



HAL
open science

IA & réseaux sociaux

Arnaud Martin, Philippe Morignot, Dominique Longin

► **To cite this version:**

Arnaud Martin, Philippe Morignot, Dominique Longin. IA & réseaux sociaux. Bulletin de l'Association Française pour l'Intelligence Artificielle, 103, 2019, Association Française d'Intelligence Artificielle. hal-04571327

HAL Id: hal-04571327

<https://ut3-toulouseinp.hal.science/hal-04571327>

Submitted on 7 May 2024

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

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



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



AfIA

Association française
pour l'Intelligence Artificielle

Bulletin N° 103

Association française pour l'Intelligence Artificielle

AfIA



PRÉSENTATION DU BULLETIN

Le [Bulletin](#) de l'Association française pour l'Intelligence Artificielle vise à fournir un cadre de discussions et d'échanges au sein de la communauté universitaire et industrielle. Ainsi, toutes les contributions, pour peu qu'elles aient un intérêt général pour l'ensemble des lecteurs, sont les bienvenues. En particulier, les annonces, les comptes rendus de conférences, les notes de lecture et les articles de débat sont très recherchés. Le [Bulletin](#) de l'AfIA publie également des dossiers plus substantiels sur différents thèmes liés à l'IA. Le comité de rédaction se réserve le droit de ne pas publier des contributions qu'il jugerait contraire à l'esprit du bulletin ou à sa politique éditoriale. En outre, les articles signés, de même que les contributions aux débats, reflètent le point de vue de leurs auteurs et n'engagent qu'eux-mêmes.

■ Édito

Les réseaux sociaux constituent un phénomène de société majeur depuis quelques années : que ce soit pour s'informer, publier une information ou une opinion, ou tout simplement s'exprimer, les médias sociaux sont omniprésents dans nos vies d'internautes connectés. L'analyse et la compréhension de ces médias au moyen d'outils informatiques est plus difficile : comment représenter cette masse de données brutes ? comment détecter les mouvements d'opinions au sein de ce volume énorme de données ? comment détecter les faux sites, les comptes malveillants ?

Notre domaine scientifique de l'Intelligence Artificielle s'est naturellement tournée vers ces formes d'expression écrite. C'est l'objet du dossier de ce numéro 103 du [Bulletin](#) de l'AfIA, préparé par Arnaud MARTIN (IRISA). Huit équipes françaises y sont recensées et auto-présentent leurs thèmes de recherche et résultats dans ce dossier.

Ce numéro contient également les rubriques traditionnelles de comptes rendus d'événements, avec pour le présent numéro un compte rendu de la journée Perspective et Défis de l'IA. (PDIA) consacrée au « Véhicule autonome et intelligence artificielle » concocté par Olivier SIMONIN (INSA Lyon), un compte-rendu des Journées Francophones sur les Systèmes Multi-Agents offertes par Gauthier PICARD (Mines Saint-Etienne) et un compte rendu de la Nuit de l'Info préparé par Florence BANNAY (IRIT). Qu'ils en soient remerciés !

Toute l'équipe de rédacteurs du [Bulletin](#) de l'AfIA vous souhaite une bonne lecture de ce numéro !

Philippe MORIGNOT
Rédacteur



SOMMAIRE

DU BULLETIN DE L'AFIA

3	Dossier « Intelligence artificielle et réseaux sociaux »	
	Édito	4
	Collecter ne suffit pas : que nous racontent les données? ou pourquoi les médias sociaux constituent un champ d'investigation à long terme...	4
	Fouille de graphes complexes.	7
	Fouille de graphes enrichis.	11
	Analyse d'activités dans les réseaux sociaux	14
	Analyse des réseaux sociaux pour des applications liées à la santé.	17
	Autodétermination des utilisateurs dans les réseaux sociaux.	19
	Recherche en réseaux sociaux dans l'équipe DM2L	22
	LIA : Laboratoire Informatique d'Avignon	25
30	Comptes rendus de journées, événements et conférences	
	26 ^{es} Journées Francophones sur les Systèmes Multi-Agents (JFSMA 2018)	31
	5 ^e journée Perspectives et Défis de l'IA (PDIA 2018)	33
	Nuit de l'Info	35
37	Thèses et HDR du trimestre	
	Thèses de Doctorat	38
	Habilitations à Diriger les Recherches	45



AfIA
Association française
pour l'Intelligence Artificielle

Dossier

« Intelligence artificielle et réseaux sociaux »

Dossier réalisé par

Arnaud MARTIN
IRISA/DRUID
Université de Rennes, CNRS, IRISA
Arnaud.Martin@univ-rennes1.fr



■ Édito

Ce dossier regroupe des travaux actuels sur les réseaux sociaux et leurs liens avec l'intelligence artificielle. Florence SÈDES montre tout d'abord en quoi les médias sociaux constituent un domaine d'applications, source de verrous scientifiques importante. Les travaux de son équipe portent essentiellement sur la définition d'un profil social des utilisateurs de différents médias sociaux. Ainsi extraire des informations issues des réseaux sociaux revient bien souvent à faire de la fouille de données sur des graphes complexes qui permettent de représenter les réseaux sociaux. Christine LARGERON montre comment, dans son équipe, les recherches se sont tournées de la détection de communautés vers la génération de graphes attribués pour valider les algorithmes développés notamment en considérant la dynamique des réseaux. Il est aussi possible d'enrichir les graphes attribués par l'extraction de motifs abstraits sur les réseaux tel que le présente Henry SOLDANO. Une autre problématique importante motivée par la quantité des données sur les réseaux sociaux est l'analyse des activités sur ces réseaux. Laurent VERCOUTER et son équipe cherchent ainsi à détecter les faux comptes, les

comptes malveillants ou encore les personnes influentes sur les réseaux. Jérôme AZÉ et Sandra BRINGAY présentent ensuite les activités de leur équipe tournées vers l'analyse des réseaux sociaux pour des applications liées à la santé pour lesquelles il est essentiel de considérer la fiabilité des contenus dans l'analyse de ceux-ci et de tenir compte de la temporalité des données. Toujours motivés par les données, Anne BOYER et l'entreprise Sailendra, ont étudié l'apport de l'algorithme *mutual vote* pour l'autodétermination des utilisateurs dans les réseaux sociaux. Rémy CAZABET et Jean-François BOULICAULT décrivent les activités de leur équipe qui tournent autour de l'analyse de réseaux sociaux, notamment de leur dynamique, des multi-graphes attribués et de l'organisation des réseaux complexes. Pour finir ce dossier, Jean-François BONASTRE expose les différentes activités de son laboratoire orientées autour des réseaux sociaux.

Sans être exhaustif, le présent dossier montre l'étendue des travaux en intelligence artificielle en France, motivés par les données issues des réseaux sociaux.

■ Collecter ne suffit pas : que nous racontent les données ? ou pourquoi les médias sociaux constituent un champ d'investigation à long terme...

IRIT / SIG
Université Paul Sabatier Toulouse 3
<https://www.irit.fr/-Equipe-SIG> | Florence SÈDES
florence.sedes@irit.fr

La donnée n'est qu'un point de départ !

L'analyse et la gestion des données concourent à un processus qui tend à les « traiter », leur donnant du sens pour qu'en résulte

une information sur laquelle l'humain peut se baser pour prendre des décisions effectives : analysée, transformée et rendue accessible à l'utilisateur, la donnée doit aussi répondre à des besoins, profils, audiences diverses. Ainsi com-



Afia

Association française
pour l'Intelligence Artificielle

préhensible et « intelligente », la valeur (ajoutée) de la donnée réside dans l'histoire que celle-ci raconte.

Les travaux de notre équipe interviennent « des données à la décision » : ils concernent les aspects liés à la gestion, la modélisation, l'analyse et l'exploration des données, dans divers contextes, comme base pour la prise de décision.

La donnée n'est pas inerte : outre les traditionnelles problématiques liées à son évolution (gestion de versions, réification, identité vs égalité, ré-identification, etc.), la dynamique des applications et des profils des usagers nous oblige à nous confronter à de nouveaux indicateurs, tels la notion de *température* (buzz, tendances, etc.), qui traduit l'intérêt de la communauté, ou de *ligne de flottaison* (niveau de visibilité plancher), qui matérialise la limite en deçà de laquelle la pertinence n'est plus de mise.

Les médias sociaux constituent un domaine d'application, d'expérimentation, un appui à la formalisation et une source de verrous scientifiques et technologiques dont nous sommes loin d'atteindre les limites. La gestion, la modélisation, l'analyse, la fouille et l'exploration de cet écosystème numérique ont été mises en avant dans différentes contributions : à partir de l'analyse de la dynamique des réseaux sociaux, par exemple pour la détection de communautés, sont étudiés les processus de diffusion de l'information, polémique, buzz, spam, fake, etc. La détection de rôle (influence sociale, autorité, propagande) induit différents modes de coordination ou de régulation pour et par ces réseaux.

La contribution initiale de notre équipe réside dans une construction du profil de l'utilisateur, appelé profil social, à partir de médias sociaux en ligne (Online Social Media - OSM). Le processus CoBSP (Community Based Social Profile) a été proposé dans l'objectif d'extraire les informations les plus pertinentes caracté-

sant un profil utilisateur, en se basant sur son réseau égocentrique et ses communautés.

La caractéristique dynamique des OSM évoquée plus tôt induit l'intégration de critères temporels dans le processus de construction du profil, selon une approche de détection des intérêts et d'enrichissement temporel dudit profil.

Cette approche utilise non seulement le réseau égocentrique et les « communautés » de chaque utilisateur, mais aussi des informations comme les documents et les « tags » partagés et/ou annotés. Les expérimentations sur les réseaux de co-auteurs de publications scientifiques DBLP et Mendeley et sur Delicious ont montré l'efficacité de la contribution pour construire un profil social de l'utilisateur pertinent et à jour. Ces travaux, appliqués dans une démarche générique, ont également été mis en œuvre sur la base des archives ouvertes HAL vue comme un réseau de co-auteurs à grande échelle : réseau de mots-clés, thèmes, institutions, communautés, équipes, usages, etc. Ces résultats permettent de suivre les évolutions thématiques de toute communauté (campagne électorale, scientifiques, femmes vs hommes, etc.) et peuvent être utilisés par des mécanismes de recommandation (d'auteurs, d'articles, de communautés, etc.).

Le jeu de données de HAL, avec 300 000 articles, plus d'un million de notices méta-données ainsi que les données d'usage (avec textes, auteurs, affiliations, figures, organismes, expérimentations), a servi de base à nos évaluations.

Le modèle social générique de profil utilisateur issu de ces travaux est susceptible d'être réutilisé dans différents domaines d'application, comme processus de dérivation des éléments du profil social de l'utilisateur, non pas à partir des individus autour de l'utilisateur, mais à partir des communautés autour de lui, dans une portion de son réseau social (réseau égocentrique).



Afia

Association française
pour l'Intelligence Artificielle

Profil utilisateur vs profil social ?

Intégrer des critères temporels dans le processus de construction du profil de l'utilisateur s'est imposé comme une évidence compte tenu des usages. Le poids « temporel », qui reflète la pertinence d'un intérêt à un instant donné, est calculé, d'une part, à partir de la pertinence des individus du réseau égocentrique de l'utilisateur en prenant en compte « la fraîcheur » de leurs liens avec l'utilisateur et, d'autre part, à partir de la pertinence des informations qu'ils partagent en prenant en compte la « fraîcheur » de ces informations.

Les expérimentations effectuées dans deux réseaux de publications scientifiques, DBLP et Mendeley, ont montré la pertinence de la proposition sur cette vérité-terrain.

Une comparaison des performances du modèle générique déjà testé sur les réseaux mentionnés auparavant a été menée avec la base de données sociales Delicious, qui a aussi servi de jeu de données pour proposer une approche de détection des intérêts et d'enrichissement temporel du profil utilisateur. Cette approche utilise non seulement le réseau égocentrique et les communautés de chaque utilisateur, mais aussi des informations comme les documents (métadonnées classiques) et les tags (métadonnées sociales) partagés et/ou annotés par les utilisateurs. L'aspect temporel a été introduit afin d'enrichir périodiquement chaque profil utilisateur.

De par leur démarche générique, ces travaux peuvent être mis en œuvre et expérimentés sur différents types de réseaux/données, par exemple pour un processus continu sur les archives ouvertes HAL : les résultats visés sont les évolutions de thématiques de recherche des chercheurs, des équipes, des communautés scientifiques, *etc.*, mais aussi des mécanismes de recommandation en ligne (d'auteurs, d'articles, *etc.*).

Des domaines d'applications variés : réseaux sociaux mobiles, Social IoT, gestion de risques, qualité de l'information...

Les domaines d'application de ces travaux, en lien avec les évolutions des infrastructures, depuis les Réseaux Sociaux Mobiles (véhiculaires, *etc.*) jusqu'à l'loT social reflètent les recherches en cours autour de la confiance et de la « privacité », comme les récentes polémiques autour de l'application STRAVA. Parmi ceux-ci, la gestion de crise (alerte, organisation des ressources, diffusion, *etc.*) nous permet de valoriser nos savoir-faire en gestion de Big Data et en métriques de Qualité de l'Information. En effet, la popularité des médias sociaux en ligne (OSM) est fortement liée à la qualité du contenu généré par l'utilisateur et à la protection de la vie privée des utilisateurs. En se basant sur la définition de la qualité de l'information, comme son aptitude à être exploitée, la facilité d'utilisation des OSM soulève de nombreux problèmes en termes de qualité de l'information qui impacte les performances des applications exploitant ces OSM (individus mal intentionnés (*i.e.* « spammeurs ») qui les utilisent pour disséminer de fausses informations et/ou des informations indésirables telles que les contenus commerciaux illégaux). La propagation et la diffusion de telles informations (les « spams ») entraînent d'énormes problèmes affectant la qualité des services proposés par les OSM. La majorité d'entre eux (Facebook, Twitter, *etc.*) étant quotidiennement attaqués par quantité d'utilisateurs mal intentionnés, les techniques de filtrage adoptées, nécessitant plusieurs semaines voire plusieurs mois pour filtrer l'information ainsi bruitée, s'avèrent inefficaces. Plusieurs défis doivent être surmontés pour réaliser une méthode de filtrage : (i) données de masse ; (ii) vie privée et sécurité ; (iii) hétérogénéité des structures dans les réseaux sociaux ; (iv) diversité des formats, sé-



mantiques, granularité, etc. des contenus générés par les utilisateurs ; (v) subjectivité vs objectivité.

Notre travail s'inscrit dans le cadre de l'amélioration de la qualité des contenus en termes de messages partagés et de profils des utilisateurs des OSM en abordant en détail les défis susmentionnés. Le spam social étant le problème le plus récurrent, nous avons proposé deux approches génériques pour détecter et filtrer le contenu spam : i) la première approche consiste à le détecter dans un flux en temps réel (par exemple les tweets) ; ii) la seconde approche est dédiée au traitement d'un grand volume des données relatives aux profils utilisateurs des spammeurs (par exemple, les comptes Twitter). Pour le contenu, nous introduisons une approche d'apprentissage non supervisée qui permet le filtrage en temps réel des tweets spams dans laquelle la fonction de classification est adaptée automatiquement. La fonction de

classification, entraînée de manière itérative, ne requiert pas de collection de données annotées manuellement.

Dans la deuxième approche, nous traitons le problème de classification des profils utilisateurs dans le contexte d'une collection de données à grande échelle. Nous proposons de faire une recherche dans un espace réduit de profils utilisateurs (une communauté d'utilisateurs) au lieu de traiter chaque profil d'utilisateur. Chaque profil appartenant à cet espace réduit est analysé pour prédire sa classe à l'aide d'un modèle de classification binaire.

Les expériences menées sur Twitter ont montré que le modèle de classification collective non supervisé proposé est capable de générer une fonction efficace de classification binaire des tweets en temps réel qui s'adapte à l'évolution des stratégies des spammeurs sociaux sur Twitter.

