



HAL
open science

Actes de la 4e conférence nationale sur les Applications Pratiques de l'Intelligence Artificielle

Sandra Bringay, Juliette Mattioli

► **To cite this version:**

Sandra Bringay, Juliette Mattioli. Actes de la 4e conférence nationale sur les Applications Pratiques de l'Intelligence Artificielle: APIA 2018. Plate-Forme Intelligence Artificielle, Association Française pour l'Intelligence Artificielle, 2018. hal-04590059

HAL Id: hal-04590059

<https://ut3-toulouseinp.hal.science/hal-04590059>

Submitted on 28 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

APIA & CNIA & IC & JFPDA & RJCIA

IA pour l'éducation IA & Santé TAL & IA Ethique & IA
Robotique & IA France@IJCAI2018

PFIA 2018

**11^e Plate-forme
Intelligence Artificielle**

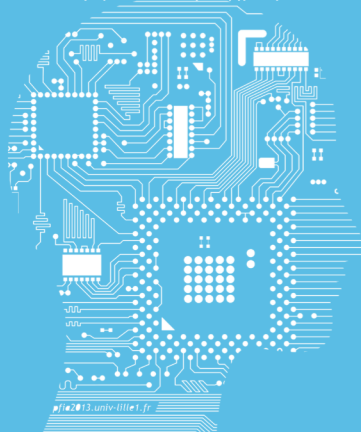
2 au 6 juillet 2018 - Nancy

Campus Sciences - Université de Lorraine
Vandœuvre-lès-Nancy

APIA 2018

ACTES
de la
**Conférence Nationale
sur les Applications Pratiques
de l'Intelligence Artificielle**

Présidentes du comité de programme
Sandra Bringay & Juliette Mattioli



sindup

cdiscount
RECONNU PAR VOTRE CLASSE

DGA

masa

AfIA
Association Française
pour l'Intelligence Artificielle

Pacte Novation

Pacte Novation

orange

SNCF

ERPI

PFIA2018.LORIA.FR

Loria

CNRS

nexter
SYSTEMS

ARDANS
Knowledge Consulting & Software Engineering

Google

UNIVERSITÉ
DE LORRAINE

métropole
GrandNancy

sopra

STERIA
SOCIÉTÉ GÉNÉRALE

ENGIE
Lab

Comité de programme

Co-Présidentes

- Sandra Bringay, LIRMM, Université Paul-Valéry Montpellier 3
- Juliette Mattioli, Thales

Comité de programme

- Carole Adam, LIG CNRS UMR 5217 - Université Joseph Fourier
- Carole Adam, LIG CNRS UMR 5217 - Université Grenoble-Alpes
- Florence Amardeilh, créatrice d'entreprise en Ingénierie des Connaissances
- Ghislain Atemezing, mondeca
- Jérôme Azé, LIRMM-UM-CNRS
- Alain Berger, Ardans
- Matthieu Boussard, Craft AI
- Bertrand Braunschweig, Inria Saclay
- Stéphan Brunessaux, AIRBUS
- Moez Bouchouicha, Université du Sud Toulon-Var
- Nathalie Chaignaud, LITIS - INSA Rouen Normandie
- Caroline Chopinaud, Craft AI
- Jan Ciger, Reviatech SAS
- Laurent Cosserat, Renault
- Jean-Marc David, Renault
- Philippe David, SNCF
- Nicolas Delaforge, mnemotix
- Etienne de Sevin, SANPSY - University of Bordeaux
- Yves Demazeau, CNRS - LIG
- Christian de Sainte Marie, IBM
- Sylvie Despres, LIMICS - Université Paris 13
- Sébastien Destercke, CNRS, UMR Heudiasyc
- Jean-Yves Donnart, THALES
- Béatrice Fuchs, LIRIS, IAE - Université Lyon 3
- Christophe Guettier, Safran Electronique et Défense
- Céline Hudelot, Ecole Centrale Paris
- Philippe Laborie, IBM
- Christophe Labreuche, thales
- Arnaud Lallouet, huawei
- Christine Largouët, Irisa /Agrocampus Ouest
- Vincent Lemaire, Orange

- Dominique Lenne, Heudiasyc - Université de Technologie de Compiègne
- Domitile Lourdeaux, Heudiasyc UMR7253 - Université de Technologie de Compiègne
- Sylvain Mahe, EDF Recherche et Développement
- Philippe Mathieu, University of Lille 1
- Nada Matta, University of Technology of Troyes
- Eunika Mercier-Laurent, CRESTIC- Université Reims Champagne Ardennes
- Christophe Menichetti, IBM
- Philippe Morignot, SAFRAN
- Selmin Nurcan, Université Paris 1 Panthéon - Sorbonne
- Jean-Marc Ogier, L3i - University of La Rochelle
- Benoit Ricard, EDF R&D
- Jean Rohmer, De Vinci Research Center (DVRC)
- Dominique Romary, ODEFI
- Catherine Tessier, Onera
- Erwan Tranvouez, LSIS - Polytech'Marseille - Université d'Aix-Marseille
- Brigitte Trousse, Université Côte d'Azur, INRIA Sophia Antipolis - Méditerranée
- Amel Yessad, LIP 6 - Université Pierre et Marie Curie

Avant-propos

La sélection d'articles publiés dans ce recueil constitue les actes de la 4ème Conférence Nationale sur les Applications Pratiques de l'Intelligence Artificielle (APIA 2018). Cette conférence s'est déroulée du 2 au 3 juillet 2018 à Nancy, dans le cadre de la plateforme PFIA 2018.

Les recherches en IA menées ces dernières années ont abouti à des résultats prometteurs. L'IA se trouve au cœur de nombreuses applications, très performantes, qui révolutionnent notre vie quotidienne et d'autres très prometteuses sont en passe de le devenir. L'objectif de la conférence a été de faire un tour d'horizon des applications concrètes qui couronnent de succès l'opérationnalisation de l'IA.

Quelles sont les applications qui s'appuient sur de l'IA ou qui nécessitent de l'IA ? Du système de surveillance militaire au système d'aide au diagnostic médical, du climatiseur à l'assistant personnel, du système d'aide à la conduite à l'analyse de données massives, etc., les applications sont nombreuses. Qu'elles soient industrielles, sociétales, économiques, politiques, environnementales, artistiques ou autres, la conférence est l'occasion de présenter des applications concrètes et des travaux dont l'objet d'étude n'est pas uniquement des cas de laboratoire. Nous attendons des contributions dans des domaines très divers : automobile, robotique, militaire, logistique, télécommunication, domotique, agronomie, réseaux sociaux, risque, big data, médical, aide à la personne, jeux vidéo, réalité virtuelle/mixte, musées, EIAH, serious games, récit interactif et bien d'autres encore. . .

L'objectif est également de voir comment ces applications concrètes font remonter des verrous scientifiques que l'IA doit résoudre pour démocratiser encore plus son utilisation. Par exemple, comment la prise en compte des réalités concrètes vues dans leur globalité amènent la prise en compte par l'IA de problèmes complexes décuplés qu'il n'est pas possible de rendre compte dans des approches réductionnistes de problèmes de laboratoire ? L'IA est-elle suffisamment expressive et intelligible pour être utilisée ? Est-elle fiable et robuste ? Est-elle capable de passer à l'échelle ? Quels sont les problèmes éthiques liés à son utilisation ? Nous nous demanderons comment les applications concrètes apportent des problématiques fondamentales ou encore comment des recherches fondamentales apportent des solutions à des problèmes complexes difficiles à résoudre sans IA ?

La conférence favorise l'échange entre chercheurs académiques, industriels et autres pour qu'ils puissent partager leurs expériences, débattre des différents verrous qu'ils rencontrent, des méthodes qu'ils mettent en œuvre pour enrichir le potentiel applicatif des modèles et des outils de l'IA et les besoins naissants, en mettant en valeur l'IA de ces applications.

Parmi 22 articles soumis, 12 articles ont été retenus (54,5% d'acceptation). Parmi 4 démonstrations soumises, toutes ont été retenues (100% d'acceptation). Les 4 démonstrations ont été présentées dans une session commune de la plateforme PFIA (APIA/JFPDA).

Dans cette édition, les thèmes abordés par les auteurs couvrent différents aspects de l'IA. Ainsi, les approches décrites s'intéressent à des thèmes variés comme l'ingénierie de la connaissance, les méthodes d'apprentissage, la modélisation statistique, la théorie des jeux, les solveurs SAT, le web des données et les données liées. Les domaines d'application sont également très diversifiés : les robots compagnons, l'habitat intelligent, l'analyse de réseaux sociaux, la planification, la maintenance, l'e-commerce...

Nous remercions vivement les auteurs pour leurs contributions, le comité de programme et les relecteurs additionnels pour la qualité de leurs recommandations ainsi que les organisateurs de la plateforme pour leur investissement.

