



UMR 5205 CNRS

Thèses et HDR

2019

Laboratoire d'InfoRmatique
en Image et Systèmes d'information



INSA



université
LUMIÈRE
LYON 2



HDR

Weakly Supervised and On-line Machine Learning for Object Tracking and Recognition in Images and Videos

Stefan Duffner 7

Thèses

Un canevas efficace pour le traitement et l'analyse des textes non structurés afin de découvrir les retards de livraison et d'optimiser la planification de routes en temps réel

Mohammad Alshaer 13

Skyline : un marqueur pour la réalité augmentée mobile en contexte urbain

Mehdi Ayadi 15

Analyse locale de surface avec la base des Wavejets : Définition de nouveaux invariants intégraux et application à l'amplification de détails géométriques

Yohann Bearzi 17

Fouille de Modèles Exceptionnels dans les Données Comportementales

Adnene Belfodil 19

Sur la découverte de sous-groupes en utilisant la théorie de l'ordre

Aimene Belfodil 23

Contributions à la gestion des processus métiers configurables : une approche orientée base de connaissances, fragmentation, et mesure d'entropies

Wehbi Benallal 25

Détection non-supervisée de motifs dans des partitions musicales manuscrites

Riyadh Benammar 27

Mining Useful Patterns in Attributed Graphs

Ahmed Anes Bendimerad 29

Linked Service Integration on the Semantic Web

Mahdi Bennara 33

Gérer et analyser les grands graphes des entités nommées Jocelyn Bernard	35
Contrôle physique de mouvement de personnages virtuels en environnement complexe Samuel Carensac	37
Suivi de l'engagement des apprenants lors de la construction de cartes mentales à partir de traces d'interaction. Rubiela Carrillo Roso	39
Interactive mapping specification and repairing in the presence of policy views Ugo Comignani	43
Reconnaissance d'expressions corporelles dans des mouvements de personnes en vue de la synthèse de style Arthur Crenn	45
Deep Similarity Metric Learning for Multiple Object Tracking Bonan Cuan	47
Intégration de données liées respectueuse de la confidentialité Rémy Delanaux	49
Simulation of calcium signaling in fine astrocytic processes Audrey Denizot	51
Représentation, modélisation et génération procédurale de paysages de rivières naturelles Thibault Dupont	53
Modelling of a Privacy Language and Efficient Policy-based De-identification Armin Gerl	55
Power-index based Management of Fraud Detection Rules: Supervised and Semi-supervised Approches Léopold Ghemmogne Fossi	57
Couverture de sommets sous contraintes Valentin Gledel	59
Generic instance segmentation for object-oriented bin-picking Matthieu Grard	61

Recherche et agrégation dans les graphes massifs Abdelmalek Habi	63
Reconfiguration et Jeux Combinatoires Marc Heinrich	65
Utilisation de l'Eye-tracking pour l'Interaction Mobile dans un Environnement Réel Augmenté Qinjie Ju	69
Capitaliser les processus d'analyse de traces d'apprentissage : modélisation ontologique et assistance à la réutilisation Alexis Lebis	71
Une approche ascendante pour la gestion énergétique d'une Smart-Grid : Modèle adaptatif et réactif fondé sur une architecture décentralisée pour un système générique centré sur l'utilisateur permettant un déploiement à grande échelle Victor Lequay	75
Recent Hidden Markov Models for Lower Limb Locomotion Activity Detection and Recognition using IMU sensors Haoyu Li	77
Towards a new data privacy-based approach for IoT Faiza Loukil	79
Préservation contre les attaques de ré-identification sur des données de mobilité Mohamed Maouche	83
Management of Selfish and Malicious Behaviors In Distributed Collaborative Systems Hayam Mousa	85
Architecture générique pour la co-construction de politiques urbaines basée sur l'apprentissage par renforcement multi-agent Simon Pageaud	87
Models and Algorithms for Online Stochastic Vehicle Routing Problems Michael Saint Guillain	89
Méthodes de sélection de voisinage et de prévision à court-terme pour l'analyse du trafic urbain Julien Salotti	91

Calcul de dose 4D à l'aide des structures tétraédriques déformables pour l'hadronthérapie	
Yazid Touileb	93
Génération procédurale d'effets atmosphériques	
Antoine Webanck.....	95
Building A Secure Infrastructure for IoT Systems in Distributed Environments	
Xiaoyang Zhu	97

Jury

Michel PAINDAVOINE, PR, Université de Bourgogne, Dijon	Rapporteur
François BREMOND, Docteur Ingénieur, INRIA Sophia Antipolis	Rapporteur
Nicolas THOME, PR, CNAM, Paris	Rapporteur
Jean-Marc ODOBEZ, DR, Idiap Research Institute, Suisse	Examineur
Carole LARTIZIEN, Docteur Ingénieur, CNRS, Creatis, Lyon	Examinatrice
Mohand-Saïd HACID, PR, Université Claude Bernard Lyon 1	Examineur
Christophe GARCIA, PR, INSA Lyon	Président

Abstract

This manuscript summarises the work that I have been involved in for my post-doctoral research and in the context of my PhD supervision activities during the past 11 years. I have conducted this work partly as a post-doctoral researcher at the Idiap Research Institute in Switzerland, and partly as an associate professor at the LIRIS laboratory and INSA Lyon in France. The technical section of the manuscript comprises two main parts: the first part being on on-line learning approaches for visual object tracking in dynamic environments, and the second part on similarity metric learning algorithms and Siamese Neural Networks (SNN). I first present our work on on-line multiple face tracking in a dynamic indoor environment, where we focused on the aspects of track creation and removal for long-term tracking. The automatic detection of the faces to track is challenging in this setting because they may not be detected for long periods of time, and false detections may occur frequently. Our proposed algorithm consisted in a recursive Bayesian framework with a separate track creation and removal step based on Hidden Markov Models including observation likelihood functions that are learnt off-line on a set of static and dynamic features related to the tracking behaviour and the objects' appearance. This approach is very efficient and showed superior performance to the state of the art in on-line multiple object tracking. In the same context, we further developed a new on-line algorithm to estimate the Visual Focus of Attention from videos of persons sitting in a room. This unsupervised on-line learning approach is based on an incremental k-means algorithm and is able to automatically extract, from a video stream, the targets that the persons are looking at in a room. I further present our research on on-line learnt robust appearance models for single-object tracking. In particular we focused on the problem of model-free, on-line tracking of arbitrary objects, where the state and model of the object to track is initialised in the first frame and updated throughout the rest of the video. Our first approach, called PixelTrack, consists in a combined detection and segmentation framework that robustly learns the appearance of the object to track and avoids drift by an effective on-line co-training algorithm. This method showed excellent tracking performance on public benchmarks, both in terms of robustness and speed, and is particularly suitable for tracking deformable objects. The second tracking approach, called MCT, employs an on-line learnt discriminative classifier that stochastically samples the training instances from a dynamic probability density function that is computed from moving and possibly distracting image background regions. The use of this motion context showed to be very effective and lead to a significant gain in the overall tracking robustness and performance. We extended this idea by designing a set of features that concisely describe the visual context of the overall scene shown in a video at a given point in time. Then, we applied several complementary tracking algorithms on a set of training videos and computed the corresponding context features for each frame. Finally, we trained a discriminative classifier off-line that estimates the most suitable

Stefan Duffner

HDR

Soutenue le 05/04/2019

Établissement :
INSA Lyon

Contact :
Stefan.Duffner@liris.cnrs.fr

tracker for a given context, and applied it on-line in an effective tracker-selection framework. Evaluated on several different “pools” of individual trackers, the combined model lead to an increased performance in terms of accuracy and robustness on challenging public benchmarks. In the second part of the manuscript, I present several contributions related to SNNs for similarity metric learning. First, we proposed a new objective function and training algorithm called Triangular Similarity Metric Learning that enhances the convergence behaviour and achieved state-of-the-art results on pairwise verification tasks, like face, speaker or kinship verification. Then, I present our work on SNNs for gesture classification from inertial sensor data, where we proposed a new class-balanced learning strategy operating on tuples of training samples and an objective function based on a polar sine formulation. Finally, I present several contributions on SNN with deeper and more complex Convolutional Neural Network models applied to the problem of person re-identification in images. In this context, we proposed different neural architectures and triplet learning methods that include semantic prior knowledge, e.g. on pedestrian attributes, body orientation and surrounding group context, using a combination of supervised and weakly supervised algorithms. Also, a new learning-to-rank algorithm for SNN, called Rank-Triplet, has been introduced and successfully applied to person re-identification. These recent works achieved state-of-the-art re-identification results on challenging pedestrian image datasets and opened new perspectives for future similarity metric approaches.

Publications

Revue internationale avec comité de lecture

- Lilei Zheng, Stefan Duffner, Khalid Idrissi, Christophe Garcia & Atila Baskurt (2018). « Pairwise Identity Verification via Linear Concentrative Metric Learning ». IEEE Transactions on Cybernetics, pp. 1-12. doi : 10.1109/TCYB.2016.2634011. HAL : hal-01435368.
- Renato Cintra, Stefan Duffner, Christophe Garcia & André Leite (2018). « Low-complexity Approximate Convolutional Neural Networks ». IEEE Transactions on Neural Networks and Learning Systems. doi : 10.1109/TNNLS.2018.2815435. HAL : hal-01727219.
- Yiqiang Chen, Stefan Duffner, Andrei Stoian, Jean-Yves Dufour & Atila 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.
- Stefan Duffner & Christophe Garcia (2017). « Fast Pixelwise Adaptive Visual Tracking of Non-Rigid Objects ». IEEE Transactions on Image Processing, vol. 26, n°5, pp. 2368-2380. doi : 10.1109/TIP.2017.2676346. HAL : hal-01574513.
- Samuel Berlemont, Grégoire Lefebvre, Stefan Duffner & Christophe Garcia (2017). « Class-Balanced Siamese Neural Networks ». Neurocomputing. HAL : hal-01580527.
- Stefan Duffner & Christophe Garcia (2015). « Using discriminative motion context for on-line visual object tracking ». IEEE Transactions on Circuits and Systems for Video Technology. doi : 10.1109/TCSVT.2015.2504739. HAL : hal-01340308.
- Stefan Duffner & Christophe Garcia (2015). « Visual Focus of Attention estimation with unsupervised incremental learning ». IEEE Transactions on Circuits and Systems for Video Technology. doi : 10.1109/TCSVT.2015.2501920. HAL : hal-01153969.
- Lilei Zheng, Stefan Duffner, Khalid Idrissi, Christophe Garcia & Atila Baskurt (2015). « Siamese Multi-layer Perceptrons for Dimensionality Reduction and Face Identification ». Multimedia Tools and Applications,. doi : 10.1007/s11042-015-2847-3. HAL : hal-01182273.
- Stefan Duffner & Jean-Marc Odobez (2014). « Leveraging Colour Segmentation for Upper-Body Detection ». Pattern Recognition, vol. 47, pp. 2222-2230. doi : 10.1016/j.patcog.2013.12.014. HAL : hal-01270720.
- Toru Nakashika, Toshiya Yoshioka, Tetsuya Takiguchi, Yasuo Ariki, Stefan Duffner & Christophe Garcia (2014). « Convolutional Bottleneck Network with Dropout for Dysarthric Speech Recognition ». Transactions on Machine Learning and Artificial Intelligence, vol. 2, pp. 1-15. doi : 10.14738/tmlai.22.150. HAL : hal-01301040.
- Petr Motlicek, Stefan Duffner, Danil Korchagin, Hervé Bourlard, Carl Scheffler, Jean-Marc Odobez, Giovanni Del Galdo, Markus Kallinger & Oliver Thiergart (2013). « Real-Time Audio-Visual Analysis for Multiperson Videoconferencing ». Advances in Multimedia, vol. 2013, 175745:1-175745:21. doi : 10.1155/2013/175745. HAL : hal-01339256.

Conférences internationales avec comité de lecture

- Paul Compagnon, Grégoire Lefebvre, Stefan Duffner & Christophe Garcia (2019). « Personalized Posture and Fall Classification with Shallow Gated Recurrent Units ». 32nd IEEE CBMS International Symposium on Computer-Based Medical Systems, 7 juin 2019, Cordoue (Espagne). HAL : hal-02095355.
- Dario Lucente, Stefan Duffner, Corentin Herbert, Joran Rolland & Freddy Bouchet (2019). « MACHINE LEARNING OF COMMITTOR FUNCTIONS FOR PREDICTING HIGH IMPACT CLIMATE EVENTS ». Climate Informatics, 4 octobre 2019, Paris (France). ArXiv : 1910.11736. HAL : hal-02322370.
- Paul Compagnon, Grégoire Lefebvre, Stefan Duffner & Christophe Garcia (2019). « Routine Modeling with Time Series Metric Learning ». 28th International Conference on Artificial Neural Networks, 17 septembre 2019, Munich (Allemagne). ArXiv : 1907.04666. HAL : hal-02165265.
- Yiqiang Chen, Stefan Duffner, Andrei Stoian, Jean-Yves Dufour & Atila 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 & Atila 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 & Atila 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 & Atila 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.
- Smrity Bhattarai, Arjuna Madanayake, Renato J. Cintra, Stefan Duffner & Christophe Garcia (2017). « Digital architecture for real-time CNN-based face detection for video processing ». IEEE Cognitive Communications for Aerospace Applications Workshop (CCAA), 28 juin 2017, Cleveland (États-Unis). doi : 10.1109/CCAAW.2017.8001608. HAL : hal-01574551.

- Yiqiang Chen, Stefan Duffner, Andrei Stoian, Jean-Yves Dufour & Atila 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.
- Nawel Medjkoune, Frédéric Armetta, Mathieu Lefort & Stefan Duffner (2017). « Autonomous object recognition in videos using Siamese Neural Networks ». EUCognition Meeting (European Society for Cognitive Systems) on "Learning: Beyond Deep Neural Networks", 22 novembre 2017, Zurich (Suisse), p. 4. HAL : hal-01630163.
- Samuel Berlemont, Grégoire Lefebvre, Stefan Duffner & Christophe Garcia (2016). « Polar Sine Based Siamese Neural Network for Gesture Recognition ». International Conference on Artificial Neural Networks, 9 septembre 2016, Barcelona (Espagne). doi : 10.1007/978-3-319-44781-0_48. HAL : hal-01369302.
- Salma Moujtahid, Stefan Duffner & Atila Baskurt (2015). « Coherent Selection of Independent Trackers for Real-time Object Tracking ». International Conference on Computer Vision Theory and Applications (VISAPP), Berlin (Allemagne), pp. 584-592. HAL : hal-01161859.
- Samuel Berlemont, Grégoire Lefebvre, Stefan Duffner & Christophe Garcia (2015). « Siamese Neural Network based Similarity Metric for Inertial Gesture Classification and Rejection ». International Conference on Automatic Face and Gesture Recognition, 8 mai 2015, Ljubljana (Slovénie), pp. 1-6. doi : 10.1109/FG.2015.7163112. HAL : hal-01179993.
- Lilei Zheng, Khalid Idrissi, Christophe Garcia, Stefan Duffner & Atila Baskurt (2015). « Logistic Similarity Metric Learning for Face Verification ». 40th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2015, 24 avril 2015, Brisbane (Australie). HAL : hal-01158949.
- Jiwen Lu, Junlin Hu, Venice Erin Liong, Xiuzhuang Zhou, Andrea Bottino, Ihtesham Ul Islam, Tiago Figueiredo Vieira, Xiaoqian Qin, Xiaoyang Tan et al. (2015). « The FG 2015 Kinship Verification in the Wild Evaluation ». 11th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2015), 8 mai 2015, Ljubljana (Slovénie). HAL : hal-01158942.
- Lilei Zheng, Khalid Idrissi, Christophe Garcia, Stefan Duffner & Atila Baskurt (2015). « Triangular Similarity Metric Learning for Face Verification ». 11th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2015), 8 mai 2015, Ljubljana (Slovénie). HAL : hal-01158908.
- Salma Moujtahid, Stefan Duffner & Atila Baskurt (2015). « Classifying Global Scene Context for On-line Multiple Tracker Selection ». British Machine Vision Conference (BMVC), 10 septembre 2015, Swansea (Royaume-Uni). HAL : hal-01208200.
- Matej Kristan, Jiri Matas, Aleš Leonardis, Michael Felsberg, Luka Cehovin, Gustavo Fernández, Tomas Vojir, Gustav Häger, Georg Nebehay et al. (2015). « The Visual Object Tracking VOT2015 Challenge Results ». Visual Object Tracking Workshop 2015 at ICCV2015, 12 décembre 2015, Santiago (Chili). doi : 10.1109/ICCVW.2015.79. HAL : hal-01336773.
- Toru Nakashika, Toshiya Yoshioka, Tetsuya Takiguchi, Yasuo Ariki, Stefan Duffner & Christophe Garcia (2014). « Dysarthric Speech Recognition Using a Convolutional Bottleneck Network ». IEEE International Conference on Signal Processing (ICSP), 19 octobre 2014, Hangzhou (Chine), pp. 505-509. HAL : hal-01301122.

- Matej Kristan, Roman Pflugfelder, Ales Leonardis, Jiri Matas, Luka Cehovin, Georg Nebehay, Tomas Vojir, Gustavo Fernández, Alan Lukezic et al. (2014). « The Visual Object Tracking VOT2014 challenge results ». Workshop on Visual Object Tracking Challenge (VOT2014) - ECCV, 6 septembre 2014, Zurich (Suisse), pp. 1-27. HAL : hal-01301090.
- Stefan Duffner & Christophe Garcia (2014). « Exploiting contextual motion cues for visual object tracking ». Workshop on Visual Object Tracking Challenge (VOT2014) - ECCV, 6 septembre 2014, Zurich (Suisse), pp. 1-12. HAL : hal-01301089.
- Stefan Duffner, Samuel Berlemont, Grégoire Lefebvre & Christophe Garcia (2014). « 3D gesture classification with convolutional neural networks ». International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 9 mai 2014, Florence (Italie), pp. 5432-5436. doi : 10.1109/ICASSP.2014.6854641. HAL : hal-01180542.
- Stefan Duffner & Christophe Garcia (2013). « Unsupervised Online Learning of Visual Focus of Attention ». 10-th IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS 2013), 27 août 2013, Krakow (Pologne), pp. 25-30. HAL : hal-00976391.
- Stefan Duffner & Christophe Garcia (2013). « PixelTrack: a fast adaptive algorithm for tracking non-rigid objects ». International Conference on Computer Vision (ICCV 2013), 1 décembre 2013, Sydney (Australie), pp. 2480-2487. HAL : hal-00976387.

Conférences nationales avec comité de lecture

- Yiqiang Chen, Stefan Duffner, Andrei Stoian, Jean-Yves Dufour & Atila Baskurt (2017). « Ré-identification de personne avec perte min-triplet ». 16èmes Journées francophones des jeunes chercheurs en vision par ordinateur (ORASIS 2017), 16 juin 2017, Colleville-sur-Mer (France). HAL : hal-01625496.

Un canevas efficace pour le traitement et l'analyse des textes non structurés afin de découvrir les retards de livraison et d'optimiser la planification de routes en temps réel

Mohammad Alshaer

Jury

Daniela GRIGORI, PR, université Paris-Dauphine	Rapporteure
Amel BOUZEGHOUB, PR, Télécom SudParis	Rapporteure
Engelbert MEPHU NGUIFO, PR, Université Blaise Pascal	Examineur
Khalid BENABDESLEM, MC, Université Lyon 1	Examineur
Hala NAJA, PR, Université Libanaise	Examinatrice
Mohand-Saïd HACID, PR, Université Lyon 1	Co-Directeur de thèse
Mohamed DBOUK, PR, Université Libanaise	Co-Directeur de thèse
Yehia TAHER, MC, Université de Versailles	Co-Encadrant de thèse
Rafiqul HAQUE AKM, DR, Cognitus	Invité

Résumé

Dans cette thèse, nous avons développé et expérimenté deux méthodes pour le traitement des données: SANA et IBRIDIA. SANA est basée sur un classificateur multinomial Naïve Bayes, tandis qu'IBRIDIA s'appuie sur l'algorithme de classification hiérarchique (CLH) de Johnson, qui est une technologie hybride permettant la collecte et le traitement de données par lots et en temps réel. SANA est une solution de service qui traite les données non structurées. Cette méthode sert de système polyvalent pour extraire les événements pertinents, y compris le contexte (tel que le lieu, l'emplacement, l'heure, etc.). En outre, il peut être utilisé pour effectuer une analyse de texte sur les événements ciblés. IBRIDIA a été conçu pour traiter des données inconnues provenant de sources externes et les regrouper en temps réel afin d'acquérir une connaissance / compréhension des données permettant d'extraire des événements pouvant entraîner un retard de livraison. Selon nos expériences, ces deux approches montrent une capacité unique à traiter des données logistiques. Cependant, SANA semble plus prometteur puisque la technologie sous-jacente (classificateur Naïve Bayes) a surpassé IBRIDIA du point de vue performance. Il est clairement indiqué que SANA était censé générer une connaissance graphique des événements collectés immédiatement en temps réel sans avoir à attendre, permettant ainsi de tirer le meilleur parti de ces événements. IBRIDIA est importante dans le domaine de la logistique pour identifier la catégorie d'événements la plus influente qui affecte la livraison. Malheureusement, pour IBRIDIA, nous devrions attendre qu'un nombre minimum d'événements se présente et nous avons toujours un démarrage à froid. Étant donné que nous sommes intéressés par la ré-optimisation de l'itinéraire en temps réel, nous avons adopté SANA comme cadre de traitement de données. Nous avons implémenté une application d'optimisation d'itinéraire afin de démontrer comment notre solution est utilisée pour extraire des informations d'événements susceptibles d'entraîner un retard de livraison et comment les itinéraires peuvent être optimisés pour éviter ce retard.

Abstract

In this thesis, we developed and experimented two data processing solutions: SANA and IBRIDIA. SANA is built on Multinomial Naïve Bayes classifier whereas IBRIDIA relies on Johnson's hierarchical clustering (HCL) algorithm which is hybrid technology that enables data collection and processing in batch style and realtime. SANA is a service-based solution which deals with unstructured data. It serves as a multi-purpose system to extract the relevant events including the context of the event (such as place, location, time, etc.). In addition, it can be used to perform text analysis over the targeted events. IBRIDIA was designed to process unknown data stemming from

Thèse

Soutenue le 13/09/2019

Établissement :
Université Claude Bernard Lyon 1

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

external sources and cluster them on-the-fly in order to gain knowledge/ understanding of data which assist in extracting events which may lead to delivery delay. According to our experiments, both of these approaches show a unique ability to process logistics data. However, SANA is found more promising since the underlying technology (Naïve Bayes classifier) outperformed IBRIDIA from performance measuring perspectives. It is clearly said that SANA was meant to generate a graph knowledge from the events collected immediately in realtime without any need to wait, thus reaching maximum benefit from these events. Whereas, IBRIDIA have an important influence within the logistics domain for identifying the most influential category of events that are affecting the delivery. Unfortunately, in IBRIDIA, we should wait for a minimum number of events to arrive and always we have a cold start. Due to the fact that we are interested in re-optimizing the route on the fly, we adopted SANA as our data processing framework. We implemented a route optimization application to demonstrate how our solution is used in extracting information of events that may lead to delivery delay and how the routes can be optimized to prevent the delay.

Skyline : un marqueur pour la réalité augmentée mobile en contexte urbain

Mehdi Ayadi

Jury

Serge MIGUET, PR, Université Lumière Lyon 2	Directeur de thèse
Mihaela SCUTURICI, MC, Université Lumière Lyon 2	Co-directrice de thèse
Chokri BEN AMAR, PR, Université de Sfax	Co-directeur de thèse
Patrick SAYD, DR, Commissariat à l'énergie atomique CEA -	Examineur
Adel ALIM, PR, Université de Sfax	Examineur
Anne RUAS, DR, IFSTTAR	Rapporteuse
Luce MORIN, PR, Université Rennes 2	Rapporteuse

Résumé

L'objectif principal de cette thèse est de permettre à un utilisateur se déplaçant en milieu urbain, de visualiser en Réalité Augmentée, à l'aide d'un smartphone, l'incidence d'un projet de construction sur le paysage urbain, une fois le bâtiment réalisé.

Nous proposons une approche dans laquelle les différentes données issues des instruments embarqués dans le dispositif sont fusionnées, afin d'estimer la pose de l'appareil : le compas magnétique permet d'estimer la direction d'observation, soit un premier degré de liberté en rotation ; l'accéléromètre permet d'évaluer grossièrement les deux autres degrés de liberté en rotation ; le GPS permet d'estimer le positionnement (trois degrés de liberté en translation). Cette première estimation de la pose nous permet de placer une caméra virtuelle dans le modèle 3D de la ville. Ainsi, à l'aide des supports de rendu 2D et 3D des dispositifs, nous générons une image de synthèse de ce que l'utilisateur visualise théoriquement depuis cette position. Néanmoins, l'utilisation de ces seuls instruments pour insérer un objet de synthèse dans le flux vidéo donne à l'utilisateur une impression très peu réaliste de la scène qu'il visionne : les objets de synthèse semblent « flotter » au gré des mouvements en raison des imprécisions des instruments. Pour pallier ce défaut, nous montrons comment le skyline peut jouer le rôle d'un marqueur de réalité augmentée. Le skyline est extrait automatiquement des images réelles acquises par la caméra et des images virtuelles générées à partir d'un modèle 3D de la ville. Une étape d'appariement de ces deux skylines permet de recalibrer le virtuel sur le réel et ainsi permettre une incrustation en temps réel des projets de construction au sein du flux vidéo et une amélioration selon deux critères : précision et stabilité avec une impression d'immersion bien meilleure. Plusieurs approches de recalage des skylines à différentes complexités algorithmiques sont proposées, basées sur des mesures de similarité et une approche d'optimisation par descente de gradient.