■ Fouille de graphes complexes

Laboratoire Hubert Curien
Université de Saint-Etienne

<https://laboratoirehubertcurien.univ-st-etienne.fr/>

Christine LARGERON

Christine.Largeron@univ-st-etienne.fr

La thématique Data Intelligence du Laboratoire Hubert Curien

Au sein du Laboratoire Hubert Curien (UMR CNRS 5516), la thématique *Data Intelligence* aborde les problématiques de *machine learning* dont l'objectif est d'apprendre automatiquement des modèles à partir de données, et de *fouille de données* dont le but est d'extraire de l'information pertinente de grands volumes de données. Ces problématiques sont extrêmement porteuses aujourd'hui en raison de leurs nombreuses applications, et s'inscrivent dans des thèmes plus larges comme le *big data*

ou la science des données. En *machine learning*, il s'agit notamment de développer de nouveaux algorithmes de *representation learning*, de *transfer learning/domain adaptation*, d'apprentissage multi-vues, adaptés notamment au contexte de la détection d'anomalies et de fraude, de concevoir de nouvelles fonctions de pertes à optimiser, adaptées à des contextes très déséquilibrés et non indépendants et identiquement distribués, ou encore de dériver des garanties théoriques sur les algorithmes développés. En fouille de données, les travaux portent plus particulièrement sur le traitement de données complexes et hétérogènes (docu-



ments multimédias, XML, Tweets, images, séquences vidéos, graphes, réseaux sociaux) et visent à construire des modèles et des espaces de représentation adaptés, et à élaborer des méthodes de découvertes de motifs à partir de ces données. Une partie de ces recherches s'attache à exploiter simultanément ces différents types de données, notamment relationnelles et textuelles, ou textuelles et image, dans le but de traiter plus efficacement des tâches classiques de fouille de données comme le classement supervisé ou non, ou encore d'améliorer les performances de systèmes de recommandation ou de recherche d'information.

Un type de données complexes : les graphes attribués

Parmi les données complexes particulièrement étudiées au sein de la thématique figurent les réseaux d'information, largement utilisés ces dernières années pour modéliser et simuler divers types de systèmes : sociaux, biologiques, informatiques, etc. Dans le cadre de nos travaux, nous nous sommes plus particulièrement intéressés aux graphes attribués où les sommets du graphe représentent les entités du système, les liens correspondent à leurs relations et où un vecteur d'attributs est associé à chaque sommet pour décrire ses caractéristiques. Ainsi par exemple, dans un réseau de co-publications scientifiques, chaque sommet correspond à un auteur auquel est associé un vecteur construit à partir du contenu de ses publications et il existe un lien entre deux auteurs s'ils ont écrit au moins un article ensemble.

De la détection de communautés...

La prise en compte de ces attributs en plus des informations relationnelles nous a conduit à revisiter les indicateurs classiques de caractérisation des sommets du graphe tels que les mesures de centralité [2]. Elle nous a aussi amené

à étendre les méthodes de détection de communautés. L'objectif de la détection de communautés est de créer une partition des sommets de telle sorte que les sommets au sein d'une même communauté soient fortement reliés entre eux et faiblement connectés au reste du réseau. En intégrant les attributs qui les décrivent, il s'agit en plus de placer dans un même groupe les sommets qui ont également approximativement les mêmes caractéristiques. Ainsi, on peut dire que la tâche consiste à résoudre à la fois un problème de classement non supervisé et de détection de communautés. Pour ce faire, nous avons tout d'abord introduit un nouveau critère de qualité d'une partition d'éléments représentés dans un espace réel. Ce critère, appelé $Q_{Inertia}$ car il est basé sur l'inertie, présente plusieurs propriétés intéressantes, d'une part il est borné et d'autre part il conserve la même valeur quelle que soit la transformation affine appliquée aux vecteurs représentant les éléments. Une des applications immédiates du critère $Q_{Inertia}$ a été d'étendre l'algorithme de Louvain, un des algorithmes les plus efficaces à ce jour pour détecter des communautés dans un graphe. En itérant une phase de migration et de fusion, ILouvain, reprend le même processus glouton que Louvain mais il vise à optimiser une combinaison linéaire de $Q_{Inertia}$ et de la modularité alors que Louvain considère uniquement ce dernier critère. Comme Louvain, un des avantages de ILouvain est que l'optimisation repose sur une approche locale [5].

... à la génération de graphes attribués

Lors de ces recherches, nous avons été confrontés au manque de benchmark et de jeux de données réelles avec vérité terrain nous permettant d'évaluer nos algorithmes. Pour y remédier, nous avons donc développé un générateur de réseaux dynamiques attribués avec structure communautaire [7, 4]. Ce générateur, est disponible sous licence libre GNU GPL V3.0



Afia

Association française
pour l'Intelligence Artificielle

sur le site du Laboratoire Hubert Curien. Deux types d'opérations, contrôlables à l'aide de paramètres par l'utilisateur, permettent de guider l'évolution du réseau, d'une part des micro-opérations appliquées sur les liens et les sommets du graphe et d'autre part des macro-opérations appliquées sur les communautés. De plus, la génération est menée de façon à produire des graphes respectant les propriétés classiques des réseaux réels telles que l'attachement préférentiel ou l'homophilie. Ainsi, à partir d'un ensemble de paramètres fixés par l'utilisateur, tels que le nombre de sommets, le nombre minimum de liens dans et hors de la communauté, le nombre de communautés, etc., le logiciel produit une séquence de graphes, avec pour chaque graphe la partition des sommets en communautés et les attributs réels caractérisant chacun des sommets. De plus, une version graphique permet de visualiser l'évolution du réseau et de sa structure communautaire au cours du temps, ou bien chacun des graphes le composant. L'interface affiche également différents indicateurs pour contrôler les propriétés vérifiées par le graphe comme par exemple la modularité, le coefficient de *clustering* et l'inertie intra-classe pour sa structure communautaire ou encore la distribution des degrés pour l'attachement préférentiel.

... ou encore à l'étude de graphes dynamiques incomplets

Nous nous sommes aussi intéressés à la détection de communautés dans des graphes dynamiques lorsque la structure du graphe est partiellement connue ; ce qui dans la pratique est le plus souvent le cas lorsque toutes les interactions entre les entités du réseau ne peuvent pas être observées. Nous avons notamment comparé des méthodes prévues pour les graphes dynamiques et basées sur la décomposition de tenseurs, et des méthodes conçues pour traiter un graphe simple mais que nous

avons adaptées au cas dynamique soit en fusionnant les données des graphes d'entrée lors du prétraitement, soit en fusionnant les partitions générées sur chacun d'eux en post-traitement. Nos résultats expérimentaux ont montré que ces secondes approches sont suffisantes lorsque la structure communautaire est clairement définie mais que les méthodes capables de gérer les graphes dynamiques sont plus performantes lorsque cette structure est plus complexe [3].

Nous poursuivons actuellement l'étude de plusieurs familles de modèles bien adaptés pour la détection de communautés ou la prédiction de liens. En particulier, en collaboration avec le LIG, dans le cadre de la thèse d'A. DULAC, nous étudions le comportement de modèles probabilistes dits *mixed-membership* vis-à-vis de l'homophilie et de l'attachement préférentiel. Nous avons notamment démontré qu'ils satisfont l'homophilie avec la similarité naturelle qui les sous-tend. Pour l'attachement préférentiel la situation est plus contrastée : ces modèles ne satisfont pas l'attachement préférentiel global. En revanche, le respect de l'attachement préférentiel local est possible selon que l'appartenance aux classes latentes est stricte ou partielle et, dans ce dernier cas si la distribution sur les classes latentes est *bursty* ou non [6].

Avec le LIRIS, dans le cadre du projet *Network Embedding with Dynamic* soutenu par la Fédération Informatique de Lyon, nous nous intéressons aussi aux méthodes de projection de graphe dans un espace réel (*embedding* de graphes), étudions leurs capacités à préserver certaines propriétés du graphe ou sa structure, et nous poursuivrons ces travaux dans le cadre du projet *ACADEMICS* soutenu par l>IDEX Lyon.



Afia

Association française
pour l'Intelligence Artificielle

Exploitation des données relationnelles pour la recherche d'information ou la recommandation

Enfin, nous avons aussi exploité les informations relationnelles pour améliorer les performances de systèmes de recommandation ou de recherche d'information. Ainsi, par exemple, dans le cadre du projet NeoTex (Exploration de documents Textuels d'un domaine par un Néophyte), un des chantiers d'usage d'ISTEX, nous avons conçu un système dont le but est d'assister la recherche documentaire d'un néophyte qui souhaite explorer la littérature d'un domaine scientifique. À partir d'une requête composée de quelques mots-clés fournis par l'utilisateur, ce système permet de classer les articles scientifiques de la base ISTEX en tenant compte non seulement de leur contenu mais aussi de scores calculés sur le graphe de citations des articles de cette base ainsi que des mesures externes. Nos expérimentations ont confirmé l'apport de ces mesures par rapport à une recherche basée uniquement sur le contenu [1].

Ces différents travaux ont été menés avec des membres du projet *Data mining and Information retrieval* : B. AUDEH, M. BEIGBEDER, O. BENYAHIA, D. COMBE, M. GÉRY, B. JEUDY, P.N. MOUGEL, des collègues du LIG : A. DULAC, E. GAUSSIER, du Liris : R. CAZABET, E. EGYED-ZSIGMOND, R. VAUDAINÉ, et de l'Université d'Alberta : O. ZAÏANE.

Références

- [1] Bissan Audeh, Michel Beigbeder, and Christine Largeron. A machine learning system for assisting neophyte researchers in digital libraries. In *14th IAPR International Conference on Document Analysis and Recognition, ICDAR 2017, Kyoto, Japan, November 9-15, 2017*, pages 317–323, 2017.
- [2] Oualid Benyahia and Christine Largeron. Centrality for graphs with numerical attributes. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2015, Paris, France, August 25-28, 2015*, pages 1348–1353, 2015.
- [3] Oualid Benyahia, Christine Largeron, and Baptiste Jeudy. Community detection in dynamic graphs with missing edges. In *11th International Conference on Research Challenges in Information Science, RCIS 2017, Brighton, United Kingdom, May 10-12, 2017*, pages 372–381, 2017.
- [4] Oualid Benyahia, Christine Largeron, Baptiste Jeudy, and Osmar R. Zaïane. Dancer : Dynamic attributed network with community structure generator. In *Machine Learning and Knowledge Discovery in Databases - European Conference, ECML PKDD 2016, Riva del Garda, Italy, September 19-23, 2016, Proceedings, Part III*, pages 41–44, 2016.
- [5] David Combe, Christine Largeron, Mathias Géry, and Elöd Egyed-Zsigmond. I-louvain : An attributed graph clustering method. In *Advances in Intelligent Data Analysis XIV - 14th International Symposium, IDA 2015, Saint Etienne, France, October 22-24, 2015, Proceedings*, pages 181–192, 2015.
- [6] Adrien Dulac, Éric Gaussier, and Christine Largeron. A study of stochastic mixed membership models for link prediction in social networks. In *2017 IEEE International Conference on Data Science and Advanced Analytics, DSAA 2017, Tokyo, Japan, October 19-21, 2017*, pages 706–715, 2017.
- [7] Christine Largeron, Pierre-Nicolas Mougel, Oualid Benyahia, and Osmar R. Zaïane. Dancer : dynamic attributed networks with community structure generation. *Knowl. Inf. Syst.*, 53(1) :109–151, 2017.



■ Fouille de graphes enrichis

LIPN | **Henry SOLDANO**

Université Paris 13 | henry.soldano@lipn.univ-paris13.fr
<https://lipn.univ-paris13.fr/> | lipn.univ-paris13.fr

Des réseaux aux réseaux enrichis

Les données structurées sous forme de graphes sont des objets devenus prédominants ces dernières années, en particulier en sciences de la vie (réseaux biologiques) et sur internet (réseaux sociaux en ligne), auxquels s'ajoutent les réseaux sociaux à proprement parler, objets d'étude de la sociologie. L'étude de ces objets s'est longtemps limitée à des questions purement topologiques, mesurant des indices et modélisant ces réseaux pour identifier leurs caractéristiques, leur structure, en extraire des communautés. Les informations complémentaires éventuelles concernant les sommets et les liens étant utilisées *a posteriori*. Depuis quelques années cependant il y a un intérêt accru pour les réseaux enrichis : réseaux attribués dans lesquels les sommets ou les liens sont étiquetés par des descriptions, réseaux multiplexes lorsque les liens qui relient les sommets sont de différentes natures, réseaux multimodaux lorsque les sommets sont de différents types, réseaux orientés lorsque les liens sont des arcs, réseaux géographiques dont les sommets sont localisés sur un espace physique. Ces différents enrichissements peuvent être de plus combinés en un *réseau hétérogène*. Ceci conduit d'une part techniquement à adapter et étendre les méthodes existantes spécifiques aux réseaux, d'autre part à importer les techniques provenant de l'apprentissage et de la fouille de données, en considérant que le réseau est une structure des données que l'on examine. Nous allons de ce fait vers la possibilité d'exploiter de la manière la plus large la nature relationnelle des données.

Dans l'équipe, nous avons récemment développé différentes approches méthodologiques concernant l'apprentissage et la fouille de données sur les réseaux sociaux multiplexes ou attribués. En particulier, nous nous sommes intéressés à la détection de communautés [4] ainsi qu'à l'énumération et à la sélection de motifs portant sur les attributs, *via* la réduction de sous-graphes à des sous-graphes *noyaux* ou *core subgraphs* en anglais. Nous nous intéressons, formellement et d'un point de vue algorithmique et expérimental, à l'extension de ces méthodes de détection de communauté et d'énumération de motifs aux réseaux enrichis. Chaque aspect du réseau requiert une manière particulière de l'aborder, que ce soit pour l'analyser, le construire, le réviser, en extraire des communautés ou des sous-graphes *noyaux*. En particulier nous nous intéresserons aux types et aux rôles des différents éléments constituant un réseau hétérogène.

La fouille de motifs abstraits dans les graphes attribués

Avant de présenter le cœur de notre projet sur l'extraction de motifs multiples dans les réseaux hétérogènes, il est nécessaire de présenter brièvement quelques éléments méthodologiques :

- La notion de *sous-graphe noyau* d'un graphe : la partie essentielle d'un graphe matérialisée par le plus grand sous-graphe dont les sommets satisfont une certaine propriété [1]. Ces sommets constituent le noyau du graphe. Il s'agit d'une généralisation du *k*-noyau proposé par SEIDMAN dans lequel



tous les sommets doivent avoir un degré au moins k .

- La notion de *motif fermé* : on suppose qu'on a un ensemble d'objets chacun décrit dans un langage partiellement ordonné par le motif le plus spécifique de ce langage qui apparaît dans l'objet. Le cas le plus connu est celui des *itemsets* : chaque objet est décrit par un sous-ensemble d'items, par exemple *abc*, et le motif *ab* apparaît dans l'objet car *ab* est inclus dans la description *abc* de l'objet. Un motif est dit fermé si c'est le motif le plus spécifique parmi tous les motifs ayant la même *extension*, c'est-à-dire apparaissant dans le même sous-ensemble d'objets. En énumérant les motifs fermés on évite d'avoir plusieurs motifs ayant même extension et on réduit ainsi l'ensemble des solutions. Il existe des algorithmes efficaces d'énumération de motifs fermés [2].

Nous avons développé une méthodologie d'énumération et sélection de motifs dans les graphes attribués, qui énumère des motifs fermés, mais au sens d'un opérateur de fermeture différent, on parle alors de *fermés abstraits*, permettant une plus grande réduction de l'ensemble des motifs solutions. En pratique, lors de l'énumération on calcule l'extension d'un motif donné, c'est-à-dire un sous-ensemble de sommets, puis on réduit ce sous-ensemble à son noyau, et on cherche le motif le plus spécifique commun aux sommets du noyau [7]. Enfin après avoir énuméré ces motifs parfois nombreux, on peut les ordonner selon une mesure d'intérêt. Nous avons ainsi proposé de mesurer à quel point ces motifs pouvaient être inattendus. Pour cela, on a recours à la simulation : si on considère un sous-graphe de x sommets quelle est en moyenne la taille du noyau ? un motif sera jugé d'autant plus intéressant que son sous-graphe noyau a une taille s'écartant de l'attendu.