Sandra Bringay et Juliette Mattioli *Présidentes du comité de programme*

Sommaire

Index des auteurs	7
Sessions plénières	9
Dounia Lahoual and Myriam Fréjus. <i>Usages des systèmes d'interaction vocale dans l'habitat : quels retours utilisateur et quels critères d'interaction à intégrer en conception ?</i>	11
Guillaume Gadek, Alexandre Pauchet, Stephan Brunessaux, Khaled Khelif and Bruno Grilhères. <i>AI techniques to analyse a social network on text, user and group level : application on Galaxy2</i>	23
Jacques Kadima Kazaku. <i>Contribution à l'identification par logique floue d'un échangeur de chaleur à tube concentrique</i>	33
Juliette Mattioli, Pierre-Olivier Robic and Thomas Reydelle. <i>L'intelligence artificielle au service de la maintenance prévisionnelle</i>	39
Michael Kirsten and Olivier Cailloux. <i>Towards automatic argumentation about voting rules</i>	43
Molka Tounsi Dhoub, Catherine Faron Zucker and Andrea Tettamanzi. <i>Categorization of B2B Service Offers : Lessons learnt from the Silex Use case</i>	53
Nicolas Gutowski, Tassadit Amghar, Fabien Chhel and Olivier Camp. <i>Bandits-Manchots Contextuels : Précision Globale Versus Individuelle</i>	61
Sarah Amar, Omar Bentahar, Lionel Cordesses, Thomas Ehrmann, Aude Laurent, Julien Page and Kevin Poulet. <i>Apprentissage par Transfert reformulé via la Théorie des Catégories</i>	69
Sarah Amar, Omar Bentahar, Lionel Cordesses, Thomas Ehrmann, Aude Laurent, Julien Page and Kevin Poulet. <i>Théorie des Catégories et Critère de Contrôlabilité pour Atari 2600</i>	79
Thomas Delacroix and Hadi Sadighyan. <i>Planifier l'épreuve E5 à l'aide d'un solveur SAT</i>	89
Valentina Pozzoli and François Ramond. <i>Optimisation en temps réel de la gestion des convergences ferroviaires</i>	97
Zhiwei Li, Elie Guàrdia Sebaoun, Marco Bevilacqua, Christelle Grauer, Matthieu Cornec and Bruno Goutorbe. <i>Classification d'Images pour la Catégorisation de Produits sur un Site de E-Commerce</i>	105
Démonstrations et Posters	113
Anthony Brage, Stéphanie Jean Daubias, Elodie Loisel and Thierry Basset. <i>JOE : le robot-compagnon des enfants asthmatiques</i>	115
Sébastien Cossin, Vianney Jouhet, Luc Lebrun, Aymerick Niamkey, Fleur Mougin, Mathieu Lambert, Gayo Diallo, Frantz Thiessard. <i>SmartCRF, un prototype pour visualiser, rechercher et annoter les informations d[Pleaseinsertintopreamble]un dossier patient dans I2B2</i>	119
Sarah Madeleine, Sylvain Maillot, Sara Thomas, Maximilien Servajean. <i>Digital Humanities : a Practical Case with Hieroglyphs Identification.</i>	123
Thinh Dong, Chan Le Duc, Myriam Lamolle. <i>ONTOREV : un moteur de révision d'ontologies OWL</i>	127

Index des auteurs

- A -

Amar, S., 69, 79
Amghar, T., 61

- B -

Basset, T., 115
Bentahar, O., 69, 79
Bevilacqua, M., 105
Brage, A., 115
Brunessaux, S., 23

- C -

Cailloux, O., 43
Camp, O., 61
Chhel, F., 61
Cordesses, L., 69, 79
Cornec, M., 105
Cossin, S., 119

- G -

Goutorbe, B., 105

- D -

Delacroix, T., 89
Diallo, G., 119
Dong, T., 127

- E -

Ehrmann, T., 69, 79

- F -

Faron Zucker, C., 53
Fréjus, M., 11

- G -

Gadek, G., 23
Grauer, C., 105
Grilhères, B., 23
Guàrdia Sebaou, E., 105
Gutowski, N., 61

- D -

Jean Daubias, S., 115

- J -

Jouhet, V., 119

- K -

Kadima Kazaku, J., 33
Khelif, K., 23
Kirsten, M., 43

- L -

Lahoual, D., 11
Lambert, M., 119
Lamolle, M., 127
Laurent, A., 69, 79
Le Duc, C., 127
Lebrun, L., 119
Li, Z., 105
Loisel, E., 115

- M -

Madeleine, S., 123
Maillot, S., 123
Mattioli, J., 39
Mougin, F., 119

- N -

Niamkey, A., 119

- P -

Page, J., 69, 79
Pauchet, A., 23
Poulet, K., 69, 79
Pozzoli, V., 97

- R -

Ramond, F., 97
Reydellet, T., 39
Robic, P-O., 39

- S -

Sadighiyan, H., 89
Servajean, M., 123

- T -

Tettamanzi, A., 53
Thiessard, F., 119
Thomas, S., 123
Tounsi Dhoub, M., 53

Sessions plénières

Utilisation des systèmes d'interaction vocale dans l'habitat : quels retours d'expérience utilisateur et quels critères d'interaction à intégrer en conception ?

D. Lahoual¹ M. Fréjus¹

¹EDF Recherche & Développement – 7 Boulevard Gaspard Monge 91120 Palaiseau

dounia.lahoual@gmail.com

myriam.frejus@edf.fr

Résumé

L'essor de la commande vocale et des assistants vocaux intelligents transforme progressivement les utilisations, les modes d'interaction, ainsi que les expériences utilisateurs. Combinés à de l'intelligence artificielle, ces systèmes d'interaction vocale proposent un ensemble de services autour du confort, de la sécurité, du divertissement et de la santé en intégrant désormais la sphère domestique. Dans une visée de conception d'interaction vocale efficiente et adaptée aux besoins et aux activités des habitants, nous proposons d'établir une cartographie des retours d'expérience utilisateurs à partir d'une revue de la littérature.

Mots-clés

Interaction vocale, expérience utilisateur, habitat, utilisabilité, appropriabilité, acceptabilité, assistants vocaux, usages.

Abstract

The rise of the voice commands and smart speakers are progressively changing uses, ways of interacting and user experiences. Combined with artificial intelligence, these voice interaction systems are offering a set of services relating to comfort, security, entertainment and health in the domestic sphere. In order to design an efficient voice interaction suited for users' needs and activities, we propose a literature review about users experience with voice commands and smart speakers.

Keywords

Vocal interaction, user experience, housing, usability, appropriability, acceptability, smart speaker, uses.

1. Introduction

Afin de s'affranchir des dispositifs matériels contraignants et simplifier les usages, l'intelligence artificielle propose de nouveaux outils et de nouvelles formes d'interaction visant une fluidité et une instantanéité à l'aide du langage naturel [27, 18]. Grâce à ses grandes capacités d'apprentissage automatique, à la programmation de nombreux algorithmes et à l'utilisation de données massives, l'IA vise à augmenter les capacités humaines en s'appuyant sur une coopération entre l'homme et la machine [16, 27]. Ainsi, ces dernières années ont été marquées par de nombreux progrès dans le domaine de l'IA, et cela dans plusieurs secteurs : robotique, véhicules autonomes, traitement de la parole et compréhension du langage naturel, maisons intelligentes, etc. [25, 9, 20]. D'abord limités à l'utilisation du smartphone, les assistants

vocaux intelligents intègrent les habitats avec l'arrivée des enceintes intelligentes telles qu'Alexa d'Amazon ou Google Home [41]. Un assistant vocal intelligent représente un logiciel intégré dans un dispositif tangible, disposant de la faculté de dialoguer avec un humain pour réaliser des tâches sur ordre vocal des utilisateurs (répondre à des questions en sélectionnant des informations sur le web, rappeler des actions importantes aux utilisateurs, effectuer certaines tâches sur des objets connectés, etc.). Disposant d'une capacité d'apprentissage automatique, ces intelligences artificielles se développent continuellement sans l'intervention des développeurs, apprennent de leurs erreurs, et cela, tout en intégrant des informations sur les utilisateurs afin de proposer des services contextualisés et personnalisés [29, 14]. Si la question de la bonne intégration de la reconnaissance de la parole dans les interfaces et produits grand public n'est pas nouvelle [44], la maturité technologique des assistants actuels questionne l'utilisation réelle qui peut en être fait tant par ses utilisateurs finaux que par les concepteurs, en lien avec les autres formes d'interaction. Nous considérons que pour développer ces nouveaux systèmes vocaux ou les intégrer à des produits ou services, il importe d'étudier et de comprendre les situations d'utilisation dans lesquelles ils s'insèrent, ainsi que les leviers et les freins à leur appropriation. Proposer des systèmes adaptés à l'activité des utilisateurs en prenant en compte le point de vue de l'humain et de ses activités constitue un des principes fondateurs de la démarche en ergonomie que nous proposons de mobiliser.

Intégrées dans un contexte de dispositifs ambiants, la commande et la reconnaissance vocale constituent ainsi de nouveaux défis en apportant de nouvelles situations et formes d'utilisation. De plus, face à la multiplication récente des services digitaux et des systèmes d'interaction vocale, l'ensemble des attitudes, des besoins, et des utilisations provenant des utilisateurs demeurent encore méconnus [51, 53]. Ces problématiques vocales questionnent ainsi en profondeur les principes d'utilisabilité, d'appropriabilité et d'acceptabilité classiques pour diverses raisons. En premier lieu, ces dispositifs vocaux créent une rupture dans les représentations, les utilisations et les routines des usagers. Ces systèmes diffèrent grandement des interfaces tangibles et traditionnelles. En second lieu, l'interface vocale privilégie l'audition et l'usage de la parole : ces capacités habituellement mobilisées en situation d'interaction humaine sont dorénavant au cœur des interactions entre l'humain et la machine. Ces

interactions se trouvent aussi modifiées en ce que le système se trouve doté d'une autonomie et d'une capacité d'adaptation qui rompt avec l'interaction homme machine connue. En dernier lieu, ce qui prévaut comme critères de conception pour les IHM traditionnelles n'est pas assuré de fonctionner avec des interfaces vocales.