Abstract

The main objective of this thesis is to propose an approach that allows a user, moving in an urban environment, to visualize in augmented reality, using a smartphone, the impact of a construction project on the urban landscape, once the building is completed. In particular, with the current attractiveness for tower construction projects in major cities (London, Paris and Lyon in particular), it is becoming a very important issue for neighborhood users to be able to visualize the impact of a construction project on the skyline and to give an objective opinion on it. For geographers and neighborhood users, this innovative solution makes it possible to take the opposite view from architectural firms, which only offer very qualitative renderings of their construction projects, with a limited number of views, to the detriment of more objective visualizations. We do not necessarily propose a photo-

Thèse

Soutenue le 02/07/2019

Établissement :

Université Lumière Lyon 2

Contact :

Serge.Miguet@iris.cnrs.fr

Analyse locale de surface avec la base des Wavejets : Définition de nouveaux invariants intégraux et application à l'amplification de détails géométriques

Yohann Bearzi

realistic rendering, but rather a more exact geometric consistency of the scene, throughout the user's movements. Visualizations at different positions and orientations correspond better to common use cases in neighborhoods.

We propose an approach in which data from smartphone's embedded instruments are merged to estimate a first user's pose: the magnetic compass estimates the viewing direction; the GPS, gyroscope and accelerometer roughly evaluate the parameters of motion (three degrees of freedom in translation, and three degrees of freedom in rotation). This first estimated pose, associated with the 3D model of the city, allows, using the rendering supports (2D and 3D) of the devices, to generate a synthetic image of what the user theoretically visualizes at this position. Nevertheless, the unique use of these instruments to insert a synthetic object into the video stream gives the user a very unrealistic impression of the viewed scene: due to instrument's data inaccuracies, synthetic objects appear to "hover" and "float" with the user's movements. For this, the skyline is automatically extracted from real (acquired by the camera) and virtual (generated) images. A matching step between the two skylines allows to realign the virtual skyline onto the real one, allowing thus a real time insertion of the 3D object in the real-time video stream, and an improvement according to two criteria: precision and stability, giving with a much better immersion impression. In this matching step, several similarity metrics are proposed that are used in an optimization approach based on a gradient descent.

Publications

Conférences internationales avec comité de lecture

- Mehdi Ayadi, Leo Valque, Mihaela Scuturici, Serge Miguet & Chokri Amar (2018). « The skyline as a marker for augmented reality in urban context ». 13th International Symposium on Visual Computing, 19 novembre 2018, Las Vegas (États-Unis). HAL : hal-01872631.
- Mehdi Ayadi, Loreta Suta, Mihaela Scuturici, Serge Miguet & Chokri Amar (2016). « A Parametric Algorithm for Skyline Extraction ». Advanced Concepts for Intelligent Vision Systems: 17th International Conference, ACIVS 2016, Lecce, Italy, October 24-27, 2016, Proceedings, 27 octobre 2016, Lecce (Italie), pp. 604-615. doi : 10.1007/978-3-319-48680-2_53. HAL : hal-01385542.

Jury

Marie CHABERT, ENSEEIHT
Georges-Pierre BONNEAU, Université Grenoble-Alpes
Victor OSTROMOUKHOV, Université Claude Bernard
Enric MEINHARDT LLOPIS, ENS Cachan
Julie DIGNE, CNRS
Raphaëlle CHAINE, Université Claude Bernard

Rapportrice
Rapporteur
Examineur
Examineur
Directrice de thèse
co-Directrice de thèse

Résumé

L'analyse de surface est un domaine de recherche difficile, qui a été un sujet de recherche très actif ces dernières décennies. Quand une surface est représentée par un ensemble de points, typiquement issus de scanners laser 3D, le manque de structure entre ces points rend leur traitement compliqué. Dans cette thèse, on propose une méthode d'analyse de surface en introduisant une nouvelle base de fonctions : les Wavejets. Cette base permet de décomposer localement une surface radialement en polynômes et angulairement en fréquences. Des propriétés de stabilité en fonction d'une mauvaise direction de normal sont démontrées. En liant les coefficients des Wavejets à des tenseurs différentiels à hauts ordres, on définit aussi des directions principales à haut ordre sur la surface. De plus, séparer localement les surfaces fréquemment nous amène à la définition de nouveaux invariants intégraux, permettant de décrire localement la surface. De tels descripteurs sont assez robustes car ils sont calculés par intégration. Enfin, Nous proposons une application à ces invariants intégraux pour l'amplification de détails géométriques, soit en changeant la position des points de la surface, soit en changeant la direction des normales, créant dans ce dernier cas l'illusion d'un changement de géométrie sur la surface.

Abstract

Surface analysis is a challenging research topic, which has gathered a lot of interest over the last few decades. When surface data is given as a set of points, which are the typical output of 3D laser scanners, the lack of structure makes it even more challenging. In this thesis, we tackle surface analysis by introducing a new function basis: the Wavejets. This basis allows to decompose locally the surface into a radial polynomial component and an angular frequency component. Stability properties with regards to a bad normal direction are demonstrated. By linking Wavejets coefficients to a high order differential tensor, we also define high order principal directions on the surface. Furthermore, locally splitting surfaces with respect to frequencies leads us to define new integral invariants, permitting to locally describe the surface. Such descriptors are quite robust since they result from an integration process. Finally, we develop an application of these new integral invariants for geometric detail amplification, either based on point position or on normal direction modification, creating in this case the illusion of a surface change.

Publications

Revues internationales avec comité de lecture

- Yohann 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.

Thèse

Soutenue le 08/11/2019

Établissement :

Université Claude Bernard Lyon 1

Contact :

Julie.Digne@liris.cnrs.fr

Autres

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

Jury

Siham AMER-YAHIA, DR, CNRS	Rapporteuse
Arno SIEBES, PR, Université d'Utrecht	Rapporteur
Arno KNOBBE, MC, Université de Leiden	Examineur
Ioana MANOLESCU, DR, INRIA	Examinatrice
Amedeo NAPOLI, DR, CNRS	Examineur
Philippe LAMARRE, PR, INSA-Lyon	Directeur de thèse
Sylvie CAZALENS, MC, INSA-Lyon	Co-directrice de thèse
Marc PLANTEVIT, MC, Université Lyon 1	Co-directeur de thèse

Résumé

Avec la prolifération rapide des plateformes de données qui récoltent des données relatives à plusieurs domaines tels que les données de gouvernements, d'éducation, d'environnement ou les données de notations de produits, plus de données sont disponibles en ligne. Ceci représente une opportunité sans égal pour étudier le comportement des individus et les interactions entre eux. Sur le plan politique, le fait de pouvoir interroger des ensembles de données de votes peut fournir des informations intéressantes pour les journalistes et les analystes politiques. En particulier, ce type de données peut être exploité pour l'investigation des sujets exceptionnellement conflictuels ou consensuels. Considérons des données décrivant les sessions de votes dans le parlement Européen (PE). Un tel ensemble de données enregistre les votes de chaque député (MPE) dans l'hémicycle en plus des informations relatives aux parlementaires (e.g., genre, parti national, parti européen) et des sessions (e.g., sujet, date). Ces données offrent la possibilité d'étudier les accords et désaccords de sous-groupes cohérents, en particulier pour mettre en évidence des comportements inattendus. Par exemple, il est attendu que sur la majorité des sessions, les députés votent selon la ligne politique de leurs partis politiques respectifs. Cependant, lorsque les sujets sont plutôt d'intérêt d'un pays particulier dans l'Europe, des coalitions peuvent se former ou se dissoudre. À titre d'exemple, quand une procédure législative concernant la pêche est proposée devant les MPE dans l'hémicycle, les MPE des nations insulaires du Royaume-Uni peuvent voter en accord sans être influencés par la différence entre les lignes politiques de leurs alliances respectives, cela peut suggérer un accord exceptionnel comparé à la polarisation observée habituellement. Dans cette thèse, nous nous intéressons à ce type de motifs décrivant des (dés)accords exceptionnels, pas uniquement sur les données de votes mais également sur des données similaires appelées données comportementales. Nous élaborons deux méthodes complémentaires appelées Debunk et Deviant. La première permet la découverte de (dés)accords exceptionnels entre groupes tandis que la seconde permet de mettre en évidence les comportements exceptionnels qui peuvent au sein d'un même groupe. Idéalement, ces deux méthodes ont pour objectif de donner un aperçu complet et concis des comportements exceptionnels dans les données comportementales. Dans l'esprit d'évaluer la capacité des deux méthodes à réaliser cet objectif, nous évaluons les performances quantitatives et qualitatives sur plusieurs jeux de données réelles. De plus, nous motivons l'utilisation des méthodes proposées dans le contexte du journalisme computationnel.

Adnene Belfodil

Thèse

Soutenue le 24/10/2019

Établissement :
INSA Lyon

Contact :
Philippe.Lamarre@liris.cnrs.fr

Abstract

With the rapid proliferation of data platforms collecting and curating data related to various domains such as governments data, education data, environment data or product ratings, more and more data are available online. This offers an unparalleled opportunity to study the behavior of individuals and the interactions between them. In the political sphere, being able to query datasets of voting records provides interesting insights for data journalists and political analysts. In particular, such data can be leveraged for the investigation of exceptionally consensual/controversial topics. Consider data describing the voting behavior in the European Parliament (EP). Such a dataset records the votes of each member (MEP) in voting sessions held in the parliament, as well as information on the parliamentarians (e.g., gender, national party, European party alliance) and the sessions (e.g., topic, date). This dataset offers opportunities to study the agreement or disagreement of coherent subgroups, especially to highlight unexpected behavior. It is to be expected that on the majority of voting sessions, MEPs will vote along the lines of their European party alliance. However, when matters are of interest to a specific nation within Europe, alignments may change and agreements can be formed or dissolved. For instance, when a legislative procedure on fishing rights is put before the MEPs, the island nation of the UK can be expected to agree on a specific course of action regardless of their party alliance, fostering an exceptional agreement where strong polarization exists otherwise. In this thesis, we aim to discover such exceptional (dis)agreement patterns not only in voting data but also in more generic data, called behavioral data, which involves individuals performing observable actions on entities. We devise two novel methods which offer complementary angles of exceptional (dis)agreement in behavioral data: within and between groups. These two approaches called Debunk and Deviant, ideally, enables the implementation of a sufficiently comprehensive tool to highlight, summarize and analyze exceptional comportments in behavioral data. We thoroughly investigate the qualitative and quantitative performances of the devised methods. Furthermore, we motivate their usage in the context of computational journalism.

Publications

Revue internationale avec comité de lecture

- Adnene Belfodil, Sylvie Cazalens, Philippe Lamarre & Marc Plantevit (2019). « Identifying exceptional (dis)agreement between groups ». Data Mining and Knowledge Discovery. doi : 10.1007/s10618-019-00665-9. HAL : hal-02383776.

Conférences internationales avec comité de lecture

- Adnene Belfodil, Aimene Belfodil, Anes Bendimerad, Philippe Lamarre, Céline Robardet, Mehdi Kaytoute & Marc Plantevit (2019). « FSSD - A Fast and Efficient Algorithm for Subgroup Set Discovery ». IEEE International Conference on Data Science and Advanced Analytics (DSAA), 5 octobre 2019, Washington DC (États-Unis). HAL : hal-02355503.
- Aimene Belfodil, Adnene Belfodil & Mehdi Kaytoute (2019). « Mining Formal Concepts using Implications between Items ». International Conference on Formal Concept Analysis (ICFCA 2019), Frankfurt (Allemagne). HAL : hal-02080577.

- Aimene Belfodil, Adnene Belfodil & Mehdi Kaytoute (2018). « Anytime Subgroup Discovery in Numerical Domains with Guarantees ». Machine Learning and Knowledge Discovery in Databases - European Conference, ECML PKDD 2018, 10 septembre 2018, Dublin (Irlande), pp. 500-516. HAL : hal-02117627.
- 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.

Aimene Belfodil

Jury

Bruno CRÉMILLEUX, PR, Université de Caen	Rapporteur
Bernhard GANTER, PR, Technische Universitaet Dresden	Rapporteur
Céline ROBARDET, PR, INSA Lyon	Directrice de thèse
Mehdi KAYTOUE, MC, Infologic	Co-directeur de thèse
Peggy CELLIER, MC, INSA Rennes	Examinatrice
Miguel COUCEIRO, PR, Université de Lorraine	Examineur
Arno SIEBES, PR, Universiteit Utrecht	Examineur
Julien ZARKA, Full Professor, Mobile Devices Ingenierie	Invité

Résumé

Comme le titre pourrait le suggérer, l'objectif principal de cette thèse est de fournir une meilleure compréhension de la tâche de la découverte de sous-groupes à travers la théorie de l'ordre. La découverte de sous-groupes (Subgroup Discovery - SD) est la tâche automatique dont le but est la découverte d'hypothèses intéressantes dans les bases de données. Autrement dit, étant donnée une base de donnée, l'espace de recherche de toutes les hypothèses que l'analyste voudra tester ainsi qu'un moyen formel pour évaluer la qualité de ces hypothèses ; la tâche automatique de la découverte de sous-groupe s'efforce de trouver les meilleurs hypothèses quant à ces trois paramètres. Afin d'élaborer des algorithmes efficaces et efficients pour cette tâche, il est important de comprendre les propriétés des espaces de recherche d'une part et les propriétés de la mesure de qualité d'autre part. Dans cette thèse, nous étendons l'état de l'art par: (i) fournir une vue unifiée sur les espaces d'hypothèses derrière la tâche de découverte de sous-groupes en utilisant la théorie de l'ordre, (ii) proposer l'espace d'hypothèses de conjonctions d'inégalités linéaires dans les bases de données numériques ainsi que différents algorithmes permettant de les énumérer et (iii) proposer un algorithme anytime - fournit progressivement des résultats - pour la tâche particulière de fouille de sous-groupe discriminants dans les bases de données numériques. Ce dernier fournit des garanties sur la qualité des sous-groupes extraits même si l'algorithme est interrompu.

Abstract

As the title of this dissertation may suggest, the aim of this thesis is to provide an order-theoretic point of view on the task of subgroup discovery. Subgroup discovery is the automatic task of discovering interesting hypotheses in databases. That is, given a database, the hypothesis space the analyst wants to explore and a formal way of how the analyst gauges the quality of the hypotheses (e.g. a quality measure); the automated task of subgroup discovery aims to extract the interesting hypothesis w.r.t. these parameters. In order to elaborate fast and efficient algorithms for subgroup discovery, one should understand the underlying properties of the hypothesis space on the one hand and the properties of its quality measure on the other. In this thesis, we extend the state-of-the-art by: (i) providing a unified view of the hypotheses space behind subgroup discovery using the well-founded mathematical tool of order theory, (ii) proposing the new hypothesis space of conjunction of linear inequalities in numerical databases and the algorithms enumerating its elements and (iii) proposing an anytime algorithm for discriminative subgroup discovery on numerical datasets providing guarantees upon interruption.

Thèse

Soutenue le 30/09/2019

Établissement :
INSA Lyon

Contact :
Céline.Robardet@liris.cnrs.fr

Publications

Conférences internationales avec comité de lecture

- Adnene Belfodil, Aimene Belfodil, Anes Bendimerad, Philippe Lamarre, Céline Robardet, Mehdi Kaytoute & Marc Plantevit (2019). « FSSD - A Fast and Efficient Algorithm for Subgroup Set Discovery ». IEEE International Conference on Data Science and Advanced Analytics (DSAA), 5 octobre 2019, Washington DC (États-Unis). HAL : hal-02355503.
- Aimene Belfodil, Adnene Belfodil & Mehdi Kaytoute (2019). « Mining Formal Concepts using Implications between Items ». International Conference on Formal Concept Analysis (ICFCA 2019), Frankfurt (Allemagne). HAL : hal-02080577.
- Aimene Belfodil, Sergei Kuznetsov & Mehdi Kaytoute (2018). « Pattern Setups and Their Completions ». The 14th International Conference on Concept Lattices and Their Applications (CLA 2018), 14 juin 2018, Olomouc (République Tchèque), pp. 243-253. HAL : hal-01818740.
- Aimene Belfodil, Adnene Belfodil & Mehdi Kaytoute (2018). « Anytime Subgroup Discovery in Numerical Domains with Guarantees ». Machine Learning and Knowledge Discovery in Databases - European Conference, ECML PKDD 2018, 10 septembre 2018, Dublin (Irlande), pp. 500-516. HAL : hal-02117627.
- Aimene Belfodil, Sergei Kuznetsov, Céline Robardet & Mehdi Kaytoute (2017). « Mining Convex Polygon Patterns with Formal Concept Analysis ». The Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI), 25 août 2017, Melbourne (Australie), pp. 1425-1432. doi : 10.24963/ijcai.2017/197. HAL : hal-01573841.

Conférences nationales avec comité de lecture

- Aimene Belfodil, Mehdi Kaytoute, 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.
- Autres
- Aimene Belfodil, Sergei Kuznetsov & Mehdi Kaytoute (2019). « On Pattern Setups and Pattern Multistructures ». ArXiv : 1906.02963. HAL : hal-02147757.

Jury

Franck MORVAN, PR, Université Toulouse Paul Sabatier	Rapporteur
Allel HADJALI, PR, ENSMA, Poitiers	Rapporteur
Marie-Hélène ABEL, PR, UTC de Compiègne	Examinatrice
Parisa GHODOUS, PR, Université Lyon 1	Examinatrice
Marinette SAVONNET, MC, Université de Bourgogne à Dijon	Examinatrice
Mahmoud BARHAMGI, MC, Université Lyon 1	Co-encadrant
Djamal BENSLIMANE, PR, Université Lyon 1	Directeur de thèse
Jean-Claude DOUTHE, PDG, Valessentia, Paris	Invité

Résumé

L'approche de gestion des processus métier configurables (cBP) proposée dans ce travail offre une vue consolidée des variantes de ces processus. Une variante est le résultat de l'ajustement d'un cBP, en réponse à des besoins fonctionnels et / ou structurels. Cette approche utilise une base de connaissances pour suivre les spécificités de chaque variante représentée sous la forme d'arborescence de structure de processus configurable (cPST). Une implémentation de l'approche consistant à générer un cPST à partir d'un cBP est proposée. L'application de cette approche permet de calculer la complexité des processus métier configurables et de prouver son avantage pour améliorer la qualité du cBP.

Abstract

We propose an approach for managing configurable Business Processes (cBP) that offer a consolidated view over these processes' variants. A variant is the result of adjusting a BP in response to functional and/or structural needs. The approach uses a knowledge base to track the specificities of each variant that is represented as configurable Process Structure Tree (cPST). An implementation of the approach that consists of generating a cPST from a cBP is proposed. We apply this approach to compute configurable business process complexity and to prove its advantage for improving cBP quality.

Thèse

Soutenue le 28/08/2019

Établissement :
Université Claude Bernard Lyon 1

Contact :
Djamal.Benslimane@liris.cnrs.fr

Riyadh Benammar

Jury

Philippe RIGAUX, PR, CNAM- Paris
Bertrand COÛASNON, MC, INSA-Rennes
Marc PLANTEVIT, MC, UCB Lyon 1
Nicole VINCENT, PR, Université Paris Descartes
Gilles VENTURINI, PR, Université de Tours
Véronique EGLIN, PR, INSA-Lyon
Christine LARGERON, PR, UJM- Saint-Etienne
Mylène PARDOEN, IR, MSH Lyon-Saint-Etienne

Rapporteur
Rapporteur
Examinateur
Examinatrice
Examinateur
Directrice de thèse
Co-directrice de thèse
Co-directrice de thèse

Résumé

Cette thèse s'inscrit dans le contexte de la fouille de données appliquées aux partitions musicales manuscrites anciennes et vise une recherche de motifs mélodiques ou rythmiques fréquents définis comme des séquences de notes répétitives aux propriétés caractéristiques. On rencontre un grand nombre de déclinaisons possibles de motifs : les transpositions, les inversions et les motifs dits « miroirs ». Ces motifs permettent aux musicologues d'avoir un niveau d'analyse approfondi sur les oeuvres d'un compositeur ou d'un style musical. Dans un contexte d'exploration de corpus de grande taille où les partitions sont juste numérisées et non transcrites, une recherche automatisée de motifs vérifiant des contraintes ciblées devient un outil indispensable à leur étude. Pour la réalisation de l'objectif de détection de motifs fréquents sans connaissance a priori, nous sommes partis d'images de partitions numérisées. Après des étapes de prétraitements sur l'image, nous avons exploité et adapté un modèle de détection et de reconnaissance de primitives musicales (tête de notes, hampes..) de la famille de réseaux de neurones à convolutions de type Region-Proposal CNN (RPN). Nous avons ensuite développé une méthode d'encodage de primitives pour générer une séquence de notes en évitant la tâche complexe de transcription complète de l'oeuvre manuscrite. Cette séquence a ensuite été analysée à travers l'approche CSMA (Constraint String Mining Algorithm) que nous avons conçue pour détecter les motifs fréquents présents dans une ou plusieurs séquences avec une prise en compte de contraintes sur leur fréquence et leur taille, ainsi que la taille et le nombre de sauts autorisés (gaps) à l'intérieur des motifs. La prise en compte du gap a ensuite été étudiée pour contourner les erreurs de reconnaissance produites par le réseau RPN évitant ainsi la mise en place d'un système de post-correction des erreurs de transcription des partitions. Le travail a été finalement validé par l'étude des motifs musicaux pour des applications d'identification et de classification de compositeurs.

Abstract

This thesis is part of the data mining applied to ancient handwritten music scores and aims at a search for frequent melodic or rhythmic motifs defined as repetitive note sequences with characteristic properties. There are a large number of possible variations of motifs: transpositions, inversions and so-called "mirror" motifs. These motifs allow musicologists to have a level of in-depth analysis on the works of a composer or a musical style. In a context of exploring large corpora where scores are just digitized and not transcribed, an automated search for motifs that verify targeted constraints becomes an essential tool for their study. To achieve the objective of detecting frequent motifs without prior knowledge, we started from images of digitized scores. After pre-processing steps on the image, we exploited and adapted a model

Thèse

Soutenue le 28/11/2019

Établissement :
INSA Lyon

Contact :
Véronique.Eglin@liris.cnrs.fr

for detecting and recognizing musical primitives (note-heads, stems..) from the family of Region-Proposal CNN (RPN) convolution neural networks. We then developed a primitive encoding method to generate a sequence of notes without the complex task of transcribing the entire manuscript work. This sequence was then analyzed using the CSMA (Constraint String Mining Algorithm) approach designed to detect the frequent motifs present in one or more sequences, taking into account constraints on their frequency and length, as well as the size and number of gaps allowed within the motifs. The gap was then studied to avoid recognition errors produced by the RPN network, thus avoiding the implementation of a post-correction system for transcription errors. The work was finally validated by the study of musical motifs for composers identification and classification.

Jury

Aristides GIONIS, PR, Aalto University	Rapporteur
Marie-Christine ROUSSET, PR, LIG Grenoble	Rapporteuse
Tijl DE BIE, PR, Ghent University	Examineur
Alexandre TERMIER, PR, Université de Rennes 1	Examineur
Siegfried NIJSSEN, MC, Université catholique de Louvain	Examineur
Céline ROBARDET, PR, INSA-Lyon	Directrice de thèse
Marc PLANTEVIT, MC, Université Lyon 1	co-Directeur de thèse.

Résumé

Nous adressons le problème de découverte de motifs dans les graphes attribués. Cette structure de données correspond à un graphe augmenté par des attributs associés aux sommets. Ils permettent de modéliser efficacement et intuitivement une large variété de bases de données réelles. L'analyse de ce type de graphes peut offrir une grande opportunité pour extraire des informations utiles et actionnables, par exemple, l'analyse des réseaux sociaux, réseaux biologiques, réseaux internet, etc. Plusieurs méthodes ont été proposées pour identifier des motifs dans les graphes attribués. La fouille de graphes attribués nécessite des méthodes qui prennent en compte au même temps la structure du graphe et les attributs décrivant les sommets, et cela génère deux défis. Premièrement, il est important de définir un langage de motifs intuitif sur lequel on peut appliquer des stratégies de recherche efficaces. Un motif étant intuitif signifie qu'il peut être facilement interprété et compris par l'utilisateur. Sachant qu'un motif est généralement défini sur un sous-graphe, il peut donc être immense en nombre de sommets, ce qui le rend difficile à comprendre. Le coût d'assimilation du motif est donc une question importante qui doit être adressée. Le deuxième défi est la formalisation de la mesure de qualité (pertinence) des motifs. Un motif est généralement pertinent s'il décrit des propriétés locales qui sont exceptionnelles, autrement, ce motif serait déjà attendu en regardant les propriétés globales du graphe. Par ailleurs, la qualité d'un motif est subjective, i.e., elle dépend significativement de l'utilisateur final, de ses connaissances antérieures sur les données et de ses préférences. Généralement, un utilisateur considère qu'un motif est utile s'il lui fournit de nouvelles connaissances, particulièrement si ce motif lui informe sur des caractéristiques ou des sujets qui intéressent habituellement l'utilisateur. Un autre problème lié à la qualité des motifs est la redondance. En d'autres termes, une méthode de fouille de données peut retourner un ensemble de motifs qui donnent des informations redondantes, car ces motifs couvrent des parties significativement superposées de sommets et d'attributs. Dans cette thèse, nous adressons ces différents défis pour le problème de la fouille de graphes attribués. Plus précisément, nous définissons de nouveaux langages de motifs, des mesures de qualités, des algorithmes pour la fouille de graphes attribués. On réalise aussi des études empiriques approfondies pour évaluer la pertinence de ces contributions.