Énumérer et sélectionner des motifs multiples dans un réseau hétérogène

Le réseau hétérogène le plus simple est un réseau bimodal, les arêtes reliant, par exemple, des acteurs aux films dans lesquels ils ont joué. Les sommets étant de deux natures différentes, ils sont décrits dans deux langages différents, ce qui pose la question de ce qu'est un motif dans ce cas. Considérons alors un bimotif comme constitué de deux parties, l'une sélectionnant des acteurs et l'autre les films. L'extension d'un bimotif est alors elle-même une paire constituée d'un sous-ensemble d'acteurs et d'un sous-ensemble de films. Des travaux récents ayant défini une notion de noyau spécifique aux réseaux bimodaux [3], on peut alors réduire l'extension d'un bimotif à son noyau. On a alors tous les éléments pour définir et énumérer les bimotifs fermés abstraits d'un réseau bimodal. Cette idée peut-être étendue à des situations où les sommets sont de même nature mais où le noyau est constitué de sommets pouvant avoir deux rôles différents. C'est le cas du noyau hub-autorité HA pour les graphes dirigés que nous avons proposé récemment [9] pour analyser un réseau de conseil entre avocats étudié par Emmanuel LAZEGA [6]. L'idée initiale proposée par Jon KLEINBERG [5] et concrétisée dans ce noyau, est d'ordonner les sommets selon deux indices inter-dépendants, un haut d'indice d'autorité correspondant à un degré entrant élevé provenant de sommets ayant un indice Hub élevé, eux-même ayant un degré sortant élevé vers des sommets avec un indice autorité élevé. Dans ce cas on a un seul langage, mais on définira des motifs différents, voire opposés, pour chacun des deux rôles. Ceci permet par exemple d'observer le sous-graphe hub-autorité des avocats de Boston demandant conseil à des avocats de Providence [8].

Notre projet en cours est d'étendre cette idée de motifs multiples associés à des types et des rôles différents de sommets mais aussi



d'arêtes, par exemple si on s'intéresse à des réseaux multiples. Ceci suppose de développer de nouveaux noyaux adaptés à cette diversité de rôles et d'en tirer les conséquences algorithmiques en termes d'énumération et d'ordonnement de motifs. L'idée générale est de déployer une méthodologie permettant aisément de s'adapter à des graphes enrichis de différentes natures. Toute structure relationnelle doit ainsi pouvoir être abordée en abstrayant ce qui est observé de manière à en avoir une perception simplifiée, idée d'abstraction qui est naturelle et a un long passé en intelligence artificielle.

Références

- [1] Vladimir Batagelj and Matjaz Zaversnik. Fast algorithms for determining (generalized) core groups in social networks. *Adv. Data Analysis and Classification*, 5(2) :129–145, 2011.
- [2] Mario Boley, Tamás Horváth, Axel Poigné, and Stefan Wrobel. Listing closed sets of strongly accessible set systems with applications to data mining. *Theor. Comput. Sci.*, 411(3) :691–700, 2010.
- [3] Monika Cerinsek and Vladimir Batagelj. Generalized two-mode cores. *Social Networks*, 42 :80 – 87, 2015.
- [4] Issam Falih, Nistor Grozavu, Rushed Kanawati, and Younès Bennani. Community detection in attributed network. In *WWW (Companion Volume)*, pages 1299–1306. ACM, 2018.
- [5] Jon M Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, 46(5) :604–632, 1999.
- [6] Emmanuel Lazega. *The collegial phenomenon : The social mechanisms of cooperation among peers in a corporate law partnership*. Oxford University Press, 2001.
- [7] Henry Soldano, Guillaume Santini, and Dominique Bouthinon. Motifs abstraits et sous-communautés dans les réseaux attribués. *Revue d'Intelligence Artificielle*, 30(4) :441–468, 2016.
- [8] Henry Soldano, Guillaume Santini, Dominique Bouthinon, Sophie Bary, and Emmanuel Lazega. Bi-pattern mining of two mode and directed networks. In Pierre-Antoine Champin, Fabien L. Gandon, Mounia Lalmas, and Panagiotis G. Ipeirotis, editors, *Companion of the The Web Conference 2018 on The Web Conference 2018, WWW 2018, Lyon , France, April 23-27, 2018*, pages 1287–1294. ACM, 2018.
- [9] Henry Soldano, Guillaume Santini, Dominique Bouthinon, and Emmanuel Lazega. Hub-authority cores and attributed directed network mining. In *29th IEEE International Conference on Tools with Artificial Intelligence, ICTAI 2017, Boston, MA, USA, November 6-8, 2017*, pages 1120–1127. IEEE Computer Society, 2017.



■ Analyse d'activités dans les réseaux sociaux

LITIS
INSA Rouen Normandie
www.litislab.eu | Laurent VERCOUTER
laurent.vercouter@insa-rouen.fr

Chercheurs impliqués

- Nour EL MAWASS, Paul HONEINE, Nicolas MALANDAIN, Alexandre PAUCHET, Julien SAUNIER, Zaher YAMAK (LITIS)
- Guillaume GADEK (Airbus/LITIS)
- Stephan BRUNESSAUX, Khaled KHELIF (Airbus)

Introduction

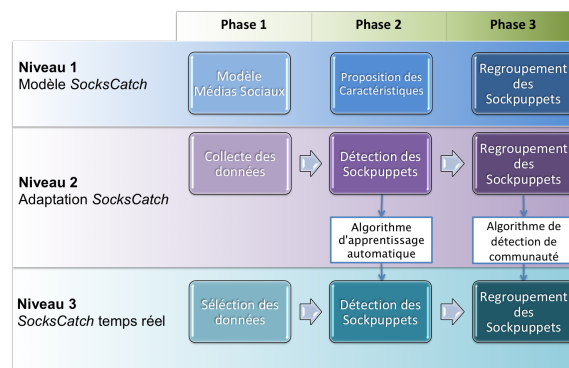
Le développement des médias sociaux en ligne, en constante croissance depuis une quinzaine d'années, a considérablement intensifié et accéléré les échanges d'informations entre individus. Ce phénomène est une formidable avancée pour un accès à l'information quasiment en temps réel et sans contrainte géographique. Mais il s'accompagne aussi de dangers liés à la diffusion de fausses informations (*fake news*), de spams ou de cyber-attaques dont le contrôle ou la vérification des sources est impossible à réaliser humainement étant donné la masse d'informations et leur vitesse de diffusion.

Les équipes MIND (Multi-agent, Interaction, Décision) et App (Apprentissage) du LITIS, en collaboration avec Airbus Defence and Space, se sont intéressées à ce problème dans le cadre de plusieurs projets visant à utiliser des méthodes issues de l'intelligence artificielle pour l'analyse d'activités dans les médias sociaux (principalement Twitter et Wikipedia). Trois de ces projets sont ici décrits visant respectivement à détecter les faux comptes, les comptes ayant un comportement de spammers et l'existence de groupes d'influence.

Détection de faux comptes sur Wikipedia

Dans le cadre de la thèse de Zaher YAMAK [3] soutenue en février 2018, nous nous sommes intéressés à la détection et la caractérisation de faux comptes créés par un même utilisateur, appelés *sockpuppets*, sur Wikipedia. Les *sockpuppets* sont fréquemment créés dans le but de manipuler l'information de cette encyclopédie en donnant l'impression qu'une modification est poussée et approuvée par plusieurs comptes.

Nous avons proposé un processus, nommé *SocksCatch*, pour détecter ces faux comptes puis rapprocher ceux créés par un même utilisateur. La figure ci-dessous illustre ce processus en trois phases. La première phase a pour objectif la sélection et le prétraitement des données. La deuxième phase a pour objectif la détection des comptes *sockpuppets* à l'aide d'un algorithme d'apprentissage automatique. La troisième phase a pour objectif le regroupement des comptes *sockpuppets* créés par un même utilisateur à l'aide d'un algorithme de détection de communauté.





Une étude expérimentale sur 10 000 comptes issus du site Wikipedia anglophone nous a permis de comparer six algorithmes d'apprentissage automatique pour la phase de détection des *sockpuppets* et de montrer que *SocksCatch*, avec les caractéristiques de description de comptes que nous avons proposées, détecte 91% des *sockpuppets* avec un algorithme de forêt aléatoire. Nous avons également comparé cinq algorithmes de détection de communauté pour la phase de regroupement, et montré que les *sockpuppets* regroupés par *SocksCatch* et les groupes réels sont similaires à 87%, selon les mesures de comparaison des clusters : *variation de l'information normalisée* (NVI) et de l'information mutuelle normalisée (NMI).

À court terme, nous allons poursuivre le développement de *SocksCatch* dans deux directions : une exploitation en ligne avec des données provenant en temps réel de l'activité d'un média social et une détection et un regroupement des *sockpuppets* automatisé, et une adaptation à d'autres médias sociaux que Wikipedia.

Détection adaptative des comptes malveillants sur les réseaux sociaux numériques

Le spam sur les réseaux sociaux (dit spam social) est un problème actif et récent. La plupart des systèmes de détection proposés dans la littérature sont basés sur l'apprentissage statistique supervisé et ont une conception rigide qui ne prend pas en compte les signaux dynamiques du réseau social et le changement constant dans la nature du spam. Ceci limite la durée de vie et l'adaptabilité du système de détection conçu. Contrairement à une tâche de classification classique, il s'agit ici d'un adversaire qui s'adapte aux filtres de détection en changeant de caractéristiques et en créant constamment de nouvelles méthodes de camouflage.

La thèse de Nour EL MAWASS s'intéresse à ce problème en proposant d'augmenter la complexité du modèle de détection pour y intégrer des signaux dynamiques liés aux applications utilisées, et de les exploiter pour concevoir un système capable de s'adapter automatiquement aux changements dans les caractéristiques du spam. Nous avons dans un premier temps proposé d'utiliser une approche probabiliste graphique (fondée sur les *Markov Random Field*) pour modéliser la tâche de détection des comptes malveillants[1]. Le modèle graphique, établi à partir de la similarité entre les comptes, permet de raisonner globalement sur la tâche de classification. Contrairement à une classification au niveau individuel, la classe attribuée à un compte ne dépend pas seulement de ses caractéristiques mais aussi des caractéristiques des autres comptes classifiés. Nous partons des *a priori* calculés individuellement sur chaque compte en utilisant des classifieurs supervisés traditionnels. Notre modèle utilise ensuite une propagation de croyance sur le graphe d'utilisateurs pour calculer l'*a posteriori* de chaque compte, corrigeant ainsi les erreurs de classification. Nous avons montré une amélioration des meilleures performances de classifieurs issus de l'état de l'art.

Détection et identification de groupes d'influence dans les médias sociaux

La thèse de Guillaume GADEK [2] vise à reconnaître des opinions, détecter des influenceurs, et identifier des groupes d'influence dans les médias sociaux et en particulier sur Twitter. Ce travail se décompose en 3 grandes problématiques : analyse des messages publiés, caractérisation des différents utilisateurs et identification des groupes sociaux.

La première problématique concerne l'analyse des opinions portées par les messages émis sur Twitter. La reconnaissance d'opinion dans des tweets est un défi scientifique : les textes



sont courts, les sujets implicites et l'orthographe inventive. Nous cherchons ici à déterminer efficacement l'orientation (positive, négative ou neutre) des tweets envers une cible donnée. Une méthode novatrice basée sur les contextonymes et contextosets a été utilisée afin de désambiguïser le contenu implicite et améliorer les classifieurs d'opinion. Les contextosets sont extraits d'un graphe de co-occurrence des mots, et permettent de représenter le sens d'un mot en fonction de ses mots environnants en contexte d'utilisation. Cette méthode a été évaluée sur le corpus de tweets annotés SemEval2016.

Le second axe de ce travail concerne l'analyse des profils utilisateurs, membres d'un réseau social. Les utilisateurs sont ainsi caractérisés par 1) l'impact qu'ils peuvent avoir par leurs messages émis et 2) par leur comportement. Ainsi, une mesure de l'impact d'un utilisateur a été proposée, ainsi que la construction de profils utilisateurs, permettant d'analyser le comportement d'une grande population. Enfin, une méthode permettant d'étiqueter automatiquement le rôle adopté par un utilisateur selon sa position dans le graphe social a été proposée.

Dans le troisième aspect traité dans ces

travaux, nous soutenons l'hypothèse que les communautés recherchées sur les réseaux sociaux devraient être topiquement cohérentes. Un modèle de représentation de l'interaction entre utilisateurs sur Twitter a donc été proposé, ainsi qu'un ensemble de métriques permettant d'évaluer la cohésion des communautés détectées. Ces éléments ont été utilisés lors d'une analyse détaillée sur un grand jeu de données de tweets.

Références

- [1] N. El Mawass, P. Honeine, and L. Vercoouter. Supervised classification of social spammers using a similarity-based markov random field approach. In *5th multidisciplinary international social networks conference*, 2018.
- [2] G. Gadek, A. Pauchet, N. Malandain, K. Khelif, L. Vercoouter, and S. Brunessaux. Topical cohesion of communities on twitter. *Procedia Computer Science*, 112 :584–593, 2017.
- [3] Z. Yamak, J. Saunier, and L. Vercoouter. Sockscatch : Automatic detection and grouping of sockpuppets in social media. *Knowledge-Based Systems*, 149 :124–142, 2018.



■ Analyse des réseaux sociaux pour des applications liées à la santé

LIRMM - ADVANSE

CNRS - Université de Montpellier

www.lirmm.fr/recherche/equipes/advanse

LIRMM - ADVANSE

CNRS - Université de Montpellier, Université Paul Valéry
Montpellier

www.lirmm.fr/recherche/equipes/advanse

Jérôme AZÉ

prenom.nom@lirmm.fr

Sandra BRINGAY

prenom.nom@lirmm.fr

Introduction

Depuis sa création en 2014, dans le contexte des Sciences des Données, l'équipe projet ADVANSE (Advanced ANalytics for Data SciencE), s'intéresse aux approches d'extraction de connaissances dans les données complexes : structurées, semi-structurées, multidimensionnelles, qualitatives, quantitatives, textuelles, spatio-temporelles, approximatives, bruitées, etc. Ses travaux allient, de manière complémentaire, des approches de fouille de données, d'apprentissage automatique et de visualisation analytique (Visual Analytics).

Un axe de recherche important dans l'équipe consiste à combiner ces différentes approches pour analyser des réseaux sociaux pour différentes applications du domaine de la santé. Notre objectif a tout d'abord été d'extraire de la connaissance à partir des données hétérogènes des médias sociaux de santé et notamment à partir de la structure des réseaux et des textes. L'originalité a été de prendre en compte les données produites directement par les patients dans le processus d'extraction de connaissances. Généralement, seules les données médicales produites par les professionnels de santé sont étudiées. Or, nous avons montré qu'il est très intéressant d'intégrer les données produites par les patients. Ces données par nature même hétérogènes, incertaines et entachées d'erreurs sont difficiles à traiter. Nous avons donc développé de nouvelles ressources

et méthodes dédiées aux réseaux sociaux pour le domaine de la santé.

Analyse semi-automatique des contenus

Nous avons mis en relation des connaissances avérées et des connaissances coconstruites par les internautes. Pour cela, dans le cadre d'une collaboration avec l'ICM et l'IMAG, nous avons exploré des méthodes supervisées, semi-supervisées et non supervisées pour décrire les productions des patients selon différentes dimensions (qui, quand, où et comment ?). Ces méthodes se sont avérées efficaces grâce à la construction originale et automatique d'une ressource mettant en relation le vocabulaire des patients et celui des professionnels de santé. Au cours du projet SIFR, cette ressource a été intégrée dans le portail [Bioportal](#) qui permet d'annoter des textes avec des concepts médicaux [5]. Cette ressource a été utilisée pour comparer les productions des patients dans les médias sociaux et les connaissances avérées issues des questionnaires de qualité de vie utilisés par les professionnels de santé dans les essais cliniques [6].

Nous avons également croisé ces caractéristiques avec d'autres dimensions liées au risque, à l'incertitude, et aux sentiments, etc. Nous avons produit un nouveau lexique dédié à l'analyse de polarités et d'émotions en français [2] qui a permis de positionner l'équipe en



première place sur une des tâches du challenge DEFT 2015.

Dernièrement, dans le cadre du projet #AIDS en collaboration avec le LERASS-CERIC, nous avons proposé un nouveau modèle de classification des textes des médias sociaux basé sur une architecture de *Deep Learning* efficace et une visualisation originale et interactive appelée *MultiStream* [3]. Cette approche a été utilisée pour explorer les messages du site Sida Info Service, et identifier les acteurs et les thèmes clés pour la compréhension des prises de risque. Cette approche a également été adaptée pour traiter la volumétrie des Tweets dans le cadre de l'analyse de la campagne de publicité SmearForSmear [4] pour la prévention du cancer du col utérin.

Fiabilité des contenus

L'équipe s'est également intéressée à la fiabilité des connaissances issues du Web Social. S'il est difficile d'empêcher les internautes de consulter des informations non pertinentes ou non fiables dans les médias sociaux, il est possible, en revanche, de concevoir des outils pour mettre en évidence des informations de qualité dans ces médias. Nous avons proposé différentes approches notamment basées sur la structure en graphe des réseaux pour détecter automatiquement le rôle (expert, non expert) et la confiance que l'on peut accorder aux internautes dans les médias sociaux selon les thématiques et au cours du temps [1].