Afin de développer des connaissances sur ces nouvelles formes d'utilisation vocale, nous proposons dans cet article de réaliser un état des lieux sur les systèmes d'interaction vocale en situation domestique. Cette première étape de veille constitue un préalable dans notre projet de recherche orienté vers la conception de services et de dispositifs vocaux dans l'habitat. Ces éléments récoltés nous permettront de répondre à un ensemble de questions :

- *Quels sont les expériences d'utilisation réussies et empêchées avec ces technologies vocales ?*
- *Comment améliorer et optimiser les interactions entre les systèmes d'interaction vocale et les utilisateurs à partir de la connaissance du contexte et de l'activité de ces derniers et en lien avec les autres formes d'interaction présentes ?*

Tout d'abord, nous présenterons un ensemble de critères d'interaction entre l'humain et les systèmes d'IA au cœur des démarches de conception. Ensuite, nous établirons un état des lieux sur les systèmes d'interaction vocale en nous focalisant sur les enjeux de la parole, les utilisations appréciées et les difficultés rencontrées par les utilisateurs. Puis, nous proposerons des principes d'interaction à intégrer lors de la conception. Enfin, nous conclurons et discuterons des résultats en précisant nos perspectives de recherche.

2. Repères théoriques sur les critères d'interaction Homme-IA

En mobilisant une démarche ergonomique, nous nous intéressons aux multiples dimensions de l'activité, à savoir ; ses aspects situés, finalisés, systémiques, diachroniques, instrumentés, individuels et collectifs. L'identification et la compréhension de ces dimensions contribuent à développer des connaissances sur l'activité réelle des utilisateurs. À terme, il importe de les intégrer dans la conception des dispositifs à base d'IA pour garantir l'utilité, l'utilisabilité, l'appropriabilité et l'acceptabilité de ces systèmes par les utilisateurs.

2.1 Une interaction efficace et opérante sur le temps : utilité, utilisabilité, appropriabilité et acceptabilité

Pour atteindre une situation durable intégrant des systèmes d'IA, il demeure indispensable de se pencher vers les critères d'utilité, d'utilisabilité, d'acceptabilité et d'appropriabilité qui sont au cœur de la démarche de conception et qui conditionnent l'adoption des systèmes. L'**utilité** renvoie à l'adéquation du dispositif et de ses services aux besoins et aux attentes des utilisateurs [50]. Ce critère tend le plus souvent à s'articuler à celui de l'**utilisabilité** qui vise à « *atteindre*

des buts définis avec efficacité, efficience et satisfaction, dans un contexte d'utilisation spécifié » d'après la norme ISO 9241-11 [26]. En d'autres termes, il s'agit d'assurer une simplicité d'utilisation pour des situations et des utilisateurs spécifiques. Quand bien même ces critères seraient intégrés dès la conception, le dispositif ne serait toujours pas assuré d'être accepté, ni approprié [13] car l'activité incorpore un aspect dynamique et longitudinal imprévisible qu'il importe d'investiguer en situation réelle. Cette caractéristique est davantage mise en lumière avec la notion d'**appropriation**, et sa déclinaison en conception, l'**appropriabilité** [23]. L'appropriation représente l'intégration, partielle ou totale, d'un objet, d'un outil ou d'un dispositif au corps propre ou la culture propre de l'acteur. L'appropriation est souvent accompagnée d'une individuation de son usage et éventuellement, de transformations plus ou moins importantes de cet objet, de cet outil ou de ce dispositif lui-même [52]. L'appropriation s'inscrit notamment sur une perspective temporelle longue intégrant les histoires d'utilisations, les représentations des utilisateurs, et leurs nombreux épisodes de succès et d'échecs colorant ce parcours singulier [40, 45]. Par ailleurs, la dynamique de cette trajectoire d'utilisation implique une confrontation itérative de l'utilisateur avec le dispositif dans laquelle les univers de l'utilisateur et du concepteur se rencontrent [5, 40] et peuvent être susceptibles non seulement de susciter des besoins inédits, mais aussi de provoquer l'évolution des schèmes d'usage et la transformation du dispositif par l'utilisateur [30, 46]. Cette dimension de l'**appropriabilité** renvoie ici à la capacité d'un dispositif à s'adapter et à réagir de façon pertinente selon le contexte, les préférences et le développement des utilisateurs sur le long terme [13]. Et c'est pourquoi les nouvelles perspectives d'autonomie des systèmes à base d'IA re-questionnent l'**appropriabilité**. Quant au critère d'**acceptabilité**, bien que déterminant pour comprendre les problématiques d'adoption ou non d'un dispositif, ce critère reste encore vivement débattu auprès des chercheurs, notamment dans ses liens avec l'**appropriabilité**. La technologie doit être perçue en adéquation avec les pratiques, les ressources, les objectifs, le système de valeurs et les normes éthiques des utilisateurs pour être acceptable [13, 36].

2.2 Une déclinaison des critères de conception pour favoriser l'appropriation

Dans un contexte d'innovation technologique en lien avec l'émergence de l'IA, nous avons pu identifier un ensemble de dimensions déterminantes à intégrer dès la conception afin de garantir une meilleure appropriation des dispositifs. Résultant des critères de conception précédemment évoqués, ces dimensions visent l'augmentation des capacités humaines à partir d'une coopération entre l'homme et la machine pour une interaction efficace et opérante en situation.

Parmi ces dimensions, la **complémentarité** évoque tout d'abord la possibilité pour l'utilisateur de s'appuyer sur des dispositifs complémentaires, fonctionnellement, pour pallier aux limites de chaque système, mais aussi – ~~en~~ termes de modalités d'interactions. L'interaction

multi-modale permet à l'utilisateur de privilégier une interaction vocale, tactile, télécommandée ou via le clavier selon ses préférences, ses objectifs et les activités en cours [15].

La multi-modalité permet aussi de répondre à un autre objectif : l'**adaptabilité** du système à l'activité des utilisateurs. Celle-ci s'appuie sur une compréhension des contextes pour s'adapter aux situations et agir efficacement, et cela sans perturber le flux de l'activité des utilisateurs : cela peut être illustré par l'allumage des lumières provoqué par des capteurs de mouvement lors de l'entrée d'un usager. Ce système peut notamment être orienté vers l'anticipation des besoins par exemple, en déclenchant le chauffage de certaines pièces avant l'arrivée d'un habitant (op. cit.). Il s'agit donc d'un système « *flexible* » pouvant réagir de façon implicite aux besoins et aux préférences des utilisateurs, tout en prenant en compte leurs expériences [3].

Ces différentes formes d'interaction couplées à leur bonne intégration aux activités en cours doivent permettre de répondre à une recherche d'**invisibilité**, preuve de leur appropriabilité. Un dispositif devient transparent dans l'usage grâce à un processus « *d'incorporation* ». Ce dernier conduit à intégrer un objet au « *corps propre* » de l'utilisateur comme un instrument pour penser, percevoir et agir. Cette incorporation peut être illustrée par un conducteur qui fait corps avec sa voiture, de telle façon qu'il pourra percevoir les aspérités de la chaussée et adapter sa conduite [52].

Le développement de systèmes intelligents doit aussi intégrer des critères favorisant le dialogue et la coopération afin de garantir à l'utilisateur un niveau de contrôle, de reprise en main et de visibilité sur les actions du système. Pour ces raisons, l'**intelligibilité** du dispositif [4] demeure cruciale pour informer l'utilisateur des situations en cours, rendre visible la logique de fonctionnement et les capacités du système, ainsi que les conséquences des actions passées et à venir des usagers [16].

Enfin, le critère d'**évolutivité** permet d'intégrer les dimensions temporelles et développementales inhérentes à l'homme. Ce critère « *renvoie à la possibilité du système de proposer des fonctionnalités en contexte et renouvelées en réponse à l'évolution des besoins et capacités* » [17], tout en s'adaptant au contexte d'utilisation des usagers [10]. Cette évolutivité du système peut consister par exemple à proposer des fonctionnalités selon le développement de l'utilisateur ; à prendre en compte les systèmes de ressources développées par les utilisateurs au fil de ses expériences (technologiques, humaines, organisationnelles) ; ou encore, l'évolutivité peut intégrer l'évolution des routines individuelles et collectives sur diverses temporalités et espaces [17].

2.3 L'interaction vocale : quels enjeux ?

La parole représente une modalité fondamentale dans la communication humaine. Dans toutes les cultures, les individus persuadent, informent et construisent des relations à travers la parole principalement. Ainsi, la sensibilité à la parole et à la voix joue ainsi un rôle

critique dans les interactions au sein du collectif. De plus, les humains ont développé les instruments de parole et d'écoute exprimant une large variété de signaux sociaux appropriés auxquels le cerveau humain est apte à analyser et à répondre. Parmi les paramètres de la parole, le rythme, le volume, la cadence transmettent des caractéristiques humaines subtiles telles que la personnalité, l'émotion, ou encore la ville d'origine d'une personne. Les humains sont si accoutumés aux caractéristiques vocales qu'ils peuvent rapidement et précisément distinguer une voix d'une autre [34]. Ainsi, dans le contexte de l'affective computing, la voix devient un facteur déterminant car il constitue un indicateur fiable pour détecter l'émotion des utilisateurs, tout comme le choix des mots peut également signaler une humeur particulière [22].