Abstract

We address the problem of pattern discovery in vertex-attributed graphs. This kind of structure consists of a graph augmented with attributes associated to vertices. Vertex-attributed graphs provide a powerful abstraction that can be used to represent many datasets in an intuitive manner. Mining these graphs can be very useful for many applications, such as analyzing social networks, biological networks, the World Wide Web, etc. Several methods have been proposed to identify patterns in these structures. Generally, these methods

**Ahmed Anes
Bendimerad**

Thèse

Soutenue le 05/09/2019

Établissement :
INSA Lyon

Contact :
Celine.Robardet@liris.cnrs.fr

define a pattern as a subgraph whose vertices satisfy some structural constraints (e.g., density, connectivity) and have a subset of attributes with homogeneous values.

When mining vertex-attributed graphs, the principled integration of both graph and attribute data poses two important challenges. First, we need to define a pattern syntax (the abstract form of patterns) that is intuitive and lends itself to efficient search. A pattern being intuitive means that it can be easily interpreted and assimilated by the user. Considering that a pattern is generally defined over a subgraph, a pattern can be often huge in terms of vertices, which makes it difficult to grasp. Thus, the assimilation cost of a pattern is an important question that needs to be addressed. The second challenge is the formalization of the pattern interestingness. A pattern is generally relevant if it depicts some local properties that are somehow exceptional, otherwise, it will be already expected from the overall properties of the graph. Furthermore, the interestingness of patterns is subjective in practice, i.e., it significantly depends on the final user, her background knowledge and her preferences. A user would consider that a pattern is useful if it brings some new knowledge to her, especially if this pattern informs about some features or topics that usually interest this user. Another common problem related to the interestingness of patterns is the redundancy issue in the result set. In other terms, a data mining approach may return a set of patterns that give redundant information, because these patterns cover very overlapping parts of vertices and attributes. Consequently, the quality of patterns depends on many different factors. In this thesis, we address these challenges for the problem of mining attributed graphs. More precisely, we define new pattern languages, interestingness models, and algorithms for attributed graphs, and we perform extensive experimental evaluations to assess the relevance of these contributions and their ability to solve the addressed challenges.

Publications

Revues internationales avec comité de lecture

- Anes Bendimerad, Marc Plantevit, Céline Robardet & Sihem Amer-Yahia (2019). « User-driven geolocated event detection in social media ». IEEE Transactions on Knowledge and Data Engineering. doi : 10.1109/TKDE.2019.2931340. HAL : hal-02272082.
- 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 Kaytue, 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.

Conférences internationales avec comité de lecture

- Anes Bendimerad, Jefrey Lijffijt, Marc Plantevit, Céline Robardet & Tijl De Bie (2019). « Contrastive antichains in hierarchies ». SIGKDD 2019, 8 août 2019, Anchorage, Alaska (États-Unis), 294:304. doi : 10.1145/3292500.3330954. HAL : hal-02114775.
- Anes Bendimerad, Ahmad Mel, Jefrey 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.

- 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.
- 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.

Mahdi Bennara

Jury

Elisabeth MURISASCO, PR, Université de Toulon	Rapporteuse.
Agnès FRONT, MC, Université de Grenoble	Rapporteuse.
Pierre MARET, PR, Université de Saint-Étienne	Examineur.
Marinette SAVONNET, MC, Université de Bourgogne	Examinatrice.
Youssef AMGHAR, PR, Université de Lyon INSA de Lyon	Directeur de thèse.
Michaël MARISSA, PR, Univ. de Pau et des Pays de l'Adour	Directeur de thèse.

Résumé

L'informatique orientée services facilite l'interopérabilité entre les systèmes distribués. Depuis quelques années, l'émergence du Web sémantique a posé de nouveaux défis pour la communauté de recherche dans les calculs et la compatibilité sémantique des données. L'approche «services» et le Web sémantique constituent une piste prometteuse pour remédier aux problèmes qui entravent les deux domaines. D'une part l'orientation services permet d'assurer l'interopérabilité des données et des traitements au niveau sémantique, et d'autre part le Web sémantique permet d'automatiser les tâches de manipulation de services à un haut niveau. Dans le cadre de notre travail de recherche, nous avons détaillé les défis que rencontre la communauté de chercheurs dans l'intégration des pratiques de l'orientation services dans le Web sémantique, et plus particulièrement l'intégration des services REST dans l'implémentation du Web qui repose sur les principes du «Linked Data» pour constituer ce que l'on appelle les «RESTful Linked Services». Les défis en question sont : La description, la découverte, la sélection et la composition. Nous avons proposé une solution pour chacun de ces défis. Les contributions que nous avons proposées sont : la structure de descripteur, un algorithme de découverte sémantique, un algorithme de sélection basé sur Skyline et les répertoires de composition. Nous pensons que l'ensemble de contributions que nous avons proposées peut être adopté par les fournisseurs de services sur le Web afin de faciliter l'intégration des pratiques du sémantique Web avec les technologies des services et de REST en particulier. Ceci permettra donc d'automatiser les tâches de manipulation de services à un haut niveau, tel que la découverte sur la base de concepts sémantiques, la sélection sur la base de propriétés non-fonctionnelles et de qualité de services et la composition de plusieurs services hétérogènes, sur le plan des données ainsi que sur le plan des traitements, afin d'obtenir des services composites avec de la valeur ajoutée.

Abstract

Service Oriented Computing allows interoperability between distributed systems. In the last years, the emergence of the semantic Web opened new challenges for the research community regarding semantic interoperability on the data and processing levels. The convergence of service orientation and the semantic Web together is a promising effort to solve the problems that hampered both research fields. On the one hand, service orientation allows interoperability on the data and processing levels, and on the other hand, semantic Web allows the automation of high-level service manipulation tasks. In our research, we detail the challenges encountered by the research community to integrate the service orientation practices with the semantic Web, more precisely, integrating REST-based services with the semantic Web implementation based on Linked Data principles to obtain RESTful Linked Services. The challenges in question are: description, discovery, selection and composition. We proposed a solution for each of these challenges. The

Thèse

Soutenue le 18/07/2019

Établissement :
INSA Lyon

Contact :
Youssef.Amghar@iris.cnrs.fr

Jocelyn Bernard

contributions we proposed are: The descriptor structure, a semantically-enabled discovery algorithm, a Skyline-based selection algorithm and composition directories. We think that these contributions can be adopted by service providers on the Web in order to allow a seamless integration of semantic Web practices with the service technologies and REST in particular. This allows the automation of high-level service manipulation tasks, such as semantically-enabled discovery, QoS-based selection and the composition of heterogeneous services, be it on the data or processing level, in order to create value-added composite services.

Jury

Jean-Luc BARIL, PR, Université de Bourgogne	Rapporteur
Thi-Bich-Hanh DAO, MC, Université d'Orléans	Rapporteuse
Angela BONIFATI, PR, Université Lyon 1	Examinatrice
Daniela GRIGORI, PR, Université Paris-Dauphine	Examinatrice
Hamamache KHEDDOUCI, PR, Université Lyon 1	Directeur de thèse
Julien GONCALVES, Directeur R&D, société ReportLinker, Lyon	invité

Résumé

Dans cette thèse nous étudierons des problématiques de graphes. Nous allons étudier des problématiques théoriques en recherche de motifs (patterns) et des problématiques appliquées en diffusion d'informations. Nous proposons deux études théoriques sur la recherche et l'énumération de sous-graphes denses que sont les cliques et quasi-cliques. Ensuite nous proposons une étude appliquée sur la propagation d'information dans un graphe d'entités nommées. Dans un premier temps, nous allons étudier la recherche de cliques dans des graphes compressés. Les problèmes MCE (pour l'anglais Maximal Clique Enumeration) et MCP (pour l'anglais Maximum Clique Problem) sont des problèmes rencontrés dans l'analyse des graphes de données. Ce sont des problèmes difficiles (NP-difficile pour MCE et NP-Complet pour MCP) pour lesquels des solutions adaptées doivent être conçues pour les grands graphes. Nous proposons de répondre à ces problèmes en travaillant sur une version compressée du graphe initial. Nous montrons les bons résultats obtenus par notre méthode pour l'énumération de cliques maximales sur des graphes compressés. Dans un second temps, nous étudierons l'énumération de quasi-cliques maximales. Nous proposons un algorithme distribué qui énumère l'ensemble des quasi-cliques maximales du graphe. Nous démontrons que cet algorithme liste l'ensemble des quasi-cliques maximales du graphe. Nous proposons également une heuristique qui liste un ensemble de quasi-cliques plus rapidement. Nous montrons l'intérêt de l'énumération de ces quasi-cliques par une évaluation des relations en regardant la co-occurrence des noeuds dans l'ensemble des quasi-cliques énumérées. Dans un troisième temps, nous travaillerons sur la diffusion d'événements dans un graphe d'entités nommées. De nombreux modèles existent pour simuler des problèmes de diffusion de rumeurs ou de maladies dans des réseaux sociaux ainsi que des problèmes de propagation de faillites dans les milieux bancaires. Nous proposons de répondre au problème de diffusion d'événements marquant dans des réseaux hétérogènes représentant un environnement économique du monde. Nous proposons un problème de diffusion, nommé problème de classification de l'infection, qui consiste à déterminer quelles entités sont concernées par un événement. Pour résoudre ce problème, nous proposons deux modèles inspirés du modèle de seuil linéaire auxquels nous ajoutons différentes fonctionnalités. Finalement, nous testons et validons nos modèles sur un ensemble d'événements.

Abstract

In this thesis we will study graph problems. We will study theoretical problems in pattern research and applied problems in information diffusion. We propose two theoretical studies on the identification/detection and enumeration of dense subgraphs, such as cliques and quasi-cliques. Then we propose an applied study on the propagation of information in a named entities graph. First, we will study the identification/detection of cliques in compressed graphs. The MCE and MCP are problems that are encountered in the analysis

Thèse

Soutenue le 06/06/2019

Établissement :
Université Claude Bernard Lyon 1

Contact :
Hamamache.Kheddouci@liris.cnrs.fr

of data graphs. These problem are difficult to solve (NP-Hard for MCE and NP-Complete for MCP), and adapted solutions must be found for large graphs. We propose to solve these problems by working on a compressed version of the initial graph. We show the correct results obtained by our method for the enumeration of maximal cliques on compressed graphs. Secondly, we will study the enumeration of maximal quasi-cliques. We propose a distributed algorithm that enumerates the set of maximal quasi-cliques of the graph. We show that this algorithm lists the set of maximal quasi-cliques of the graph. We also propose a heuristic that lists a set of quasi-cliques more quickly. We show the interest of enumerating these quasi-cliques by an evaluation of relations by looking at the co-occurrence of nodes in the set of enumerated quasi-cliques. Finally, we work on the event diffusion in a named entities graph. Many models exist to simulate diffusion problems of rumors or diseases in social networks and bankruptcies in banking networks. We address the issue of significant events diffusion in heterogeneous networks, representing a global economic environment. We propose a diffusion problem, called infection classification problem, which consists to determine which entities are concerned by an event. To solve this problem we propose two models inspired by the linear threshold model to which we add different features. Finally, we test and validate our models on a set of events

Publications

Conférences internationales avec comité de lecture

- Jocelyn Bernard, Julien Goncalves & Hamamache Kheddouci (2019). « Propagation d'événements dans un graphe économique ». Extraction et Gestion des connaissances (EGC), 25 janvier 2019, Metz (France), pp. 315-320. HAL : hal-02096635.
- Jocelyn Bernard, Sicong Shao, Cihan Tunc, Hamamache Kheddouci & Salim Hariri (2018). « Quasi-cliques Analysis for IRC Channel Thread Detection ». Complex Networks, 11 décembre 2018, Cambridge (France), pp. 578-589. doi : 10.1007/978-3-030-05411-3_47. HAL : hal-01997430.
- Bernard Jocelyn & Hamida Seba (2018). « Solving the Maximal Clique Problem on Compressed Graphs ». 24th International Symposium on Methodologies on Intelligent (ISMIS 2018), 29 octobre 2018, LImassol (Chypre), pp. 45-55. doi : 10.1007/978-3-030-01851-1_5. HAL : hal-01886654.

Conférences nationales avec comité de lecture

- Jocelyn Bernard & Hamida Seba (2018). « Résolution de problèmes de cliques dans les grands graphes ». EGC 2018, 22 janvier 2018, Paris (France). HAL : hal-01886724.

Jury

Franck MULTON, PR, Université Rennes 2	Rapporteur
Lionel REVERET, CR, INRIA Rhône-Alpes	Rapporteur
Marie-Paule CANI, PR, Ecole Polytechnique	Examinatrice
Sylvie GIBET, PR, Université de Bretagne Sud	Examinatrice
Atila BASKURT, PR, INSA-LYON	Directeur de thèse
Saïda BOUAKAZ, PR, Univ. Claude Bernard Lyon 1	Co-directrice de thèse
Nicolas PRONOST, MC, Univ. Claude Bernard Lyon 1	Co-encadrant de thèse

Résumé

Cette thèse traite de l'animation de personnages virtuels composés de corps rigides reliés par des articulations et contrôlés par des interactions physiques (forces et moments). Le contrôleur est le système qui calcule dynamiquement ces interactions. Notre objectif est d'étudier et de réaliser un contrôleur pour la simulation de mouvements d'un personnage en interaction avec un fluide. La complexité du comportement de tels milieux ne permet pas de prédire les interactions entre le personnage et le fluide. Il en découle que le contrôleur proposé doit être capable de réagir à celles-ci.

Nous avons focalisé nos travaux sur la conception d'un contrôleur de type SIMBICON capable de s'adapter aux perturbations apportées par la présence d'un fluide simulé physiquement. Ce choix est motivé par notre contrôleur précédent qui proposait un contrôleur en interaction avec un fluide représenté à travers l'utilisation de formules simples de dynamique des fluides. L'utilisation d'une véritable simulation physique du fluide nous permettrait d'améliorer le réalisme physique de la simulation en prenant en compte l'impact du déplacement du personnage sur le fluide. Ayant pour objectif un contrôleur interactif nous nous sommes focalisés sur deux axes principaux. Le premier est la conception d'un contrôleur capable de supporter des fréquences de simulation faibles tout en conservant la vitesse de calcul apporté par l'utilisation du modèle SIMBICON. Nous proposons de réduire les instabilités introduites par l'utilisation de fréquences de simulation faibles par un système de feedback utilisant une optimisation en ligne permettant d'obtenir une meilleure stabilité des contacts. Ce système, associé à une étude des paramètres du système en fonction de la fréquence de simulation, nous a permis de proposer un contrôleur capable de supporter des fréquences de simulation allant jusqu'à 225Hz. Le second axe de recherche visait à proposer une implémentation entièrement GPU et interactive d'une simulation lagrangienne de fluide. Nous avons étudié l'impact sur les performances de notre implémentation GPU de plusieurs optimisations proposées par des travaux proposant des implémentations parallèles CPU. Nous proposons également une solution permettant de déplacer la zone de fluide simulé en cours de simulation pour limiter l'espace de simulation du fluide à la proximité immédiate du personnage au cours de son déplacement pour assurer une simulation du fluide en temps interactif.

Abstract

This thesis deals with the animation of virtual characters composed of rigid bodies linked by joints and controlled by physic interactions (forces and torques). The controller is the system that dynamically calculates these interactions. Our goal is to study and create a controller that is able to control the character interacting with a fluid. The complexity of the behavior of such environment renders predicting the interactions between the fluid

Thèse

Soutenue le 05/07/2019

Établissement :
Université Claude Bernard Lyon 1

Contact :
Nicolas.Pronost@liris.cnrs.fr

**Rubiela Carrillo
Roso**

and the character impossible. Therefore, the controller must be able to react to such interactions. We have focused our works on the conception of a SIMBICON typed controller that is able to handle the perturbations caused by the physically simulated fluid. This choice is motivated by our previous controller that was able to handle the interactions with a simplified fluid based on simplified fluid dynamics equations. The use of the complete fluid dynamic equations should allow for a higher realism by taking into account to impact of the character motion on the fluid. Since our objective is to obtain an interactive controller, we focus our works on two main axes. The first one is the conception of a controller able to handle low simulation frequencies while keeping the high computation speed brought by the use of the SIMBICON framework. We propose to use a feedback system using an online optimization to reduce the instabilities caused by the of the low simulation frequencies. On top of this system, we study the evolution of the value of the PD-controllers depending on the simulation frequency to be able to propose a controller able to handle simulation frequencies as low as 225Hz. The second research axis aims to conceive a fully GPU implementation of a Lagrangian fluid simulation. We study the impact of various optimization proposed by previous work on our GPU implementation. We also present a system allowing us to move the simulation area of the fluid to be able to keep the character centered in the fluid simulation. This system allows us to only use a small area for the fluid simulation allowing us to propose a system showing interactive execution times.

Publications

Conférences internationales avec comité de lecture

- Samuel Carensac, Nicolas Pronost & Saïda Bouakaz (2018). « Physics-based control of walking virtual characters in low frequency simulations ». 31st International Conference on Computer Animation and Social Agents CASA 2018, 23 mai 2018, Pekin (Chine), pp. 77-82. doi : 10.1145/3205326.3205361. HAL : hal-01884827.
- Samuel Carensac, Nicolas Pronost & Saïda Bouakaz (2017). « Physics-based control of virtual characters in low frequency simulations ». journées françaises d'informatique graphique (j.FIG), 23 octobre 2017, rennes (France). HAL : hal-01765240.
- Samuel Carensac, Nicolas Pronost & Saïda Bouakaz (2015). « Real-time gait control for partially immersed bipeds ». ACM Motion in Games, 18 novembre 2015, Paris (France), pp. 177-182. HAL : hal-01263675.

Jury

Vanda LUENGO, PR, Sorbonne Université	rapporteur
Sébastien GEORGE, PR, Le Mans Université	rapporteur
Nathalie GUIN, MC, Université Lyon 1	examineur
Stéphanie MAILLES-VIARD METZ, MC, Univ. Montpellier 3	examineur
Mohand-Saïd HACID, PR, Université Lyon 1	directeur de thèse
Élise LAVOUÉ, MC, Université Lyon 3	co-directeur
Yannick PRIÉ, PR, Université de Nantes	co-directeur

Résumé

À la différence de l'apprentissage par mémorisation (rote learning), l'apprentissage significatif (meaningful learning) vise à associer de nouvelles connaissances à des connaissances déjà acquises. La construction de cartes mentales exige et supporte la mise en place de stratégies d'apprentissage significatif, et permet de rendre visible la structure de connaissances de l'apprenant. Cependant, les enseignants qui intègrent la construction de cartes mentales dans leurs activités pédagogiques doivent se contenter du rendu final des cartes mentales, et risquent de faire de mauvaises interprétations et évaluations de celles-ci par manque d'information sur leur processus de construction. Dans cette thèse, nous nous intéressons à l'observation a posteriori de l'engagement de l'apprenant dans ses dimensions comportementale et cognitive, afin de proposer des indicateurs orientés processus qui permettent de comprendre ses actions et ses choix de construction de carte mentale. Nous avons suivi la méthodologie de recherche orientée par la conception (Design Based Research), qui nous a permis de proposer trois niveaux de contributions : 1) un modèle théorique d'engagement (comportemental et cognitive) pour les activités de construction de cartes mentales, 2) un ensemble d'indicateurs d'engagement de l'apprenant à partir de traces capturées automatiquement lors de la construction de cartes mentales, et 3) un tableau de bord appelé MindMap Monitor présentant différents indicateurs à l'enseignant pour le suivi de la classe et des apprenants. Le modèle a été obtenu à partir d'une étude de littérature sur les théories de l'engagement de l'apprenant issues notamment de la recherche en psychologie de l'éducation. Les indicateurs ont été définis en croisant le modèle avec les résultats de plusieurs études de terrain avec les enseignants. Le tableau de bord implémentant les indicateurs a été conçu en trois itérations. Son interface présente des visualisations synthétiques permettant de comparer les élèves de la classe et d'identifier ceux en difficulté lors de la construction de leur carte, et des visualisations détaillées décrivant l'activité de construction d'une carte pour chaque élève. Notre tableau de bord a été évalué au cours d'une expérimentation avec 12 enseignants en comparant son utilisation avec celle des cartes mentales finales associées aux vidéos de leur processus de construction. Les résultats montrent que nos indicateurs sur MindMap Monitor permettent de mieux identifier les élèves en difficulté, les difficultés partagées, ainsi que les difficultés pour un élève. Les résultats concernant la compréhension du processus de construction de la carte mentale sont plus nuancés. Nous avons également pu identifier plusieurs pistes d'amélioration sur le contenu du tableau de bord et sa présentation. Les perspectives de notre travail concernent principalement le suivi de l'engagement des apprenants en temps réel pour l'intervention et l'adaptation de la stratégie pédagogique.

Thèse

Soutenue le 11/03/2019

Établissement :
Université Claude Bernard Lyon 1

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

Abstract

In contrast to rote learning, meaningful learning aims to associate new knowledge with knowledge already acquired. Mind mapping activities require and support the implementation of meaningful learning strategies and enlighten the knowledge structure of the learner. However, teachers who integrate mind mapping into their educational activities have to deal with the final rendering of maps, and risk to misinterpret and wrongly evaluate them due to the lack of information about their construction process. In this thesis, we are interested in a posteriori observation of the engagement of learners along its behavioral and cognitive dimensions, in order to propose process oriented indicators that help to understand actions and construction choices of mind maps. We followed the Design Based Research methodology, that allowed us to propose three levels of contributions: 1) a theoretical model of engagement (behavioral and cognitive) for mind mapping activities, 2) a set of indicators of learner engagement constructed from automatically captured map building traces, and 3) a dashboard called MindMap Monitor presenting various indicators to teachers for class and learners monitoring. The model was obtained from a literature review on theories of engagement, including research in educational psychology. The indicators have been defined by comparing the model with the results of several field studies with teachers. The dashboard implementing the indicators was developed following three iterations. Its interface presents synthetic views allowing the comparison of students in the class, the identification of those in difficulty, and detailed views describing the mind mapping activity for each student. Our dashboard was evaluated with an experiment involving 12 teachers. We compared its use with that of final mind maps associated with videos of their construction process. Results show that our indicators on MindMap Monitor are useful to better identify students in difficulty, shared difficulties, as well as difficulties for individual students. Results concerning the understanding of the mind maps construction process are more balanced. We were also able to identify several ways to improve both the content and the visualizations of the dashboard. The perspectives of our work are mainly related to monitoring learners' engagement in real time for the intervention and adaptation of the teachers' educational strategies

Publications

Revue internationale avec comité de lecture

- Elise Lavoué, Maedeh Kazemitabar, Tenzin Doleck, Susanne Lajoie, Rubiela Carrillo & Gaëlle Molinari (2019). « Towards emotion awareness tools to support emotion and appraisal regulation in academic contexts ». *Educational Technology Research and Development*, pp. 1-24. doi : 10.1007/s11423-019-09688-x. HAL : hal-02187319.

Revue nationale avec comité de lecture

- Rubiela Carrillo, Yannick Prié, Leslie Guillaume & Elise Lavoué (2018). « Suivre l'engagement des apprenants dans l'activité de construction de cartes mentales ». *Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation*, vol. 25, n°1, pp. 1-29. doi : 10.23709/sticf.25.1.5. HAL : hal-01939672.

Conférences internationales avec comité de lecture

- Rubiela Carrillo, Clément Renaud, Yannick Prié & Élise Lavoué (2017). « Dashboard for monitoring student engagement in mind mapping activities ». 2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT), 3 juillet 2017, Timisoara (Roumanie), pp. 433-437. doi : 10.1109/ICALT.2017.100. HAL : hal-01631926.
- Rubiela Carrillo, Elise Lavoué & Yannick Prié (2016). « Towards Qualitative Insights for Visualizing Student Engagement in Web-based Learning Environments ». Workshop on Web Science and Technology for Education (WEBED 2016), in conjunction with the 25th International Conference Companion on World Wide Web (WWW 2016), 11 avril 2016, Montreal (Canada), pp. 893-898. doi : 10.1145/2872518.2891076. HAL : hal-01313398.

Conférences nationales avec comité de lecture

- Rubiela Carrillo, Clément Renaud, Yannick Prié & Élise Lavoué (2017). « Tableau de bord pour le suivi de l'engagement des apprenants lors de la construction de cartes mentales ». 8^e Conférence sur les Environnements Informatiques pour l'Apprentissage Humain (EIAH 2017), 6 juin 2017, Strasbourg (France), pp. 113-124. HAL : hal-01631931.
- Rubiela Carrillo, Elise Lavoué & Yannick Prié (2015). « Vers l'analyse de l'engagement de l'apprenant : une approche par la visualisation multi-échelle de traces hétérogènes ». Deuxième journée EIAH & IA, 3 juillet 2015, Rennes (France), pp. 13-22. HAL : hal-01231811.

Autres

- Rubiela Carrillo, Yannick Prié & Elise Lavoué (2019). « Observing Learner Engagement on Mind Mapping Activities Using Learning Analytics ». 14th European Conference on Technology Enhanced Learning (EC-TEL 2019), 19 septembre 2019, Delft (Pays-Bas). Poster. doi : 10.1007/978-3-030-29736-7_63. HAL : hal-02314973.
- Maedeh A. Kazemitabar, Elise Lavoué, Susanne P. Lajoie, Rubiela Carrillo & Tenzin Doleck (2019). « Examining the Role of Emotion Awareness and Sharing Emotions during Collaborative Learning ». 13th international conference on Computer Supported Collaborative Learning (CSCL 2019), 17 juin 2019, France (France). Poster. HAL : hal-02053182.