Analyse temporelle des réseaux sociaux

Par ailleurs, les aspects temporels ont également été étudiés pour capturer l'évolution de thématiques au cours du temps et de l'histoire des patients. Ainsi, nous avons, par exemple, pu mettre en avant des sujets d'interrogation des patients associés à des états d'esprit et à un événement de leur histoire médicale, jusqu'alors méconnus des professionnels

de santé [7]. Ces travaux sont menés dans le cadre du projet DontDolt sur la détection des personnes à risque suicidaires en collaboration avec le département des Urgences Psychiatrique du CHU de Montpellier. Nous avons combiné différentes méthodes pour capter les changements brusques ou plus insidieux (e.g. votes entre différents classifieurs, détection de *concepts drift*, *active learning* pour prendre en compte les réactions des experts en charge du *monitoring*, etc.).

Conclusion

Dans cet article, nous avons présenté différentes méthodes développées dans l'équipe ADVANSE pour monitorer de manière globale la santé des populations *via* l'analyse des données des réseaux sociaux. Ces données sont extrêmement variables du point de vue inter et intra-individu. Les fouiller afin de mieux les interpréter et obtenir des modèles prédictifs permettant de déclencher efficacement des alertes restent des questions de recherche d'actualité. Les défis consistent ici à déterminer, pour chaque application, les caractéristiques pertinentes à repérer dans les textes des messages ou dans les comportements repérables dans le graphe des interactions dans les réseaux sociaux afin de produire des modèles robustes sur ces données particulières (hétérogènes, volumineuses, etc.). Ici, les questions éthiques et juridiques seront centrales (e.g. vie privée, consentement).

Références

- [1] Amine Abdaoui, Jérôme Azé, Sandra Bringay, and Pascal Poncelet. Collaborative content-based method for estimating user reputation in online forums. In *Web Information Systems Engineering - WISE 2015 - 16th International Conference, Miami, FL, USA, November 1-3, 2015, Proceedings, Part II*, pages 292–299, 2015.



- [2] Amine Abdaoui, Jérôme Azé, Sandra Bringay, and Pascal Poncelet. FEEL : a french expanded emotion lexicon. *Language Resources and Evaluation*, 51(3) :833–855, 2017.
- [3] E. Cuenca, A. Sallaberry, F. Y. WANG, and P. Poncelet. Multistream : A multiresolution streamgraph approach to explore hierarchical time series. *IEEE Transactions on Visualization and Computer Graphics*, 2018.
- [4] Philippe Lenoir, Bilel Moulahi, Jérôme Azé, Sandra Bringay, Gregoire Mercier, and François Carbonnel. Raising awareness about cervical cancer using twitter : Content analysis of the 2015 #smearforsmear campaign. *J Med Internet Res*, 19(10) :e344, Oct 2017.
- [5] Mike Donald Tapi Nzali, Jérôme Azé, Sandra Bringay, Christian Lavergne, Caroline Mollevi, and Thomas Optiz. Reconciliation of patient/doctor vocabulary in a structured resource. *Health Informatics Journal*, 0(0) :1460458217751014, 0. PMID : 29332530.
- [6] Donald Mike Tapi Nzali, Sandra Bringay, Christian Lavergne, Caroline Mollevi, and Thomas Opitz. What patients can tell us : Topic analysis for social media on breast cancer. *JMIR Med Inform*, 5(3) :e23, Jul 2017.
- [7] Mia Johnson Vioulès, Bilel Moulahi, Jérôme Azé, and Sandra Bringay. Detection of suicide-related posts in twitter data streams. *IBM Journal of Research and Development*, 62(1) :7, 2018.

■ Autodétermination des utilisateurs dans les réseaux sociaux

LORIA/KIWI | **Anne BOYER**
Université de Lorraine | anne.boyer@loria.fr
kiwi.loria.fr

Membres

- Charif HAYDAR, responsable Recherche et Innovation de l'entreprise Sailendra SAS, ex-doctorant de KIWI

Thématique générale de l'équipe

L'activité de recherche de l'équipe KIWI (Knowledge, Information & Web Intelligence) du laboratoire LORIA (UMR 7503) s'inscrit dans le domaine de l'apprentissage automatique et de la fouille de données. En particulier, elle concerne la modélisation explicative, prédictive et prescriptive du comportement d'un utilisateur en interaction avec un système numérique, à partir de l'analyse au-

tomatique des traces numériques collectées. Il s'agit par exemple de déterminer les facteurs qui expliquent le comportement observé d'un utilisateur à un moment donné, de prédire quel devrait être le comportement futur d'un utilisateur en fonction des observations réalisées, ou de définir quelle(s) action(s) ou séquence(s) d'actions recommander à un utilisateur pour atteindre un objectif donné. La modélisation d'un comportement utilisateur permet donc de le comprendre, d'identifier des comportements-type, de rapprocher les différents utilisateurs sur leur comportement (détermination de communautés), de prédire les comportements futurs des utilisateurs en vue d'une personnalisation ou de recomman-



Afia

Association française
pour l'Intelligence Artificielle

dations, etc. Pour atteindre ces objectifs, l'approche retenue dans KIWI est à base de traces numériques. Ces traces concernent des observés (l'objet d'étude), par exemple les utilisateurs, en lien avec des objets (ou items) relatifs au phénomène, par exemple des contenus en ligne, des produits d'un catalogue de vente, etc. Nous faisons l'hypothèse implicite que les traces collectées sont représentatives du phénomène à modéliser.

Initiés dans le domaine applicatif du e-commerce et des systèmes de recommandation, les travaux de KIWI trouvent de nombreux domaines d'applications (e-santé, héritage culturel, analyse des politiques publiques, etc.) et se focalisent depuis environ cinq ans sur le domaine de l'éducation (notamment sur les *Learning Analytics* ou Analyse des données de l'apprentissage).

Les traces exploitées peuvent être de toute nature (logs, textes, *clickstream*, audio, données d'*eye trackers*, etc.), du moment qu'elles peuvent être collectées, stockées et exploitées sous un format numérique. Cela conduit à considérer des données complexes, multisources, multiéchelles, incertaines, bruitées, etc. pour modéliser le comportement d'un utilisateur. Ces traces, bien que présentes en très gros volume, sont présentes en relativement faible nombre pour chaque observé (utilisateur ou objet). Ainsi les algorithmes proposés doivent prendre en compte le paradoxe dû à la fois la parcimonie des données à l'échelle individuelle, et la massivité des données à l'échelle globale.

Parmi les sources de données possibles, figurent les réseaux sociaux. Les données collectées sont les traces laissées par les utilisateurs au travers des actions qu'ils entreprennent pour consulter, commenter, ou partager des contenus, établir des liens sociaux avec d'autres utilisateurs. L'extraction de connaissance à partir de ces données permet de mo-

déliser des comportements sociaux, d'améliorer l'expérience d'utilisateurs, d'anticiper certains dangers ou attaques, de prédire des comportements, voire de les influencer. C'est pourquoi depuis plusieurs années, KIWI a considéré différentes problématiques comme la détection de communautés dans les réseaux sociaux (thèse d'I. ESSLIMANI), la détection de leaders ou d'utilisateurs représentatifs (HDR d'A. BRUN), la propagation de la confiance dans les réseaux sociaux (thèse de C. HAYDAR).

Dans la suite de cet article, nous allons nous centrer sur un aspect spécifique des travaux menés dans l'équipe KIWI : la détection automatique de communautés dans les réseaux sociaux et l'identification des patterns comportementaux qui les caractérisent, au travers d'un travail mené dans le contexte des systèmes de recommandation avec Charif HAYDAR, ancien doctorant de KIWI et actuellement ingénieur de recherche dans la société Sailendra.

Description des travaux

Ce travail repose sur l'exploitation du clustering. Le clustering n'est certes pas une technique propre à la modélisation de comportement mais il représente un outil important pour la modélisation de comportement collectif [4]. Il permet notamment de détecter les communautés et les patterns comportementaux qui les identifient, ce qui permet par la suite, de trouver un équilibre entre besoin de personnalisation et nécessité d'ouverture pour éviter la sclérose des communautés. Plusieurs approches existent pour accomplir une tâche de clustering, comme le clustering hiérarchique, les K-means, ou le clustering à base de densité [2]. On peut définir un algorithme de clustering comme un moyen de regrouper des vecteurs dans un espace multidimensionnel en fonction de la proximité entre eux [2].

Dans le domaine de la recommandation



comme dans beaucoup d'autres, l'entrée de l'algorithme est une matrice qui représente les préférences observées des utilisateurs pour les items proposés. Par exemple, les lignes de la matrice représentent les utilisateurs que l'on souhaite segmenter, et les colonnes représentent les produits (ou items) du catalogue. Chaque valeur de la matrice représente l'intérêt prononcé d'un utilisateur à l'égard d'un item. Puisqu'on est souvent face à d'immenses catalogues, et que l'utilisateur ne peut consulter l'ensemble du catalogue, beaucoup de valeurs restent inconnues. La matrice est très creuse (de 95 à 99% de données manquantes le plus souvent), ce qui impacte la qualité de clustering, et encourage à trouver des algorithmes adaptés à ce type de matrice. De même, l'intensité d'utilisation des réseaux sociaux et donc la quantité des actions tracées varie largement suivant les utilisateurs ou les communautés. Il est donc intéressant de disposer d'un algorithme de clustering capable de s'adapter à différentes densités de données disponibles. L'algorithme Mutual Vote (MV) [1] est un algorithme de clustering conçu pour s'adapter au manque de données, et à la variance de densité de données dans un espace multidimensionnel. De plus, les algorithmes à base de prototype (K-means) sont bons pour trouver des clusters à format spécifique. Les algorithmes sensibles à la densité sont capables de détecter différentes formes de clusters, sauf qu'il faut les paramétrer à une valeur fixe de granularité, ce qui les rend incapables de détecter les clusters de granularités différentes [3].

Considérons que chaque utilisateur est représenté par un vecteur, qui contient l'ensemble des données disponibles sur l'utilisateur. L'idée de l'algorithme MV est d'associer à chaque vecteur une liste de priorités qui contient les vecteurs (autres utilisateurs) qu'il aimerait inviter à rejoindre son cluster. Un système de négociation à base de la logique sub-

jective se met en place pour gérer le regroupement des vecteurs pour former les clusters.

Pour commencer, MV calcule les distances entre le vecteur actif et l'ensemble des vecteurs disponibles. Nous utilisons la distance de Jaccard pour gérer la différence potentielle de taille entre les vecteurs. Les vecteurs dont la distance est inférieure à un seuil donné s sont considérés comme membres de la liste des priorités (s est souvent fixé à la moyenne moins un écart-type des distances). La volonté d'un vecteur d'inviter un vecteur de sa liste de priorité à son cluster n'est pas la même pour tous les membres de la liste. Une fonction de mapping est utilisée pour convertir les distances en probabilités de façon que plus la distance est petite plus la probabilité est élevée. Concrètement, ces probabilités expriment à quel point le vecteur actif aimerait être en un seul cluster avec chacun de membre de sa liste de priorité. La logique subjective (LS) est utilisée pour gérer les probabilités. LS est une extension de la logique probabiliste qui associe un propriétaire à chaque probabilité et l'appelle opinion. L'intérêt de la logique subjective réside dans sa capacité à agréger des opinions venant de différents propriétaires. La négociation commence par l'envoi par chaque vecteur d'une invitation au premier membre de sa liste de priorité. Le récepteur donne un avis favorable seulement si le demandeur figure à la tête de sa liste de priorité (la liste de récepteur), sinon il rejette la demande. Si la demande est acceptée, les deux vecteurs forment désormais un cluster, et leurs listes de priorités sont agrégées par la fonction d'agrégation de la logique subjective, ce qui change l'ordre des priorités. Les vecteurs qui veulent rejoindre le cluster doivent désormais négocier avec les deux membres de cluster à la fois. Ainsi, l'opération continue jusqu'à ce qu'aucune négociation ne soit possible (aucun cluster n'a plus envie de se fusionner avec un autre).



MV a été testé sur de données réelles d'un CRM pour regrouper les clients en fonction de leurs domaines d'intérêt, dans le contexte des systèmes de recommandation. L'objectif était de pouvoir proposer des produits d'ouverture issus de la communauté, et de trouver des clients potentiels pour un produit ou une gamme de produits. Selon nos expérimentations, MV s'est montré supérieur à K-means et à DBSCAN (l'algorithme à base de densité le plus répandu qui a servi de baseline) en termes de précision et de qualité d'étiquetage.

Références

[1] A. Boyer C. Haydar. A new statistical density clustering algorithm based on mutual

vote and subjective logic applied to recommender systems. pages 59–66, 2017.

[2] V Kumar PN Tan, M Steinbach. Data mining cluster analysis : basic concepts and algorithms. 2013.

[3] J. Cao S. Zhou, A. Zhou. A data-partitioning-based dbscan algorithm. *Journal of computer research and development*, 37(10) :1153–1159, 2000.

[4] Y. A. Wendong. Social media user partitioning based on ensemble clustering. pages 1–6, 2016.

■ Recherche en réseaux sociaux dans l'équipe DM2L

LIRIS/DM2L

Université Lyon 1

<https://projet.liris.cnrs.fr/dm2l/>

LIRIS/DM2L

INSA

<https://projet.liris.cnrs.fr/dm2l/>

Rémy CAZABET

remy.cazabet@gmail.com

Jean-François BOULICAUT

Thème général de l'équipe

L'équipe DM2L se positionne sur les aspects théoriques de la fouille de données et de l'apprentissage statistique, avec notamment une expertise sur la fouille de motifs sous contraintes, la fouille de graphes, la détection de communautés, les techniques de coclustering, les réseaux bayésiens, les méthodes ensemblistes et la sélection de variables.

Description des travaux en lien avec les réseaux sociaux.

Nous travaillons depuis de nombreuses années sur les aspects méthodologiques et algorithmiques de la découverte de motifs sous

contraintes et nous avons notamment étudié la découverte d'ensembles de nœuds *a priori* intéressants dans des graphes qui pouvaient être dynamiques [11, 6], attribués [13, 10] (sur les nœuds ou les relations encodées) ou les deux [7, 12]. Naturellement, ce type de travail est souvent validé qualitativement et quantitativement sur des graphes d'interactions sociales. Récemment, l'un de nos axes principaux d'étude de réseaux sociaux repose sur l'*Exceptional contextual subgraph mining*, problème consistant à découvrir dans un graphe donné des sous-graphes exceptionnels, c'est-à-dire différant sensiblement du reste du graphe, à la fois en terme de topologie mais également en terme de propriétés des composantes du



graphe (nœuds, liens). L'équipe a notamment proposé un algorithme passant à l'échelle [9], dont l'efficacité a été montrée par l'application à des données géolocalisées extraites de plateformes de réseaux sociaux [2] et des données de déplacements en vélos en libre service [3].

L'équipe a une forte expertise dans l'analyse de la dynamique de réseaux sociaux. Récemment, elle a proposé la méthode DSC (*Discovering Similarities Change*, [1]) qui, appliquée à des réseaux de covotation, permet de mettre en évidence l'évolution de groupes de votants pouvant être reconnus comme les membres de formations politiques en évolution. Nous avons aussi travaillé à la détection d'événements dans les plateformes de réseaux sociaux, en utilisant des séries temporelles [8] et des méthodes de détection de communautés [5]. Une des facettes intéressantes de la dynamique des réseaux sociaux concerne la détection de communautés dynamiques avec des contributions sur l'état de l'art [15] mais aussi le développement de méthodes originales [4].

Projets marquants

Le projet ANR BITUNAM qui a débuté en septembre 2018 propose de transposer les outils et techniques d'analyse des réseaux sociaux aux réseaux de transactions financières en crypto-monnaies, en particulier Bitcoin. Il est mené en collaboration avec des socio-économistes (Jérôme BLANC, Marie FARE, Laboratoire Triangle, Université de Lyon), avec pour objectif de caractériser la nature des transactions, des acteurs et donc de ces *monnaies* elles-mêmes. Les objectifs sont notamment d'utiliser la détection de communautés pour ré-identifier les adresses d'un même utilisateur afin de construire un réseau de transaction entre acteurs et non plus seulement entre adresses anonymes, puis d'identifier les acteurs clefs, les phénomènes de diffusion, la structure du réseau, etc. Ce projet va continuer des tra-

voux préliminaires de l'équipe sur le sujet [14].