Pour ces raisons, la parole demeure une forme d'interaction intuitive, invisible et rapide pour réaliser une diversité de tâches à partir des commandes vocales. Non seulement les interfaces vocales libèrent les utilisateurs des contraintes « *WIMP* » (Fenêtre, icônes, menus et curseur), mais elles visent également à accompagner ceux et celles dont les yeux et les mains sont affairés à d'autres tâches [34, 22]. Elles permettent aussi de surmonter certaines contraintes individuelles et situationnelles, telles que les personnes portant des handicaps et des déficiences visuelles ou encore les enfants, qui n'ont pas la dextérité ou une taille de main suffisante pour les claviers et les pointeurs. Et tous les environnements ne sont pas tous disposés à proposer des commandes via une manipulation directe [19, 22].

Afin d'encourager une coopération et une meilleure acceptabilité de ces formes d'interaction, certains travaux démontrent l'importance de concevoir une interaction qui puisse s'accorder de façon appropriée à l'état et au timbre émotionnel de l'utilisateur afin de communiquer de l'empathie, du soutien et de la sincérité dans l'échange [34]. Dans cette visée, ces auteurs invitent notamment à créer un sentiment de similarité entre l'utilisateur et l'interface à travers l'adéquation entre le genre, la personnalité et l'accent qui encourageraient des sentiments de sympathie et de confiance pour assurer le succès de ces technologies. Ces recommandations pour la conception s'appuient sur un résultat majeur sur la façon dont l'humain interagit vocalement avec une machine : le cerveau humain ferait rarement la distinction entre le fait de parler à une machine (même si elle possède une faible compréhension de la parole et une qualité médiocre de production langagière) et parler à une personne. Ainsi, les humains mobiliseraient les mêmes parties du cerveau pour interagir avec les machines et les humains. Les machines ne seraient donc pas exemptées des attentes sociales activées lors des échanges et de l'écoute (op. cit.).

Ces éléments soulignent ainsi l'essor d'une forme d'interaction qui pourrait s'intégrer de façon harmonieuse dans l'activité réelle des utilisateurs. De plus, les caractéristiques évolutives et adaptatives de l'IA sous-tendant ces systèmes laissent présager une meilleure adéquation pour leur appropriabilité par un

meilleur respect des dimensions que nous avons évoquées. Toutefois, la combinaison souhaitable avec d'autres modalités ainsi que la question de l'intelligibilité des systèmes à base d'IA, souvent présentés comme une boîte noire, restent ouvertes. Ces questions et premières suppositions devront être mises en perspective avec l'activité réelle et les points de vue des utilisateurs.

3. État de l'art sur l'activité réelle des systèmes d'interaction vocale

Nous présenterons dans les prochaines sections un état des lieux sur les systèmes d'interaction vocale. Pour cela, nous nous appuyerons sur une diversité de sources telles que les travaux scientifiques, les enquêtes menées par des instituts spécialisés, ainsi qu'une diversité d'articles et de billets récents sur les pages du web. Parmi nos références bibliographiques, notre webographie constitue une source importante compte tenu du développement des technologies vocales et de leur intégration récente dans les foyers. Afin de garantir une validité des données récoltées dans cette revue de l'état de l'art sur les interactions vocales, nous nous appuyons sur la triangulation des données visant une multiplication et une combinaison des matériaux récoltés. Cette triangulation nous permettra notamment d'obtenir une connaissance plus complète et différenciée du même phénomène étudié [15, 42].

3.1 Les technologies vocales : des expériences d'utilisation réussies et appréciées

Parmi les utilisations des technologies vocales en situation domestique et professionnelle, nous retrouvons des expériences réussies et appréciées par les utilisateurs que nous avons catégorisés en trois catégories principales : les qualités des formes d'interaction vocale, les qualités du système vocal et les effets des interactions vocales sur l'activité des utilisateurs.

3.1.1 Les qualités de l'interaction : accélérateur opératoire, multi-modalité, adaptabilité intra-individuelle et courtoisie

Parmi les qualités retrouvées dans les interactions vocales, nous retrouvons tout d'abord **l'instantanéité, la facilité et le confort d'utilisation** de mobiliser la parole comme un moyen d'interaction. Il intègre une rapidité de commande et de contrôle dépassant largement les actions consistant à presser un bouton ou cliquer à l'aide d'une souris « *Alexa m'a gâtée par le confort qu'elle propose, mais elle me fait réellement gagner du temps* » [21]. Les retours vocaux sont notamment rapidement compréhensibles, ils évitent toute ambiguïté que nous pourrions rencontrer via les alarmes sonores et les voyants lumineux [51]. De plus, la commande vocale semble avantageuse pour réaliser des actions complexes nécessitant plusieurs touches et de multiples interfaces. Une étude réalisée par Turunen et al. [53] s'est appuyée sur une évaluation multimodale d'un questionnaire de dispositifs domestiques (un centre

multimédia) auprès de 26 participants. Les résultats ont démontré que la commande vocale avait été très appréciée par les usagers et qu'elle apportait notamment des bénéfices supplémentaires pour réaliser des commandes complexes nécessitant plusieurs touches ou l'utilisation de multiples interfaces (exemple d'une commande dite complexe sans le vocal : « *Enregistrer tous les épisodes de Friends* »).

En définitive, le vocal constitue une façon naturelle, fluide, intuitive et aisée d'interagir avec son environnement : « *Si je travaille tard, en montant les escaliers, je demande : Alexa, allume la chambre à coucher 1. Et quand j'arrive dans la pièce, la lumière est allumée. C'est une expérience plus naturelle que d'utiliser des applications* » [48]. L'interaction vocale agit ainsi en tant qu'accélérateur opératoire en simplifiant ou en supprimant certains modes opératoires [6].

Nous apprenons notamment que la **complémentarité des formes d'interaction** assure une meilleure acceptabilité du vocal lorsqu'il est associé à d'autres formes d'interaction : certains utilisateurs apprécieraient par exemple de combiner le vocal à la manipulation télécommandée [53]. Des travaux ont été réalisés par DeRenzi et al. [12] pour évaluer une interface professionnelle combinant le vocal et l'affichage web auprès d'agents de la santé. Les auteurs révèlent que certains professionnels préfèrent associer les canaux vocaux et visuels pour mieux ajuster les commandes vocales, enrichir les données récoltées, ou encore, pour vérifier la concordance des informations apportées [12]. Dans cette visée, nous observons que les modalités se soutiennent les unes aux autres [51]. De plus, certaines modalités se prêtent mieux à certaines opérations telles que les interactions gestuelles pour exprimer une commande télécommandée fréquente en situation de visionnage de film dans l'obscurité [53].

L'adaptation à l'utilisateur et au contexte ambiant représente également un facteur clé dans l'utilisation du vocal. Diverses caractéristiques sont appréciées en situation : les utilisateurs n'ont guère besoin de connaître la multitude des services et des applications intégrées dans l'assistant vocal pour utiliser ses principales fonctions. Le haut-parleur intelligent s'ajuste généralement à la demande formulée et au vocabulaire de l'utilisateur tout en continuant à cumuler des connaissances fines sur les utilisateurs cibles, de façon à contextualiser chaque proposition de service [29]. Par ailleurs, l'expérience utilisateur s'en trouve améliorée lorsque les interactions vocales continuent d'être efficaces en situation de bruits de fond. Whitenton affirme ainsi que l'assistant vocal Echo d'Alexa privilégie les interactions vocales sur tout le reste : intégrant sept microphones, il met l'accent sur la distinction entre la commande vocale et les bruits de fonds tels que les sources de musique [54].

Enfin, certains utilisateurs semblent apprécier le fait de privilégier certaines **formes d'adressage et d'interaction** avec ces interfaces **intégrant une forme de courtoisie et de respect des genres**. Contrairement à l'assistant vocal Alexa proposant des formes

d'interaction sur le ton de l'injonction (« *Alexa, éteins les lumières* » ; « *Alexa, programme mon réveil pour 8h* »), Google Home invite davantage à converser, tout en proposant une interaction plus respectueuse (« *Ok Google, lance l'actualité* » ; « *Hey Google, programme mon réveil pour 8h* »). Une utilisatrice attire ainsi l'attention sur l'impact émotionnel des technologies sur l'humain et sur le modèle et les valeurs sociales véhiculées par certains assistants vocaux : « *La différence d'expérience vécue ici est énorme. Lorsque j'utilise Alexa après la Google Home, je me sens autoritaire (...) Ce n'est pas isolé de la façon dont les hommes considèrent les femmes, mais cela s'étend aussi sur la manière dont les femmes se considèrent entre elles. Les moindres détails d'expérience utilisateur ont de l'importance (...) Je souhaite remercier Google. Merci d'avoir porté une attention à ces détails, et aux utilisatrices* » [37].