Ugo Comignani

Jury

Reinhard PICHLER, PR, Technische Universität Wien	Rapporteur
Pierre SENELLART, PR, ENS/Université PSL	Rapporteur
Laure BERTI-ÉQUILLE, DR, Aix-Marseille Université	Examinatrice
Marie-Laure MUGNIER, PR, Université de Montpellier	Examinatrice
Hamamache KHEDDOUCI, PR, Université Lyon 1	Examineur
Angela BONIFATI, PR, Université Lyon 1	Directrice de thèse
Emmanuel COQUERY, MC, Université Lyon 1	Co-directeur
Romuald THION, MC, Université Lyon 1	Co-directeur

Résumé

La migration de données entre des sources aux schémas hétérogènes est un domaine en pleine croissance avec l'augmentation de la quantité de données en accès libre, et le regroupement des données à des fins d'apprentissage automatisé et de fouilles. Cependant, la description du processus de transformation des données d'une instance source vers une instance définie sur un schéma différent est un processus complexe même pour un utilisateur expert dans ce domaine. Cette thèse aborde le problème de la définition de mapping par un utilisateur non expert dans le domaine de la migration de données, ainsi que la vérification du respect par ce mapping des contraintes d'accès ayant été définies sur les données sources. Pour cela, dans un premier temps nous proposons un système dans lequel l'utilisateur fournit un ensemble de petits exemples de ses données, et est amené à répondre à des questions booléennes simples afin de générer un mapping correspondant à ses besoins. Dans un second temps, nous proposons un système permettant de réécrire le mapping produit de manière à assurer qu'il respecte un ensemble de vues de contrôle d'accès définis sur le schéma source du mapping. Plus précisément, le premier grand axe de cette thèse est la formalisation du problème de la définition interactive de mappings, ainsi que la description d'un cadre formel pour la résolution de celui-ci. Cette approche formelle pour la résolution du problème de définition interactive de mappings est accompagnée de preuves de bonnes propriétés. À la suite de cela, basés sur le cadre formel défini précédemment, nous proposons des algorithmes permettant de résoudre efficacement ce problème en pratique. Ces algorithmes visent à réduire le nombre de questions auxquelles l'utilisateur doit répondre afin d'obtenir un mapping correspondant à ces besoins. Pour cela, les mappings possibles sont ordonnés dans des structures de treillis imbriqués, afin de permettre un élagage efficace de l'espace des mappings à explorer. Nous proposons également une extension de cette approche à l'utilisation de contraintes d'intégrité afin d'améliorer l'efficacité de l'élagage. Le second axe majeur vise à proposer un processus de réécriture de mapping qui, étant donné un ensemble de vues de contrôle d'accès de référence, permet d'assurer que le mapping réécrit ne laisse l'accès à aucune information n'étant pas accessible via les vues de contrôle d'accès. Pour cela, nous définissons un protocole de contrôle d'accès permettant de visualiser les informations accessibles ou non à travers un ensemble de vues de contrôle d'accès. Ensuite, nous décrivons un ensemble d'algorithmes permettant la réécriture d'un mapping en un mapping sûr vis-à-vis d'un ensemble de vues de contrôle d'accès. Comme précédemment, cette approche est complétée de preuves de bonnes propriétés. Afin de réduire le nombre d'interactions nécessaires avec l'utilisateur lors de la réécriture d'un mapping, une approche permettant l'apprentissage des préférences de l'utilisateur est proposée, cela afin de permettre le choix entre un processus interactif ou automatique. L'ensemble des algorithmes décrits dans cette thèse ont fait l'objet d'un

Thèse

Soutenue le 19/09/2019

Établissement :
Université Claude Bernard Lyon 1

Contact :
Angela.Bonifati@liris.cnrs.fr

prototypage et les expériences réalisées sur ceux-ci sont présentées dans cette thèse.

Abstract

Data exchange between sources over heterogeneous schemes is an ever-growing field of study with the increased availability of data, oftentimes available in open access, and the pooling of such data for data mining or learning purposes. However, the description of the data exchange process from a source to a target instance defined over a different schema is a cumbersome task, even for users acquainted with data exchange. In this thesis, we address the problem of allowing a non-expert user to specify a source-to-target mapping, and the problem of ensuring that the specified mapping does not leak information forbidden by the security policies defined over the source. To do so, we first provide an interactive process in which users provide small examples of their data, and answer simple boolean questions in order to specify their intended mapping. Then, we provide another process to rewrite this mapping in order to ensure its safety with respect to the source policy views. As such, the first main contribution of this thesis is to provide a formal definition of the problem of interactive mapping specification, as well as a formal resolution process for which desirable properties are proved. Then, based on this formal resolution process, practical algorithms are provided. The approach behind these algorithms aims at reducing the number of boolean questions users have to answers by making use of quasi-lattice structures to order the set of possible mappings to explore, allowing an efficient pruning of the space of explored mappings. In order to improve this pruning, an extension of this approach to the use of integrity constraints is also provided. The second main contribution is a repairing process allowing to ensure that a mapping is “safe” with respect to a set of policy views defined on its source schema, i.e., that it does not leak sensitive information. A privacy-preservation protocol is provided to visualize the information leaks of a mapping, as well as a process to rewrite an input mapping into a safe one with respect to a set of policy views. As in the first contribution, this process comes with proofs of desirable properties. In order to reduce the number of interactions needed with the user, the interactive part of the repairing process is also enriched with the possibility of learning which rewriting is preferred by users, in order to obtain a completely automatic process. Last but not least, we present experiments over the open source prototypes built from two contributions of this thesis.

Jury

Boubakeur BOUFAMA, PR, Université de Windsor, Canada	Rapporteur
Renaud SEGUIER, PR, Centrale Supélec, Rennes	Rapporteur
Edmond BOYER, DR, INRIA Grenoble Rhône-Alpes	Examinateur
Valérie BURDIN, PR, Université de Bretagne Occidentale	Examinatrice;
Salima HASSAS, PR, Université Claude Bernard Lyon 1	Examinatrice
Saida BOUAKAZ, PR, Université Claude Bernard Lyon 1	Directrice de thèse;
Hubert KONIK, MC, Université Jean Monnet	Co-Encadrant;
Alexandre MEYER, MC, Université Claude Bernard Lyon 1	Co-Encadrant;

Résumé

Le thème de ma thèse concerne la reconnaissance et la synthèse d'expressions faciales et corporelles. Notre problématique est d'étudier, de comprendre et d'extraire les éléments qui traduisent l'état émotionnel d'une personne à partir des expressions de son visage et de son corps, dans le but de la reconnaissance et également de la synthèse de l'émotion ou du style dans un geste. Ce double objectif de reconnaissance et de synthèse permettra de généraliser de nouveaux modes d'interactions dans des applications comme les jeux-vidéos, l'interaction homme-machine, etc. En effet, ces applications pourront s'enrichir de scénario qui pourraient d'adapter à l'état émotionnel de l'utilisateur. Pour répondre à ce problème, le point crucial est de comprendre « où » se situe l'information de style dans une action ou une animation. En effet, un humain sait en quelques secondes caractériser une expression qu'il voit alors que les algorithmes de reconnaissance en sont encore loin notamment pour les expressions corporelles où à notre connaissance, peu de travaux ont été réalisés comparé à la reconnaissance des expressions faciales. Concernant la reconnaissance des expressions corporelles, notre objectif est, dans un premier temps, de proposer des descripteurs capables de reconnaître l'expression portée par une posture. Dans un second temps, nous souhaitons également que ces descripteurs permettent de séparer le mouvement réalisé de l'expression perçue, afin d'être réutilisé pour la synthèse d'animations. En effet, dans le domaine de l'animation, la création d'une action portant une émotion ou un style nécessite énormément de travail, de savoir-faire et de temps à un animateur afin de proposer des animations stylisées. Par exemple, dans un jeu-vidéo, pour créer de telles animations, il est souvent nécessaire de disposer d'une énorme base de données de mouvements incluant chaque style pour chaque personnage virtuel. Pour disposer d'une telle base de données, on a souvent recours à l'une des 2 méthodes suivantes. La première consiste à procéder à la capture de tous les mouvements réalisés par différents acteurs jouant différents styles. Dans la seconde, c'est le graphiste qui doit réaliser les différentes animations à la main en utilisant un logiciel d'animation. Notre objectif dans ce cadre est de mettre à profit les descripteurs quantifiant l'expression détectée en reconnaissance d'expressions faciales et corporelles afin de développer des outils capables de changer / éditer le style ou l'expression d'une animation. Ces outils permettront d'assister et de faciliter le travail des graphistes en leur permettant de synthétiser rapidement une « animation primale » stylisée. Ces animations stylisées pourront être affinées en post-processing en apportant une touche artistique à l'animation générée.

Thèse

Soutenue le 17/12/2019

Établissement :
Université Claude Bernard Lyon 1

Contact :
Saida.Bouakaz@liris.cnrs.fr

Jury

Thierry CHATEAU, PR, UCA
Alice CAPLIER, PR, Grenoble-INP
Michel PAINDAVOINE, PR, Université de Bourgogne
Christophe GARCIA, PR, INSA Lyon
Khalid IDRISSE, MC, INSA Lyon

Rapporteur
Rapporteur
Examineur
Co-directeur de thèse
Directeur de thèse

Résumé

Le suivi d'objets multiples dans une scène est une tâche importante dans le domaine de la vision par ordinateur, et présente toujours de très nombreux verrous. Les objets doivent être détectés et distingués les uns des autres de manière continue et simultanée. Les approches «suivi par détection» sont largement utilisées, où la détection des objets est d'abord réalisée sur toutes les frames, puis le suivi est ramené à un problème d'association entre les détections d'un même objet et les trajectoires identifiées. La plupart des algorithmes de suivi associent des modèles de mouvement et des modèles d'apparence. Dans cette thèse, nous proposons un modèle de ré-identification basé sur l'apparence et utilisant l'apprentissage de métrique de similarité. Nous faisons tout d'abord appel à un réseau siamois profond pour apprendre un mapping de bout en bout, des images d'entrée vers un espace de caractéristiques où les objets sont mieux discriminés. De nombreuses configurations sont évaluées, afin d'en déduire celle offrant les meilleurs scores. Le modèle ainsi obtenu atteint des résultats de ré-identification satisfaisants comparables à l'état de l'art. Ensuite, notre modèle est intégré dans un système de suivi d'objets multiples pour servir de guide d'apparence pour l'association des objets. Un modèle d'apparence est établi pour chaque objet détecté s'appuyant sur le modèle de ré-identification. Les similarités entre les objets détectés sont alors exploitées pour la classification. Par ailleurs, nous avons étudié la coopération et les interférences entre les modèles d'apparence et de mouvement dans le processus de suivi. Un couplage actif entre ces 2 modèles est proposé pour améliorer davantage les performances du suivi, et la contribution de chacun d'eux est estimée en continue. Les expérimentations menées dans le cadre du benchmark «Multiple Object Tracking Challenge» ont prouvé l'efficacité de nos propositions et donné de meilleurs résultats de suivi que l'état de l'art.

Abstract

Multiple object tracking, i.e. simultaneously tracking multiple objects in the scene, is an important but challenging visual task. Objects should be accurately detected and distinguished from each other to avoid erroneous trajectories. Since remarkable progress has been made in object detection field, "tracking-by-detection" approaches are widely adopted in multiple object tracking research. Objects are detected in advance and tracking reduces to an association problem: linking detections of the same object through frames into trajectories. Most tracking algorithms employ both motion and appearance models for data association. For multiple object tracking problems where exist many objects of the same category, a fine-grained discriminant appearance model is paramount and indispensable. Therefore, we propose an appearance-based re-identification model using deep similarity metric learning to deal with multiple object tracking in mono-camera videos. Two main contributions are reported in this dissertation: First, a deep Siamese network is employed to learn an end-to-end mapping from input images to a discriminant embedding space. Different metric learning

Thèse

Soutenue le 12/09/2019

Établissement :
INSA Lyon

Contact :
Khalid.Idrissi@liris.cnrs.fr

Rémy Delanaux

configurations using various metrics, loss functions, deep network structures, etc., are investigated, in order to determine the best re-identification model for tracking. In addition, with an intuitive and simple classification design, the proposed model achieves satisfactory re-identification results, which are comparable to state-of-the-art approaches using triplet losses. Our approach is easy and fast to train and the learned embedding can be readily transferred onto the domain of tracking tasks. Second, we integrate our proposed re-identification model in multiple object tracking as appearance guidance for detection association. For each object to be tracked in a video, we establish an identity-related appearance model based on the learned embedding for re-identification. Similarities among detected object instances are exploited for identity classification. The collaboration and interference between appearance and motion models are also investigated. An online appearance-motion model coupling is proposed to further improve the tracking performance. Experiments on Multiple Object Tracking Challenge benchmark prove the effectiveness of our modifications, with a state-of-the-art tracking accuracy.

Publications

Conférences internationales avec comité de lecture

- Bonan Cuan, Khalid Idrissi & Christophe Garcia (2019). « Online Appearance-Motion Coupling for Multi-Person Tracking in Videos ». ECDMO 2019, 17 février 2019, Amsterdam (Pays-Bas). HAL : hal-01989056.
- Bonan Cuan, Khalid Idrissi & Christophe Garcia (2018). « Deep Siamese Network for Multiple Object Tracking ». IEEE 20th International Workshop on Multimedia Signal Processing (MMSP 2018), 29 août 2018, Vancouver (Canada). HAL : hal-01807982.

Jury

Benjamin NGUYEN, PR, INSA Centre-Val de Loire	rapporteur
Hala SKAF-MOLLI, MC, Université de Nantes	rapporteuse
Nicolas ANCIAUX, DR, INRIA Saclay - Île-de-France	examineur
Hamamache KHEDDOUCI, PR, Université Lyon 1	examineur
Nathalie PERNELLE, MC, Université Paris-Sud	examinatrice
Angela BONIFATI, PR, Université Lyon 1	directrice de thèse
Marie-Christine ROUSSET, PR, Univ. Grenoble Alpes	co-directrice de thèse
Romuald THION, MC, Université Lyon 1	co-directeur de thèse

Résumé

La confidentialité des données personnelles est un souci majeur et un problème peu étudié dans le cas publication de données dans le Web des données ouvertes (ou LOD cloud, pour Linked Open Data cloud). Ce nuage formé par le LOD est un réseau d'ensembles de données interconnectés et accessibles publiquement sous la forme de graphes de données modélisés dans le format RDF, et interrogés via des requêtes écrites dans le langage SPARQL. Ce cadre fortement standardisé est très utilisé de nos jours par des organismes publics comme des entreprises. Mais certains acteurs (notamment du secteur privé) sont toujours réticents à la publication de leurs données, découragés par de potentiels soucis potentiel de confidentialité. Pour pallier cela, nous présentons et développons un cadre formel déclaratif pour la publication de données liées respectant la confidentialité, dans lequel les contraintes de confidentialité et d'utilité des données sont spécifiées sous forme de politiques (des ensembles de requêtes SPARQL). Avec ce cadre formel, nous résolvons un problème de construction de séquences d'opérations d'anonymisation afin de construire un graphe anonymisé satisfaisant des politiques d'utilité et de confidentialité. Cette approche est indépendante des données et du graphe considéré, et consiste en l'analyse statique des deux politiques pour déterminer des séquences d'opérations d'anonymisation à appliquer à n'importe quel graphe RDF pour satisfaire les politiques fournies. Un autre aspect à prendre en compte est qu'un nouveau graphe publié dans le nuage LOD est évidemment exposé à des failles de confidentialité car il peut être relié à des données déjà publiées dans d'autres données liées. Dans le second volet de cette thèse, nous nous concentrons donc sur le problème de construction d'anonymisations sûres d'un graphe RDF garantissant que relier le graphe anonymisé à un graphe externe quelconque ne causera pas de brèche de confidentialité. En prenant un ensemble de requêtes de confidentialité en entrée, nous étudions le problème de sûreté indépendamment des données du graphe, et la construction d'une séquence d'opérations d'anonymisation permettant d'assurer cette sûreté. Nous détaillons des conditions suffisantes sous lesquelles une instance d'anonymisation est sûre pour une certaine politique de confidentialité fournie et montrons que nos algorithmes sont robustes même en présence de liens de type sameAs (liens d'égalité entre entités en RDF), qu'ils soient explicites ou inférés par de la connaissance externe. Enfin, nous évaluons l'impact de cette solution en la testant sur divers graphes. Nous montrons que cette approche est efficace même sur de gros graphes (plusieurs millions de triplets) et, via différentes métriques adaptées, que ces algorithmes ne sont que peu destructeurs, cela même quand les politiques de confidentialité couvrent une grosse partie du graphe. Via une méthode simple et efficace pour assurer la confidentialité et l'utilité de graphes RDF dans des cas d'utilisation plausibles, cette nouvelle approche suggère de nombreuses extensions possibles et, sur le long terme, plus de travail sur la

Thèse

Soutenue le 13/12/2019

Établissement :

Université Claude Bernard Lyon 1

Contact :

Angela.Bonifati@liris.cnrs.fr

publication de données liées respectueuse de la confidentialité.

Abstract

The recent development of Linked Data on the Web enabled the fast publication of huge amounts of open and linked data, structured as graphs using the RDF format structure; Integration of data from the Web has to make possible the unified interrogation of various sources, but also the enrichment of an organizational database, the augmentation of applications using content, or even a way to provide a global reference format to publish proprietary data. The absence of a common pattern makes the integration of Web data rely on instances themselves, matching entities referenced in the nodes of RDF graphs. Yet there may be many references, and they can be erroneous, which prevents a reliable use for integrating the matching data. Entity resolution is the automated process linking the two identifiers in order to access every information related to this point of interest. The entity resolution process can reveal confidentiality problems if unexpected links between open data and proprietary data are created. The goal of this PhD thesis is to provide a formal and algorithmical solution to define rules for entity resolution and check the compatibility of these rules with security and confidentiality policies. Scientific publications are considered, along with collaborations for future calls for projects regarding new challenges in Europe and in France on Big Data. The city of Grenoble already committed to provide the data from its Open Data portal for this project.

Publications

Conférences internationales avec comité de lecture

- Rémy Delanaux, Angela Bonifati, Marie-Christine Rousset & Romuald Thion (2018). « Query-based Linked Data Anonymization ». The 17th International Semantic Web Conference (ISWC 2018), 12 octobre 2018, Monterey (États-Unis), pp. 530-546. HAL : hal-01896276.

Jury

Geneviève DUPONT, DR, ULB	rapporteuse
Leonid SAVTCHENKO, PR, UCL	rapporteur
Christian HENNENBERGER, Assistant Professor, Uni. Bonn	examinateur
Laurent VENANCE, DR, CNRS	examinateur
Guillaume BESLON, PR, INSA Lyon	examinateur
Aude PANATIER, CR, INSERM	examinatrice
Hugues BERRY, DR, INRIA	directeur de thèse
Hédi SOULA, PR, UMPC	co-directeur de thèse

Résumé

Les astrocytes sont des cellules gliales du système nerveux central, essentielles à la formation des synapses, à la barrière hémato-encéphalique ainsi qu'au maintien de l'homéostasie. Récemment, les astrocytes ont été identifiés comme éléments clés du traitement de l'information dans le système nerveux central. Les astrocytes peuvent communiquer avec les neurones au niveau des synapses et moduler la communication neuronale en libérant des gliotransmetteurs et en absorbant des neurotransmetteurs. L'utilisation de nouvelles techniques comme la microscopie à super-résolution et les indicateurs calciques encodés génétiquement a permis de révéler une grande diversité spatio-temporelle des signaux calciques astrocytaires. La majorité de ces signaux sont observés au sein de leurs prolongements cellulaires, qui sont le site de communication entre neurones et astrocytes. Ces prolongements sont trop fins pour être observés en microscopie optique conventionnelle, de sorte que la microscopie à super-résolution et la modélisation informatique sont les seules méthodes adaptées à leur étude. Les travaux présentés dans cette thèse ont pour but d'étudier l'effet des propriétés spatiales (telles que la géométrie cellulaire, les distributions moléculaires et la diffusion) sur les signaux calciques dans les prolongements astrocytaires. Historiquement, les signaux calciques ont été modélisés à l'aide d'approches déterministes non spatiales. Ces modèles ont permis l'étude des signaux calciques à l'échelle de la cellule entière voire à l'échelle du réseau de cellules. Ces méthodes ne prennent cependant pas en compte la stochasticité inhérente aux interactions moléculaires ainsi que les effets de diffusion, qui jouent un rôle important dans les petits volumes. Cette thèse présente un modèle stochastique et spatial qui a été développé dans le but d'étudier les signaux calciques dans les prolongements fins astrocytaires. Ce travail a été réalisé en collaboration avec des expérimentateurs, qui nous ont fourni des données de microscopie électronique et à super-résolution. Ces données ont permis de valider le modèle. Les simulations du modèle suggèrent que (1) la diffusion moléculaire, fortement influencée par la concentration et la cinétique des buffers calciques endogènes et exogènes, (2) l'organisation spatiale intracellulaire des molécules, notamment le co-clustering des canaux calciques, (3) la géométrie du reticulum endoplasmique et sa localisation dans la cellule, (4) la géométrie cellulaire influencent fortement les signaux calciques et pourraient être responsables de leur grande diversité spatio-temporelle. Ces travaux contribuent à une meilleure compréhension du traitement de l'information par les astrocytes, un prérequis pour une meilleure compréhension de la communication entre les neurones et les astrocytes ainsi que de son influence sur le fonctionnement du cerveau.

Audrey Denizot

Thèse

Soutenue le 08/11/2019

Établissement :
INSA Lyon

Contact :
Hugues.Berry@liris.cnrs.fr

Abstract

Astrocytes are predominant glial cells that are essential for the formation of synapses, participate to the blood-brain barrier, maintain the metabolic, ionic and neurotransmitter homeostasis and have recently emerged as key elements of information processing in the central nervous system [1]. Astrocytes can contact neurons at synapses and modulate neuronal communication via the release of gliotransmitters and the uptake of neurotransmitters [2]. The use of super-resolution microscopy [3] and highly sensitive genetically encoded Ca^{2+} indicators (GECIs) [4] have revealed a striking spatiotemporal diversity of Ca^{2+} signals in astrocytes. Most astrocytic signals occur in processes, which are the sites of neuron-astrocyte communication, and strongly differ from somatic signals [5]. Those processes are too fine to be resolved by conventional light microscopy so that super-resolution microscopy and computational modeling remain the only methodologies to study those compartments. The work presented in this thesis aims at investigating the effect of spatial properties (as e.g. cellular geometry, molecular distributions and diffusion) on Ca^{2+} signals in those processes, which are deemed essential in such small volumes. Historically, Ca^{2+} signals were modeled with deterministic well-mixed approaches, which enabled the study of Ca^{2+} signals in astrocytic networks or whole-cell events [6]. Those methods however ignore the stochasticity inherent to molecular interactions as well as diffusion effects, which both play important roles in small volumes. In this PhD, I have developed a spatially-extended stochastic model in order to investigate Ca^{2+} signals in fine astrocytic processes. This work was performed in collaboration with experimentalists that performed electron as well as super-resolution microscopy. The model was validated against experimental data. Simulations of the model suggest that (1) molecular diffusion, strongly influenced by the concentration and kinetics of endogenous and exogenous buffers, (2) intracellular spatial organization of molecules, notably the co-clustering of Ca^{2+} channels, (3) ER geometry and localization within the cell and (4) cellular geometry strongly influence Ca^{2+} dynamics and can be responsible for the striking diversity of astrocytic Ca^{2+} signals. This work contributes to a better understanding of astrocyte Ca^{2+} signals, a prerequisite for understanding neuron-astrocyte communication and its influence on brain function.

Keywords: Central nervous system, astrocytes, signaling, calcium, computational neuroscience, stochastic modeling, reaction-diffusion.

[1] A. Verkhratsky and M. Nedergaard, "Physiology of Astroglia," *Physiological Reviews*, vol. 98, pp. 239–389, Jan. 2018.

[2] N. Bazargani and D. Attwell, "Astrocyte calcium signaling: the third wave," *Nature Neuroscience*, vol. 19, pp. 182–189, Feb. 2016.

[3] J. P. Heller and D. A. Rusakov, "The Nanoworld of the Tripartite Synapse: Insights from Super-Resolution Microscopy," *Frontiers in Cellular Neuroscience*, vol. 11, 2017.

[4] S. Berlin, E. C. Carroll, Z. L. Newman, H. O. Okada, C. M. Quinn, B. Kallman, N. C. Rockwell, S. S. Martin, J. C. Lagarias, and E. Y. Isacoff, "Photoactivatable genetically encoded calcium indicators for targeted neuronal imaging," *Nature Methods*, vol. 12, pp. 852–858, Sept. 2015.

[5] R. Srinivasan, B. S. Huang, S. Venugopal, A. D. Johnston, H. Chai, H. Zeng, P. Golshani, and B. S. Khakh, " Ca^{2+} signaling in astrocytes from $\text{Ipf3r2}^{-/-}$ mice in brain slices and during startle responses in vivo," *Nature Neuroscience*, vol. 18, pp. 708–717, May 2015.

[6] T. Manninen, R. Havela, and M.-L. Linne, "Computational Models for Calcium-Mediated Astrocyte Functions," *Frontiers in Computational Neuroscience*, vol. 12, Apr. 2018.