L'équipe est également impliquée dans le projet IDEX Breakthrough Academics *Data Science & Machine Learning for Complex and Dynamical Models*. L'un des deux objectifs de ce projet est la compréhension quantitative des systèmes sociaux, et en particulier des réseaux sociaux, et leur description dynamique, à partir des traces numériques de relations sociales et économiques. L'équipe DM2L sera particulièrement impliquée sur des méthodes mêlant avancées récentes en *machine learning* et en analyse de réseaux. Les sujets de l'équipe seront plus particulièrement l'étude de systèmes à base de réseaux de neurones profonds intégrant des connaissances et des contraintes métiers, et le développement de méthodes basées sur le plongement de graphes dans des espaces en faible dimension (*network embedding*) pour la détection de structures mésoscopiques et d'anomalies dans des réseaux dynamiques.

Enfin, plusieurs doctorants de l'équipe travaillent également sur des sujets fortement orientés sur l'analyse de réseaux sociaux :

- *Découverte automatique de l'organisation des réseaux complexes*. Aujourd'hui, plusieurs modèles statistiques de réseaux peuvent être utilisés pour comprendre son organisation (Communautés, Assortativité, Organisation, Spatiale, Hiérarchique, etc.), mais chacun de ces modèles entraîné sur un graphe en donnera une interprétation différente. Nous cherchons à trouver, pour un graphe donné, le modèle le plus à même de l'expliquer selon les principes du rasoir d'Occam, ainsi qu'à combiner ces modèles afin d'expliquer plus précisément les facteurs sous-jacents à la structure d'un réseau donné.
- *Analyse de multi-graphes attribués dynamiques : application aux réseaux sociaux et réseaux d'entreprise*. L'objectif de cette thèse en collaboration avec le Data R&D



Institute de EMLyon Business School est de développer de nouveaux outils théoriques pour l'analyse de données massives hétérogènes et structurées, dans un contexte d'applications business (marketing quantitatif, etc.)

Références

- [1] Adnene Belfodil, Sylvie Cazalens, Philippe Lamarre, and Marc Plantevit. Flash points : Discovering exceptional pairwise behaviors in vote or rating data. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 442–458. Springer, 2017.
- [2] Ahmed Anes Bendimerad, Marc Plantevit, and Céline Robardet. Unsupervised exceptional attributed sub-graph mining in urban data. In *Data Mining (ICDM), 2016 IEEE 16th International Conference on*, pages 21–30. IEEE, 2016.
- [3] Anes Bendimerad, Rémy Cazabet, Marc Plantevit, and Céline Robardet. Contextual subgraph discovery with mobility models. In *International Workshop on Complex Networks and their Applications*, pages 477–489. Springer, 2017.
- [4] Souâad Boudebza, Rémy Cazabet, Faïçal Azouaou, and Omar Nouali. Olcpm : An online framework for detecting overlapping communities in dynamic social networks. *Computer Communications*, 123 :36–51, 2018.
- [5] Rémy Cazabet, Hideaki Takeda, Masahiro Hamasaki, and Frédéric Amblard. Using dynamic community detection to identify trends in user-generated content. *Social Network Analysis and Mining*, 2(4) :361–371, 2012.
- [6] Loïc Cerf, Tran Bao Nhan Nguyen, and Jean-François Boulicaut. Mining constrained cross-graph cliques in dynamic networks. In *Inductive Databases and Constraint-Based Data Mining.*, pages 199–228. 2010.
- [7] Elise Desmier, Marc Plantevit, Céline Robardet, and Jean-François Boulicaut. Trend mining in dynamic attributed graphs. In *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, pages 654–669. Springer, 2013.
- [8] Pierre Houdyer, Albrecht Zimmerman, Mehdi Kaytoue, Marc Plantevit, Joseph Mitchell, and Céline Robardet. Gazouille : Detecting and illustrating local events from geolocalized social media streams. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 276–280. Springer, 2015.
- [9] Mehdi Kaytoue, Marc Plantevit, Albrecht Zimmermann, Anes Bendimerad, and Céline Robardet. Exceptional contextual subgraph mining. *Machine Learning*, 106(8) :1171–1211, 2017.
- [10] Nicolas Mougél, Christophe Rigotti, Marc Plantevit, and Olivier Gandrillon. Finding maximal homogeneous clique sets. *Knowledge and Information Systems*, 35(1) :1–30, 2013.
- [11] Kim Ngan Tan Nguyen, Loïc Cerf, Marc Plantevit, and Jean-François Boulicaut. Discovering descriptive rules in relational dynamic graphs. *Intelligent Data Analysis*, 17(1) :49–69, 2013.
- [12] Gunce Keziban Orman, Vincent Labatut, Marc Plantevit, and Jean-François Boulicaut. Interpreting communities based on the evolution of a dynamic attributed network. *Social Network Analysis and Mining*, 5(1) :1–22, 2015.
- [13] Adriana Prado, Marc Plantevit, Céline Robardet, and Jean-François Boulicaut. Mining graph topological patterns : Fin-



- ding co-variations among vertex descriptors. *IEEE Transactions on Data and Knowledge Engineering*, 25(9) :2090–2104, 2013.
- [14] Cazabet Remy, Baccour Rym, and Latapy Matthieu. Tracking bitcoin users activity using community detection on a network of weak signals. In *International Workshop on Complex Networks and their Applications*, pages 166–177. Springer, 2017.
- [15] Giulio Rossetti and Rémy Cazabet. Community discovery in dynamic networks : A survey. *ACM Computing Surveys (CSUR)*, 51(2) :35, 2018.

■ LIA : Laboratoire Informatique d'Avignon

LIA | **Jean-François BONASTRE**
Université d'Avignon et des Pays de Vaucluse | jean-francois.bonastre@univ-lia.univ-avignon.fr

Chercheurs impliqués :

- Jean-François BONASTRE, PR
- Richard DUFOUR, MCF
- Rachid EL-AZOUZI, PR
- Rosa FIGUEIREDO, MCF
- Corinne FREDOUILLE, MCF HDR
- Serigne GUEYE, MCF HDR
- Majed HADDAD, MCF
- Yezekael HAYEL, MCF HDR
- Stéphane HUET, MCF
- Bassam JABAIAH, MCF
- Tania JIMENEZ, IR
- Pierre JOURLIN, MCF
- Vincent LABATUT, MCF
- Fabrice LEFEVRE, PR
- Georges LINARES, PR
- Driss MATROUF, MCF HDR
- Mohamed MORCHID, MCF
- Mickael ROUVIER, MCF
- Éric SAN-JUAN, MCF
- Juan-Manuel TORRES MORENO, MCF HDR
- Rénato DE MORI, PR émérite
- Marc EL BEZE, PR émérite
- Pierre-Michel BOUSQUET, PRAG
- Eitan ALTMAN, DR INRIA associé
- Didier JOSSELINE, DR CNRS associé

Présentation générale

Le LIA est une équipe d'accueil (EA 4128) créée en 1987 à l'Université d'Avignon et des Pays de Vaucluse. Son activité de recherche est actuellement organisée en trois axes thématiques : *Langage, Optimisation & recherche opérationnelle*, et *Réseaux informatiques*. Des travaux relatifs à l'IA appliquée aux réseaux sociaux apparaissent dans le cadre de chacun d'entre eux. Pour le langage, il s'agit essentiellement de l'étude des contenus échangés sur les médias sociaux et de l'extraction de connaissances issus de flux multimédia, audio et/ou vidéo. Dans le domaine de l'optimisation, le LIA s'intéresse à la résolution de problèmes définis sur des graphes, tels que leur partitionnement, qui ont de multiples applications dans l'analyse de réseaux sociaux. Enfin, dans le domaine des réseaux informatiques, le LIA manie des outils relatifs à la théorie des jeux, qui sont notamment utilisés pour étudier des processus se déroulant sur les réseaux sociaux.

Deux axes transversaux ont récemment été adjoints à ces thèmes, afin de favoriser les échanges à l'intérieur (et à l'extérieur) du laboratoire : *Société numérique* et *Systèmes complexes*. Le premier est relatif aux liens entre-



Afia

Association française
pour l'Intelligence Artificielle

nus par le LIA avec un ensemble de laboratoires de SHS, via la FR 3621 Agorantic (Culture, Patrimoines, et Sociétés Numériques). Ce type de collaboration a permis de lancer plusieurs travaux portant sur l'analyse de réseaux sociaux au sens large, aussi bien en-ligne que hors-ligne. Le deuxième axe comporte une activité de recherche sur l'extraction et l'analyse de réseaux complexes, qui incluent notamment les réseaux sociaux. Le LIA est également partie prenante du LABEX BLRI (Brain and Language Research Institute) et de l'institut de convergence ILCB (*Institute of Language, Communication and the Brain*), dans lesquels sont menés des travaux concernant les réseaux sociaux en association avec des neurologues et des psychocognitivistes.

RI & TALN

Un certain nombre de travaux récents menés dans les axes *Langage* et *Société numérique* ont porté sur la recherche d'information et/ou le traitement du langage naturel dans le contexte des réseaux sociaux. Trois types d'approches non-mutuellement exclusives peuvent être identifiés : l'extraction de réseaux sociaux (au sens sociologique) à partir de données brutes ; l'utilisation de réseaux sociaux comme modèles pour résoudre des problèmes de plus haut niveau ; et l'exploitation de contenus publiés sur les médias sociaux de type Twitter.

L'extraction de réseaux d'interactions sociales est une tâche particulièrement pertinente dans le cadre de nos collaborations au sein de la FR Agorantic, car l'objet réseau social constitue une passerelle naturelle entre informatique et sciences sociales. Un travail en cours impliquant des politistes et historiens vise ainsi à identifier des événements dans des articles de presse, billets de blogs, notices biographiques, et autres textes, afin d'en extraire des réseaux de coparticipation à ces événements. Le but est à la fois d'étudier le réseau social en soi,

mais également de l'utiliser pour répondre à des questions disciplinaires, notamment relatives à la façon dont le Web joue le rôle de miroir déformant sur l'activité des élus [9].

La FR Agorantic a financé le doctorat de X. BOST, qui a permis de développer une activité relative au résumé automatique de séries télévisées [3]. En plus des descripteurs vidéos classiques, l'originalité de la méthode proposée est de s'appuyer sur une représentation de l'intrigue prenant la forme d'un réseau social de personnages. Il s'agit d'un graphe dynamique extrait en analysant les dialogues au moyen d'une méthode baptisée *lissage narratif*, et permettant de tenir compte des spécificités du médium. Dans le cadre de la thèse d'E. PAPÉGNIES [14] (en cours), nous proposons une méthode à base de réseaux conversationnels pour identifier des situations d'abus sur des messageries en ligne. Ces travaux montrent que l'évolution de la structure de la conversation est plus révélatrice que sa teneur pour détecter ces situations.

Le LIA a participé au projet [ANR Imagiweb](#) (*Analyse de la dynamique des images sur le Web 2.0*), porté par ERIC (Université Lumière Lyon 2). Son objectif était d'étudier comment le Web construit, diffuse, et affecte l'image de marque. Le rôle du LIA a été de préparer un corpus de tweets et de l'analyser pour identifier les opinions et leur évolution. Dans le cadre de la thèse de J.-V. COSSU, nous avons proposé une méthode semi-supervisée d'annotation des tweets, utilisant l'*active learning* [4], et une approche à base de *Partial Least Squares Path Modeling* (PLS-PM) pour leur exploitation [6]. Nous avons également mené plusieurs travaux ponctuels relatifs à la classification de tweets et de tweetos : prédiction d'influence sur la base des données du challenge Replab (CLEF 2014) [5] ; estimation du lieu de publication via une méthode de type LDA (*Latent Dirichlet Allocation*) [12] ; prédiction de retweet [11].



Afia

Association française
pour l'Intelligence Artificielle

Le projet [ANR GaFes](#) (*Galerie des festivals*), qui est porté par le LIA, vise au développement d'approches pluridisciplinaires pour le traitement de données culturelles. Outre le LIA, il implique également le Centre Norbert Elias (Sciences de l'information et de la communication), l'institut Eurecom, la société Syl-labs et le cabinet d'enquêtes GECE. Le projet concerne l'analyse et la ré-éditorialisation des traces que les festivals culturels laissent sur Internet, et tout particulièrement les médias sociaux. Le LIA intervient notamment sur des problèmes de structuration, visualisation et exploitation des données hétérogènes collectées [16]. Les données extraites constituent un terrain d'enquête numérique pour des analyses sociologiques consacrées à l'étude des publics des festivals et des nouvelles pratiques festiva-lières. Le corpus collecté est aussi le support de tâches des campagnes d'évaluation CLEF 2017 et 2018 (*Cultural Microblog Contextua-lization*).

Jeux évolutionnaires

Une bonne partie de l'activité dans les axes *Réseaux* et *Systèmes complexes* porte sur les jeux évolutionnaires, notamment ceux ayant lieu sur des graphes modélisant des réseaux sociaux.

[Congas](#) (*Dynamics and Coevolution in Multi-Level Strategic Interaction Games*) était un projet FP7 visant à développer des modèles et outils fondés sur la théorie des jeux pour analyser, prédire, et contrôler des systèmes complexes multi-niveaux. Il regroupait des partenaires prestigieux (TuDelft, CREATENET, IN-RIA, University of Pisa, Technion, Imperial College London), parmi lesquels le LIA a joué un rôle important en élaborant des outils basés sur les jeux évolutionnaires et multirésolution, et intégrant des concepts de théorie des graphes, mécanique statistique, contrôle et optimisation. Nous avons notamment étudié les dyna-

miques de réplication dans une population d'individus organisée sous la forme de communautés avec différents niveaux d'interactions [2], et montré l'impact de cette structure d'interaction sur les dynamiques de réplication et la stabilité des équilibres.

L'étude de ces processus dynamiques de prise de décision a été poursuivie dans des travaux plus récents, avec comme objectif de comprendre les stratégies optimales de publication dans les réseaux sociaux lorsque plusieurs générateurs de contenu sont en compétition [17]. Nous avons étudié l'impact de ces stratégies individuelles sur la diversité de contenu proposée aux usagers. Des outils issus de l'approximation stochastique sont utilisés pour modéliser le comportement des sources de contenu en fonction de la popularité des contenus postés sur les réseaux sociaux. Enfin, dans un travail récent [20], nous avons étudié un modèle stochastique de dynamique d'opinion dans les réseaux sociaux, et l'impact de la structure de communautés sur les solutions quasi-stationnaires de ces dynamiques. Nous avons également proposé des modèles d'inspiration biologique pour prédire l'évolution de la popularité de vidéos sur YouTube [19], et un modèle à base de marche aléatoire pour étudier le mécanisme de recommandation sur la même plate-forme web-sociale [15].

Analyse de graphes

Un certain nombre de travaux récents menés dans le cadre des axes *Optimisation*, *Systèmes complexes* et *Société numérique* concernent le développement d'outils pour l'analyse de graphes modélisant notamment des systèmes sociaux, au sens large. On peut distinguer ici deux types de travaux : ceux cherchant à caractériser des nœuds de façon individuelle, et ceux visant à partitionner le graphe pour identifier et exploiter sa structure de communautés.



Plusieurs travaux ponctuels ont été menés pour développer des mesures de centralité nodales. La mesure pour réseaux multiplexes (*i.e.*, à plusieurs couches) définie dans [18] est basée sur un modèle de diffusion d'opinion dans plusieurs médias sociaux simultanément. La généralisation de la mesure de rectitude proposée dans [8], qui s'applique aux réseaux spatiaux, est issue d'une collaboration avec les géographes de la FR Agorantic.

Une partie de notre activité actuelle porte sur la description de réseaux complexes à un niveau intermédiaire, *via* l'analyse de leur structure de communautés. Nous proposons pour cela une approche consistant à rechercher des motifs séquentiels fréquents dans des réseaux dynamiques à attributs (*i.e.*, dont les nœuds sont décrits de façon individuelle) [13], avec application à différents réseaux sociaux. Nous avons également défini une méthode non supervisée permettant de caractériser la position d'un nœud relativement à la structure de communautés, qui s'est révélée efficace pour détecter certains types d'utilisateurs abusifs sur Twitter [7].

Enfin, un travail en cours dans le cadre de la thèse de N. ARINIK porte sur le partitionnement de graphes signés (*i.e.*, chaque lien est décrit par poids, pouvant être négatif). Ce type de graphe permet de représenter des systèmes contenant des relations antagonistes. Nous explorons plusieurs variantes du concept d'*équilibre structurel*, en les adaptant à plusieurs cas d'application tels que la description du comportement des députés européens [1] ou la détection de corruption dans les marchés publics. Nous formulons ces tâches de partitionnement sous la forme de problèmes d'optimisation, et proposons des méthodes de résolution efficaces [10].

