approach ». IEEE Transactions on Dependable and Secure Computing, pp. 1-14. doi : 10.1109/TDSC.2018.2884470. HAL : hal-01910687.

Conférences internationales avec comité de lecture

- Sophie Cerf, Vincent Primault, Antoine Boutet, Sonia Ben Mokhtar, Sara Bouchenak, Nicolas Marchand & Bogdan Robu (2017). « Données de mobilité : protection de la vie privée vs. utilité des données ». Conférence francophone d'informatique en parallélisme, architecture et système (CompAS), Sophia Antipolis (France). HAL : hal-01527666.
- Sonia Ben Mokhtar, Antoine Boutet, Louafi Bouzouina, Patrick Bonnel, Olivier Brette, Lionel Brunie, Mathieu Cunche, Stephane D'alu, Vincent Primault et al. (2017). « PRIVA'MOV: Analysing Human Mobility Through Multi-Sensor Datasets ». NetMob 2017, 7 avril 2017, Milan (Italie). HAL : hal-01578557.
- Sophie Cerf, Vincent Primault, Antoine Boutet, Sonia Ben Mokhtar, Robert Birke, Sara Bouchenak, Lydia Chen, Nicolas Marchand & Bogdan Robu (2017). « PULP: Achieving Privacy and Utility Trade-off in User Mobility Data ». SRDS 2017 - 36th IEEE International Symposium on Reliable Distributed Systems, 29 septembre 2017, Hong Kong (Hong-Kong), pp. 164-173. HAL : hal-01578635.
- Sophie Cerf, Bogdan Robu, Nicolas Marchand, Antoine Boutet, Vincent Primault, Sonia Ben Mokhtar & Sara Bouchenak (2016). « Toward an Easy Configuration of Location Privacy Protection Mechanisms ». ACM/IFIP/USENIX Middleware conference, 16 décembre 2016, Trente (Italie). HAL : hal-01376640.
- Vincent Primault, Sonia Ben Mokhtar & Lionel Brunie (2015). « Privacy-preserving Publication of Mobility Data with High Utility ». 2015 35th IEEE International Conference on Distributed Computed Systems, 2 juillet 2015, Columbus, Ohio (États-Unis), pp. 802-803. ArXiv : 1506.09074. HAL : hal-01159291.
- Vincent Primault, Sonia Ben Mokhtar, Cédric Lauradoux & Lionel Brunie (2015). « Time Distortion Anonymization for the Publication of Mobility Data with High Utility ». 14th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, 22 août 2015, Helsinki (Finlande). ArXiv : 1507.00443. HAL : hal-01170060.
- Nicolas Haderer, Vincent Primault, Patrice Raveneau, Christophe Ribeiro, Romain Rouvoy & Sonia Ben Mokhtar (2014). « Towards a Practical Deployment of Privacy-preserving Crowd-sensing Tasks ». Middleware Posters and Demos '14, 8 décembre 2014, Bordeaux (France). doi : 10.1145/2678508.2678530. HAL : hal-01095787.
- Vincent Primault, Sonia Ben Mokhtar, Cédric Lauradoux & Lionel Brunie (2014). « Differentially Private Location Privacy in Practice ». Third Workshop on Mobile Security Technologies (MoST) 2014, 17 mai 2014, San Jose (États-Unis). ArXiv : 1410.7744. HAL : hal-01148230.

Conférences nationales avec comité de lecture

- Sabina Surdu, Vincent Primault & Yann Gripay (2013). « SQuIF: Subtile Quête d'Informations personnelles issues de Fichiers (dém) ». BDA'13, Bases de Données Avancées 2013, 22 octobre 2013, Nantes (France), pp. 1-5. HAL : hal-01339309.

Jury

Monique GRANDBASTIEN, PR, Université de Lorraine	Rapporteuse
Ronan CHAMPAGNAT, MC, Université de la Rochelle	Rapporteur
Amine CHIKH, PR, Université Abou Bekr Belkaid de Tlemcen	Examineur
Sébastien GEORGE, PR, Université du Mans	Directeur de thèse
Karim BOUAMRANE, PR, Université d'Oran,	Directeur de thèse
Karim SEHABA, MC, Université Lumière Lyon 2	Co-directeur de thèse
Azeddine CHIKH, PR, Université Abou Bekr Belkaid de Tlemcen	Invité
Sidi Mohammed BENSLIMANE, PR, ESI de Sidi Bel Abbès	Invité