3.1.2 Les qualités du système : efficacité, adaptabilité inter-individuelle et évolutivité

La concision et l'adaptation des feedbacks aux requêtes facilitent l'acceptabilité des technologies vocales. En effet, la pertinence et la justesse des réponses apportées par ces interfaces contribue grandement à la réussite de l'expérience utilisateur. Cette expérience positive est renforcée si l'utilisateur utilise son propre vocabulaire, il s'agit d'une commande vocale libre [53]. Il a d'ailleurs été observé que certains assistants vocaux, tels qu'Alexa, disposent d'une plus grande précision de réponse face aux requêtes des utilisateurs [54]. Alexa proposerait un traitement sémantique plus performant que Siri : « *Si vous mettez en place un chronomètre et que vous souhaitez savoir combien de temps il reste : Hey Siri, combien de temps il reste ? ==> J'ai trouvé un article sur le Times. Dois-je te le lire ? (...) Alexa, combien de temps reste-t-il ? ==> Environ 6min et 10 sec (...) Alexa interprète le temps comme lié au chronomètre et Siri le relie à une recherche du web, jusqu'à ce que le mot-clé « Timer » soit ajouté* » (op. cit.). Ainsi, éviter les biais d'interprétation en s'appuyant sur la commande vocale libre de l'utilisateur contribue à améliorer grandement l'expérience utilisateur.

Par ailleurs, nous pouvons remarquer que les technologies vocales intègrent naturellement une forme d'interaction pouvant s'adapter à une grande variété d'utilisateurs et d'individus, dont les personnes en situation de handicap ou « déficientes ». La commande et la reconnaissance vocale intègrent ainsi une **adaptation inter-individuelle** consistant à aménager des produits, des services ou des équipements susceptibles d'être utilisés par toute personne, peu importe le sexe, l'âge, la situation ou le handicap. L'interaction devient ainsi utilisable et mobilisable par une diversité de publics telles que les enfants, les illettrés, les personnes âgées, les malvoyants, ou les publics ayant des difficultés à interpréter certains graphiques ou certaines données complexes [43, 33, 12].

L'évolution continue de l'IA représente un autre facteur clé non négligeable dans l'intégration des technologies vocales dans les situations quotidiennes

des utilisateurs. Les technologies de « *deep learning* » s'appuient sur un apprentissage automatique continu favorisant un cumul de connaissances contextualisées sur les utilisateurs qui sont constamment mis à jour [29]. De plus, les techniques vocales intégrées aux assistants vocaux évoluent sans cesse au rythme du web et non au rythme des cycles de conception, ce qui permet de renouveler de façon fluide de nouvelles applications et de nouveaux services [48].

3.1.3 Les effets favorables sur l'activité : cumul, continuité et personnalisation

La **capacité à cumuler des tâches** est l'une des conséquences et des opportunités permises par l'interaction vocale qui est grandement plébiscitée par les utilisateurs. Lors de situations contraintes où les mains et le regard sont affairés, les utilisateurs sont ainsi en capacité de réaliser des doubles tâches durant les routines matinales par exemple : « *Pendant que je m'habille, je lui demande : « Alexa, lance le carburant » pour démarrer la machine à café (...) En continuant ma routine dans la salle de bain et ma chambre, je demande à Alexa de lancer n'importe quelle chanson, artiste, genre ou playlist dont j'ai envie* » [21]. Ce cumul de tâche révèle notamment des utilisations qui s'ancrent profondément dans les routines en faisant de la commande vocale un élément incontournable du quotidien : « *Avant même que j'ouvre les yeux, je discute avec Alexa. Éteins l'alarme. Allume la radio. Quel temps fait-il aujourd'hui ?* » ; « *D'habitude le matin, je peux mettre de la musique avant de sortir du lit. Ou bien, lorsque je sors de la douche avec les mains mouillées, plutôt que d'aller vers téléphone avec les mains humides, je peux juste lancer la musique pendant que je me prépare pour aller travailler* » [47]. Il est donc possible de gérer une diversité de tâches à partir des commandes vocales depuis n'importe quelle pièce sans l'utilisation des mains [51, 31, 21].

Dans le cours d'activité des utilisateurs et des habitants, nous observons que les technologies vocales favorisent une **mobilité et une continuité de l'expérience**. Ainsi, la possibilité de gérer des commandes vocales depuis n'importe quelle pièce du foyer sans les mains, ni du regard, favorise non seulement une grande mobilité dans l'espace domestique, mais elle permet notamment de soutenir la dynamique et le passage d'une activité à l'autre dans le cours d'action en situation domestique ou à l'extérieur (automobile, transports, etc.) [21, 47]. Nous pouvons illustrer par un verbatim cette forme de continuité de l'expérience d'un espace à un autre encouragée par les systèmes d'interaction vocale : « *Si je veux continuer mon livre audio pendant que je m'occupe du linge, du tricot ou autre chose qui ne demande un certain niveau d'attention, je dirai : Alexa, lit mon livre, pour reprendre là où j'en étais dans la voiture* ». L'utilisatrice n'a ainsi guère besoin de préciser de quel ouvrage il s'agit, car le dispositif intelligent comprend la demande sous-jacente [21].

La **création d'un écosystème personnalisé** représente une autre expérience réussie. En effet, les interactions vocales intégrées dans les haut-parleurs intelligents proposent un cadre dans lequel les utilisateurs sont

invités à programmer des fonctions et des services adaptés à leurs besoins et leurs activités. Cet écosystème personnalisé s'accompagne par exemple d'une personnalisation des objets connectés environnants lorsque certains utilisateurs les commandent à distance : « *Alexa, allume la salle de contrôle* » (chambre) ; « *Alexa, lance le carburant* » (café) ; « *Alexa, éteint le repaire* » (salon) ==> « *Ça m'amuse, le nom est unique (...) ce n'est pas une mauvaise chose pour démarrer sa journée en ricanant* » [21]. Ce système constitué par les utilisateurs vise à contrôler un ensemble de services et de dispositifs via cet instrument pivot représenté par l'assistant vocal à l'aide des interactions vocales. Ce qui semble faire émerger une forme de centralisation des tâches via cette « tour de contrôle » qui facilite la diversité des commandes : « *Nous ne faisons pas des choses nouvelles. Nous réalisons juste un peu plus des tâches que nous faisons déjà. Et c'est plus facile comme ça. Nous écoutions déjà la radio, mais maintenant c'est plus facile de le faire. Tout est réuni au même endroit* » [47].

Cette structuration en écosystème semble également évoluer vers une **distribution spatiale et inter-individuelle des dispositifs** dans les foyers. Une enquête menée par Edison Research a démontré que 42% des foyers disposant d'un assistant vocal possède au moins deux haut-parleurs intelligents dans leurs foyers [28]. Lorsque l'utilisation est satisfaisante, ces foyers multiplient ces dispositifs afin de partager cette expérience utilisateur et ces services avec d'autres membres du foyer d'une part, et pour favoriser une continuité des activités via la commande vocale dans une diversité de pièces d'autre part : « *Notre premier Echo (Alexa), nous l'avons placé dans le salon pour le rendre accessible à n'importe qui. Nous l'avons tellement apprécié que nous avons acheté un Dot (Alexa) pour la chambre aussi* ». Ainsi, les dispositifs, initialement localisés dans les pièces collectives, sont progressivement intégrés vers les pièces plus intimes telles que les chambres à coucher, les salles de bain, ou encore les toilettes [47].

3.2 Les technologies vocales : difficultés et problèmes éthiques soulevés

3.2.1 Des formes d'interaction imposées

➤ Empêchement dans l'action : visibilité et mobilité

Des formes d'interaction contraintes ont notamment été soulevées telles que l'**absence d'écran visuel** pour soutenir le vocal afin de proposer une plus grande diversité et richesse des données¹ [31, 12, 32, 54]. Une étude réalisée par Luria et al. [31] portant sur des dispositifs domestiques intégrant une diversité d'interaction (un haut-parleur à commande vocale ; un écran tactile mural, une application mobile ; un robot social) a justement démontré l'importance du canal

visuel auprès des utilisateurs pour disposer d'informations sur la réussite ou non des commandes réalisées et pour obtenir des informations sur la situation en cours et le contexte de la machine : « *Au final, je ne sais rien avec le vocal. C'était si facile de réaliser les commandes, mais finalement vous ne saviez pas ce qui n'allait pas* ». Ce qui peut contribuer à créer une impression de manque de contrôle (que nous aborderons dans la section suivante). D'autres empêchements opératoires ont été recensés tels que l'impossibilité pour un utilisateur d'annuler des achats via la commande vocale (l'utilisateur doit passer une application), de paramétrer ou de personnaliser des fonctions via l'interaction vocale exclusivement [48]. L'**absence de mobilité** de ces dispositifs vocaux (la plupart des haut-parleurs intelligents doivent rester branchés électriquement, excepté le modèle Tap d'Amazon par exemple) empêche son transport d'une pièce à l'autre [54, 35]. Ce qui peut être susceptible de créer des ruptures dans la continuité des activités lorsque l'habitant se déplace d'une pièce à l'autre. Toutefois, pour pallier à cette contrainte, nous avons vu précédemment que de plus en plus d'habitants installent une multiplicité d'assistants vocaux par pièce du foyer [47].