Références

- [1] Nejat Arinik, Rosa Figueiredo, and Vincent Labatut. Signed graph analysis for the interpretation of voting behavior. In *i-KNOW/SNANDIG*, 2017.
- [2] Nesrine Ben Khalifa, Rachid El-Azouzi, Yezekael Hayel, and Issam Mabrouki. Evolutionary games in interacting communities. *Dyn. Games Appl.*, 7(2) :131–156, 2016.
- [3] Xavier Bost, Vincent Labatut, Serigne Gueye, and Georges Linarès. Extraction and analysis of dynamic conversational networks from tv series. In *Social Network Based Big Data Analysis and Applications*, pages 55–84. 2018.
- [4] Jean-Valère Cossu, Killian Janod, Emmanuel Ferreira, Julien Gaillard, and Marc El Bèze. NLP-based classifiers to generalize experts assessments in e-reputation. In *CLEF*, pages 340–351, 2015.
- [5] Jean-Valère Cossu, Vincent Labatut, and Nicolas Dugué. A review of features for the discrimination of twitter users. *Soc. Netw. Anal. Min.*, 6(1) :25, 2016.
- [6] Jean-Valère Cossu, Eric Sanjuan, Juan-Manuel Torres-Moreno, and Marc El Bèze. Automatic classification and PLS-PM modeling for profiling reputation of corporate entities on twitter. In *NLDB*, volume 9103, pages 282–289, 2015.
- [7] Nicolas Dugué, Vincent Labatut, and Anthony Perez. A community role approach to assess social capitalists visibility in the twitter network. *Soc. Netw. Anal. Min.*, 5 :26, 2015.
- [8] Vincent Labatut. Continuous average straightness in spatial graphs. *J. Complex Netw.*, 6(2) :269–296, 2018.
- [9] Vincent Labatut and Guillaume Marrel. La visibilité politique en ligne : Contribution à



- la mesure de l'e-reputation politique d'un maire urbain. In *Conférence Big Data et visibilité en ligne*, 2017.
- [10] Mario Levorato, Lucia Drummond, Yuri Frota, and Rosa Figueiredo. An ILS algorithm to evaluate structural balance in signed social networks. In *SAC*, pages 1117–1122, 2015.
- [11] Mohamed Morchid, Richard Dufour, Pierre-Michel Bousquet, Georges Linarès, and Juan-Manuel Torres-Moreno. Feature selection using PCA for massive retweet detection. *Pattern Recogn. Lett.*, 49 :33–39, 2014.
- [12] Mohamed Morchid, Yonathan Portilla, Didier Josselin, Richard Dufour, Eitan Altman, Marc El Bèze, Jean-Valère Cossu, Georges Linarès, and Alexandre Reiffers. An author-topic based approach to cluster tweets and mine their location. In *Spatial Stat.*, volume 27, pages 26–29, 2015.
- [13] Günce Keziban Orman, Vincent Labatut, Marc Plantevit, and Jean-François Boulicaut. Interpreting communities based on the evolution of a dynamic attributed network. *Soc. Netw. Anal. Min.*, 5 :20, 2015.
- [14] Etienne Papegnies, Vincent Labatut, Richard Dufour, and Georges Linarès. Graph-based features for automatic online abuse detection. In *SLSP*, pages 70–81, 2017.
- [15] Yonathan Portilla, Alexandre Reiffers, Eitan Altman, and Rachid El-Azouzi. A study of YouTube recommendation graph based on measurements and stochastic tools. In *BDSN*, pages 430–435, 2015.
- [16] Mathias Quillot, Cassandre Ollivier, Richard Dufour, and Vincent Labatut. Exploring temporal analysis of tweet content from cultural events. In *SLSP*, pages 82–93, 2017.
- [17] Alexandre Reiffers, Yezekael Hayel, and Eitan Altman. Posting behaviour dynamics and active filtering for content diversity in social networks. *IEEE T. Signal Inf. Process. Netw.*, 3(2), 2017.
- [18] Alexandre Reiffers and Vincent Labatut. Opinion-based centrality in multiplex networks : A convex optimization approach. *Netw. Sci.*, 5(2) :213–234, 2017.
- [19] Cédric Richier, Eitan Altman, Rachid El-Azouzi, Tania Jimenez, Georges Linarès, and Yonathan Portilla. Bio-inspired models for characterizing youtube view-count. In *ASONAM*, pages 297–305, 2014.
- [20] Vineeth Varma, Yezekael Hayel, and Constantin Morarescu. Continuous time opinion dynamics of agents with multi-levelled opinions and binary actions. In *IEEE Infocom*, 2018.



AfIA
Association française
pour l'Intelligence Artificielle

Comptes rendus de journées, événements et conférences



■ 26èmes Journées Francophones sur les Systèmes Multi-Agents (JFSMA 2018)

Par

Gauthier PICARD

Laboratoire Hubert Curien
UMR CNRS 5516
Mines Saint-Étienne
Institut Henri Fayol
picard@emse.fr

Christophe LANG

FEMTO-ST / DEODIS
Université de Franche-Comté
christophe.lang@univ-fcomte.fr

Nicolas MARILLEAU

UMMISCO
IRD
nicolas.marilleau@ird.fr

La 26^e édition des « Journées Francophones sur les Systèmes Multi-Agents », présidées par Gauthier PICARD (Mines-Saint-Étienne) a été organisée à Métabief par Christophe LANG (Université de Franche-Comté) et Nicolas MARILLEAU (IRD), du 10 au 12 octobre 2018. Cet événement a reçu le soutien de l'Afia, de la région Bourgogne Franche-Comté, du laboratoire FEMTO-ST, de l'IRD et de l'université Bourgogne Franche-Comté.

Nourri de pluridisciplinarité, le paradigme multiagent fournit un cadre conceptuel pour l'étude et la conception de systèmes dont la dynamique globale est le fruit d'entités autonomes (les agents) qui interagissent dans un environnement commun. Depuis 1993, les Journées Francophones sur les Systèmes Multi-Agents (JFSMA) réunissent chaque année, trois jours durant, des chercheurs qui étudient, utilisent et font évoluer ce paradigme pour adresser des problématiques issues de domaines liés à l'informatique (intelligence et vie artificielle, génie logiciel, robotique collective, etc.) et aux sciences humaines et naturelles (écono-

mie, sociologie, éthologie, etc.). Les JFSMA constituent ainsi un rendez-vous scientifique privilégié, placé sous le signe de l'échange et de l'ouverture.

Programme Journée 1 (10/10/18)

- 9h30.** « Ouverture des journées », par Gauthier PICARD
- 10h00.** Session « simulation multiagents : trafic et modèles de comportements »
- 10h00.** « Situations accidentogènes dans les générateurs de trafic routier », par Philippe MATHIEU, Antoine NONGAILLARD et Benenice REFFET
- 10h40.** « Simulation du comportement latéral des véhicules fondée sur une approche à base de forces », par Julien SAUNIER
- 11h20.** « Modélisation du changement de voie de véhicules autonomes à différents niveaux d'abstraction », par Jérémy SOBIE-RAJ, Guillaume HUTZLER, Hanna KLAUDEL et Lydie NOUVELIÈRE
- 12h15.** Pause déjeuner
- 14h00.** Session « Résolution collective de problème et optimisation multiagents »
- 14h00.** « Planification de vidage d'images satellitaires par systèmes multiagents auto-adaptatifs », par Timothée JAMMOT, Elsy KADDOUM et Serge RAINJONNEAU
- 14h15.** « DeciMaxSum : Décimer pour résoudre des DCOP cycliques plus efficacement », par Jesus CERQUIDES, Rémi EMONET, Gauthier PICARD et Juan Antonio RODRIGUEZ AGUILAR
- 14h30.** « Optimisation sous contraintes distribuée : une introduction au domaine », par Gauthier PICARD
- 15h00.** Pause café
- 15h30.** Session « Négociation multiagents »



AfIA

Association française
pour l'Intelligence Artificielle

15h30. « Négociation bilatérale par concession : un état de l'art », par Maxime MORGE

16h10. « MoCaNA, un agent de négociation automatique utilisant la recherche arborescente de Monte-Carlo », par Cedric BURON, Sylvain DUCTOR, Zahia GUES-SOUM et Olivier ROUSSEL

16h50. « Négociation multiagent one-to-many et mécanismes de coordination pour la gestion de la satisfaction des utilisateurs d'un service », par Amro NAJJAR, Gauthier PICARD, Yazan MUALLA et Kamal SINGH

17h00. Réunion de la communauté

18h30. Fin de la journée

Programme Journée 2 (11/10/18)

9h00. « Protocoles de population optimaux en espace », par Janna BURMAN (Université Paris-Sud), conférencière invitée

10h00. Pause café

10h30. Session « Interactions et coopération »

10h30. « Une approche passive pour réduire les coûts de traitement des interactions entre agents », par Nicolas DIOT, Fabrice BOUQUET et Christophe LANG

10h45. « Un modèle de norme intégrant les conditions spatio-temporelles », par Sitraka Oliva RAHARIVÉLO et Jean-Pierre MÜLLER

11h25. « Gestion d'intentions multiples pour agents ambiants coopératifs », par Arthur CASALS, Assia BELBACHIR, Amal EL FALLAH SEGHROUCHNI et Anarosa Alves Franco BRANDÃO

11h40. « Modélisations de l'interaction pollinisateurs/palmiers à l'aide d'approches déterministes et multiagents. Application sur l'estimation du taux de nouaison », par Yves DUMONT, Jean-Christophe SOULIÉ et Fabien MICHEL

12h00. Pause déjeuner

13h00. Session « Auto-organisation et ap-

proches ascendantes »

13h00. « Recrutement et auto-organisation : Vers un modèle multiagent complet d'une colonie d'abeilles », par Jeremy RIVIERE, Cedric ALAUX, Yves LE CONTE, Yves LAYEC, Andre LOZAC'H, Vincent RODIN et Frank SINGHOFF

13h40. « Découverte d'équilibres dynamiques globaux par coopération locale et sans fonction d'évaluation », par Sébastien Maignan et Carole BERNON

14h20. « Modèle et performance d'une gestion locale autonome des conflits dans le trafic aérien », par Augustin DEGAS, Elsy KADDOUM, Françoise ADREIT, Marie-Pierre GLEIZES et Arcady RANTRUA

14h45. Événement social

17h00. Session « Démonstrations »

18h30. Fin de la journée

Programme Journée 3 (12/10/18)

9h00. « Enabling decentralised decision-making in AI – an agent environments perspective », par Kostas STATHIS (Royal Holloway, University of London), conférencier invité

10h00. Pause café

10h30. Session « Simulation multiagents : couplage et multi-niveaux »

10h30. « Modélisation multiagents pour la gestion individuelle et collective d'une maladie infectieuse », par Stéphane KREBS, Sébastien PICAULT et Pauline EZANNO

10h45. « Modèle de couplage multiagent pour simulations individu centrées », par Khadim NDIAYE, Flavien BALBO, Jean-Paul JAMONT et Michel OCCELLO

11h25. « Co-simulation à base d'outils multiagents : un cas d'étude avec NetLogo », par Thomas PARIS, Laurent CIARLETTA et Vincent CHEVRIER

11h40. « Couplage de simulations multiagents pour la conception de politiques urbaines »,



par Simon PAGEAUD, Veronique DES-
LANDRES, Vassilissa LEHOUX et Salima
HASSAS

12h00. Clôture des journées

Résumé des interventions

Cette année le comité de programme a reçu 52 soumissions parmi lesquelles ont été sélectionnées 11 présentations longues, 9 présentations courtes, et 6 démonstrations, pour nous proposer un programme scientifique riche et varié. Les verrous affrontés dans plusieurs de ces articles sont ceux de la complexité, du passage à l'échelle, et de la dynamique de l'environnement. Les modèles et méthodes qui sont proposés, plaçant au centre les notions de distribution et de décentralisation des décisions des agents, abordent de manière renouvelée des techniques de simulation individu-centrée, de résolution collective de problèmes, de négociation multiagents, d'interaction, et d'auto-organisation.

L'excellent travail sélectionné par le comité de programme (« MoCaNA, un agent de négociation automatique utilisant la recherche arborescente de Monte-Carlo » par Cedric BURON, Sylvain DUCTOR, Zahia GUESSOUM et Olivier ROUSSEL) recevra le prix du meilleur article des journées financé par l'AfIA. Les

meilleurs articles sélectionnés par le comité de programme feront également l'objet d'un numéro spécial de la Revue d'Intelligence Artificielle (<https://ria.revuesonline.com/>).

Conclusion

Cette année est également marquée par l'émergence ou la confirmation de champs d'applications pour lesquels les apports des systèmes multiagents sont notables, comme l'Internet des objets, le trafic routier, les véhicules autonomes, ou bien la simulation de systèmes complexes. Ces domaines applicatifs sont autant d'arguments en faveur d'une approche multiagents de la conception et de contrôle de systèmes où la décentralisation des décisions est un pré-requis ou une exigence. La présence de démonstrations lors de ces journées témoignent également de la maturité technologique des contributions francophones.

Les actes de ces journées ont été publiés chez Cépaduès [1].

Références

- [1] G. Picard, C. Lang, and N. Marilleau, editors. *Journées Francophones sur les Systèmes Multi-Agents (JFSMA'18) - Distribution et décentralisation*. Cépaduès, 2018.

■ 5^e journée Perspectives et Défis de l'IA (PDIA 2018)

Par **Olivier SIMONIN**
CITI / Chroma
INSA Lyon, Inria
olivier.simonin@insa-lyon.fr

L'AfIA a organisé sa cinquième journée PDIA (Perspectives et Défis de l'IA) sur le thème « véhicule autonome et intelligence artificielle » le 11 octobre 2018, à l'Université Paris-Descartes.

Résumé de la thématique

On parle beaucoup de véhicule autonome, sans toujours faire la connexion avec l'IA. On parle encore plus d'IA, mais sans forcément



penser à toutes les possibilités pour les véhicules autonomes. Les défis sont nombreux, parmi lesquels la sûreté et la validation des algorithmes embarqués, la perception et la localisation, l'interaction avec l'humain, la prise de décision temps réel, la navigation en trafic dense, etc. Ainsi, la journée PDIA 2018 visait à dresser un panorama de ces enjeux au travers d'exposés et de tables rondes.

Participation

La journée a attiré plus de 75 personnes (inscrites). Huit présentations invitées étaient proposées, ainsi que deux tables rondes. La journée a débuté par un mot de bienvenue de F. DARDEL, Président de l'Université Paris-Descartes, qui nous accueillait gracieusement dans l'amphithéâtre Fourier.

La journée a donné lieu à de nombreux échanges entre le public et les intervenants, mais également lors des tables rondes. Nous avons recueillis lors de la journée, et après, de nombreux messages de remerciements pour la qualité des exposés et des échanges.

Conférenciers invités et tables rondes

9h00. « L'IA au service du transport autonome : défis, limites et opportunités », par Patrice AKNIN, (IRT SystemX).

L'IA recouvre une myriade de domaines et technologies dont les niveaux de pénétration dans l'industrie sont contrastés. Nous exposons les avancées et promesses de l'IA pour le transport autonome dont chacun perçoit le potentiel de rupture, surtout si couplée à d'autres évolutions des modèles de mobilité, électrique, connectée ou encore partagée.

9h30. « Systèmes multi-agents pour la simulation et la modélisation du trafic », par Philippe MATHIEU, (Université Lille).

Avec l'avènement du véhicule autonome et la mutation que va connaître le secteur automobile, la simulation de trafic routier a

pris un nouvel essor. Nous montrons comment une approche à base d'agents peut être utilisée pour modéliser un trafic routier ferroviaire ou fluvial, et comment elle peut servir de base au test du véhicule autonome.

10h30. « Le défi de la validation en sécurité de l'IA pour la perception et la décision en conduite autonome de train », par Philippe DAVID, (SNCF).

Après une présentation de notre approche de la détection, nous montrons la nécessité d'une intelligence artificielle décisionnelle dont le niveau de sécurité est démontrable. Au-delà des exigences juridiques, nous montrons la nécessaire diminution de la sensibilité structurelle de l'architecture du système de décision face aux exigences non clairement définies de sécurité.

11h00. « Validation et vérification du logiciel pour les systèmes autonomes », par Félix Ingrand, (CNRS, Toulouse).

L'application des techniques de validation et de la vérification (V&V) aux systèmes autonomes, qui recourent à des fonctions décisionnelles, pose de nouveaux challenges, tant pour la V&V des fonctions de base et de leur intégration, que pour des fonctions décisionnelles où l'on distinguera les approches basées sur des modèles explicites vs appris

11h30. « Table ronde Validation/vérification dans les différents secteurs », par Philippe DAVID (SNCF), Félix INGRAND (LAAS), Alai PIPERNO (UTAC-CERAM) et animée par Philippe MATHIEU.