Jury

Venceslas BIRI, PR, Université Paris-Est-Marne-La-Vallée

Jean-Michel DISCHHLER, PR, Université de Strasbourg

Saïda BOUAKAZ, PR, Université Claude Bernard Lyon 1

Emmanuelle DARLES, MC, Université de Poitiers

Éric GALIN, PR, Université Claude Bernard Lyon 1

Adrien PEYTAVIE, MC, Univ. Claude Bernard Lyon 1

Rapporteur

Rapporteur

Examinatrice

Examinatrice

Directeur de thèse

Co-Directeur de thèse

Résumé

La création de paysages virtuels réalistes est un défi majeur de l'informatique graphique. Le but de cette thèse est d'amplifier un terrain de grande taille et d'ajouter du dynamisme en animant des phénomènes pseudopériodiques. Nous nous sommes alors intéressés aux scènes comprenant des rivières: la modélisation de ce type de scène inclut la génération de la géométrie du réseau hydraulique mais aussi l'animation du mouvement de l'eau correspondante. Bien que ces deux éléments aient été étudiés individuellement, il n'existe pas de méthode pour générer de manière concomitante le réseau et l'animation. Afin de répondre à cette problématique, nous proposons un nouveau modèle de représentation des rivières permettant l'animation de la surface de l'eau ainsi qu'une méthode de génération procédurale pour sculpter et animer un réseau cohérent de rivières sur un terrain existant. Notre modèle de rivière utilise les arbres de construction hiérarchique de terrain pour former la géométrie du lit de la rivière et nous proposons un nouveau modèle appelé arbre d'écoulement pour représenter la surface de l'eau. Les feuilles de l'arbre sont des primitives de flux à support compact spatialement localisées et variables dans le temps. Les fonctions utilisées représentent des caractéristiques spécifiques comme des ondulations, des turbulences ou encore des tourbillons. Elles sont procédurales et paramétrées afin de générer de la variété. Les nœuds internes de l'arbre sont des opérateurs de combinaison permettant de former la surface de l'eau. Ce modèle permet de créer une animation temps réel de l'eau sur des rivières de plusieurs kilomètres de long tout en comportant localement du détail. Nous proposons également une méthode de génération de rivière qui utilise l'élévation d'un terrain de base pour obtenir le réseau d'écoulement. Il creuse ensuite le lit des rivières dans le terrain et génère de manière concomitante l'animation de la surface de l'eau. Les caractéristiques, telles que la largeur, la profondeur et la forme du lit de la rivière, ainsi que l'élévation et l'écoulement de la surface du fluide, sont fonction du terrain et du type de la rivière. Le lit de la rivière est creusé dans le terrain en combinant des primitives d'élévation à support compact échantillonnées le long de la trajectoire de la rivière. La génération permet la production d'une vaste gamme de formes de rivières, allant des rivières à méandres dans les plaines aux rapides de montagnes. Le modèle permet une édition interactive et intuitive des trajectoires des rivières mais aussi du choix des primitives. Notre méthode de création de rivière permet d'amplifier un grand terrain en sculptant un lit de rivière précis mais aussi de créer l'animation de la surface de l'eau correspondante.

Abstract

The creation of realistic virtual landscapes is a major challenge in computer graphics. The aim of this thesis is to amplify a large terrain and add dynamism by animating pseudoperiodic phenomena. That is why we became interested in scenes involving rivers: modeling this kind of scene includes the generation of the river network geometry, but also the animation of

Thèse

Soutenue le 11/12/2019

Établissement :

Université Claude Bernard Lyon 1

Contact :

Eric.Galin@liris.cnrs.fr

the corresponding water movement. Although these two elements have been studied individually, there is no method to generate the network and the animation concomitantly. In order to solve this problem, we propose a new river model allowing the animation of the water surface as well as a procedural generation method to sculpt and animate a coherent river network on an existing site. Our river model uses a hierarchical construction tree to form the riverbed geometry and we propose a new model called blend-flow tree to represent the water surface. The leaves of the tree, the flow primitives, are spatially localized and time-varying on compact support. The functions represent specific characteristics such as ripples, turbulences or eddies. They are procedural and configured so they can generate variety. The internal nodes of the tree are combination operators that form the final surface of the water. This model allows to create a real time animation of the water on rivers several kilometers long while including details locally. We also propose a river generation method that uses the elevation of a bare terrain to obtain the flow network. It then digs the riverbeds in the ground and simultaneously generates the animation of the water surface. Characteristics, such as the width, depth and shape of the riverbed, as well as the elevation and flow of the fluid surface, depend on the terrain and the type of the river. The riverbed is excavated in the ground by combining compact supported elevation primitives sampled along the river's path. The generation allows the production of a wide range of river shapes, from meandering rivers on the plains to mountain rapids. The model allows an interactive and intuitive editing of river trajectories but also of the choice of primitives. Our river creation method allows us to amplify a large area of land by sculpting a precise riverbed but also to create the animation of the corresponding water surface.

Publications

Revues internationales avec comité de lecture

- Adrien Peytavie, Thibault Dupont, Eric Guérin, Yann Cortial, Benes Benes, James Gain & Eric Galin (2019). « Procedural Riverscapes ». Computer Graphics Forum. HAL : hal-02281637.

Jury

Elisa BERTINO, PR
Abdelmalek BENZEKRI, PR
Lionel BRUNIE, PR Directeur de thèse
Nadia BENNANI, MC
Harald KOSCH, PR
Richard LENZ, PR
Frédéric CUPPENS, PR
Michael GRANITZER, PR

Rapporteur
Rapporteur
Co-Directrice de thèse
Directeur de thèse
Examineur
Examineur
Examineur

Abstract

The processing of personal information is omnipresent in our data driven society enabling personalized services, which are regulated by privacy policies. Although privacy policies are strictly defined by the General Data Protection Regulation (GDPR), no systematic mechanism is in place to enforce them. Especially if data is merged from several sources into a data set with different privacy policies associated, the management and compliance to all privacy requirements is challenging during the processing of the data set. Privacy policies can vary hereby due to different policies for each source or personalization of privacy policies by individual users. Thus, the risk for negligent or malicious processing of personal data due to defiance of privacy policies exists. To tackle this challenge, we propose a privacy-preserving framework. Within this framework privacy policies are expressed in the proposed Layered Privacy Language (LPL) which allows to specify legal privacy policies and privacy-preserving de-identification methods. The policies are enforced by a Policy-based De-identification (PD) process. The PD process enables efficient compliance to various privacy policies simultaneously while applying pseudonymization, personal privacy anonymization and privacy models for deidentification of the data-set. Thus, the privacy requirements of each individual privacy policy are enforced filling the gap between legal privacy policies and their technical enforcement.

Publications

Conférences internationales avec comité de lecture

- Armin Gerl & Dirk Pohl (2018). « Critical Analysis of LPL according to Articles 12 - 14 of the GDPR ». Proceedings of International Conference on Availability, Reliability and Security, 30 août 2018, Hamburg (Allemagne), 26:1-26:9. doi : 10.1145/3230833.3233267. HAL : hal-01881867.
- Armin Gerl & Florian Pre (2018). « LPL Personal Privacy Policy User Interface: Design and Evaluation ». Mensch und Computer 2018 - Workshopband, Dresden (Allemagne), pp. 1-10. HAL : hal-01881865.
- Armin Gerl (2018). « Extending Layered Privacy Language to Support Privacy Icons for a Personal Privacy Policy User Interface ». Proceedings of British HCI 2018, 2 juin 2018, Belfast (Irlande), pp. 1-5. HAL : hal-01881862.
- Éditions scientifique d'ouvrages (livres, chapitres, colloques, congrès, n° spéciaux) (1)
- Armin Gerl, Nadia Bennani, Harald Kosch & Lionel Brunie (2018). « LPL, Towards a GDPR-Compliant Privacy Language: Formal Definition and Usage ». Transactions on Large-Scale Data- and Knowledge-Centered Systems XXXVII, Springer, Berlin, Heidelberg, pp. 41-80. doi : 10.1007/978-3-662-57932-9_2. HAL : hal-01866050.

Thèse

Soutenue le 05/12/2019

Établissement :
Université de Passau
(Passau, Allemagne)

Contact :
Nadia.Bennani@liris.cnrs.fr

Power-index based Management of Fraud Detection Rules: Supervised and Semi-supervised Approches

Léopold

Ghemmogne Fossi

Jury

Sylvie CALABRETTO, PR, INSA de Lyon, France
Mariagrazia FUGINI, PR, Politecnico di Milano, Italy
Elvinia M. RICCOBENE, PR, Univ. degli Studi di Milano, Italy
Jacques SAVOY, PR, Université de NEUCHÂTEL, Switzerland
Bogdan GABRYS, PR, Univ. of Technology Sydney, Australia
Alberto TROMBETTA, Dr, Univ. degli Studi dell'Insubria, Italy
Paul YOO, Dr, Cranfield University, UK

Examinatrice
Examinatrice
Examinatrice
Examinatrice
Rapporteur
Rapporteur
Rapporteur

Résumé

Cette thèse, intitulée "Une approche basée sur la théorie des jeux pour la sélection de fonctionnalités pour une prise de décision multicritère efficace: Quelques cas d'utilisation de la classification", traite de la détection de fraude par carte de crédit. Selon la Banque Centrale Européenne, la valeur des fraudes utilisant des cartes émises dans l'espace unique de paiements en euros (SEPA) en 2016 s'élevait à 1,8 milliard d'euros. Ainsi le défi pour les institutions financières est celui de réduire la fraude sur les cartes de crédit. En règle générale, les systèmes de détection de la fraude sont constitués d'un système automatique construit à base de règles "si-alors" qui contrôlent toutes les transactions en entrée et déclenchent une alerte si la transaction est considérée suspecte. Ensuite, un groupe de personnel expert vérifie l'alerte et décide si cette dernière est un vrai positif ou un faux positif. Les critères utilisés dans la sélection des règles maintenues opérationnelles sont principalement basés sur la performance individuelle des règles. Cette approche ignore en effet la non-additivité des règles. Nous proposons une nouvelle approche utilisant des indices de puissance, concept développé dans le cadre de la théorie des jeux coopératifs (CGT). Cette approche attribue aux règles un score normalisé qui quantifie l'influence de la règle sur les performances globales du groupe de règles. Les indices que nous utilisons sont le Shapley Value (SV) et le Banzhaf Value (BV). Les principales applications de ces indices sont: 1) l'aide à la décision de conserver ou de supprimer une règle du groupe; 2) la sélection du nombre k de règles les mieux classées, afin de travailler avec un ensemble de règles plus compact. En utilisant des données réelles de fraude par carte de crédit contenant environ 300 règles et $3,5 \cdot 10^5$ transactions, nous montrons que: 1) Cette approche permet de mieux exécuter les performances du groupe que celle qui évalue les règles isolément. 2) La performance de l'ensemble des règles peut être atteinte en conservant un dixième seulement des règles. Nous observons ensuite que cette application peut être considérée comme une tâche de sélection de caractéristiques pour un classificateur: nous montrons que notre approche est comparable aux algorithmes courants de référence en sélection des caractéristiques (FS). De plus, il présente un avantage dans la gestion des règles, en ce sens qu'il attribue un score normalisé à chaque règle. Ce qui n'est pas le cas pour la plupart des algorithmes de sélection des caractéristiques, qui se concentrent uniquement sur une solution d'ensemble pour obtenir des fonctionnalités hautes performances. Dans une autre contribution, nous proposons une nouvelle version du Banzhaf Value, à savoir le k -Banzhaf; cette nouvelle version surclasse la première en terme de temps de calcul et possède des performances comparables. Alors que pour un ensemble de N éléments, le Banzhaf normal calcule 2^{N-1} différences, le k -Banzhaf quant à lui calcule seulement $\binom{n-1}{k-1}$. Enfin, nous mettons en œuvre un processus d'auto-apprentissage (sorte de bootstrap) afin de renforcer le processus d'apprentissage dans un algorithme d'apprentissage automatique (Random Forest Classifier). Nous comparons ces derniers avec nos trois indices

Thèse

Soutenue le 30/09/2019

Établissement :
Université de Milan

Contact :
Elod.Egyed-Zsigmond@liris.cnrs.fr

de puissance pour effectuer une classification sur les données de fraude par carte de crédit du monde réel utilisées dans la première partie du manuscrit. En conclusion, nous observons que la sélection de caractéristiques basée sur les indices de puissance a des résultats comparables avec les algorithmes de référence en FS ainsi que dans le processus d'auto-apprentissage. Keywords: Détection de Fraud à la Carte, Théorie des Jeux de Coalition, Indice de Pouvoir, Valeur de Shapley, Indice de Banzhaf, Indice de Banzhaf restreint, Apprentissage Semi-supervisé, Apprentissage supervisé, Auto-apprentissage.+

Abstract

In the thesis, we provide two main contributions addressing the Feature Selection task, based on the conceptual tools of power index derived from coalitional Game Theory. Both contributions are based on a real-world use case rely on fraud alarm generation. We observe that typically the aggregation of the rules in their process of the governance of a pool of Near-Real-Time classification rules are operated in OR. Thus, it is remarkably non-additive, and we argue that in a non-additive context, assessing the performance in isolation might not guarantee optimal performance of the ensemble. We propose in a first time to support the rule management process and then we suggest an approach to the assessment of specific rules based on their contribution to the overall pool performance. Our method ranks the rules based on their Shapley Value with respect to a target pool performance metrics and the top-k rules are kept operational, while the reminders are considered for deletion from the pool. The second contribution is divided into two parts and aimed at self-training technique (the classifier is first trained on labelled data and the result is used to annotate unlabeled data; then unlabeled data classified with a high threshold confidence are added to the training set). This technique uses its predictions to improve at each iteration. In the first part, deal with the well-known power indices Shapley value, Banzhaf index we also propose a new power index, the restricted Banzhaf index/ k-Banzhaf index an improved version of Banzhaf index. This new power index was obtained in collaboration with some members of the team. In the second part we use the three power indices studied and two greedy-based methods, apply them to the semi-supervised machine learning technique, with the scope to improve the quality of our feature selection process.

Jury

Nicolas NISSE, CR, INRIA Sophia Antipolis	Rapporteur
Eric SOPENA, PR, Université de Bordeaux	Rapporteur
Angela BONIFATI, PR, Université Lyon	Examinatrice
Miloš STOJAKOVIĆ, PR, Université de Sciences de Novi Sad	Examineur
Eric DUCHÊNE, MC, Université Lyon 1	Directeur de thèse
Aline PARREAU, CR, Université Lyon 1	Co-directrice de thèse

Résumé

Cette thèse porte sur le problème de la couverture d'ensembles finis dans une structure discrète. Cette problématique très générale permet de nombreuses approches et nous faisons l'étude de certaines d'entre elles. Le premier chapitre introduit les notions qui seront indispensables à la bonne compréhension de cette thèse et fait un bref état de l'art sur certains problèmes de couvertures, en particulier le problème de domination dans les graphes. Le second chapitre aborde la domination de puissance, une variante du problème de domination qui a la particularité qu'on lui adjoint un phénomène de propagation. Nous étudions ce problème pour les grilles triangulaires et les grilles carrées de dimension 3. Dans le troisième chapitre, nous revenons à la domination classique mais dans un contexte ludique, avec le jeu de domination Maker-Breaker. Nous étudions la complexité du problème consistant à décider quel joueur gagne, la durée minimale d'une partie si les deux joueurs jouent parfaitement, et dérivons ce jeu pour la domination totale et dans une version Avoier-Enforcer. Le quatrième chapitre traite du nombre géodésique fort, un problème qui a la particularité de se couvrir à l'aide de plus courts chemins dans le graphe. Nous étudions le nombre géodésique fort de plusieurs classes de graphes ainsi que son comportement en relation avec le produit cartésien. Enfin, dans le cinquième chapitre, nous quittons le domaine des graphes pour étudier l'identification de points dans le plan par des disques. En plus de couvrir chaque point d'un certain ensemble par des disques, nous souhaitons que l'ensemble des disques couvrant chaque point soit unique et non vide. Nous donnons des résultats dans certains cas particuliers, des bornes dans le cas général et étudions la complexité du problème quand le rayon des disques est fixé.

Abstract

This PhD thesis concerns the problem of covering finite sets in a discrete structure. This very general issue allows numerous approaches and we study some of them. The first chapter introduces the notions that are essentials to the understanding of this thesis and makes a brief state of the art on some covering problems, including the domination problem. The second chapter addresses the power dominating problem, a variation of the dominating problem with a propagation process. We study this problem on triangular grids and square grids of dimension 3. In the third chapter, we come back to the classical domination but in the context of a game, with the Maker-Breaker domination game. We study the complexity of the problem of deciding which player has a winning strategy and the minimum duration of a game if both players play perfectly. We also derive this problem for total domination and for an Avoier-Enforcer version. The fourth chapter is about the strong geodetic number: a problem with the distinctive characteristic that the covering is made by shortest paths in the graph. We study the strong geodetic number of several graph classes and its behaviour for the Cartesian

Thèse

Soutenue le 24/09/2019

Établissement :

Université Claude Bernard Lyon 1

Contact :

Eric.Duchene@liris.cnrs.fr

Matthieu Grard

product. Lastly, in the fifth chapter, we leave the realm of graphs to study the identification of points using disks. More than just covering every point of a certain set, the subset of disks covering each point must be unique to that point. We give results on particular configurations, bounds on the general case and we study the complexity of the problem when the radius of the disks is fixed

Publications

Revues internationales avec comité de lecture

- Valentin Gledel, Vesna Iršič & Sandi Klavžar (2019). « Maker-Breaker domination number ». Bulletin of the Malaysian Mathematical Sciences Society. doi : 10.1007/s40840-019-00757-1. ArXiv : 1810.04397. HAL : hal-01935244.
- Valentin Gledel & Aline Parreau (2019). « Identification of points using disks ». Discrete Mathematics, vol. 342, n°1, pp. 256-269. doi : 10.1016/j.disc.2018.10.002. ArXiv : 1705.11116. HAL : hal-01531101.
- Antoine Dailly, Valentin Gledel & Marc Heinrich (2019). « A generalization of Arc-Kayles ». International Journal of Game Theory, pp. 1-21. doi : 10.1007/s00182-018-0639-5. ArXiv : 1709.05219. HAL : hal-01587921.
- Autres
- Valentin Gledel, Michael A. Henning, Vesna Iršič & Sandi Klavžar (2019). « Maker-Breaker total domination game ». ArXiv : 1902.00204. HAL : hal-02021678.
- Valentin Gledel & Vesna Iršič (2018). « Strong geodetic number of complete bipartite graphs, crown graphs and hypercubes ». ArXiv : 1810.04004. HAL : hal-01892332.
- Prosenjit Bose, Valentin Gledel, Claire Pennarun & Sander Verdonschot (2018). « Power domination on triangular grids with triangular and hexagonal shape ». HAL : hal-01814346.
- Valentin Gledel, Vesna Iršič & Sandi Klavžar (2018). « Strong geodetic cores and Cartesian product graphs ». ArXiv : 1803.11423. HAL : hal-01756450.
- Eric Duchene, Valentin Gledel, Aline Parreau & Gabriel Renault (2018). « Maker-Breaker domination game ». HAL : hal-01848805.

Jury

Marie-Odile BERGER
Liming CHEN
Emmanuel Dellandréa
Jean-Marc ODOBEZ
Laetitia LEYRIT
Dimitris SAMARAS
Thierry CHÂTEAU

présidente
directeur de thèse
co-directeur de thèse
examineur
examinatrice
rapporteur
rapporteur

Résumé

Le dévracage robotisé est une tâche industrielle en forte croissance visant à automatiser le déchargement par unité d'une pile d'instances d'objet en vrac pour faciliter des traitements ultérieurs tels que la formation de kits ou l'assemblage de composants. Cependant, le modèle explicite des objets est souvent indisponible dans de nombreux secteurs industriels, notamment alimentaire et automobile, et les instances d'objet peuvent présenter des variations intra-classe, par exemple en raison de déformations élastiques. Les techniques d'estimation de pose, qui nécessitent un modèle explicite et supposent des transformations rigides, ne sont donc pas applicables dans de tels contextes. L'approche alternative consiste à détecter des prises sans notion explicite d'objet, ce qui pénalise fortement le dévracage lorsque l'enchevêtrement des instances est important. Ces approches s'appuient aussi sur une reconstruction multi-vues de la scène, difficile par exemple avec des emballages alimentaires brillants ou transparents, ou réduisant de manière critique le temps de cycle restant dans le cadre d'applications à haute cadence. En collaboration avec Siléane, une entreprise française de robotique industrielle, l'objectif de ce travail est donc de développer une solution par apprentissage pour la localisation des instances les plus prenables d'un vrac à partir d'une seule image, en boucle ouverte, sans modèles d'objet explicites. Dans le contexte du dévracage industriel, notre contribution est double. Premièrement, nous proposons un nouveau réseau pleinement convolutionnel (FCN) pour délinéer les instances et inférer un ordre spatial à leurs frontières. En effet, les méthodes état de l'art pour cette tâche reposent sur deux flux indépendants, respectivement pour les frontières et les occultations, alors que les occultations sont souvent sources de frontières. Plus précisément, l'approche courante, qui consiste à isoler les instances dans des boîtes avant de détecter les frontières et les occultations, se montre inadaptée aux scénarios de dévracage dans la mesure où une région rectangulaire inclut souvent plusieurs instances. A contrario, notre architecture sans détection préalable de régions détecte finement les frontières entre instances, ainsi que le bord occultant correspondant, à partir d'une représentation unifiée de la scène. Deuxièmement, comme les FCNs nécessitent de grands ensembles d'apprentissage qui ne sont pas disponibles dans les applications de dévracage, nous proposons une procédure par simulation pour générer des images d'apprentissage à partir de moteurs physique et de rendu. Plus précisément, des vracs d'instances sont simulés et rendus avec les annotations correspondantes à partir d'ensembles d'images de texture et de maillages auxquels sont appliquées de multiples déformations aléatoires. Nous montrons que les données synthétiques proposées sont vraisemblables pour des applications réelles au sens où elles permettent l'apprentissage de représentations profondes transférables à des données réelles. A travers de nombreuses expériences sur une maquette réelle avec robot, notre réseau entraîné sur données synthétiques surpasse la méthode industrielle de référence, tout en obtenant des performances

Thèse

Soutenue le 20/05/2019

Établissement :
Ecole Centrale de Lyon

Contact :
Liming.Chen@liris.cnrs.fr

temps réel. L'approche proposée établit ainsi une nouvelle référence pour le dévrage orienté-objet sans modèle d'objet explicite.

Abstract

Referred to as robotic random bin-picking, a fast-expanding industrial task consists in robotizing the unloading of many object instances piled up in bulk, one at a time, for further processing such as kitting or part assembling. However, explicit object models are not always available in many bin-picking applications, especially in the food and automotive industries. Furthermore, object instances are often subject to intra-class variations, for example due to elastic deformations. Object pose estimation techniques, which require an explicit model and assume rigid transformations, are therefore not suitable in such contexts. The alternative approach, which consists in detecting grasps without an explicit notion of object, proves hardly efficient when the object geometry makes bulk instances prone to occlusion and entanglement. These approaches also typically rely on a multi-view scene reconstruction that may be unfeasible due to transparent and shiny textures, or that reduces critically the time frame for image processing in high-throughput robotic applications. In collaboration with Siléane, a French company in industrial robotics, we thus aim at developing a learning-based solution for localizing the most affordable instance of a pile from a single image, in open loop, without explicit object models. In the context of industrial bin-picking, our contribution is two-fold. First, we propose a novel fully convolutional network (FCN) for jointly delineating instances and inferring the spatial layout at their boundaries. Indeed, the state-of-the-art methods for such a task rely on two independent streams for boundaries and occlusions respectively, whereas occlusions often cause boundaries. Specifically, the mainstream approach, which consists in isolating instances in boxes before detecting boundaries and occlusions, fails in bin-picking scenarios as a rectangle region often includes several instances. By contrast, our box proposal-free architecture recovers fine instance boundaries, augmented with their occluding side, from a unified scene representation. As a result, the proposed network outperforms the two-stream baselines on synthetic data and public real-world data sets. Second, as FCNs require large training datasets that are not available in bin-picking applications, we propose a simulation-based pipeline for generating training images using physics and rendering engines. Specifically, piles of instances are simulated and rendered with their ground-truth annotations from sets of texture images and meshes to which multiple random deformations are applied. We show that the proposed synthetic data is plausible for real-world applications in the sense that it enables the learning of deep representations transferable to real data. Through extensive experiments on a real-world robotic setup, our synthetically trained network outperforms the industrial baseline while achieving real-time performances. The proposed approach thus establishes a new baseline for model-free object-oriented bin-picking.

Publications

Conférences internationales avec comité de lecture

- Matthieu Grard, Romain Brégier, Florian Sella, Emmanuel Dellandréa & Liming Chen (2017). « Object segmentation in depth maps with one user click and a synthetically trained fully convolutional network ». HFR - 10th International Workshop on Human-Friendly Robotics, 6 novembre 2017, Napoli (Italie), pp. 207-221. doi : 10.1007/978-3-319-89327-3_16. ArXiv : 1801.01281. HAL : hal-01674511.

Jury

Zoubida KEDAD, MC, Univ. Versailles Saint Quentin en Yveline	Rapporteuse
Olivier TOGNI, PR, Université de Bourgogne	Rapporteur
Angela BONIFATI, PR, Université Lyon 1	Examinatrice
Mohand BOUGHANEM, PR, Université Paul Sabatier	Examineur
Hamamache KHEDDOUCI, PR, Université Lyon 1	Directeur de thèse
Brice EFFANTIN, MC, Université Lyon 1	Co-Directeur de thèse

Résumé

Ces dernières années ont connu un regain d'intérêt pour l'utilisation des graphes comme moyen fiable de représentation et de modélisation des données, et ce, dans divers domaines de l'informatique. En particulier, pour les grandes masses de données, les graphes apparaissent comme une alternative prometteuse aux bases de données relationnelles. Plus particulièrement, le recherche de sous-graphes s'avère être une tâche cruciale pour explorer ces grands jeux de données. Dans cette thèse, nous étudions deux problématiques principales. Dans un premier temps, nous abordons le problème de la détection de motifs dans les grands graphes. Ce problème vise à rechercher les k-meilleures correspondances (top-k) d'un graphe motif dans un graphe de données. Pour cette problématique, nous introduisons un nouveau modèle de détection de motifs de graphe nommé la Simulation Relaxée de Graphe (RGS), qui permet d'identifier des correspondances de graphes avec un certain écart (structurel) et ainsi éviter le problème de réponse vide. Ensuite, nous formalisons et étudions le problème de la recherche des k-meilleures réponses suivant deux critères, la pertinence (la meilleure similarité entre le motif et les réponses) et la diversité (la dissimilarité entre les réponses). Nous considérons également le problème des k-meilleures correspondances diversifiées et nous proposons une fonction de diversification pour équilibrer la pertinence et la diversité. En outre, nous développons des algorithmes efficaces basés sur des stratégies d'optimisation en respectant le modèle proposé. Notre approche est efficiente en terme de temps d'exécution et flexible en terme d'applicabilité. L'analyse de la complexité des algorithmes et les expérimentations menées sur des jeux de données réelles montrent l'efficacité des approches proposées. Dans un second temps, nous abordons le problème de recherche agrégative dans des documents XML. Pour un arbre requête, l'objectif est de trouver des motifs correspondants dans un ou plusieurs documents XML et de les agréger dans un seul agrégat. Dans un premier temps nous présentons la motivation derrière ce paradigme de recherche agrégative et nous expliquons les gains potentiels par rapport aux méthodes classiques de requête. Ensuite nous proposons une nouvelle approche qui a pour but de construire, dans la mesure du possible, une réponse cohérente et plus complète en agrégeant plusieurs résultats provenant de plusieurs sources de données. Les expérimentations réalisées sur plusieurs ensembles de données réelles montrent l'efficacité de cette approche en termes de pertinence et de qualité de résultat.