Résumé

Avec la quantité croissante du contenu pédagogique produit chaque jour par les utilisateurs, il devient très difficile pour les apprenants de trouver les ressources les plus adaptées à leurs besoins. Les systèmes de recommandation sont utilisés dans les plateformes éducatives pour résoudre le problème de surcharge d'information. Ils sont conçus pour fournir des ressources pertinentes à un apprenant en utilisant certaines informations sur les utilisateurs et les ressources. Le présent travail s'inscrit dans le contexte des systèmes de recommandation des ressources pédagogiques, en particulier les systèmes qui utilisent des informations sociales. Nous avons défini une approche de recommandation de ressources éducatives en se basant sur les résultats de recherche dans le domaine des systèmes de recommandation, des réseaux sociaux et des environnements informatiques pour l'apprentissage humain. Nous nous appuyons sur les relations sociales entre apprenants pour améliorer la précision des recommandations. Notre proposition est basée sur des modèles formels qui calculent la similarité entre les utilisateurs d'un environnement d'apprentissage pour générer trois types de recommandation, à savoir la recommandation des 1) ressources populaires, 2) ressources utiles et 3) ressources récemment consultées. Nous avons développé une plateforme d'apprentissage, appelée *lcraa*, qui intègre nos modèles de recommandation. La plateforme *lcraa* est un environnement d'apprentissage social qui permet aux apprenants de télécharger, de visualiser et d'évaluer les ressources éducatives. Dans cette thèse, nous présentons les résultats d'une expérimentation menée pendant deux ans qui a impliqué un groupe de 372 apprenants d'*lcraa* dans un contexte éducatif réel. L'objectif de cette expérimentation est de mesurer la pertinence, la qualité et l'utilité des ressources recommandées. Cette étude nous a permis d'analyser les retours des utilisateurs concernant les trois types de recommandations. Cette analyse a été basée sur les traces des utilisateurs enregistrées avec *lcraa* et sur un questionnaire. Nous avons également effectué une analyse hors ligne en utilisant un jeu de données afin de comparer notre approche avec quatre algorithmes de référence.

Abstract

With the increasing amount of educational content produced daily by users, it becomes very difficult for learners to find the resources that are best suited to their needs. Recommendation systems are used in educational platforms to solve the problem of information overload. They are designed to provide relevant resources to a learner using some information about users and resources. The present work fits in the context of recommender systems for educational resources, especially systems that use social information. We have defined an educational resource recommendation approach based on research findings in the area of recommender systems, social networks, and

Thèse

Soutenu le 03/07/2018

Établissement :
INSA Lyon

Contact :
Karim.Sehaba@iris.cnrs.fr

Technology-Enhanced Learning. We rely on social relations between learners to improve the accuracy of recommendations. Our proposal is based on formal models that calculate the similarity between users of a learning environment to generate three types of recommendation, namely the recommendation of 1) popular resources; 2) useful resources; and 3) resources recently consulted. We have developed a learning platform, called Icraa, which integrates our recommendation models. The Icraa platform is a social learning environment that allows learners to download, view and evaluate educational resources. In this thesis, we present the results of an experiment conducted for almost two years on a group of 372 learners of Icraa in a real educational context. The objective of this experiment is to measure the relevance, quality and usefulness of the recommended resources. This study allowed us to analyze the user's feedback on the three types of recommendations. This analysis is based on the users' traces which was saved with Icraa and on a questionnaire. We have also performed an offline analysis using a dataset to compare our approach with four base line algorithms.

Publications

Conférences internationales avec comité de lecture

- Mohammed Tadlaoui, Karim Sehaba & Sébastien George (2015). « Recommendation of Learning Resources Based on Social Relations ». CSEDU 2015 - Proceedings of the 7th International Conference on Computer Supported Education, 25 mai 2015, Lisbon (Portugal), pp. 425-432. doi : 10.5220/0005452304250432. HAL : hal-01205730.
- Mohammed Tadlaoui (2014). « Recommandation de ressources pédagogiques basée sur les relations sociales ». RJC EIAH 2014, 17 juin 2014, La Rochelle (France). HAL : hal-01283750.

Conférences nationales avec comité de lecture

- Mohammed Tadlaoui, Karim Sehaba & Sébastien Georges (2014). « Recommandation de ressources pédagogiques dans les réseaux sociaux en ligne ». Journée Scientifique «Systèmes de recommandation et réseaux sociaux», 10 octobre 2014, Chambéry (France). HAL : hal-01497876.

Éditions scientifique d'Ouvrages (livres, chapitres, colloques, congrès, n° spéciaux)

- Mohammed Tadlaoui, Azeddine Chikh & Karim Bouamrane (2013). « ALEM: A Reference Model for Educational Adaptive Web Applications ». Intelligent and Adaptive Educational-Learning Systems : Achievements and Trends, Springer Berlin Heidelberg, pp. 25-48. doi : 10.1007/978-3-642-30171-1_2. HAL : hal-01339317.

Autres

- Mohammed Tadlaoui, Karim Sehaba & Sébastien George (2014). « Recommandation de ressources pédagogiques dans les réseaux sociaux en ligne ». Journée Scientifique «Systèmes de recommandation et réseaux sociaux», 10 octobre 2014, Chambéry (France). Poster. HAL : hal-01283753.

Laboratoire d'Informatique en Image et Systèmes d'information

Contact

Université Claude Bernard Lyon 1

Bâtiment Nautibus

25 avenue Pierre de Coubertin

F-69622 Villeurbanne Cedex

☎ +33 472 43 36 10

✉ secretariat@liris.cnrs.fr

<http://liris.cnrs.fr/>