14h00. « Validation des systèmes de conduite autonomes », par Xavier DELACHE, (Sous-directeur des études et de la prospective, DGITM).

14h30. « Les aspects éthiques du déploiement du véhicule autonome », par Ebru BURCU DOGAN, (VEDECOM).

Les aspects éthiques du véhicule autonome



(VA) peuvent avoir un impact disproportionné sur leur acceptabilité par le grand public, et engendrer le scepticisme. Il ne faut néanmoins pas réduire les questions éthiques aux situations de dilemmes. Nous exposons les conséquences du déploiement du VA et des nouvelles mobilités pour la société en général.

15h00. « IA intégrée / à intégrer au sein des systèmes robotiques militaires pour l'autonomie », par Bruno STEUX, Jean-François MASSOL, Bruno RICAUD (Nexter Systems).

Nous présentons la mobilité autonome et des briques d'intelligence artificielle intégrées dans les produits robotiques développés chez Nexter. Nous ouvrons notre discours sur la mutation des conflits et nous exposons les implications sur les besoins technologiques futurs.

15h30. « Navigation sociale des véhicules autonomes », par Anne SPALANZANI, (Uni-

versité Grenoble Alpes, INRIA).

Le problème central est de garantir la sécurité de tous les véhicules autonomes qui se déplacent dans un espace dynamique peuplé d'humains, et de s'assurer de l'acceptabilité des usagers. Nous proposons des stratégies intégrant perception et prédiction de l'état de l'environnement mais aussi conventions sociales et anticipation des intentions des humains.

16h30. « Table ronde Plates-formes d'intégration et relations avec l'humain », par Ebru BURCU DOGAN (VEDECOM), Jérôme LEMAIRE (DGA) et Olivier SIMONIN (INSA Lyon), et animée par Yves DEMAZEAU.

Organisation

La journée était organisée conjointement par Yves DEMAZEAU et Olivier SIMONIN, avec le soutien du bureau de l'AfIA.

■ Nuit de l'Info

Par

Florence BANNAY
IRIT / Équipe ADRIA
Université Toulouse 3
bannay@irit.fr

La Nuit de l'Info 2018 a eu lieu de 16h39 à 8h05 dans la nuit du 6 au 7 décembre sur le thème de *L'assistance numérique d'un explorateur en conditions extrêmes*. L'AfIA a proposé pour la 4^e année un défi *Mettez de l'IA dans votre moteur*. Le principe de la Nuit de l'Info est très simple : Le jeudi 6, au coucher du Soleil, 16h39, en séance plénière : les organisateurs remettent un sujet (le même pour toute la France) aux participants. Les étudiants s'organisent en groupes (sur un ou plusieurs sites) : ils développent un projet (informatique, marketing, rédactionnel, etc.) tout en ciblant un ou

plusieurs défis. Le vendredi 7, au lever du Soleil, 08h05, les développements sont figés. Le vendredi matin, pendant que les étudiants dorment des jurys se réunissent (un jury par défi) et examinent les travaux réalisés par les différentes équipes durant la nuit.

La compétition

Le défi qu'a proposé l'AfIA s'intitulait *Mettez de l'intelligence dans votre moteur* : « Vous mettez en œuvre une ou plusieurs méthodes d'intelligence artificielle (IA) dans votre projet et vous indiquerez en quoi ces méthodes rendent votre réalisation plus performante ou pertinente. Vous pouvez faire appel à des techniques classiques, en cours de développement ou futuristes. L'équipe qui aura mis le plus en



avant les avantages de l'utilisation de l'IA dans son projet remportera ce défi ». Les réponses devaient inclure :

- une description des problématiques IA rencontrées dans le projet ;
- une explication de l'intégration de l'IA dans l'application (description de ce qui a été réalisé qui relève de l'IA ou de ce qui aurait pu être fait avec des outils d'IA) ;
- une analyse des avantages et inconvénients de cette intégration.

La récompense est un montant de 400€ assorti d'invitations à présenter une démonstration lors de la compétition organisée par l'AfIA au sein de la Plate-Forme IA 2019 à Toulouse.

Les équipes

Parmi les douze équipes inscrites pour relever ce défi, cinq ont réussi à remettre à temps un projet : GIFCOMS, ISITCOM-Shadow-Brokers2, ISIT'COM-Techno-Knights, Lord of the Ping et MLG.

Jury

Le jury était composé de Florence BAN-NAY (coordinatrice), Anne-Gwenn BOSSER, Philippe MORIGNOT et Sylvie SAGET. Les projets ont été jugés sur cinq critères : fonctionnement (est-ce que la proposition peut être mise en œuvre?), innovation, qualité de l'IA, spectre IA couvert, qualité des explications.

Les résultats

Nous avons eu du mal à départager les deux meilleures équipes : l'équipe Lord of the Ping a finalement été déclarée vainqueur pour le recul de ses explications et recevra un prix de

200€. Néanmoins vu la qualité des résultats de l'équipe MLG, elle recevra la même somme de 200€. Le classement final est le suivant :

1. Lord of the Ping a développé un chatbot sarcastique avec détection de serpent par apprentissage. L'équipe est allée jusqu'au bout de deux implémentations utilisant de l'IA. Elle a bien expliqué les problèmes et les limites des approches utilisées.
2. MLG a construit un site web permettant de saisir des données, de visualiser les courbes correspondantes et de donner les positions des robots. Le jury a apprécié les résultats présentés au travers de vidéos qui montrent deux applications (robots de rangement et géolocalisation) mais aurait aimé une explication globale.
3. ISITCOM_Techno_Knights propose de développer une montre connectée avec des capteurs de température (peau, air), rythme cardiaque, respiration. L'idée novatrice est de se servir des données captées pour détecter l'état psychologique par apprentissage afin d'envoyer une alerte adaptée. Le jury a jugé l'idée intéressante et réalisable mais déplore l'absence des détails techniques IA et concrets d'implémentation.
4. ISITCOM_Shadow_Brokers2 et
5. GIFCOMS : le premier propose un système de capteurs permettant de détecter les dangers de nuit, le deuxième présente une plateforme intelligente de survie basée sur l'*Internet of thing*. Le point de vue IA est peu développé dans les deux projets qu'ont soumis ces équipes.

Merci à tous les participants pour le travail fourni et rendez-vous l'année prochaine pour un nouveau défi de l'AfIA !



Afia
Association française
pour l'Intelligence Artificielle

Thèses et HDR du trimestre

Si vous êtes au courant de la programmation de soutenances de thèses ou HDR en Intelligence Artificielle cette année, vous pouvez nous les signaler en écrivant à redacteur@afia.asso.fr.



■ Thèses de Doctorat

Mohammed Yacine TALEB

« [Optimizing Distributed In-memory Storage Systems Fault-tolerance, Performance, Energy Efficiency](#) »

Supervision : *Gabriel ANTONIU*
Toni CORTES

Le 02/10/2018, à l'Université de Rennes,
Ecole normale supérieure

Joseph EHOUNOU

« [Algorithmes de graphes pour la découverte de la topologie d'un réseau énergétique par la connaissance de ses flots](#) »

Supervision : *Dominique BARTH*

Le 02/10/2018, à l'Université de Paris Saclay

Aline HUFSCMITT

« [Decomposition des jeux dans le domaine du General Game Playing.](#) »

Supervision : *Nicolas JOUANDEAU*
Jean noel VITTAUT

Le 04/10/2018, à l'Université de Paris 8

Juan carlos Munoz FERNANDEZ

« [A generic and extensible requirements engineering framework for self adaptative software systems.](#) »

Supervision : *Camille SALINESI*
Raul Mazo PENA
Gabriel TAMURA

Le 08/10/2018, à l'Université de Paris 1

Gabrielle PARIS

« [Resolution of some optimisation problems on graphs and combinatorial games](#) »

Supervision : *Eric DUCHENE*
Paul DORBEC

Le 09/10/2018, à l'Université de Lyon

Elinirina Irena ROBINSON

« [Filtering and uncertainty propagation methods for model-based prognosis](#) »

Supervision : *Tarek RAISSI*

Le 10/10/2018, au CNAM, Paris

Mikael MONET

« [Combined complexity of probabilistic query evaluation](#) »

Supervision : *Pierre SENELLART*

Le 12/10/2018, à l'Université de Paris Saclay

Damien LIGIER

« [Functional encryption applied to privacy-preserving classification : practical use, performances and security](#) »

Supervision : *Caroline FONTAINE*

Le 15/10/2018, à l'École nationale supérieure Mines-Telecom Atlantique Bretagne Pays de la Loire

Clement GAUTRAIS

« [Signatures : detecting and characterizing complex recurrent behavior in sequential data](#) »

Supervision : *Alexandre TERMIER*

Le 16/10/2018, à l'Université de Rennes 1

Nicolas AUDEBERT

« [Classification de données massives de télédétection](#) »

Supervision : *Sebastien LEFEVRE*

Bertrand Honore Henri Le SAUX

Le 17/10/2018, à l'Université de Lorient

Quentin DELMEE

« [Resolution exacte de problèmes de localisation de services bi-objectifs en variables mixtes](#) »

Supervision : *Xavier GANDIBLEUX*

Anthony PRZYBYLSKI

Le 19/10/2018, à l'Université de Nantes



AfIA

Association française
pour l'Intelligence Artificielle

Anil GOYAL

« Un cadre PAC-Bayésien pour l'apprentissage multivues »

Supervision : *Massih reza AMINI*
Emilie MORVANT

Le 23/10/2018, à l'Université de Lyon

Vinicius Ferraris pignataro MAZZEI

« Detection de changement par fusion d'images de teledetection de resolutions et modalites differentes »

Supervision : *Marie CHABERT*
Nicolas DOBIGEON

Le 26/10/2018, à l'INP, Toulouse

Jessie CARBONNEL

« L'analyse formelle de concepts : un cadre structurel pour l'etude de la variabilite de familles de logiciels »

Supervision : *Marianne HUCHARD*

Le 29/10/2018, à l'Université de Montpellier

Amina ANNANE

« Utilisation des ressources de connaissances externes pour ameliorer l'alignement d'ontologies biomédicales »

Supervision : *Zohra BELLAHSENE*
Faical AZOUAOU

Le 29/10/2018, à l'Université de Montpellier

Khaled MAAFA

« Jeux et treillis, aspects algorithmiques »

Supervision : *Lhouari NOURINE*
Mohammed said RADJEF

Le 29/10/2018, à l'Université de Clermont Auvergne

Yann BARSAMIAN

« Pic-Vert : une implementation de la methode particulaire pour architectures multi-coeurs »

Supervision : *Eric VIOLARD*

Le 31/10/2018, à l'Université de Strasbourg

Mahieddine DELLABANI

« Methodes formelles pour les systemes distribues temps-reel »

Supervision : *Saddek BENSALÉM*

Le 31/10/2018, à l'Université de Grenoble Alpes

Thomas MEYSSONNIER

« Vers des systemes et outils de notation et de composition pour la musique electroacoustique »

Supervision : *Myriam Desainte CATHERINE*

Le 02/11/2018, à l'Université de Bordeaux

Romain RINCE

« Behavior Recognition on Noisy Data-streams Constrained by Complex Prior Knowledge »

Supervision : *Philippe LERAY*

Romain KERVARC

Le 07/11/2018, à l'Université de Nantes

Gabin PERSONENI

« Apport des ontologies de domaine pour l'extraction de connaissances a partir de donnees biomédicales Contribution of domain ontologies for knowledge discovery in biomedical data »

Supervision : *Marie dominique DEVIGNES*
Adrien COULET

Le 09/11/2018, à l'Université de Lorraine

Laurent LUCIEN

« Contribution a une modelisation globale de la collaboration dans les systemes multi-agents : Application aux entites mobiles intelligentes »

Supervision : *Laurent PHILIPPE*

Christophe LANG

Nicolas MARILLEAU

Le 09/11/2018, à l'Université de Bourgogne Franche-Comte



Afia

Association française
pour l'Intelligence Artificielle

Yann HODE

« Contribution a l'interpretation d'images et verification de la consistance d'un graphe »

Supervision : *Aline DERUYVER*

Le 12/11/2018, à l'Université de Strasbourg

Anne MORVAN

« Contributions to unsupervised learning from massive high-dimensional data streams : structuring, hashing and clustering »

Supervision : *Jamal ATIF*

Le 12/11/2018, à l'Université de Paris Sciences et Lettres

Lydia Ould OUALI

« Modele de negociation collaborative base sur la relation interpersonnelle de dominance »

Supervision : *Nicolas SABOURET*

Le 12/11/2018, à l'Université de Paris Saclay

Firas abdulmajeed AI QASEER

« Politiques de planification tenant compte de criteres de duree de production et de consommation d'energie pour la gestion environnementale. interpretation de donnees objectives et subjectives pour le management environnemental dans les systemes industriels »

Supervision : *Denis GIEN*

Le 15/11/2018, à l'Université de Clermont Auvergne

Olivier SANS

« Methodes a intervalles et strategies de parcours d'arbre pour l'optimisation globale »

Supervision : *Gilles TROMBETTONI*

Le 19/11/2018, à l'Université de Montpellier

Fadoua CHAKCHOUK

« Contribution a la robustesse dans les CSPs Distribues par replication locale »

Supervision : *Rene MANDIAU*

Le 19/11/2018, à l'Université de Valenciennes

Mounir CHADLI

« Analyse des systemes temps-reel de point de vue ordonnancement »

Supervision : *Axel LEGAY*

Le 21/11/2018, à l'Université de Rennes 1

Guillaume GADEK

« Detection d'opinions, d'acteurs-cles et de communautes thematiques dans les medias sociaux »

Supervision : *Alexandre PAUCHET*

Laurent VERCOUTER

Nicolas MALANDAIN

Le 22/11/2018, à l'Université de Normandie

Ahmad MAZYAD

« Optimisation de profils numeriques a l'aide d'algorithmes a essai »

Supervision : *Cyril FONLUPT*

Le 22/11/2018, à l'Université de Littoral

Boukaye boubacar TRAORE

« Modelisation des informations et extraction des connaissances pour la gestion des crises »

Supervision : *Bernard Kamsu FOGUEM*

Fana TANGARA

Le 23/11/2018, à l'INP, Toulouse

Wissam SIBLINI

« Apprentissage multi label interactif : comparaisons d'approches et nouvelles propositions »

Supervision : *Pascale Kuntz COSPEREC*

Le 23/11/2018, à l'Université de Nantes



Afia

Association française
pour l'Intelligence Artificielle

Qing XU

« Modeliser les changements mineurs et majeurs d'individus en interactions : application a la conversion a l'agriculture biologique »

Supervision : *Guillaume DEFFUANT*

Le 26/11/2018, à l'Université de Clermont Auvergne

Tarek BENKHELIF

« Publication de donnees personnelles respectueuse de la vie privee. : une demarche fondee sur le co-clustering. »

Supervision : *Marc GELGON*

Guillaume RASCHIA

Le 27/11/2018, à l'Université de Nantes

Liyang XIAO

« Contributions a la planification et a l'optimisation dans le systeme de sante moderne »

Supervision : *Amir Hajjam el HASSANI*

Mahjoub DRIDI

Le 27/11/2018, à l'Université de Bourgogne Franche-Comte

Maxime BUCHER

« Apprentissage et exploitation de representations semantiques pour la classification et la recherche d'images »

Supervision : *Frederic JURIE*

Le 27/11/2018, à l'Université de Normandie

Lynda KHALI

« Fouille de donnees a partir de series temporelles dimages satellites »

Supervision : *Maguelonne TEISSEIRE*

Dino IENCO

Le 28/11/2018, à l'Université de Montpellier

El mehdi KHALFI

« Les identites au centre de la mise en uvre de comportements dans le cadre de collectifs multi-agents : application au Web des Objets »

Supervision : *Jean Paul JAMONT*

Michel OCCELLO

Le 28/11/2018, à l'Université de Grenoble Alpes

Abdeldjalil RAMOUL

« Systeme de planification en mixed-initiative pour l'assistance a la gestion des systemes informatises complexes »

Supervision : *Sylvie PESTY*

Damien PELLIER

Le 28/11/2018, à l'Université de Grenoble Alpes

Yacine ABOUD

« Fouille de motifs : entre accessibilite et robustesse »

Supervision : *Anne BOYER*

Armelle BRUN

Le 28/11/2018, à l'Université de Lorraine

Wen SUN

« Algorithmes heuristiques pour des problemes de coloration de graphes »