Abstract

Recent years have witnessed a growing renewed interest in the use of graphs as a reliable means for representing and modeling data. Thereby, graphs enable to ensure efficiency in various fields of computer science, especially for massive data where graphs arise as a promising alternative to relational databases for big data modeling. In this regard, querying data graph proves to be a crucial task to explore the knowledge in these datasets. In this

Thèse

Soutenue le 26/11/2019

Établissement :

Université Claude Bernard Lyon 1

Contact :

Hamamache.Kheddouci@liris.cnrs.

fr

dissertation, we investigate two main problems. In the first part we address the problem of detecting patterns in larger graphs, called the top-k graph pattern matching problem. We introduce a new graph pattern matching model named Relaxed Graph Simulation (RGS), to identify significant matches and to avoid the empty-set answer problem. We formalize and study the top-k matching problem based on two classes of functions, relevance and diversity, for ranking the matches according to the RGS model. We also consider the diversified top-k matching problem, and we propose a diversification function to balance relevance and diversity. Moreover, we provide efficient algorithms based on optimization strategies to compute the top-k and the diversified top-k matches according to the proposed model. The proposed approach is optimal in terms of search time and flexible in terms of applicability. The analyze of the time complexity of the proposed algorithms and the extensive experiments on real-life datasets demonstrate both the effectiveness and the efficiency of these approaches. In the second part, we tackle the problem of graph querying using aggregated search paradigm. We consider this problem for particular types of graphs that are trees, and we deal with the query processing in XML documents. Firstly, we give the motivation behind the use of such a paradigm, and we explain the potential benefits compared to traditional querying approaches. Furthermore, we propose a new method for aggregated tree search, based on approximate tree matching algorithm on several tree fragments, that aims to build, the extent possible, a coherent and complete answer by combining several results. The proposed solutions are shown to be efficient in terms of relevance and quality on different real-life datasets.

Publications

Conférences internationales avec comité de lecture

- Abdelmalek Habi, Brice Effantin & Hamamache Kheddouci (2018). « Fast top-k search with relaxed graph simulation ». The 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 28 août 2018, Barcelona (Espagne). HAL : hal-01887427.
- Abdelmalek Habi, Brice Effantin & Hamamache Kheddouci (2017). « Search and Aggregation in XML Documents ». 28th International Conference on Database and Expert Systems Applications, 31 août 2017, Lyon (France), pp. 290-304. HAL : hal-01590605

Jury

Nicolas BOUSQUET, CR, Université Grenoble Alpes	co-Encadrant
Celina DE FIGUEIREDO, PR, Univ. Fédérale de Rio de Janeiro	Rapportrice
Éric DUCHÊNE, MC, Université Lyon 1,	Directeur
Sylvain GRAVIER, DR, Université Grenoble Alpes	co-Directeur
Isabelle GUERIN-LASSOUS, PR, Université Lyon 1	Examinatrice
Claire MATHIEU, DR, Université Paris 7	Rapportrice
Jan VAN DEN HEUVEL, PR, École d'économie et de science politique de Londres	Examineur

Résumé

Cette thèse explore des problématiques liées aux jeux. Nous nous intéressons en particulier aux jeux sans information cachée et sans aléa, à un ou deux joueurs. Les jeux à un joueur que nous considérons peuvent être décrit de la façon suivante : il y a un ensemble de configurations, qui représente tous les états possibles du jeu, et le joueur est autorisé à transformer une configuration en une autre via un certain nombre de règles. Le but est d'atteindre une configuration cible donnée à partir d'une configuration initiale. Les problèmes pouvant être définis avec ce formalisme sont appelés problèmes de reconfiguration, et incluent à la fois certains puzzles combinatoires tels que le Rubik's cube ou Rush-hour, mais également des transformations sur des problèmes de recherche pour lesquels l'ensemble des configurations est l'ensemble des solutions d'une instance d'un problème donné. Les jeux à deux joueurs qui nous intéressent sont appelés jeux combinatoires. Pour ces jeux là, nous considérons le cas où les joueurs jouent parfaitement, ce qui revient à tenter de caractériser lequel des joueurs possède une stratégie gagnante (qui lui assure la victoire quels que soient les coups de l'adversaire). Dans cette thèse, nous étudions ainsi des jeux à un et deux joueurs, et tout particulièrement la complexité de ceux-ci, c'est-à-dire évaluer à quel point il est difficile (d'un point de vue algorithmique) de calculer une stratégie gagnante. Plus précisément, nous étudions la reconfiguration de deux problèmes sur les graphes : les couplages parfaits et la coloration de graphes, pour lesquels nous donnons à la fois des preuves de difficulté à l'aide de réductions, et des algorithmes polynomiaux dans des sous-cas spécifiques. Nous nous intéressons aussi à deux problèmes en lien avec la reconfiguration : la génération aléatoire de coloration, via l'application de transformations aléatoires à une coloration donnée, et des algorithmes de coloration online. Enfin, la dernière partie de ces travaux porte sur les jeux combinatoires, pour lesquels nous nous concentrons plutôt sur des aspects structurels. Nous considérons tout d'abord les propriétés d'une famille de jeux particulière, les jeux de soustractions partisans, et étudions ensuite une construction générale pour combiner plusieurs règles de jeux, et en créer de nouvelles.

Abstract

This thesis explores problems related to games. The games that we consider in this study are games for which there is no hidden information: all the players have access to all the information related to the game; there is also no randomness and everything is deterministic. A few well-known board games (such as chess or go) fall in this category and are representative of the kinds of games that we consider here. Central to the study of these games is the notion of strategy, which roughly speaking, is a way of playing which ensures a certain objective. The main question of interest, when both

Thèse

Soutenue le 09/07/2019

Établissement :

Université Claude Bernard Lyon 1

Contact :

Eric.Duchene@liris.cnrs.fr

playing and studying a game, is the problem of finding the ‘best’ strategy, which secures the victory for the player following it. In this thesis, we consider both one-player games, also called combinatorial puzzles, and two-player games. Examples of combinatorial puzzles include Rubik’s cube, Rush-Hour, Sokoban, the 15 puzzle, or peg solitaire. Recently, some types of one-player games in particular have received a strong regain of interest as part of the larger area of reconfiguration problems. The puzzles we described above can all be described in the following way: there is a set of configurations, which represents all the possible states of the game; and the player is allowed to transform a configuration using a prescribed set of moves. Starting from an initial configuration, the goal is to reach a target configuration by a succession of valid moves. Reconfiguration extends this definition to any search problem: the set of configuration becomes the set of solutions to an instance of a given problem, and we ask whether we can transform one given solution to another using only a prescribed set of moves. Hence, in addition to the combinatorial puzzles, reconfiguration problems also include many ‘games’ which are not played by humans anymore but are instead mathematical problems sharing a lot of similarities with combinatorial puzzles. The study of reconfiguration problems has been driven by many different motivations. It has algorithmic applications: it can be seen as a way to adapt a current solution already in place to reach a new one by only making small local changes. It is also connected to other problems such as random sampling, approximate counting or problems coming from statistical physics. It can also be used as a tool for understanding the performance of some heuristic algorithms based on local modifications of solutions such as local search. Two-player games, which are also called combinatorial games, have been studied since the beginning of the twentieth century, with the works of Bouton which were continued with the development of a nice theory by Berlekamp, Conway, and Guy, unifying a certain number of classical games. We focus in this study on perfect strategies (i.e., players always choosing the best possible move), and try to characterize who wins under perfect play for various games. This approach is at the heart of what is called Combinatorial Game Theory. Most of the research in this area is focused on ‘math games’ which are games invented by mathematicians, often with simple rules and almost never played by humans. The main motivation for studying these games comes from the nice, and sometimes unexpected connections these games have with other areas of mathematics, such as for example number theory, automata, or dynamical systems. In this thesis, we study one- and two-player games. The questions we consider are often related to complexity. Complexity theory consists in trying to classify problems depending on their hardness. By hardness we mean to quantify how much time it would take for a computer to solve the problem. An other aspect of this research consists in investigating structural properties that these games can satisfy. Finally, one of our main tools is the notion of graph, and we use in particular methods and techniques from graph theory to answer the different questions we just mentioned

Publications

Revue internationale avec comité de lecture

- Marthe Bonamy, Oscar Defrain, Marc Heinrich & Jean-Florent Raymond (2019). « Enumerating Minimal Dominating Sets in Triangle-Free Graphs ». Leibniz International Proceedings in Informatics. doi : 10.4230/LIPIcs.STACS.2019.16. HAL : hal-02181721.
- Antoine Dailly, Valentin Gledel & Marc Heinrich (2019). « A generalization of Arc-Kayles ». International Journal of Game Theory, pp. 1-21. doi : 10.1007/

s00182-018-0639-5. ArXiv : 1709.05219. HAL : hal-01587921.

- Eric Duchene, Marc Heinrich, Urban Larsson & Aline Parreau (2018). « The switch operators and push-the-button games: a sequential compound over rulesets ». Theoretical Computer Science, vol. 715, pp. 71-85. doi : 10.1016/j.tcs.2018.01.003. ArXiv : 1707.07966. HAL : hal-01566273.

Conférences internationales avec comité de lecture

- Marthe Bonamy, Nicolas Bousquet, Marc Heinrich, Takehiro Ito, Yusuke Kobayashi, Arnaud Mary, Moritz Mühlenhaller & Kunihiro Wasa (2019). « The Perfect Matching Reconfiguration Problem ». Mathematical Foundations of Computer Science (MFCS), Aachen (Allemagne). doi : 10.4230/LIPIcs.MFCS.2019.80. HAL : hal-02335588.
- Nicolas Bousquet & Marc Heinrich (2017). « Computing maximum cliques in B 2 EPG graphs ». WG: Workshop on Graph-Theoretic Concepts in Computer Science, 23 juin 2017, Eindhoven (Pays-Bas). HAL : hal-01557335.
- Autres
- Sylvain Gravier & Marc Heinrich (2019). « Online graph coloring with bichromatic exchanges ». HAL : hal-02167055.
- Nicolas Bousquet & Marc Heinrich (2019). « A polynomial version of Cereceda’s conjecture ». ArXiv : 1903.05619. HAL : hal-02167051.
- Michelle Delcourt, Marc Heinrich & Guillem Perarnau (2018). « The Glauber dynamics for edges colourings of trees ». ArXiv : 1812.05577. HAL : hal-02156373.

Utilisation de l'Eye-tracking pour l'Interaction Mobile dans un Environnement Réel Augmenté

Qinjie Ju

Jury

Christophe KOLSKI, MC, Univ. Polytechnique Hauts-de-France	Rapporteur
Jean VANDERDONCKT, MC, Université Catholique de Louvain	Rapporteur
Patrick GIRARD, MC, Université de Poitiers	Examineur
Francis JAMBON, MC, Université Grenoble Alpes	Examineur
Christine MICHEL, MC, INSA de Lyon	Examinatrice
Stéphane DERRODE, MC, Ecole Centrale de Lyon,	Directeur
René CHALON, PR, Ecole Centrale de Lyon	Co-directeur

Résumé

Les dispositifs d'eye-tracking ont un très fort potentiel en tant que modalité d'entrée en IHM (Interaction Homme-Machine), en particulier en situation de mobilité. Dans cette thèse, nous nous concentrons sur la mise en œuvre de cette potentialité en mettant en évidence les scénarios dans lesquels l'eye-tracking possède des avantages évidents par rapport à toutes les autres modalités d'interaction. Au cours de nos recherches, nous avons constaté que cette technologie ne dispose pas de méthodes pratiques pour le déclenchement de commandes, ce qui réduit l'usage de tels dispositifs. Dans ce contexte, nous étudions la combinaison d'un eye-tracker et des mouvements volontaires de la tête lorsque le regard est fixe, ce qui permet de déclencher des commandes diverses sans utiliser les mains ni changer la direction du regard. Nous avons ainsi proposé un nouvel algorithme pour la détection des mouvements volontaires de la tête à regard fixe en utilisant uniquement les images capturées par la caméra de scène qui équipe les eye-trackers portés sur la tête, afin de réduire le temps de calcul. Afin de tester la performance de notre algorithme de détection des mouvements de la tête à regard fixe, et l'acceptation par l'utilisateur du déclenchement des commandes par ces mouvements lorsque ses deux mains sont occupées par une autre activité, nous avons effectué des expériences systématiques grâce à l'application EyeMusic que nous avons conçue et développée. Cette application EyeMusic est un système pour l'apprentissage de la musique, capable de jouer les notes d'une mesure d'une partition que l'utilisateur ne comprend pas. En effectuant un mouvement volontaire de la tête qui fixe de son regard une mesure particulière d'une partition, l'utilisateur obtient un retour audio. La conception, le développement et les tests d'utilisabilité du premier prototype de cette application sont présentés dans cette thèse. L'utilisabilité de notre application EyeMusic est confirmée par les résultats expérimentaux : 85% des participants ont été en mesure d'utiliser tous les mouvements volontaires de la tête à regard fixe que nous avons implémentés dans le prototype. Le taux de réussite moyen de cette application est de 70%, ce qui est partiellement influencé par la performance intrinsèque de l'eye-tracker que nous utilisons. La performance de notre algorithme de détection des mouvements de la tête à regard fixe est 85%, et il n'y a pas de différence significative entre la performance de chaque mouvement de la tête testé. Nous avons également exploré deux scénarios d'applications qui reposent sur les mêmes principes de commande, EyeRecipe et EyePay, dont les détails sont également présentés dans cette thèse.

Abstract

Eye-tracking has a very strong potential in human-computer interaction (HCI) as an input modality, particularly in mobile situations. In this thesis, we concentrate in demonstrating this potential by highlighting the scenarios in which the eye-tracking possesses obvious advantages comparing with

Thèse

Soutenue le 09/04/2019

Établissement :
Ecole Centrale de Lyon

Contact :
Stéphane.Derode@liris.cnrs.fr

all the other interaction modalities. During our research, we find that this technology lacks convenient action triggering methods, which can scale down the performance of interacting by gaze. In this instance, we investigate the combination of eye-tracking and fixed-gaze head movement, which allows us to trigger various commands without using our hands or changing gaze direction. We have proposed a new algorithm for fixed-gaze head movement detection using only scene images captured by the scene camera equipped in front of the head-mounted eye-tracker, for the purpose of saving computation time. To test the performance of our fixed-gaze head movement detection algorithm and the acceptance of triggering commands by these movements when the user's hands are occupied by another task, we have implemented some tests in the EyeMusic application that we have designed and developed. The EyeMusic system is a music reading system, which can play the notes of a measure in a music score that the user does not understand. By making a voluntary head movement when fixing his/her gaze on the same point of a music score, the user can obtain the desired audio feedback. The design, development and usability testing of the first prototype for this application are presented in this thesis. The usability of our EyeMusic application is confirmed by the experimental results, as 85% of participants were able to use all the head movements we implemented in the prototype. The average success rate of this application is 70%, which is partly influenced by the performance of the eye-tracker we use. The performance of our fixed-gaze head movement detection algorithm is 85%, and there were no significant differences between the performance of each head movement. Apart from the EyeMusic application, we have explored two other scenarios that are based on the same control principles: EyeRecipe and EyePay. The details of these two applications are also presented in this thesis.

Publications

Revues internationales avec comité de lecture

- Qinjie Ju, René Chalon & Stéphane Derrode (2019). « Assisted Music Score Reading Using Fixed-Gaze Head Movement: Empirical Experiment and Design Implications ». Proceedings of the ACM on Human-Computer Interaction. HAL : hal-02115757.

Conférences internationales avec comité de lecture

- Qinjie Ju, Stéphane Derrode & René Chalon (2017). « Utilisation de l'Eye-tracking pour l'Interaction Mobile dans un Environnement Réel Augmenté ». 29ème conférence francophone sur l'Interaction Homme-Machine, 1 septembre 2017, Poitiers (France), 5 p. HAL : hal-01577862.

Autres

- Qinjie Ju, René Chalon & Stéphane Derrode (2018). « Fixed-gaze head movement detection for triggering commands ». Workshop on Models and Analysis of Eye Movements, 8 juin 2018, Grenoble (France). HAL : hal-01842324.

Jury

Michel DESMARAIS, PR, école Polytechnique Montréal	Rapporteur
Catherine FARON-ZUCKER, MC, école Polytech'Nice-Sophia	Rapporteuse
Julien BROISIN, MC, Université Paul Sabatier	Examineur
Serge GARLATTI, PR, IMT Atlantique	Examineur
Christophe MARSALA, PR, Sorbonne Université	Examineur
Vanda LUENGO, PR, Sorbonne Université	Directrice de thèse
Marie LEFEVRE, MC, Université Claude Bernard Lyon 1	Co-Encadrante
Nathalie GUIN, MC, Université Claude Bernard Lyon 1	Co-Encadrante

Résumé

Cette thèse en informatique porte sur la problématique de la capitalisation des processus d'analyse de traces d'apprentissage au sein de la communauté des Learning Analytics (LA). Il s'agit de permettre de partager, adapter et réutiliser ces processus d'analyse de traces. Actuellement, cette capitalisation est limitée par deux facteurs importants : les processus d'analyse sont dépendants des outils d'analyse qui les mettent en œuvre - leur contexte technique - et du contexte pédagogique pour lequel ils sont menés. Cela empêche de les partager, mais aussi de les réexploiter simplement en dehors de leurs contextes initiaux, quand bien même les nouveaux contextes seraient similaires. L'objectif de cette thèse est de fournir des modélisations et des méthodes permettant la capitalisation des processus d'analyse de traces d'apprentissage, ainsi que d'assister les différents acteurs de l'analyse, notamment durant la phase de réutilisation. Pour cela, nous répondons aux trois verrous scientifiques suivant : comment partager et combiner des processus d'analyse mis en œuvre dans différents outils d'analyse ? ; comment permettre de réexploiter un processus d'analyse existant pour répondre à un autre besoin d'analyse ? ; comment assister les différents acteurs lors de l'élaboration et de l'exploitation de processus d'analyse ? Notre première contribution, issue d'une synthèse de l'état de l'art, est la formalisation d'un cycle d'élaboration et d'exploitation des processus d'analyse, afin d'en définir les différentes étapes, les différents acteurs et leurs différents rôles. Cette formalisation est accompagnée d'une définition de la capitalisation et de ses propriétés. Notre deuxième contribution répond au premier verrou lié à la dépendance technique des processus d'analyse actuels, et à leur partage. Nous proposons un méta-modèle qui permet de décrire les processus d'analyse indépendamment des outils d'analyse. Ce méta-modèle formalise la description des opérations utilisées dans les processus d'analyse, des processus eux-mêmes et des traces utilisées, afin de s'affranchir des contraintes techniques occasionnées par ces outils. Ce formalisme commun aux processus d'analyse permet aussi d'envisager leur partage. Il a été mis en œuvre et évalué dans un de nos prototypes. Notre troisième contribution traite le deuxième verrou sur la réexploitation des processus d'analyse. Nous proposons un framework ontologique pour les processus d'analyse, qui permet d'introduire de manière structurée des éléments sémantiques dans la description des processus d'analyse. Cette approche narrative enrichit ainsi le formalisme précédent et permet de satisfaire les propriétés de compréhension, d'adaptation et de réutilisation nécessaires à la capitalisation. Cette approche ontologique a été mise en œuvre et évaluée dans un autre de nos prototypes. Enfin, notre dernière contribution répond au dernier verrou identifié et concerne de nouvelles pistes d'assurances aux acteurs, notamment une nouvelle méthode de recherche des processus d'analyse, s'appuyant sur nos propositions précédentes. Nous exploitons le cadre ontologique de l'approche narrative pour définir des règles d'inférence

Thèse

Soutenue le 22/05/2019

Établissement :

Université Claude Bernard Lyon 1

Contact :

Marie.Lefevre@liris.cnrs.fr

et des heuristiques permettant de raisonner sur les processus d'analyse dans leur ensemble (e.g. étapes, configurations) lors de la recherche. Nous utilisons également le réseau sémantique sous-jacent à cette modélisation ontologique pour renforcer l'assistance aux acteurs en leur fournissant des outils d'inspection et de compréhension lors de la recherche. Cette assistance a été mise en œuvre dans un de nos prototypes, et évaluée empiriquement.

Abstract

This thesis in computer science focuses on the problem of capitalizing analysis processes of elearning traces within the Learning Analytics (LA) community. The aim is to allow these analysis processes to be shared, adapted and reused. Currently, this capitalization is limited by two important factors: the analysis processes are dependent on the analysis tools that implement them - their technical context - and the pedagogical context for which they are conducted. This prevents them from being shared, but also from being simply reused outside their original contexts, even if the new contexts are similar. The objective of this thesis is to provide models and methods for the capitalisation of analysis processes of elearning traces, as well as to assist the various actors involved in the analysis, particularly during the reuse phase. To do this, we answer the following three scientific questions: how to share and combine analysis processes implemented in different analysis tools; how to reuse an existing analysis process to meet another analysis need; how to assist the different actors in the development and exploitation of analysis processes; and how to support them in the development and exploitation of analysis processes. Our first contribution, resulting from a synthesis of the state of the art, is the formalization of a cycle of elaboration and exploitation of the analysis processes, in order to define the different stages, the different actors and their different roles. This formalization is accompanied by a definition of capitalization and its properties. Our second contribution responds to the first barrier related to the technical dependence of current analysis processes and their sharing. We propose a meta-model that allows to describe the analysis processes independently of the analysis tools. This meta-model formalizes the description of the operations used in the analysis processes, the processes themselves and the traces used, in order to avoid the technical constraints caused by these tools. This formalism, common to the analysis processes, also makes it possible to consider their sharing. It has been implemented and evaluated in one of our prototypes. Our third contribution deals with the second lock on the reuse of analysis processes. We propose an ontological framework for analysis processes, which allows semantic elements to be directly introduced, in a structured way, during the description of analysis processes. This narrative approach thus enriches the previous formalism and makes it possible to satisfy the properties of understanding, adaptation and reuse necessary for capitalisation. This ontological approach was implemented and evaluated in another of our prototypes. Finally, our last contribution responds to the last lock identified and concerns new assistances to actors, in particular a new method of researching analysis processes, based on our previous proposals. We use the ontological framework of the narrative approach to define inference rules and heuristics to reason about the analysis processes as a whole (e.g. steps, configurations) during the research. We also use the semantic network underlying this ontological modeling to strengthen assistance to actors by providing them with inspection and understanding tools during the research. This assistance was implemented in one of our prototypes, and empirically evaluated.

Publications

Conférences internationales avec comité de lecture

- Alexis Lebis, Marie Lefevre, Vanda Luengo & Nathalie Guin (2018). « Capitalisation of Analysis Processes : Enabling Reproducibility, Openness and Adaptability thanks to Narration ». LAK '18 - 8th International Conference on Learning Analytics and Knowledge, 9 mars 2018, Sydney (Australie), pp. 245-254. doi : 10.1145/3170358.3170408. HAL : hal-01714184.
- Alexis Lebis, Marie Lefevre, Vanda Luengo & Nathalie Guin (2016). « Towards a Capitalization of Processes Analyzing Learning Interaction Traces ». 11th European Conference on Technology Enhanced Learning (EC-TEL 2016), 16 septembre 2016, Lyon (France), pp. 397-403. doi : 10.1007/978-3-319-45153-4_33. HAL : hal-01336851.

Conférences nationales avec comité de lecture

- Alexis Lebis, Marie Lefevre, Vanda Luengo & Nathalie Guin (2019). « Recherche intelligente de processus d'analyse de traces d'e-learning via des inférences sémantiques ». Journées francophones d'Ingénierie des Connaissances (IC), 1 juillet 2019, Toulouse (France). HAL : hal-02155496.
- Alexis Lebis (2018). « Assistance à la réutilisation de processus d'analyse de traces d'apprentissage via une approche narrative et sémantique ». Septièmes Rencontres Jeunes Chercheurs en EIAH (RJC EIAH 2018), 6 avril 2018, Besançon (France). HAL : hal-01769541.
- Alexis Lebis, Marie Lefevre, Vanda Luengo & Nathalie Guin (2017). « Capitaliser les processus d'analyse de traces d'e-learning ». Méthodologies et outils pour le recueil, l'analyse et la visualisation des traces d'interaction - ORPHEE-RDV, 30 janvier 2017, Font-Romeu (France). HAL : hal-01637556.
- Alexis Lebis, Marie Lefevre, Vanda Luengo & Nathalie Guin (2017). « Approche narrative des processus d'analyses de traces d'apprentissage : un framework ontologique pour la capitalisation ». Environnements Informatiques pour l'Apprentissage Humain, 9 juin 2017, Strasbourg (France). HAL : hal-01517145.
- Alexis Lebis (2016). « Vers une capitalisation des processus d'analyse de traces ». Rencontres Jeunes Chercheurs en EIAH (RJC-EIAH 2016), 17 juin 2016, Montpellier (France). HAL : hal-01336850.