➤ Déstabilisation et malaise des usagers : sentiment de présence

L'interaction vocale peut provoquer pour certains utilisateurs qui ne sont pas accoutumés, une **forme de déstabilisation** pouvant constituer un frein dans l'acceptation de ce type d'interaction : « *Pour utiliser l'interface vocale, on doit s'arrêter, et réfléchir à ce qu'on doit dire* » [31]. Par ailleurs, nous retrouvons une autre forme de déstabilisation plus récurrente avec ces technologies vocales : il s'agit d'un malaise et d'une crainte provoqués par un **sentiment de présence** liée à cette parole qui s'apparente à la présence d'un humain : « *Je n'aime pas le fait de me parler à moi-même, ça me donne une impression étrange et très bizarre... Je ne sais pas, c'est bizarre d'avoir une voix qui vous répond. C'est un peu effrayant* » ; « *Je ne peux pas vraiment expliquer pourquoi, mais je ne me sens pas à l'aise avec. Parce que c'est une machine et qu'il peut me répondre. Je me sens juste inconfortable* » ; « *Il pourrait devenir une sorte d'amis au bout d'un certain temps. On pourrait se sentir très à l'aise avec, comme si on était plus seul ? Peut-être que c'est un peu effrayant* » [31]. Ces technologies biomimétiques révèlent ainsi le risque de certaines technologies et intelligence artificielle à brouiller les frontières entre l'être vivant et les machines [1].

Ce malaise provoqué par ce sentiment de présence peut être d'autant plus marqué lorsque le dispositif vocal se met à fonctionner indépendamment des commandes ou du libre choix des utilisateurs. Pour illustrer cet aspect, nous pouvons citer un problème technique ayant été cité récemment dans l'actualité en Mars 2018 : un ensemble de haut-parleurs Alexa dans les foyers se mettaient à rire spontanément sans aucune raison apparente, et sans aucun lien avec une commande vocale, et cela à

1. À l'heure où nous écrivons ces lignes, de nouveaux assistants vocaux disposant d'un écran visuel ont été déployés sur le marché. Toutefois, notre étude ne s'appuie pas sur ces dispositifs multimodaux.

n'importe quel moment de la journée. Les utilisateurs ont proposé une diversité de qualificatifs pour décrire ce rire pouvant être « *sardonique* » ; « *effrayant* » ; « *inquiétant* » ; « *glaçant* » ; ou encore « *démoniaque intempestif* ». Voici quelques verbatims des utilisateurs ayant vécu cette défaillance technique : « *J'étais dans mon lit et sur le point de m'endormir quand Alexa a lancé un rire fort et terrifiant. Il y a de grandes chances que je sois assassiné cette nuit* » [7] ; « *Donc, Alexa a décidé de rire aléatoirement pendant que j'étais dans la cuisine. Je pensais qu'un gamin riait derrière moi* » ; « *On était en train de parler de choses assez confidentielles et Alexa s'est mise à rire. Quelqu'un d'autre a déjà eu ça ? On n'avait pas l'impression de l'avoir déclenchée accidentellement, elle s'est juste mise à rire comme ça. C'était très effrayant* » [24]. Vécue et jugée comme une expérience désagréable, voire effrayante par les usagers, ce bug technique renforce également ce sentiment de malaise par les enjeux d'anthropomorphisation des technologies que cela soulève.

3.2.2 Défaillances du système : manque d'intelligibilité, d'efficacité et d'adaptabilité inter-individuelle au sein d'un même foyer

Parmi les défaillances du système vocal, nous pouvons citer une forme **d'opacité** sur les potentialités des services vocaux. Ce manque d'information peut nuire à l'utilisabilité si les utilisateurs ne connaissent pas l'étendue, ni les limites fonctionnelles de leurs appareils. Pour illustrer ce type de défaillance, nous citerons la logique de fonctionnement technique du dispositif vocal Alexa. Cet assistant vocal intègre des paramètres par défaut que bien souvent, les utilisateurs ne connaissent pas d'emblée ou n'en saisissent pas toutes les subtilités et les impacts possibles sur leurs activités. Parmi ces défauts de paramétrage : Alexa propose des services ou des produits que l'on ne cherche pas à acheter ; l'utilisateur ne peut pas demander à Alexa d'annuler un achat, il faut le faire depuis une application ; le paramétrage par défaut n'inclut pas un code pin à 4 chiffres pour éviter des requêtes d'achats qui ne proviennent pas de l'utilisateur principal. Et enfin, Alexa autorise n'importe quelle personne proche du dispositif à commander des appareils et des services depuis le compte Amazon de l'utilisateur principal, et cela inclus les enfants, les voix de la radio ou de la télévision [8]. Ce qui induit une perte de contrôle nuisible à l'utilisabilité et l'acceptabilité de tels dispositifs.

L'échec d'interprétation des commandes et du contexte constitue un autre frein à l'acceptabilité de ces technologies. Parmi cette catégorie nous retrouvons trois principales contraintes. Tout d'abord, la *mauvaise reconnaissance vocale* induit un manque de fiabilité en cas de mauvaise interprétation des commandes tels que les biais sémantiques évoqués précédemment, d'une mauvaise qualité de la reconnaissance vocale pour certaines voix retrouvées plus particulièrement auprès des personnes âgées, les enfants ou les personnes atteintes d'une maladie [43, 51]. Cet échec de reconnaissance vocale ou de détection des phrases

d'activation peut aussi intervenir en cas de perturbation selon l'ambiance sonore tels que les fonds musicaux ou les voix multiples [54]. Nous retrouvons ensuite la *dimension incommode* de la commande vocale selon des situations spécifiques. Réaliser une liste d'achats par exemple soulève certaines limites ainsi qu'un défaut d'interprétation dans la commande vocale qui entache l'expérience utilisateur : « *Ajouter un élément dans la liste de course est facile et rapide, mais ajouter plusieurs articles devient vite répétitif et coûteux temporellement. Vous devez attendre la confirmation verbale avant de répéter le procédé (...) Après le troisième article, vous commencez à penser qu'il pourrait sûrement y avoir un moyen plus rapide d'y parvenir. Vous pouvez raccourcir les étapes en ajoutant plusieurs articles en même temps, mais la liste finit par afficher une chaîne de texte comme si cela représentait un seul article* » (op. cit.). Enfin, nous identifions une dernière dimension liée au *collectif*. Si l'assistant vocal ne dispose pas de la capacité à distinguer les différentes voix du foyer, son partage auprès de plusieurs membres peut empêcher la personnalisation du système à l'utilisateur cible et ainsi modifier les futures recommandations et la compréhension du contexte de l'usager [11].

Par ailleurs, certains utilisateurs évoquent une déception et une mauvaise expérience utilisateur lorsqu'ils découvrent les **limites de certains « skills »** d'assistants vocaux ne supportant pas les requêtes complexes impliquant un ensemble de facteurs : « *En lançant l'application « tremblement de terre », on ne peut pas demander le plus gros tremblement de terre qui a eu lieu la semaine passée et l'application envoie une information erronée* » [48].

3.2.3 Renoncement à certains principes éthiques : confidentialité, capacité d'agir, contrôle et diversité servicielle

L'utilisation des interfaces vocales, et plus précisément celui des assistants vocaux, invite à questionner la dimension éthique des données. En effet, cette problématique met en exergue la confidentialité et la sécurisation des données mises à mal par ces technologies qui doivent cumuler une quantité massive de données personnelles pour fonctionner efficacement [11]. Dans un souci d'efficacité du dispositif et de contextualisation des services, les utilisateurs sont encouragés à **renoncer à la confidentialité** totale de leurs données dont on recommande un stockage permanent sur les serveurs afin d'assurer une mémoire des actions et des requêtes [43, 11, 32]. Certains utilisateurs évoquent une intrusion dans l'intimité liée à l'enregistrement continu des activités et des paroles, un inconfort à diffuser certaines requêtes et l'absence d'un bouton mute à part entière sur certains dispositifs : « *Je n'aime pas parce que j'ai la sensation d'être toujours écouté, et je dois rester prudente sur ce que je lui dis* » [31]. Cette question de la confidentialité est également soulevée en situation d'utilisation partagée d'un assistant vocal au sein du foyer : ce qui pourrait occasionner l'intervention spontanée du dispositif pour

proposer un produit ou un service dans un moment inadapté ou en présence d'autres membres ou de personnes extérieures. Ces situations d'utilisation nous amènent ainsi à distinguer deux types de confidentialité à intégrer en conception : il s'agit de la *discrétion* – renvoyant au souhait de ne pas être entendu des autres personnes lors des interactions vocales (échelle locale) – et de la *privauté* – suggérant la protection et la sécurisation des données personnelles stockées par les dispositifs (échelle plus globale) –.

Les publics âgés mettent notamment l'accent sur certains risques inhérents à la commande vocale, celui de développer une **passivité et une oisiveté** dans les tâches quotidiennes qu'ils préfèrent prendre le temps de réaliser manuellement [51, 43]. Il est donc impératif de privilégier des tentatives d'aide à la personne qui ne réduisent ni le sens du contrôle, ni leur propre estime [49]. Les travaux de Portet et al. [43] sur les systèmes vocaux en situation domestique soulignent bien l'importance de préserver auprès des personnes âgées ce sentiment de contrôle et leur capacité à agir en toute autonomie : « *J'aime bien agir plutôt que parler (...) J'aime bien fermer mes volets, etc. (...) Moi en ce moment je préfère faire les choses parce que sinon c'est glisser vers l'inactivité. Faudrait vraiment que je puisse plus le faire, parce que sinon on fait plus rien, on se couche et puis voilà* » (op. cit.).