Supervision : *Jin Kao HAO*

Alexandre CAMINADA

Le 29/11/2018, à l'Université de Angers

Gregoire PICHON

« On the use of low-rank arithmetic to reduce the complexity of parallel sparse linear solvers based on direct factorization techniques »

Supervision : *Pierre RAMET*

Mathieu FAVERGE

Le 29/11/2018, à l'Université de Bordeaux



Afia

Association française
pour l'Intelligence Artificielle

Rachel BAWDEN

« [Going beyond the sentence : Contextual Machine Translation of Dialogue](#) »

Supervision : *Sophie ROSSET*

Le 29/11/2018, à l'Université de Paris Saclay

Lucas BOCZKOWSKI

« [SEARCH AND BROADCAST IN A STOCHASTIC ENVIRONMENT, A BIOLOGICAL PERSPECTIVE.](#) »

Supervision : *Amos KORMAN*

Iordanis KERENIDIS

Le 30/11/2018, à l'Université de Sorbonne Paris Cite

Vera SHALAEVA

« [Arbre de decision temporel multi-operateur](#) »

Supervision : *Ahlame DOUZAL*

Gilles BISSON

Le 30/11/2018, à l'Université de Grenoble Alpes

Jieying CHEN

« [Extraction de connaissances a partir de terminologies en logique de description](#) »

Supervision : *Philippe DAGUE*

Le 30/11/2018, à l'Université de Paris Saclay

Nizar MHADHBI

« [Fouilles de graphes et application aux reseaux sociaux](#) »

Supervision : *Lakhdar SAIS*

Le 03/12/2018, à l'Université de Artois

Belen Baez MIRANDA

« [Generation de recits a partir de donnees ambiantes](#) »

Supervision : *Francois PORTET*

Catherine GARBAY

Le 03/12/2018, à l'Université de Grenoble Alpes

Julien FRADIN

« [Graphes complexes en biologie : problemes, algorithmes et evaluations](#) »

Supervision : *Guillaume FERTIN*

Le 04/12/2018, à l'Université de Nantes

Khaled BELAHCENE

« [Explications pour l'agregation des preferences - une contribution a l'aide a la decision responsable](#) »

Supervision : *Vincent MOUSSEAU*

Le 05/12/2018, à l'Université de Paris Saclay

Nadira Boudjani NADIRA

« [Aide a la construction et l'evaluation des preuves mathematiques deductives par les systemes d'argumentation](#) »

Supervision : *Souhila KACI*

Le 05/12/2018, à l'Université de Montpellier

Alexis CORNET

« [Algorithmes et resultats de complexite pour des problemes de graphes avec contraintes additionnelles.](#) »

Supervision : *Christian LAFOREST*

Le 05/12/2018, à l'Université de Clermont Auvergne

Indira lisa MEDROUK

« [Reseaux profonds pour la classification des opinions multilingue.](#) »

Supervision : *Jacqueline SIGNORINI*

Anna Pappa DELBANO

Le 06/12/2018, à l'Université de Paris 8

Jeremy DECERLE

« [Contribution a l'optimisation de la planification des tournées de soins pour l'hospitalisation a domicile](#) »

Supervision : *Olivier GRUNDER*

Amir Hajjam el HASSANI

Oussama BARAKAT

Le 06/12/2018, à l'Université de Bourgogne Franche-Comte



Celia CHATEL

« [Modeles de classification en classes empietantes, cas des modeles arbores](#) »

Supervision : *Francois BRUCKER*
Pascal PREA

Le 07/12/2018, à l'Université de Aix-Marseille

Silvio Domingos CARDOSO

« [MAISA - Maintenance of semantic annotations](#) »

Supervision : *Chantal REYNAUD*

Le 07/12/2018, à l'Université de Paris Saclay

Hong Phong PHAM

« [Studies on Optimal Colorful Structures in Vertex-Colored Graphs](#) »

Supervision : *Yannis MANOUSSAKIS*

Le 07/12/2018, à l'Université de Paris Saclay

Sebastien DELECRAZ

« [Approches jointes texte/image pour la comprehension multimodale de documents](#) »

Supervision : *Frederic BECHET*
Benoit FAVRE

Le 10/12/2018, à l'Université de Aix-Marseille

Feda ALMUHISEN

« [Exploitation de l'analyse formelle de concepts et de l'extraction de motifs pour l'analyse de trajectoires d'objets mobiles](#) »

Supervision : *Mohamed QUAFAROU*
Nicolas DURAND

Le 10/12/2018, à l'Université de Aix-Marseille

William SCHUELLER

« [Contrôle actif de la croissance de la complexité dans les Language Games](#) »

Supervision : *Pierre Yves OUDEYER*

Le 10/12/2018, à l'Université de Bordeaux

Vanel steve Siyou FOTSO

« [Extraction de connaissances de series temporelles cycliques et incertaines : application a l'analyse de la locomotion en Fauteuil Roulant Manuel](#) »

Supervision : *Engelbert Mephu NGUIFO*

Le 11/12/2018, à l'Université de Clermont Auvergne

Florent BONNIER

« [Algorithmes paralleles pour le suivi de particules](#) »

Supervision : *Nahid Emad PETITON*
Xavier JUVIGNY

Le 12/12/2018, à l'Université de Paris Saclay

Jose Luis Vilchis MEDINA

« [MODELISATION DES SYSTEMES RESILIENTS EN LOGIQUE NON MONOTONE APPLICATION A DRONE SOLAIRE](#) »

Supervision : *Pierre SIEGEL*
Andrei DONCESCU

Le 12/12/2018, à l'Université de Aix-Marseille

Adeline GRANET

« [Extraction d'information dans des documents manuscrits anciens](#) »

Supervision : *Emmanuel MORIN*
Harold MOUCHERE
Solen QUINIOU

Le 12/12/2018, à l'Université de Nantes

Yves MOCQUARD

« [Analyse probabiliste de protocoles de population](#) »

Supervision : *Bruno SERICOLA*

Le 13/12/2018, à l'Université de Rennes 1



Afia

Association française
pour l'Intelligence Artificielle

Jean Charles VIALATTE

« On Convolution of Graph Signals and Deep Learning on Graph Domains »

Supervision : *Gilles COPPIN*

Le 13/12/2018, à l'École nationale supérieure Mines-Telecom Atlantique Bretagne Pays de la Loire

Clement DELGRANGE

« Apprentissage base sur l'usage en interaction humaine avec un assistant adaptatif »

Supervision : *Peter DOMINEY*

Le 13/12/2018, à l'Université de Lyon

Konstantinos CHATZILYGEROUDIS

« Apprentissage micro-data pour l'adaptation en robotique Micro-Data Reinforcement Learning for Adaptive Robots »

Supervision : *Jean baptiste MOURET*

Le 14/12/2018, à l'Université de Lorraine

Elena EPURE

« Automatically Modeling Conversations as Processes of Interrelated Speech Intentions »

Supervision : *Camille SALINESI*

Le 14/12/2018, à l'Université de Paris 1

Simon MOURA

« Apprentissage multi-cibles : theorie et applications »

Supervision : *Massih Reza AMINI*

Le 17/12/2018, à l'Université de Grenoble Alpes

Oriane DERMY

« Prediction du mouvement humain pour la robotique collaborative: du geste accompagne au mouvement corps entier. Prediction of human movement for collaborative robotics: from accompanied gesture to whole-body movement. »

Supervision : *Francois CHARPILLET*

Serena IVALDI

Le 17/12/2018, à l'Université de Lorraine

Guilhem CHERON

« Modelisation structuree et reconnaissance des actions humaines dans les videos »

Supervision : *Ivan LAPTEV*

Cordelia SCHMID

Le 17/12/2018, à l'Université de Paris Sciences et Lettres

Ye XIA

« Combinaison d'heuristiques pour optimiser et dimensionner le placement d'applications IoT dans le Fog »

Supervision : *Frederic DESPREZ*

Le 17/12/2018, à l'Université de Grenoble Alpes

Valentina ZANTEDESCHI

« A Unified View of Local Learning: Theory and Algorithms for Enhancing Linear Models »

Supervision : *Marc SEBBAN*

Le 18/12/2018, à l'Université de Lyon

Marwa Hadj salah ELLOUMI

« Desambiguisation lexicale de l'arabe pour et par la traduction automatique »

Supervision : *Herve BLANCHON*

Mounir ZRIGUI

Le 18/12/2018, à l'Université de Grenoble Alpes

Florian GOLEMO

« comment dresser votre robot - nouveaux environnements de formation robotique et nouvelles methodes de transfert de strategies du simulateur au vrai robot »

Supervision : *Pierre Yves OUDEYER*

Le 19/12/2018, à l'Université de Bordeaux

Nicolas Galvez RAMIREZ

« Un cadre pour la generation autonome de strategies dans la satisfiabilite modulo des theories »

Supervision : *Frederic SAUBION*

Eric MONFROY

Le 19/12/2018, à l'Université de Angers



Afia

Association française
pour l'Intelligence Artificielle

Mohammad Noorani BAKERALLY

« Technologies du Web sémantique pour la gestion de données liées temps réel dans les villes intelligentes. »

Supervision : *Olivier BOISSIER*

Le 20/12/2018, à l'Université de Saint-Etienne, EMSE

Giulia De SANTIS

« Modeling and recognizing network scanning activities with finite mixture models and hidden markov models »

Supervision : *Olivier FESTOR*

Abdelkader LAHMADI

Le 20/12/2018, à l'Université de Lorraine

Noura HERRADI

« Representation sémantique multilingue, multiculturelle et temporelle des relations interpersonnelles, appliquée à une prothèse de mémoire »

Supervision : *Elisabeth METAIS*

Le 20/12/2018, au CNAM, Paris

Jean Ogier du TERRAIL

« Réseaux de Neurones Convolutionnels Profonds pour la détection de petits véhicules en Imagerie Aérienne. »

Supervision : *Frederic JURIE*

Le 20/12/2018, à l'Université de Normandie

Nathalie CHARBEL

« Representation sémantique de corpus de documents hétérogènes pour une nouvelle stratégie de recherche d'information : application au domaine du Bâtiment »

Supervision : *Christian SALLABERRY*

Le 21/12/2018, à l'Université de Pau

■ Habilitations à Diriger les Recherches

Nous n'avons pas reçu de notification d'HDR pour ce trimestre. Si vous êtes au courant d'une soutenance, merci de nous avvertir.



À PROPOS DE L'AfIA

L'objet de l'AfIA, Association Loi 1901 sans but lucratif, est de promouvoir et de favoriser le développement de l'Intelligence Artificielle (IA) sous ses différentes formes, de regrouper et de faire croître la communauté française en IA, et, à la hauteur des forces de ses membres, d'en assurer la visibilité.

L'AfIA anime la communauté par l'organisation de grands rendez-vous annuels. L'AfIA organise ainsi chaque année une semaine de l'IA, la Plate-forme IA (PFIA 2017 Caen, PFIA 2018 Nancy) au sein de laquelle se tiennent la Conférence Nationale d'Intelligence Artificielle (CNIA), les Rencontres des Jeunes Chercheurs en IA (CNIA) et la Conférence sur les Applications Pratiques de l'IA (APIA) ainsi que des conférences thématiques qui peuvent évoluer d'une année à l'autre sans récurrence obligée.

C'est ainsi que du 2 au 6 juillet 2018 à Nancy, lors de son édition 2018, la Plate-Forme IA hébergea, outre les 21^e CNIA, 16^e CNIA et 4^e APIA, les 29^{es} IC et les 13^{es} JFPDA. Y étaient également programmées une compétition d'intégration verticale DriveTo-Gaether ainsi que six journées communes organisées avec d'autres associations ou institutions.

Forte du soutien de ses 514 adhérents à jour de leur cotisation, l'AfIA assure :

- Le maintien d'un [Site Web](#) dédié à l'IA, reproduisant également les Brèves de l'IA,
- Une journée recherche annuelle sur les Perspectives et Défis en IA (PDIA 2017),
- Une journée industrielle annuelle ou Forum Industriel en IA (FIIA 2017),
- La remise annuelle d'un [Prix de Thèse](#) de Doctorat en IA,
- Le soutien à plusieurs [Collèges](#), actuellement au nombre de 4, ayant leur propre activité :
 - Collège Industriel (depuis janvier 2016),
 - Collège Science de l'Ingénierie des Connaissances (depuis avril 2016),

sances (depuis avril 2016),

- Collège Systèmes Multi-Agents et Agents Artificiels (depuis octobre 2016),
- Collège Représentation et Raisonnement (depuis avril 2017)

- La parution trimestrielle des [Bulletins](#) de l'AfIA, en accès libre à tous depuis le site web,
- Un lien entre adhérents sur les réseaux sociaux [LinkedIn](#), [Facebook](#) et [Twitter](#),
- Le [parrainage](#) scientifique et financier de conférences et d'écoles d'été en IA,
- La diffusion mensuelle de Brèves sur les actualités de l'IA en France (breves@afia.asso.fr),
- La réponse aux consultations officielles ou officieuses (OPECST, MESRI, MINEF, ANR, etc.),
- La réponse aux questions de la presse écrite et de la presse orale, également sur internet.

L'AfIA organise aussi mensuellement des Journées communes avec d'autres Associations (en 2018 : EGC&IA avec EGC, Droit&IA avec la SLC, TAL&IA avec ATALA) et avec d'autres institutions (en 2018 : IA pour l'Education avec la MEN/DEN, Éthique&IA avec le COERLE, Rob&IA avec le GdR Robotique).

Enfin, l'AfIA contribue à la participation de ses membres aux grands événements de l'IA. Ainsi, les membres de l'AfIA, pour leur inscription à PFIA 2018, bénéficieront d'une réduction équivalente à deux fois le coût de leur adhésion à l'AfIA, leur permettant d'assister à PFIA 2018 sur 5 jours au tarif de 95,00 € !

Rejoignez-vous vous aussi et adhérez à l'AfIA pour contribuer au développement de l'IA en France. L'adhésion peut être individuelle ou, à partir de cinq adhérents, être réalisée au titre de personne morale (institution, laboratoire, entreprise). Pour adhérer, il suffit de vous rendre sur le site des [adhésions](#) de l'AfIA.

Merci également de susciter de telles adhésions en diffusant ce document autour de vous !



CONSEIL D'ADMINISTRATION

Yves DEMAZEAU, *président*
Élise BONZON, *vice-président*
Catherine FARON-ZUCKER, *trésorière*
Sandra BRINGAY, *secrétaire*
Dominique LONGIN, *rédacteur*
Arnaud MARTIN, *webmestre*

Membres :

Carole ADAM, Emmanuel ADAM, Patrick ALBERT, Florence BANNAY, Alain BERGER, Frédéric MARIS, Juliette MATTIOLI, Engelbert MEPHU NGUIFO, Davy MONTICOLO, Philippe MORIGNOT, Bruno PATIN, Gauthier PICARD, Olivier SIMONIN, Serena VILLATA

COMITÉ DE RÉDACTION

Emmanuel ADAM
Rédacteur
Emmanuel.Adam@univ-valenciennes.fr

Claire LEFÈVRE
Rédacteur
claire.lefevre@univ-angers.fr

Dominique LONGIN
Rédacteur en chef
Dominique.Longin@irit.fr

Philippe MORIGNOT
Rédacteur
philippe.morignot@vedecom.fr

LABORATOIRES ET SOCIÉTÉS ADHÉRANT COMME PERSONNES MORALES

.....
Ardans, Berger Levrault, CRIL, CRISAL, Dassault Aviation, GRETTIA, GREYC, Huawei, I3S, IBM, INRIA Sophia Antipolis Mediterranee, IRIT, ISAE-SUPAERO, Lab-STICC, LAMSADE, LERIA, LIG2P, LHC, LIG, LIMICS, LIMSI, LIP6, LIPAPE, LIRIS, LIRMM, LITIS, MalAGE, Naver Labs, MaiAGE, Renault, Thales, Veolia.

■ Pour contacter l'Afia

Président

Yves DEMAZEAU
L.I.G./C.N.R.S., Maison Jean Kuntzmann
110, avenue de la Chimie, B.P. 53
38041 Grenoble cedex 9
Tél. : +33 (0)4 76 51 46 43
president@afia.asso.fr

Serveur WEB

<http://www.afia.asso.fr>

Adhésions, liens avec les adhérents

Catherine FARON-ZUCKER
tresorier@afia.asso.fr

■ Calendrier de parution du Bulletin de l'Afia

	Hiver	Printemps	Été	Automne
Réception des contributions	15/12	15/03	15/06	15/09
Sortie	31/01	30/04	31/07	31/10