Une approche ascendante pour la gestion énergétique d'une Smart-Grid : Modèle adaptatif et réactif fondé sur une architecture décentralisée pour un système générique centré sur l'utilisateur permettant un déploiement à grande échelle

Victor Lequay

Jury

Jacob BEAL, Chercheur, Université de l'Iowa
Marie-Pierre GLEIZES, MC, Université Toulouse 3, IRIT
Ada DIAGONESCU, MC, Telecom Paris, LTCI
Hamamache KHEDDOUCI, PR, Université Lyon 1, LIRIS
Salima HASSAS, PR, Université Lyon 1, LIRIS
Mathieu LEFORT, MC, Université Lyon 1, LIRIS

Rapporteur
Rapporteuse
Rapporteuse
Examineur
Directrice de thèse
Co-encadrant

Résumé

Le domaine de la gestion de l'énergie dans les smart grids a été largement exploré ces dernières années, de nombreuses approches différentes étant proposées dans la littérature. En collaboration avec notre partenaire industriel Ubiant, qui déploie des solutions d'optimisation énergétique du bâtiment, nous avons mis en évidence le besoin d'un système vraiment robuste et évolutif exploitant la flexibilité de la consommation résidentielle pour optimiser l'utilisation de l'énergie au sein d'une smart grid. Dans le même temps, nous avons observé que la majorité des travaux existants se concentraient sur la gestion de la production et du stockage et qu'aucune des architectures proposées n'étaient véritablement décentralisées. Notre objectif était alors de concevoir un mécanisme dynamique et adaptatif permettant de tirer parti de toute la flexibilité existante tout en garantissant le confort de l'utilisateur et une répartition équitable des efforts d'équilibrage ; mais aussi de proposer une plate-forme ouverte et modulaire avec laquelle une grande variété d'appareils, de contraintes et même d'algorithmes pourraient être interfacés. Dans cette thèse, nous avons réalisé (1) une évaluation des techniques actuelles de prévision de la consommation individuelle en temps réel, dont les résultats nous ont amenés à suivre (2) une approche ascendante et décentralisée pour l'ajustement diffus résidentiel reposant sur un mécanisme de compensation pour assurer un effacement stable. Sur cette base, nous avons ensuite construit (3) une plate-forme générique centrée sur l'utilisateur pour la gestion de l'énergie dans les réseaux intelligents, permettant une intégration aisée de plusieurs périphériques, une adaptation rapide à l'évolution de l'environnement et des contraintes, ainsi qu'un déploiement efficace.

Abstract

The field of Energy Management Systems for Smart Grids has been extensively explored in recent years, with many different approaches being described in the literature. In collaboration with our industrial partner Ubiant, which deploys smart homes solutions, we identified a need for a highly robust and scalable system that would exploit the flexibility of residential consumption to optimize energy use in the smart grid. At the same time we observed that the majority of existing works focused on the management of production and storage only, and that none of the proposed architectures are fully decentralized. Our objective was then to design a dynamic and adaptive mechanism to leverage every existing flexibility while ensuring the user's comfort and a fair distribution of the load balancing effort ; but also to offer a modular and open platform with which a large variety of devices, constraints and even algorithms could be interfaced. In this thesis we realised (1) an evaluation of state of the art techniques in real-time individual load forecasting, whose results led us to follow (2) a bottom-up and decentralized approach to distributed residential load shedding system relying on a dynamic compensation mechanism to provide a stable curtailment. On this basis,

Thèse

Soutenue le 11/12/2019

Établissement :
Université Claude Bernard Lyon 1

Contact :
Salima.Hassas@liris.cnrs.fr

we then built (3) a generic user-centered platform for energy management in smart grids allowing the easy integration of multiple devices, the quick adaptation to changing environment and constraints, and an efficient deployment.

Recent Hidden Markov Models for Lower Limb Locomotion Activity Detection and Recognition using IMU sensors

Haoyu Li

Jury

François SEPTIER, PR, Université Bretagne Sud
Faïcel CHAMROUKHI, PR, Université de Caen
Latifa OUKHELLOU, IFSTTAR
Lamia BENYOUSSEF, Docteur Ingénieur, EPITA
Stéphane DERRODE, PR, École Centrale de Lyon
Wojciech PIECZYNSKI, PR, Télécom SudParis

Rapporteur
Rapporteur
Examinatrice
Examinatrice
Directeur de thèse
Co-directeur de thèse

Résumé

Le contexte de la thèse est celui du quantified-self, un mouvement né en Californie qui consiste à mieux se connaître en mesurant les données relatives à son corps et à ses activités. Les travaux de recherche ont consisté à développer des algorithmes d'analyse des signaux d'un capteur IMU (Inertial Measurement Unit) placé sur la jambe pour reconnaître différentes activités de mouvement telles que la marche, la course, la montée d'escalier.. Ces activités sont reconnaissables grâce à la forme des signaux d'accélération et de vitesse angulaire du capteur, tous triaxiaux, pendant le mouvement des jambes lors du cycle de marche. Pour résoudre ce problème de reconnaissance, les travaux de thèse ont permis la construction d'un modèle de chaîne de Markov cachée particulier, appelé chaîne triplet semi-Markov, qui combine un modèle semi-Markov et un modèle de mélange gaussien dans un modèle de Markov triplet. Ce nouveau modèle est adapté à la fois à la nature du cycle de marche et à l'enchaînement des activités que l'on peut réaliser dans la vie quotidienne. Pour adapter les paramètres du modèle aux différences de morphologie et de comportement humain, nous avons développé des algorithmes d'estimation des paramètres en ligne et hors ligne. Pour établir les performances d'apprentissage et de classification des algorithmes, nous avons mené des expériences sur la base d'enregistrements recueillis pendant la thèse et d'un ensemble de données publiques. Les résultats sont systématiquement comparés aux algorithmes de reconnaissance actuels.

Abstract

The thesis context is that of the quantified self, a movement born in California that consists in getting to know oneself better by measuring data relating to one's body and activities. The research work consisted in developing algorithms for analyzing signals from an IMU (Inertial Measurement Unit) sensor placed on the leg to recognize different movement activities such as walking, running, stair climbing.. These activities are recognizable by the shape of the sensor's acceleration and angular velocity signals, both tri-axial, during leg movement and gait cycle. To address the recognition problem, the thesis work resulted in the construction of a particular hidden Markov chain, called semi-triplet Markov chain, which combines a semi-Markov model and a Gaussian mixture model in a triplet Markov model. This model is both adapted to the nature of the gait cycle, and to the sequence of activities as it can be carried out in daily life. To adapt the model parameters to the differences in human morphology and behavior, we have developed algorithms for estimating parameters both off-line and on-line. To establish the classification and learning performance of the algorithms, we conducted experiments on the basis of recordings collected during the thesis and on public dataset. The results are systematically compared with state-of-the-art algorithms.

Thèse

Soutenue le 03/12/2020

Établissement :

Ecole Centrale de Lyon

Contact :

Stéphane.Derrode@liris.cnrs.fr

Publications

Revues internationales avec comité de lecture

- Haoyu Li, Stéphane Derrode & Wojciech Pieczynski (2019). « Lower limb locomotion activity recognition of healthy individuals using semi-Markov model and single wearable inertial sensor ». *Sensors*, vol. 19, pp. 4242-1. doi : 10.3390/s19194242. HAL : hal-02351365.
- Haoyu Li, Stéphane Derrode & Wojciech Pieczynski (2019). « An adaptive and on-line IMU-based locomotion activity classification method using a triplet Markov model ». *Neurocomputing*, n°362, pp. 94-105. doi : 10.1016/j.neucom.2019.06.081. HAL : hal-02270579.

Conférences internationales avec comité de lecture

- Haoyu Li, Stéphane Derrode, Lamia Benyoussef & Wojciech Pieczynski (2018). « Free-walking 3D pedestrian large trajectory reconstruction from IMU sensors ». *EUSIPCO 2018: 26th European Signal Processing Conference*, 7 septembre 2018, Rome (Italie), pp. 657-661. doi : 10.23919/EUSIPCO.2018.8553462. HAL : hal-01870621.
- Stéphane Derrode, Haoyu Li, Lamia Benyoussef & Wojciech Pieczynski (2018). « Unsupervised pedestrian trajectory reconstruction from IMU sensors ». *TAIMA 2018: Traitement et Analyse de l'Information Méthodes et Applications*, 5 mai 2018, Hammamet (Tunisie). HAL : hal-01786223.

Jury

Harald KOSCH, PR, Université Passau-Allemagne	Rapporteur
Michael MARISSA, PR, Université de Primorska, Slovénie	Rapporteur
Benjamin NGUYEN, PR, INSA Centre Val de Loire, Bourges	Examinateur
Genoveva VARGAS SOLAR, CR, CNRS, Laboratoire LIG	Examinatrice
Chirine GHEDIRA-GUEGAN, PR, IAE Lyon Université Lyon3	Directrice de thèse
Aïcha-Nabila BENHARKAT, MC, INSA de Lyon	Co-directrice de thèse

Résumé

Les objets connectés collectent et partagent des données dans différents domaines tels que les maisons intelligentes, les réseaux de distribution d'électricité intelligents et la santé. Selon Cisco, le nombre d'objets connectés devrait atteindre 50 milliards d'ici 2030 avec une quantité de données produites d'environ cinq cents zettaoctets. Toutefois, ces données recueillies sont généralement très riches et comprennent souvent des métadonnées telles que l'emplacement, l'information temporelle et le contexte, rendant ainsi possible de déduire facilement les habitudes personnelles, les comportements et les préférences des individus. De plus, l'analyse de ces données recueillies nécessite la collaboration de plusieurs intervenants. Ainsi, en raison du niveau élevé de la sensibilité des données et du manque de confiance entre les parties impliquées dans un tel réseau, ces données ne doivent pas être partagées, sans que la vie privée du propriétaire des données soit respectée. En effet, la protection de la vie privée des données issues des objets connectés est devenue un défi majeur, en particulier avec la pression croissante de la législation. Nos travaux de recherche sont focalisés sur trois problématiques complémentaires qui sont la problématique de la modélisation des exigences de la protection de la vie privée, la problématique de monitoring les objets connectés et la garantie du respect d'une politique commune qui correspond à la fois aux préférences des propriétaires des données et aux conditions des consommateurs des données, et enfin la problématique de protection de la vie privée durant tout le cycle de vie des données générées par ces objets dans une architecture décentralisée qui élimine le besoin de faire confiance aux parties impliquées dans le réseau d'objets connectés. Afin de répondre à ces problématiques, nous avons proposé dans un premier lieu une ontologie appelée LloPY qui modélise la métadonnée ainsi que les contraintes de manipulation des données en adéquation avec les normes et les lois de protection de la vie privée. Puis, pour aligner sémantiquement les exigences en matière de protection de la vie privée des propriétaires des données ainsi que des consommateurs des données, nous avons étendu l'ontologie par des relations sémantiques d'arborescence et des règles sémantiques d'inférence qui génèrent une politique de protection de la vie privée commune. Cette politique décrit comment les données doivent être manipulées une fois partagées avec un consommateur donné. Afin de garantir le respect de cette politique commune, nous avons introduit le framework PrivBlockchain, un framework qui implique toutes les parties intervenantes dans un réseau d'objets connectés dans la protection des données qui en sont issues lors des phases de collecte, du transfert, du stockage jusqu'à la phase de l'utilisation ou bien l'analyse. Le framework proposé repose, d'une part, sur la technologie de la blockchain d'où le support d'une architecture décentralisée, tout en éliminant le besoin de faire confiance aux parties impliquées dans le réseau d'objets connectés et, d'autre part, sur les contrats dits « intelligents » d'où le support d'une politique auto-appliquée et lisible par la machine. Son rôle est de protéger la vie privée lors des phases de collecte, du transfert, du stockage

Faiza Loukil

Thèse

Soutenue le 15/10/2019

Établissement :
INSA LyonContact :
Chirine.Ghedira-Guegan@
liris.cnrs.fr

jusqu'à la phase de l'analyse des données issues des objets connectés. Enfin, nous avons validé notre proposition par l'élaboration et l'implantation d'un prototype afin de prouver sa faisabilité et analyser ses performances.

Abstract

The Internet of Things (IoT) connects and shares data collected from smart devices in several domains, such as smart home, smart grid, and healthcare. According to Cisco, the number of connected devices is expected to reach 500 Billion by 2030. Five hundred zettabytes of data will be produced by tremendous machines and devices. Usually, these collected data are very sensitive and include metadata, such as location, time, and context. Their analysis allows the collector to deduce personal habits, behaviors and preferences of individuals. Besides, these collected data require the collaboration of several parties to be analyzed. Thus, due to the high level of IoT data sensitivity and lack of trust on the involved parties in the IoT environment, the collected data by different IoT devices should not be shared with each other, without enforcing data owner privacy. In fact, IoT data privacy has become a severe challenge nowadays, especially with the increasing legislation pressure. Our research focused on three complementary issues, mainly (i) the definition of a semantic layer designing the privacy requirements in the IoT domain, (ii) the IoT device monitoring and the enforcement of a privacy policy that matches both the data owner's privacy preferences and the data consumer's terms of service, and (iii) the establishment of an end-to-end privacy-preserving solution for IoT data in a decentralized architecture while eliminating the need to trust any involved IoT parties. To address these issues, our work contributes to three axes. First, we proposed a new European Legal compliant ontology for supporting preserving IoT Privacy, called LloPY that describes the IoT environment and the privacy requirements defined by privacy legislation and standards. Then, we defined a reasoning process whose goal is generating a privacy policy by matching between the data owner's privacy preferences and the data consumer's terms of service. This privacy policy specifies how the data will be handled once shared with a specific data consumer. In order to ensure this privacy policy enforcement, we introduced an IoT data privacy-preserving framework, called PrivBlockchain, in the second research axis. PrivBlockchain is an end-to-end privacy-preserving framework that involves several parties in the IoT environment for preserving IoT data privacy during the phases of collection, transmission, storage, and processing. The proposed framework relied on, on the one hand, the blockchain technology, thus supporting a decentralized architecture while eliminating the need to trust any involved IoT parties and, on the other hand, the smart contracts, thus supporting a machine-readable and self-enforcing privacy policy whose goal is to preserve the privacy during the whole data life cycle, covering the collection, transmission, storage and processing phases. Finally, in the third axis, we designed and implemented the proposal in order to prove its feasibility and analyze its performances.

Publications

Conférences internationales avec comité de lecture

- Faiza Loukil, Chirine Ghedira, Khoulood Boukadi & Benharkat Aïcha-Nabila (2018). « Towards an end-to-end IoT data privacy-preserving framework using blockchain technology ». 19th International Conference on Web Information Systems Engineering (WISE), 12 novembre 2018, Dubai (Émirats Arabes Unis), pp. 68-78. doi : 10.1007/978-3-030-02922-7_5. HAL : hal-01871493.

- Faiza Loukil, Chirine Ghedira, Khoulood Boukadi & Benharkat Aïcha-Nabila (2018). « Semantic IoT Gateway: Towards Automated Generation of Privacy-Preserving Smart Contracts in the Internet of Things ». 26th International Conference on Cooperative Information Systems (CoopIS), 24 octobre 2018, Valletta (Malte), pp. 207-225. doi : 10.1007/978-3-030-02610-3_12. HAL : hal-01871490.
- Faiza Loukil, Chirine Ghedira, Khoulood Boukadi & Aïcha-Nabila Benharkat (2018). « LloPY: A legal compliant ontology to preserve privacy for the Internet of Things ». 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), 27 juillet 2018, Tokyo (Japon), pp. 701-706. doi : 10.1109/COMPSAC.2018.10322. HAL : hal-01825750.
- Faiza Loukil, Chirine Ghedira, Benharkat Aïcha-Nabila, Khoulood Boukadi & Zakaria Maamar (2017). « Privacy-Aware in the IoT Applications: A Systematic Literature Review ». International Conference on Cooperative Information Systems (CoopIS) 2017. Proceedings, Part I. Lecture Notes in Computer Science 10573, Springer 2017, ISBN 978-3-319-69461-0, 23 octobre 2017, Rhodes (Grèce), pp. 552-569. doi : 10.1007/978-3-319-69462-7_35. HAL : hal-01630125.

Mohamed Maouche

Jury

Marco FIORE, Full Professor, CNR-IEIT
Catuscia PALAMIDESSI, DR, INRIA
Claude CASTELLUCCIA, DR, INRIA
Jean-Marc PETIT, PR, INSA-Lyon
Sara BOUCHENAK, PR, INSA-Lyon
Sonia BEN MOKHTAR, DR, CNRS

Rapporteur
Rapporteuse
Examineur
Examinatrice
Directrice de thèse
Co-directrice de thèse

Résumé

De nos jours, avec la large propagation de différents appareils mobiles, de nombreux capteurs accompagnent des utilisateurs. Ces capteurs peuvent servir à collecter des données de mobilité qui sont utiles pour des urbanistes ou des chercheurs. Cependant, l'exploitation de ces données soulèvent de nombreuses menaces quant à la préservation de la vie privée des utilisateurs. En effet, des informations sensibles tel que le lieu domicile, le lieu de travail ou même les croyances religieuses peuvent être inférées de ces données. Durant la dernière décennie, des mécanismes de protections appelées "Location Privacy Protection Mechanisms (LPPM)" ont été proposés. Ils imposent des garanties sur les données (e.g., k-anonymity ou differential privacy), obfusquent les informations sensibles (e.g., efface les points d'intérêt) ou sont une contre-mesure à des attaques particulières. Nous portons notre attention à la ré-identification qui est un risque précis lié à la préservation de la vie privée dans les données de mobilité. Il consiste en a un attaquant qui des lors qu'il reçoit une trace de mobilité anonymisée, il cherche à retrouver l'identifiant de son propriétaire en la rattachant à un passif de traces non-anonymisées des utilisateurs du système. Dans ce cadre, nous proposons tout d'abords des attaques de ré-identification AP-Attack et ILL-Attack servant à mettre en exergue les vulnérabilités des mécanismes de protections de l'état de l'art et de quantifier leur efficacité. Nous proposons aussi un nouveau mécanisme de protection HMC qui utilise des heat maps afin de guider la transformation du comportement d'un individu pour qu'il ne ressemble plus au soi du passé mais à un autre utilisateur, le préservant ainsi de la ré-identification. Cette modification de la trace de mobilité est contrainte par des mesures d'utilité des données afin de minimiser la qualité de service ou les conclusions que l'on peut tirer à l'aide de ces données.

Abstract

With the wide propagation of handheld devices, more and more mobile sensors are being used by end-users on a daily basis. Those sensors could be leveraged to gather useful mobility data for city planners, business analysts and researchers. However, gathering and exploiting mobility data raises many privacy threats. Sensitive information such as one's home or workplace, hobbies, religious beliefs, political or sexual preferences can be inferred from the gathered data. In the last decade, Location Privacy Protection Mechanisms (LPPMs) have been proposed to protect user data privacy. They alter data mobility to enforce formal guarantees (e.g., k-anonymity or differential privacy), hide sensitive information (e.g., erase points of interests) or act as countermeasures for particular attacks. In this thesis, we focus on the threat of re-identification which aims at re-linking an anonymous mobility trace to the know past mobility of its user. First, we propose re-identification attacks (AP-Attack and ILL-Attack) that find vulnerabilities and stress current state-of-the-art LPPMs to quantify their effectiveness. We also propose a new protection mechanism HMC that uses heat maps to guide the transformation

Thèse

Soutenue le 06/11/2019

Établissement :
INSA Lyon

Contact :
Sara.Bouchenak@liris.cnrs.fr

of mobility data to change the behaviour of a user, in order to make her look similar to someone else rather than her past self which preserves her from re-identification attacks. This alteration of mobility trace is constrained with the control of the utility of the data to minimize the distortion in the quality of the analysis realized on this data.

Jury

David Bromberg
Lionel Brunie
Ahmed Mostefaoui
Romain Rouvoy
Sonia Ben Mokhtar
Omar Hasan

président
directeur de thèse
rapporteur
rapporteur
examinatrice
examineur

Résumé

La détection participative est un nouveau paradigme dans lequel les citoyens utilisent volontairement leur téléphone portable pour capturer et partager les données détectées de leur environnement afin de surveiller et d'analyser certains phénomènes. Divers systèmes de réputation ont été proposés pour surveiller le comportement des participants dans les applications de détection participative afin d'identifier ceux qui fournissent de mauvaises contributions. Cependant, les systèmes de réputation existants ne fournissent pas de garanties de confidentialité aux contributeurs. Ainsi, les utilisateurs hésitent généralement à se joindre à des campagnes de détection participatives car ils sont invités à fournir leurs données détectées, notamment l'heure, l'emplacement, etc. Il a été démontré dans différents travaux que jusqu'à 95% de l'identité des participants peut être réidentifiée par le partage de quatre. Contributions contenant des données multi-capteurs. Pour cette raison, l'intégration entre les systèmes de protection de la vie privée et les systèmes de réputation est un besoin crucial pour la création d'applications de détection participative sécurisées et fiables. Cette intégration nécessite la garantie d'objectifs apparemment contradictoires. En effet, les systèmes de réputation surveillent les comportements des participants tout au long des interactions. Alors que l'objectif des systèmes de préservation de la vie privée est de dissocier le lien entre les interactions ultérieures. Dans cette thèse, nous étudions des stratégies possibles pour intégrer des systèmes préservant la confidentialité dans des systèmes sensibles à la réputation. Cette intégration pose un nouveau défi en matière de vie privée en raison de la contradiction entre leurs objectifs. Ensuite, une nouvelle attaque a été définie (attaque RR). L'attaque RR exploite cette contradiction pour détecter la succession des contributions fournies par le même participant. Ensuite, un nouveau protocole préservant la confidentialité (PrivaSense) est proposé. PrivaSense définit une méthode qui garantit à la fois la confidentialité et la réputation et résout simultanément leur contradiction. Enfin, nous proposons un nouveau système de réputation DTSRS, qui repose sur un ensemble de participants fiables pour évaluer la réputation des participants. DTSRS introduit de nouvelles mesures pour évaluer la confiance des contributions des participants. DTSRS a la capacité de se défendre plus fortement contre un ensemble d'attaques qui n'ont pas été considérées dans la littérature (par exemple, Collision, On-off, etc.). DTSRS permet au serveur d'applications d'agréger les données plus précisément. De plus, le DTSRS évalue correctement la confiance et la réputation même si un grand nombre d'adversaires sont inclus dans la campagne de détection.

Abstract

Different reputation systems have been proposed to monitor participants' behavior in participatory sensing applications. The objective of these systems is to identify those who provide the system with bad contributions. Moreover, participants are usually prohibited to join sensing campaign since they are

Hayam Mousa

Thèse

Soutenue le 19/06/2019

Établissement :
INSA Lyon

Contact :
Lionel.Brunie@liris.cnrs.fr

Simon Pageaud

asked to provide their sensed data including time, location, etc. It has been shown in different works that up to 95% of participants' identities are re-identified through sharing four contributions containing multi-sensor data. Therefore, the integration between privacy preserving systems and reputation systems is a crucial need for building secure and reliable participatory sensing applications. This integration requires the assurance of seemingly conflicting objectives. Indeed, reputation systems monitor participants' behaviors along subsequent interactions. Simultaneously, privacy preserving systems tend to detach the link between these interactions. First, possible strategies to integrate privacy preserving systems within reputation-aware systems have been studied. This integration raises a new privacy challenge because of the contradiction between their objectives. Then, a new attack has been defined (RR attack). RR attack exploits this contradiction in order to detect the succession of contributions provided by the same participant. Next, a new privacy preserving reputation-aware protocol (PrivaSense) is proposed. PrivaSense defines a methodology that ensures both privacy and reputation and simultaneously solve their contradiction. Finally, we propose a novel reputation system DTSRS. DTSRS depends on a set of the most trusted participants in order to assess the reputation of participants. DTSRS improves new measures to evaluate the trust of participants' contributions. DTSRS has the ability to defend more strongly against a set of attacks that was not considered in the literature (e.g. Collision, On-off, etc.). DTSRS enables the application server to aggregate the data more accurately. Moreover, DTSRS assesses both trust and reputation correctly even if a large number of adversaries are included in the sensing campaign.

Jury

Vincent CHEVRIER, PR, LORIA	rapporteur
Stéphane GALLAND, PR, UTBM	rapporteur
Frédéric AMBLARD, PR, IRIT	examineur
Éric DUCHÊNE, MC, Université Lyon 1	examineur
Julie DUGDALE, MC, LIG	examinatrice
Salima HASSAS, PR, Université Lyon 1	directrice de thèse
Véronique DESLANDRES, MC, Université Lyon 1	co-encadrante
Vassilissa LEHOUX, Chercheure, NAVER LABS EUROPE	co-encadrante

Résumé

Dans cette thèse, nous proposons un outil SmartGov, mixant simulation multi-agents et apprentissage multi-agents par renforcement profond, pour permettre la co-construction de politiques urbaines et inscrire les acteurs de la ville dans la boucle de conception. La Smart City permet à l'outil d'intégrer les données collectées par les capteurs présents dans la ville pour la modéliser de façon réaliste. Notre première contribution est une architecture générique pour construire une simulation multi-agents représentant la ville, et étudier l'émergence de comportement globaux avec des agents réalistes capables de réagir aux décisions politiques. Grâce à une modélisation multi-niveaux, et le couplage de différentes dynamiques, le système apprend les spécificités de l'environnement pour proposer des politiques pertinentes. Notre seconde contribution concerne l'autonomie et l'adaptation de la couche décisionnelle avec un apprentissage par renforcement multi-agents et multi-niveaux. Un ensemble d'agents, regroupés en clusters, est distribué dans le périmètre étudié pour apprendre des spécificités locales sans connaissance a priori de son environnement. L'attribution d'un score de confiance et de récompenses individuelles permettent d'atténuer l'impact de la non-stationnarité sur la réutilisation d'expériences nécessaire à l'apprentissage profond. Ces contributions conduisent à un système complet de co-construction de politiques urbaines dans le contexte de la Smart City. Nous comparons notre modèle avec d'autres approches de la littérature sur une politique de tarification du stationnement urbain, afin de mettre en évidence les apports et les limites de nos contributions.