La **perte de contrôle** représente un autre frein à l'acceptabilité du vocal cité par certains utilisateurs. Ils évoquent une forme d'inconfort à perdre l'initiative et le contrôle de certaines fonctions ou utilisations d'objets domestiques : « *En utilisant le vocal, ça donne l'impression que c'est vous qui décidez mais quelqu'un d'autre décide aussi pour vous. Sur cet écran tactile, vous sentez que vous avez le contrôle à 100%, mais ici avec le vocal, c'est comme si vous attendiez qu'ils acceptent ce que vous demandez. L'inconvénient c'est qu'on ne se sent pas à 100% en situation de contrôle comme les autres interfaces. C'est comme si on sollicitait une chose à la maison et qu'on doit attendre que ce soit approuvé* » [31].

Des utilisateurs évoquent aussi un **manque de diversité de choix** dans les propositions spontanées et contextualisées de certains assistants vocaux [11, 8]. Davis [11] évoque plusieurs paramètres inhérents au dispositif vocal Alexa intégrant des recommandations (dont la logique de fonctionnement reste opaque pour les utilisateurs) sont proposées aux utilisateurs sans choix possibles : l'assistant vocal propose des achats à partir de l'historique de l'utilisateur, du choix d'Amazon (meilleures notations, meilleurs prix, etc.) et des items éligibles lors de la commande « *prime* ». Ce dispositif se retrouve donc dans une position privilégiée et unique pour dicter des choix de produits et des choix de médias (op. cit.). Ce fonctionnement s'appuie sur l'usage d'algorithmes de recommandation pouvant non seulement empêcher l'utilisateur de disposer d'une liberté de choix dont il disposerait en délaissant les interactions vocales, mais ces algorithmes peuvent notamment déployer une bulle filtrante autour des usagers. Ce qui les amènerait à consommer des produits

ou des services, à se divertir et à s'informer le plus souvent à partir des mêmes sources sans s'ouvrir à d'autres univers ou d'autres opinions [38].

4. Des principes de conception d'interaction à privilégier

En nous appuyant sur les éléments récoltés concernant la compréhension des formes d'utilisation de technologies vocales en termes de leviers et de freins dans l'utilisabilité et l'acceptabilité, nous proposons ici des lignes directrices pour alimenter la conception de systèmes d'interaction vocale à destination des utilisateurs dans de multiples contextes. Selon la diversité des utilisations appréciées et des difficultés rencontrées, nous recommandons de privilégier dans les interfaces vocales les dix propositions suivantes :

Feedback utilisateur : encourager un temps de réponse relativement rapide (entre la commande et le retour vers l'utilisateur) ; privilégier la présence d'un voyant lumineux ou d'une réponse vocale confirmant la réception de la commande vocale ; intégrer une transparence des sources utilisées pour les demandes d'information et les services de recommandation divers (pour les utilisateurs intéressés).

Mode d'interaction et multimodalité : concevoir l'interaction vocale en complémentarité avec les autres modalités d'interaction existantes pour favoriser une richesse et une diversité de commande et de feedbacks (exemple : vocal, télécommande et écran visuel) ; proposer des interactions vocales privilégiant la parole libre (mode de conversation routinier).

Adaptabilité : assurer une adaptation du système à partir de l'intégration du vocabulaire de l'utilisateur ; proposer un système qui apprend également de ses erreurs ; inclure la possibilité de former le système par l'utilisateur.

Confidentialité et sécurité : intégrer un bouton « mute » à part entière ; assurer une confidentialité des données récoltées et stockées sur un serveur protégé ; privilégier une utilisation multi-utilisateurs (profils, droits d'accès) ; permettre le contrôle, l'exploitation et la modification de ses données personnelles et rendre visible les données recueillies par le système.

Contrôle du système : inclure la possibilité aux utilisateurs de garder ou de reprendre la main sur le système en cas de dysfonctionnement (pouvoir utiliser les luminaires ou l'utilisation de volets électriques) ou selon les préférences des habitants au sein du foyer. Par exemple, il peut s'agir du débrayage, d'interagir de façon spontanée ou encore de prévoir des automatisations selon le niveau de maîtrise du système (via le mobile, le PC ou par commande vocale). Et surtout il s'agit d'informer de façon continue, anticipée ou instantanée les utilisateurs sur les changements d'états du système via un système d'alertes et de notifications pertinentes.

Personnalisation : offrir la possibilité de modifier les sources utilisées pour les recherches d'information et les services de recommandations ; programmer et

personnaliser les services avec l'interaction vocale (en plus des interactions via le mobile) ; proposer une personnalisation originale des objets connectés et des phrases d'activation ; permettre et faciliter l'utilisation de la programmation visant l'automatisation ou le déclenchement d'actions uniques ou regroupées selon des situations spécifiques avec les commandes dite IFTTT (« *If This, Then That* ») ; adapter le système aux contextes utilisateurs en intégrant la compréhension de l'implicite auprès des dispositifs (ex : inférer la position et l'activité des utilisateurs).

Qualité de capture sonore : favoriser une reconnaissance vocale selon la diversité des contextes ambiants via l'intégration de micro embarqués, de capteurs et de dispositifs d'annulation d'écho (pour compenser la position de l'utilisateur, l'acoustique de l'habitat et la présence de bruits de fond).

Ubiquité : favoriser une continuité de l'expérience entre les membres du foyer, leurs diverses préoccupations domestiques, les différents espaces empruntés, les nombreux dispositifs utilisés et les autres modalités d'interaction existantes en proposant un écosystème se déployant sur une variété de supports.

Multi-utilisateurs : privilégier un système de reconnaissance des voix au sein d'un même foyer (pour éviter le « détournement » du système par des voix extérieures : TV, radio, etc.) ; proposer un paramètre de contrôle pour l'utilisateur principal afin de garder la main sur ses propres données et déléguer s'il le souhaite une partie ou la totalité du contrôle du dispositif à d'autres utilisateurs.

Dimension humanisante : encourager la dimension humanisante lors des interactions pour les utilisateurs intéressés par cette expérience utilisateur. Cela consisterait pour l'utilisateur à choisir entre un « *assistant fonctionnel* » ou un « *assistant de compagnonnage* ». Ce dernier choix proposerait une variété de critères et d'attributs que l'utilisateur pourrait ajouter, ajuster ou enlever selon ses préférences : variation de voix ou de timbre, intégration de marques de politesse et/ou d'humour, sélectionner un type de personnalité pour l'IA, expression et reconnaissance des émotions, etc.

5. Discussion et perspectives

En définitive, les systèmes d'interaction vocale n'ont pas vocation à remplacer les autres systèmes et modalités d'interaction. En effet, nous avons observé que les utilisateurs préfèrent garder le contrôle de leurs dispositifs en ayant recours à diverses modalités selon la situation d'utilisation et les besoins exprimés. Bien que basique dans les formes d'utilisations déployées, l'interaction vocale vient surtout faciliter et renforcer des utilisations préexistantes dans les routines des habitants (écouter de la musique/actualités, allumer/éteindre les lumières ou son alarme, noter sa liste de course, etc.). Par ailleurs, nous remarquons que les interactions vocales sont majoritairement utilisées en situation de multitâches lorsque le regard et les mains sont affairés, et cela afin de poursuivre quatre principales logiques d'utilisation : la recherche

d'assistance, le divertissement, le pilotage d'objets connectés, ainsi que l'anticipation et l'organisation de son activité. Mais pour renforcer l'acceptabilité de ces systèmes, il est nécessaire de renforcer une expérience utilisateur intuitive et efficiente dans laquelle ces systèmes s'adaptent au vocabulaire propre à l'utilisateur, comprennent l'implicite de certaines commandes vocales et s'adaptent aux diverses situations d'utilisations. Les technologies vocales telles que les assistants vocaux favorisent notamment une expérience partagée par tous les membres du foyer, il est donc indispensable de penser les interactions et les utilisations vers une diversité de profils (adultes, enfants, âgés) et d'y intégrer une discrétion des données grâce à la différenciation des voix et une hiérarchisation des niveaux de contrôle de l'assistant en cas d'utilisation collective. Enfin, il est indispensable de penser le vocal comme une modalité d'interaction pouvant soutenir la continuité des activités.

La synthèse des utilisations et des retours d'expériences sur les technologies vocales et les assistants vocaux démontre bien l'importance de mettre à jour les activités réelles des utilisateurs, leurs préoccupations, ainsi que leurs logiques intrinsèques. Élaborées à partir d'une première exploration de la littérature sur les systèmes vocaux, les principes de conception d'interaction précédemment évoqués ne constituent qu'un cadre général pour intégrer les logiques d'utilisation et les besoins utilisateurs dans la conception de systèmes utilisant IA et reconnaissance vocale. Nous pointons l'importance de baser la conception et le développement de ces dispositifs sur une approche centrée utilisateurs, itérative et pluridisciplinaire intégrant :

1. L'analyse et la modélisation des activités en situation de référence où le futur dispositif sera/est déployé : celle-ci vise à contextualiser les besoins et les logiques d'utilisation des différents profils utilisateurs ;
2. Une définition des fonctionnalités afin de proposer un ensemble de scénarios d'usages (use-case) pertinents et des principes d'interactions à respecter pour le projet de conception ;
3. Des phases de maquettage et d'évaluation de ces fonctionnalités, formes d'interaction et scénarios. L'évaluation en situation réelle ou réaliste d'utilisation permet de valider les choix de conception et de connaître l'appropriation réellement faite du système.