Abstract

In this thesis, we propose the SmartGov model, coupling multi-agent simulation and multi-agent deep reinforcement learning, to help co-construct urban policies and integrate all stakeholders in the decision process. Smart Cities provide sensor data from the urban areas to increase realism of the simulation in SmartGov. Our first contribution is a generic architecture for multi-agent simulation of the city to study global behavior emergence with realistic agents reacting to political decisions. With a multi-level modeling and a coupling of different dynamics, our tool learns environment specificities and suggests relevant policies. Our second contribution improves autonomy and adaptation of the decision function with multi-agent, multi-level reinforcement learning. A set of clustered agents is distributed over the studied area to learn local specificities without any prior on the environment. Trust score assignment and individual rewards help reduce non-stationary impact on experience replay in deep reinforcement learning. These contributions bring forth a complete system to co-construct urban policies in the Smart City. We compare our model with different approaches from the literature on a parking fee policy to display the benefits and limits of our contributions.

Thèse

Soutenue le 06/09/2019

Établissement :
Université Claude Bernard Lyon 1

Contact :
Salima.Hassas@liris.cnrs.fr

**Michael Saint
Guillain**

Jury

Yves DEVILLE, PR, UCLouvain
Christine SOLNON, PR, INSA-Lyon, France
Charles PECHEUR, PR, UCLouvain
Romain BILLOT, PR, IMT-Atlantique, France
Nadia BRAUNER, PR, Université Grenoble Alpes, France
Sabine LIMBOURG, PR, HEC Liège, Belgium
Christophe GARCIA, PR, INSA Lyon

supervisor and secretary
supervisor
chairperson

Résumé

Quels seront les objectifs et défis des métropoles de demain ? La plupart des problèmes issus du monde réel sont sujets à l'inconnu, nécessitant de prendre de nouvelles décisions de façon dynamique, à la demande, en fonction des événements aléatoires qui se réalisent. Dans cette thèse, nous nous attaquons à un problème majeur, du moins en perspectives : la gestion dynamique d'une flotte de véhicules en contexte urbain. Étant donné une flotte de véhicules et un ensemble de clients, chacun pouvant potentiellement et à tout moment émettre une requête nécessitant une intervention, l'objectif de cette thèse est de fournir une réponse à la question suivante. Étant donné l'état courant à un moment donné, comment gérer notre flotte de véhicules afin de maximiser l'espérance du nombre total de requêtes satisfaites à la fin de la journée ? Ou encore, comment minimiser l'espérance du délai moyen d'intervention de nos véhicules ? La question de l'optimisation sous incertitude se pose depuis déjà plusieurs décennies. Grâce aux avancées à la fois théoriques et technologiques, nous sommes chaque jour un peu plus en mesure de palier à l'inconnu. Cependant, la plupart des problèmes intéressants restent extrêmement difficiles à résoudre, si ce n'est impossible. Cette thèse explore certains concepts fondamentaux de l'optimisation sous incertitude. En intégrant une composante stochastique aux modèles à optimiser, il est en effet possible de créer de l'anticipation.

Abstract

What will be tomorrow's big cities objectives and challenges? Most of the operational problems from the real world are inherently subject to uncertainty, requiring the decision system to compute new decisions dynamically, as random events occur. In this thesis, we aim at tackling an important growing problem in urban context: online dynamic vehicle routing. Applications of online vehicle routing in the society are manyfold, from intelligent on demand public transportation to sameday delivery services and responsive home healthcare. Provided the current state at some moment of the day, which are the best vehicle actions such that the expected number of satisfied requests is maximized by the end of the operational day? How can we minimize the expected average intervention delays of our mobile units? Optimization under uncertainty is definitely not a recent issue. Thanks to evolution of both theoretical and technological tools, our ability to face the unknown constantly grows. However, most of the interesting problems remain extremely hard, if not impossible, to solve. There is still a lot of work. Generally speaking, this thesis explores some fundamentals of optimization under uncertainty. By integrating a stochastic component into the models to be optimized, we will see how it is in fact possible to create anticipation.

Thèse

Soutenue le 13/09/2019

Établissement :
UCLouvain

Contact :
Christine.Solnon@liris.cnrs.fr

Méthodes de sélection de voisinage et de prévision à court-terme pour l'analyse du trafic urbain

Julien Salotti

Jury

Elisa FROMONT, PR, Université de Rennes	Rapporteuse
Vincent AGUILÉRA, Docteur, Direction des Routes Île-de-France	Rapporteur
Azedine BOULMAKOUL, PR, Université Hassan II de Casablanca	Examineur
Peter STURM, DR, INRIA Grenoble Rhône-Alpes	Examineur
Serge FENET, MC, Université Claude Bernard Lyon 1	Co-encadrant
Christine SOLNON, PR, INSA Lyon	Directrice de thèse
Nour-Eddin EL FAOUZI, DR, Ifsttar - ENTPE	Co-directeur de thèse
Romain BILLOT, MC, IMT Atlantique,	Invité

Résumé

Dans le contexte de la ville intelligente, le besoin d'informer, d'anticiper, et d'agir sur l'état du trafic sur le réseau devient crucial. Ce dernier est à l'origine du développement de nombreuses méthodes de prévision de trafic à court-terme (dans l'heure). À l'ère des données, on observe une augmentation de nos capacités à stocker et à traiter des données rapidement. Parallèlement, un nombre croissant de tronçons de route sont équipés de capteurs. Cette évolution technologique se reflète par une évolution des méthodes de prévision de trafic. Dans ces travaux, nous explorons plusieurs problématiques afin d'améliorer les performances de prévision des modèles sur les données de trafic. La première question concerne le voisinage spatio-temporel : quelles données doit-on considérer pour prédire l'activité future observée par un capteur ? La deuxième question concerne le choix de la meilleure méthode de prévision en fonction du type de réseau (urbain, autoroute) et de l'horizon de prévision. Enfin, nous traitons de la question du choix optimal de la résolution temporelle des données, de son impact sur la prévision. Pour répondre à ces questions, nous avons étudié de nombreuses approches de prévision (ARIMA, VAR, k-NN, SVR, réseaux de neurones) et deux mécanismes de sélection de variables (Lasso, TiGraMITe). Cette étude expérimentale a été effectuée sur des données du réseau urbain de Lyon (fournie par la Métropole) et des données d'autoroute du réseau de Marseille.

Abstract

In the context of Smart Cities, the need to inform drivers, to anticipate and to take action to regulate the traffic flow becomes critical. This need has driven the development of a large number of short-term (less than one hour) traffic flow forecasting methods. The era of big data has seen the rise in computing power, in storage capacity and in our ability to process information in real-time. At the same time, more and more road segments are equipped with traffic sensors. This evolution of technology is reflected in the evolution of traffic forecasting methods. In this work, we explore multiple questions in order to improve the accuracy of forecasting models on traffic data. The first question deals with the spatio-temporal neighborhood: what information should we consider in order to predict the future activity on a sensor? The second question is about the choice of the best forecasting methods with respect to the nature of the network (urban, freeway) and the forecast horizon. The last question concerns the choice of the optimal temporal aggregation for the data and its impact on the forecasting accuracy. In order to investigate these questions, we have studied a large panel of forecasting methods (ARIMA, VAR, k-NN, SVR, neural networks) and two variable selection mechanisms (Lasso, TiGraMITe). This experimental study has been conducted on data from the urban network of Lyon and from urban freeway of Marseille.

Thèse

Soutenue le 24/09/2019

Établissement :
INSA Lyon

Contact :
Christine.Solnon@liris.cnrs.fr

Publications

Conférences internationales avec comité de lecture

- Julien Salotti, Serge Fenet, Romain Billot, Nour-Eddin Faouzi & Christine Solnon (2018). « Comparison of traffic forecasting methods in urban and suburban context ». Internationale Conference on Tools with Artificial Intelligence (ICTAI), 7 novembre 2018, Volos (Grèce), pp. 846-853. HAL : hal-01895136.
- Serge Fenet, Yannick Perret & Julien Salotti (2014). « Extract space-time dynamics from sensor network to build urban traffic prediction model: a machine learning point of view ». Urban Modelling Symposium, 17 octobre 2014, Lyon (France). HAL : hal-01241516.

Conférences nationales avec comité de lecture

- Julien Salotti, Romain Billot, Nour-Eddin El Faouzi, Serge Fenet & Christine Solnon (2016). « Vers l'utilisation de graphes de liens causaux pour l'amélioration de la prévision court-terme du trafic routier ». RFP-IA 2016 Journée Transports Intelligents, 28 juin 2016, Clermont-Ferrand (France). HAL : hal-01318166.

Jury

Lydia MAIGNE, MC, Université Clermont Auvergne
Marc NEVEU, PR, Université de Bourgogne
Laurence CHEZE, PR, Université Lyon 1
Alamin MANSOURI, PR, Université de Bourgogne
Juliette THARIAT, PR, Centre Antoine Lacassagne
Behzad SHARIAT, PR, Université Lyon 1
Michael BEUVE, PR, Université Lyon 1
Hamid LADJAL, MC, Université Lyon 1

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

Résumé

L'estimation de la distribution de dose et d'énergie en présence du mouvement des tissus induit par la respiration, constitue un défi technologique important dans la planification du traitement en hadronthérapie. Notamment pour le cancer pulmonaire, dans lequel de nombreuses difficultés apparaissent comme la variation de densité des tissus, le changement de la forme des organes ainsi que le décalage de la position de la tumeur pendant la respiration. Tous ces paramètres affectent la portée du faisceau d'ions utilisés pendant le traitement, et, par conséquent entraînent une distribution de dose inattendue. L'objectif principal de cette thèse est de proposer une méthode de calcul de dose basée sur les structures tétraédriques, qui permet d'estimer les distributions de dose des organes en mouvement en utilisant les simulations Monte Carlo. Ces distributions de dose sont calculées en utilisant une carte de densité tétraédrique dépendante du temps, décrivant l'anatomie interne du corps humain. De plus, le mouvement interne peut être représenté à l'aide d'une modélisation biomécanique résolue par la méthode des éléments finis (MEF) ou d'une carte de déplacement issue d'un recalage d'images déformable. Contrairement aux méthodes basées sur les structures classiques à base de voxels, la dose déposée s'accumule à l'intérieur de chaque tétraèdre au cours de la déformation, surmontant ainsi le problème du suivi tissulaire puisque le tétraèdre est défini comme une partie d'un tissu dont la composition chimique et la topologie ne changent pas. Dans la première partie de la thèse, nous avons développé une méthode de calcul de dose qui génère une carte de dose 4D en utilisant un modèle tétraédrique spécifique au patient. En outre, nous étudions l'effet du niveau de détail des maillages tétraédriques sur la précision de la distribution de la dose obtenue. Dans la deuxième partie, nous nous concentrons sur l'optimisation de la géométrie tétraédrique pour réduire le temps de simulation, sachant que l'obtention d'une distribution de dose précise peut être coûteux en termes de temps. Pour surmonter ce problème, nous avons proposé une nouvelle approche qui prend en compte la direction du faisceau afin de minimiser l'erreur de l'épaisseur équivalent eau des tétraèdres avant le volume de la tumeur. Cette méthode permet d'obtenir un maillage tétraédrique grossier et, par conséquent, d'améliorer les performances de calcul dans les simulations de Monte Carlo, tout en conservant une distribution de dose précise dans le volume cible.

Mots clés : Hadronthérapie, maillages tétraédriques, structures voxéliques, dosimétrie, Monte Carlo, modélisation.

Abstract

The estimation of energy and dose distribution patterns in respiratory-induced organ motion constitutes a significant challenge in hadron therapy treatment planning and dosimetry. Notably for lung cancer in which many

Thèse

Soutenue le 30/09/2019

Établissement :

Université Claude Bernard Lyon 1

Contact :

Behzad.Shariat@liris.cnrs.fr

difficulties arise, like tissue densities variation and the tumor position shifting during respiration. All these parameters affect the ranges of protons or ions used in treatment when passing through different tissues and can easily result in unexpected dose distribution. The present work consists of calculating the dose distributions of moving organs by means of Monte Carlo simulations and patient-specific modeling tools. The dose distributions are calculated using a time-dependent tetrahedral density map, describing the internal anatomy of the human body. Additionally, the internal motion can be described using either a biomechanical modeling based on Finite Element Analysis (FEA) or deformable image registration displacement map. Unlike methods based on the conventional voxel-based structures, the deposited energy is accumulated inside each tetrahedron during deformation, thus overcoming the problem of tissue tracking since that the tetrahedron is defined as a part of a tissue whose chemical composition and topology do not change. The first part of the Ph.D. project proposes a dose calculation method that generates a 4D dose map using a patient-specific tetrahedral model. Besides, we study the effect of the level of detail of tetrahedral meshes on the accuracy of the resulted dose distribution. In the second part, we focus on the optimization of the tetrahedral geometry to address the problem of time simulation, since obtaining a precise dose distribution can be very time-consuming. To overcome this issue, we've defined a new approach that takes into account the direction of the beam to minimize the error of the water equivalent thickness of the tetrahedrons before the tumor volume. This method allows for a coarsened tetrahedral mesh and as a result, improved computational performance in Monte Carlo simulations while guaranteeing a precise dose distribution in the target volume.

Keywords: Hadron therapy, tetrahedral meshes, voxel-based structures, dosimetry, Monte Carlo, modeling.

Publications

Revue internationale avec comité de lecture

- Yazid Touileb, Hamid Ladjal, Michael Beuve & Behzad Shariat (2018). « Particle-beam-dependent optimization for Monte Carlo simulation in hadrontherapy using tetrahedral geometries ». *Physics in Medicine and Biology*, vol. 63, n°13, p. 135021. doi : 10.1088/1361-6560/aacbe5. HAL : hal-01821168.

Conférences internationales avec comité de lecture

- Yazid Touileb, Hamid Ladjal, Joseph Azencot, Michael Beuve & Behzad Shariat (2016). « 4D dose calculations: Tetrahedral meshes versus voxel-based structures ». *ICTR-PHE 2016*, 19 février 2016, CERN, Genève (Suisse). HAL : hal-01272135.
- Petru Manescu, Hamid Ladjal, Yazid Touileb, Joseph Azencot, Michael Beuve & Behzad Shariat (2016). « 4D POSITRON EMISSION TOMOGRAPHY IMAGE RECONSTRUCTION BASED ON BIOMECHANICAL RESPIRATORY MOTION ». *IEEE International Symposium on Biomedical Imaging (IEEE ISBI)*, 16 avril 2016, Prague (République Tchèque), pp. 99-102. HAL : hal-01267527.
- Yazid Touileb, Petru Manescu, Hamid Ladjal, Joseph Azencot, Michael Beuve & Behzad Shariat (2016). « MOTION-INDUCED MONTE CARLO DOSE CALCULATION USING DEFORMABLE TETRAHEDRAL MESHES ». *IEEE International Symposium on Biomedical Imaging (IEEE ISBI)*, 16 avril 2016, Prague (République Tchèque), pp. 1257-1260. HAL : hal-01267508.
- Hamid Ladjal, Nadir Skendraoui, Matthieu Giroux, Yazid Touileb, Joseph Azencot, Michael Beuve, Philippe Giraud & Behzad Shariat (2015). « Physiological and Biomechanical Model of Patient Specific Lung Motion Based on 4D CT Images ». *The 8th Biomedical Engineering International Conference (IEEE BMEiCON2015)*, 27 novembre 2015, Patt (Thaïlande), pp. 1-5. HAL : hal-01214310.

Jury

Loic BARTHE, PR, Université de Toulouse	rapporteur
Saida BOUAKAZ, PR, Université Lyon 1	examinatrice
Eric GALIN, PR, Université Lyon 1	directeur de thèse
Eric GUÉRIN, MC, INSA Lyon	co-encadrant
Jean-Michel DISCHLER, PR, Université de Strasbourg	rapporteur
Céline LOSCOS, PR, Université Reims Champagne-Ardenne	examinatrice

Résumé

Cette thèse s'intéresse à la synthèse de paysages naturels, et plus particulièrement, à leur portion céleste. L'aspect du ciel est gouverné par de nombreux phénomènes atmosphériques parmi lesquels les nuages jouent un rôle prépondérant car ils sont fréquemment présents et couvrent de grandes étendues. Même sans considérer directement le ciel, la densité des nuages leur permet de modifier intensément l'illumination globale d'un paysage. Les travaux de cette thèse se concentrent donc principalement sur l'édition, la modélisation et l'animation d'étendues nuageuses aux dimensions d'un paysage. Comme la simulation thermodynamique de la formation des nuages est difficilement contrôlable et que les détails du volume simulés sont rapidement limités, nous proposons plutôt une méthode par génération procédurale. Nous érigeons un modèle léger de paysage nuageux sous forme d'une hiérarchie de fonctions. Les détails les plus fins sont obtenus par composition de bruits procéduraux et reproduisent les formes de différents genres de nuages. La présence nuageuse à grande échelle est quant à elle décrite à haut niveau et à différents instants par des cartes dessinées par l'utilisateur. Ces cartes discrètes sont transformées en primitives implicites statiques ensuite interpolées par métamorphose en prenant en compte le relief et les vents pour produire des trajectoires cohérentes. Le champ implicite obtenu par mélange des primitives interpolantes constitue le champ spatiotemporel de densité nuageuse. Des images sont finalement synthétisées par rendu du milieu participatif atmosphérique selon notre propre implémentation exécutée en parallèle sur carte graphique.

Abstract

This thesis focusses on the synthetization of natural landscapes, and more particularly on their celestial part. The aspect of the sky is governed by plenty of atmospheric phenomena, among which clouds play a major role for they are recurrent and widespread. Even without directly considering the sky, the density of the clouds allows them to intensely modify the global illumination of a landscape. The work of this thesis thus focuses mainly on the editing, modelling and animation of cloud areas of landscape dimensions. Because simulating the thermodynamics behind the formation of clouds hardly provides any control and the details of the simulated volume are quickly limited, we propose instead a procedural generation method. We build a lightweight cloudscape model as a hierarchy of functions. The finest details are obtained by composing procedural noises and reproduce the specific shapes of different kinds of clouds. The large-scale cloud presence is described at a high level and at different times by maps drawn by the user. These discrete maps are transformed into implicit static primitives and then interpolated by morphing, accounting for relief and winds in order to produce coherent trajectories. The implicit field obtained by mixing the interpolating primitives represents the spatiotemporal field of cloud density. Images are finally synthesized by rendering of the atmospheric participative medium according to our own implementation, executed in parallel on a graphic card.

Antoine Webanck

Thèse

Soutenue le 16/07/2019

Établissement :

Université Claude Bernard Lyon 1

Contact :

Eric.Galin@liris.cnrs.fr

Xiaoyang Zhu

Jury

Maria POTOP-BUTUCARU, PR, UPMC
Khalil DRIRA, DR, Laboratoire CNRS LAAS
Christine MORIN, DR, INRIA Rennes
Lionel BRUNIE, PR, INSA-Lyon
Youakim BADR, MC, INSA-Lyon

Rapporteure
Rapporteur
Examinatrice
Examineur
Directeur de thèse

Résumé

Le principe de l'Internet des objets (IdO) est d'interconnecter non seulement les capteurs, les appareils mobiles et les ordinateurs, mais aussi les particuliers, les maisons, les bâtiments intelligents et les villes, ainsi que les réseaux électriques, les automobiles et les avions, pour n'en citer que quelques-uns. Toutefois, la réalisation de la connectivité étendue de l'IdO tout en assurant la sécurité et la confidentialité des utilisateurs reste un défi. Les systèmes IdO présentent de nombreuses caractéristiques non conventionnelles, telles que l'évolutivité, l'hétérogénéité, la mobilité et les ressources limitées, qui rendent les solutions de sécurité Internet existantes inadaptées aux systèmes basés sur IdO. En outre, l'IdO préconise des réseaux peer-to-peer où les utilisateurs, en tant que propriétaires, ont l'intention d'établir des politiques de sécurité pour contrôler leurs dispositifs ou services au lieu de s'appuyer sur des tiers centralisés. En nous concentrant sur les défis scientifiques liés aux caractéristiques non conventionnelles de l'IdO et à la sécurité centrée sur l'utilisateur, nous proposons une infrastructure sécurisée de l'IdO rendue possible par la technologie de la blockchain et pilotée par des réseaux peer-to-peer sans confiance. Notre infrastructure sécurisée pour l'IdO permet non seulement l'identification des individus et des collectifs (entreprises, familles, organisations), mais aussi l'identification fiable des objets de l'IdO (dispositifs, services) par leurs propriétaires en se référant à la blockchain dans les réseaux peer-to-peer sans confiance. La blockchain fournit à notre infrastructure sécurisée de l'IdO une base de données fiable, immuable et publique qui enregistre les identités individuelles et collectives, ce qui facilite la conception du protocole d'authentification simplifié de l'IdO sans dépendre des fournisseurs d'identité tiers. En outre, notre infrastructure sécurisée pour l'IdO adopte un paradigme d'IdO socialisé qui permet à toutes les entités de l'IdO (les individus, les collectifs, les choses) d'établir des relations et rend l'IdO extensible et omniprésent les réseaux où les propriétaires peuvent profiter des relations pour définir des politiques d'accès pour leurs appareils ou services. En outre, afin de protéger les opérations de notre infrastructure sécurisée pour l'IdO contre menaces de sécurité, nous introduisons également un mécanisme autonome de détection des menaces en complément de notre cadre de contrôle d'accès, qui peut surveiller en permanence le comportement anormal des opérations de dispositif ou de service. Enfin, nous réalisons le prototype de notre solution, nous présentons des cas d'utilisation et nous effectuons des expériences et des simulations qui montrent que l'infrastructure sécurisée que nous proposons est capable d'interconnecter efficacement toutes les entités de l'IdO grâce à nos mécanismes d'authentification et d'autorisation et de détecter les menaces connues et inconnues avec des taux de détection élevés et de faibles alarmes positives erronées.

Abstract

The premise of the Internet of Things (IoT) is to interconnect not only sensors, mobile devices, and computers but also individuals, homes, smart buildings,

Thèse

Soutenue le 24/06/2019

Établissement :
INSA Lyon

Contact :
Lionel.Brunie@liris.cnrs.fr

and cities, as well as electrical grids, automobiles, and airplanes, to mention a few. However, realizing the extensive connectivity of IoT while ensuring user security and privacy still remains a challenge. There are many unconventional characteristics in IoT systems such as scalability, heterogeneity, mobility, and limited resources, which render existing Internet security solutions inadequate to IoT-based systems. Besides, the IoT advocates for peer-to-peer networks where users as owners intend to set security policies to control their devices or services instead of relying on some centralized third parties. By focusing on scientific challenges related to the IoT unconventional characteristics and user-centric security, we propose an IoT secure infrastructure enabled by the blockchain technology and driven by trust-less peer-to-peer networks. Our IoT secure infrastructure allows not only the identification of individuals and collectives (e.g., companies, families, organizations) but also the trusted identification of IoT things (e.g., devices, services) through their owners by referring to the blockchain in trustless peer-to-peer networks. The blockchain provides our IoT secure infrastructure with a trustless, immutable and public ledger that records individuals and collectives identities, which facilitates the design of the simplified authentication protocol for IoT without relying on third-party identity providers. Besides, our IoT secure infrastructure adopts socialized IoT paradigm which allows all IoT entities (i.e., individuals, collectives, things) to establish relationships and makes the IoT extensible and ubiquitous networks where owners can take advantage of relationships to set access policies for their devices or services. Furthermore, in order to protect operations of our IoT secure infrastructure against security threats, we also introduce an autonomic threat detection mechanism as the complementary of our access control framework, which can continuously monitor anomaly behavior of device or service operations. At last, we prototype our solution, present use cases, and run experiment and simulation which show that our proposed IoT secure infrastructure can effectively interconnect all IoT entities through our authentication and authorization mechanisms and detect both known and unknown threats with high detection rates and low false positive alarms.

Publications

Revues internationales avec comité de lecture

- Xiaoyang Zhu & Youakim Badr (2018). « Identity Management Systems for the Internet of Things: A Survey Towards Blockchain Solutions ». *Sensors*, vol. 18, n°12, p. 4215. HAL : hal-01945947.

Conférences internationales avec comité de lecture

- Xiaoyang Zhu & Youakim Badr (2018). « A Survey on Blockchain-based Identity Management Systems for the Internet of Things ». 2018 IEEE Confs on Internet of Things, Green Computing and Communications, Cyber, Physical and Social Computing, Smart Data, Blockchain, Computer and Information Technology, Congress on Cybermatics, 30 août 2018, Halifax (Canada), pp. 1568-1573. HAL : hal-01873095.
- Xiaoyang Zhu & Youakim Badr (2018). « Fog Computing Security Architecture for the Internet of Things using Blockchain-based Social Networks ». 2018 IEEE Confs on Internet of Things, Green Computing and Communications, Cyber, Physical and Social Computing, Smart Data, Blockchain, Computer and Information Technology, Congress on Cybermatics, 30 août 2018, Halifax (Canada), pp. 1361-1366. HAL : hal-01873093.

- Jesus Pacheco, Xiaoyang Zhu, Youakim Badr & Salim Hariri (2017). « Enabling Risk Management for Smart Infrastructures with an Anomaly Behavior Analysis Intrusion Detection System ». IEEE 2nd International Workshops on Foundations and Applications of Self* Systems (FAS*W), 19 septembre 2017, Tuscon (États-Unis), pp. 324-328. HAL : hal-01619366.
- Xiaoyang Zhu, Youakim Badr, Jesus Pacheco & Salim Hariri (2017). « Autonomic Identity Framework for the Internet of Things ». International Conference on Cloud and Autonomic Computing (ICCAC), 22 septembre 2017, Tuscon (États-Unis), pp. 69-79. HAL : hal-01575956.

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/>