En application de cette démarche, et en complémentarité avec l'analyse de la littérature ici présentée, nous réalisons actuellement des observations en situation réelle afin de modéliser le développement et les utilisations d'écosystèmes d'objets connectés dont des systèmes à base de reconnaissance vocale. Ces travaux de recherche en ergonomie contribuent à la conception de services et de dispositifs soutenant la maîtrise de l'énergie dans l'habitat dans le secteur énergétique adaptés au grand public et s'insérant de façon fluide et durable dans les activités et les routines des habitants.

Bibliographie

- [1] Alliance des Sciences et Technologies du Numérique (ALLISTENE). *Éthique de la recherche en robotique*. Rapport n°1 de la CERNA, Commission de réflexion sur l'Éthique de la Recherche en sciences et technologies du Numérique d'Allistene, 2014.
- [2] J. Barcenilla & J. Bastien. L'acceptabilité des nouvelles technologies : quelles relations avec l'ergonomie, l'utilisabilité et l'expérience utilisateur ? *Le travail humain*, vol. 72, (4), 311-331, 2009.
- [3] J. M. C. Bastien, D.L. Scapin & C. Leulier, Une comparaison des critères ergonomiques et des principes de dialogue ISO 9241-10 dans une tâche d'évaluation d'interface. *Revue d'interaction homme-machine*, 1 (1), pp. 33-63, 1998.
- [4] V. Bellotti & K. Edwards, Intelligibility and accountability: Human consideration in context-aware systems. *Human-Computer Interaction*, 162 (2-4): pp. 193-212, 2001.
- [5] M.E. Bobiller-Chaumon, M. Dubois & M. Retour, *E-banking: Nouveaux services, nouveaux usages, nouvelles compétences*. Paris, France, 2003.
- [6] E. Brangier, A. Dufresne, S. Hammes-Adelé, Approche symbiotique de la relation humain-technologie : perspectives pour l'ergonomie informatique. *Le travail humain*, 2009/4 (Vol. 72), pp. 333-353, 2009.
- [7] A. Charney, Amazon Alexa s'esclaffe à cause d'un problème d'écoute, 2018. *01net.com*, <http://www.01net.com/actualites/quand-amazon-alexa-s-esclaffe-et-terrifie-ses-utilisateurs-1391170.html>
- [8] S. Cobb, Amazon Echo and the Alexa dollhouses: Security tips and takeaways, 2017. *We live security.com*, <https://www.welivesecurity.com/2017/01/10/amazon-echo-alexa-dollhouse-security/>
- [9] Commission Nationale Informatique & Libertés (CNIL), *Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*. Synthèse du débat public animé par la CNIL dans le cadre de la mission de réflexion confiée par la loi pour une république numérique, Décembre 2017.
- [10] J. Coutaz & L. Nigay, Multimodalité et plasticité des interfaces homme-machine en informatique ambiante : concepts et espaces de conception. Florence Sèdes, Jean-Marc Ogier, Pierre Marquis. *Information Interaction intelligence - Le point sur le i* (3), Cépaduès Editions, pp.179-214, 2012.
- [11] B. Davis, The problem with voice user interfaces like Amazon Alexa, 2016. *Econsultancy*, <https://econsultancy.com/blog/68499-the-problem-with-voice-user-interfaces-like-amazon-alexa>
- [12] B. DeRenzi, N. Dell, J. Wacksman, S. Lee & N. Lesh, Supporting Community Health Workers in India through Voice- and Web-Based Feedback. *In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 2017.
- [13] M. Dubois & M. E. Bobillier-Chaumon, L'acceptabilité des technologies : bilans et nouvelles perspectives. *Le travail humain*, 72(4), 305, 2009.
- [14] A. Edip, Siri, Alexa, Google Home, M... Les assistants numériques, révolution ou gadget dangereux ?, 2017. *Capital.fr*, *Polemik*, <https://www.capital.fr/polemik/siri-alexa-google-home-m-les-assistants-numeriques-revolution-ou-gadget-dangereux-1230739>
- [15] M-F. Fortin. *Le processus de la recherche : de la conception à la réalisation*. Montréal : Décarie éditeur, 1996.
- [16] M. Fréjus, *Construire la coopération Homme-IA*. Document interne. Saclay : EDF R&D, 2017.
- [17] M. Fréjus, Élargissement et renouvellement des questions traitées par l'ergonomie dans le domaine du développement durable : retour sur 12 ans de travaux sur les activités domestiques et la maîtrise des consommations énergétiques. *Psychologie française*, à paraître.
- [18] FUTUR.E.S/Cap Digital, Futur.e.s # Humain X Machine : 3 dimensions de l'interaction illustrées par nos démos, 2017. *Blog.futuresfestivals.com*, <https://blog.futuresfestivals.com/futur-e-s-humainxmachine-3-dimensions-de-linteraction-illustrees-nos-demos/>
- [19] D. Gardner-Bonneau, *Human Factors and Voice Interactive Systems*. Kluwer Academic Publishers, Norwell, MA, 1999
- [20] H. Guillaud, Vers une éthique pour l'intelligence artificielle ?, 2016. *Internet actu.net*, <http://www.internetactu.net/a-lire-ailleurs/vers-une-ethique-pour-lintelligence-artificielle/>
- [21] A. L. Hamilton, How i use Alexa, 2017, *Love my Echo.com*, <https://lovemyecho.com/2017/01/03/how-i-use-alexa/>
- [22] R. Harris, *Voice interaction Design, Crafting the new conversational speech systems*. San Francisco: Morgan Kaufmann Publishers, 2005.
- [23] J.-B., Haué, J.-B. Intégrer les aspects situés de l'activité dans une ingénierie cognitive centrée sur la situation d'utilisation. *Activités*, 2004.
- [24] H. Hillard, Amazon. L'assistant personnel Alexa se met à rire tout seul et effraye les utilisateurs, 2018. *Ouest-France.fr*, <https://www.ouest-france.fr/economie/entreprises/amazon/amazon-son-assistant-personnel-alexa-se-met-rire-tout-seul-et-effraye-les-utilisateurs-5610264>
- [25] Institut National de Recherche en Informatique et en Automatique (INRIA), *Intelligence Artificielle. Les défis actuels et l'action d'Inria*. Livre blanc N°01. Coordonné par B. Braunschweig, 2016.
- [26] ISO 9241-11, *Exigences ergonomiques pour travail de bureau avec terminaux à écrans de visualisation (TEV) - Partie 11: Lignes directrices relatives à l'utilisabilité*, 1998.
- [27] L. Janin, L'intelligence artificielle : mythes et réalités, 2017. *Stratégie.gouv.fr*, <http://www.strategie.gouv.fr/point-de-vue/lintelligence-artificielle-mythes-realites>
- [28] B. Kinsella, 42 Percent of amazon echo owners have 2 or more devices, 2017. *Voicebot.ai*,

- <https://www.voicebot.ai/2017/06/22/42-percent-amazon-echo-owners-2-devices/>
- [29] F. Lefevre, Interaction vocale & intelligences artificielles : état des lieux et opportunités, 2017. *Medium.com*, <https://medium.com/@francklefevre/interaction-vocale-intelligences-artificielles-%C3%A9tat-des-lieux-et-opportunit%C3%A9s-3b45239ee4a7>
- [30] É. Loup-Escande, J. Burkhardt & S. Richir (2013). Anticiper et évaluer l'utilité dans la conception ergonomique des technologies émergentes : une revue. *Le travail humain*, vol. 76, (1), 27-55, 2013.
- [31] M. Luria, G. Hoffman & O. Zuckerman, Comparing Social Robot, Screen and Voice Interfaces for Smart-Home Control. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17). ACM, New York, NY, USA, pp. 580-628, 2017.
- [32] O. Metwally & B. Gregory B, Alexa, start living healthy! Workshop on Interactive Systems in Health Care (WISH), 2017, http://wish.gatech.edu/wp-content/uploads/metwally_omar.pdf
- [33] C. Munteanu & G. Penn, Speech-based Interaction: Myths, Challenges, and Opportunities. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, New York, NY, USA, pp. 992-995, 2016.
- [34] C. Nass & S. Brave, Wired for Speech – How Voice Activates and Advances the Human–Computer Relationship. *MIT Press*, 2007.
- [35] N. Nguyen, I Finally Bought An Amazon Echo And This Is What Happened, 2016. *Buzzfeed.com*, https://www.buzzfeed.com/nicolenguyen/amazon-echo-review?utm_term=.ftMWKOVwe#.jFDRxV3Wv
- [36] J. Nielsen, *Usability engineering*. Boston: Academic Press, 1993.
- [37] J. Paolino, Google Home vs Alexa. Two Simple User Experience Design Gestures that Delighted a Female User, 2017. *Medium.com*, <https://medium.com/startup-grind/google-home-vs-alexa-56e26f69ac77>
- [38] Pariser, *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Press, 2011.
- [39] H. Pasquier, Définir l'acceptabilité sociale dans les modèles d'usage. Vers l'introduction de la valeur sociale dans la prédiction du comportement d'utilisation. *Thèse de Doctorat en psychologie sociale*, Université de Rennes 2, 2012.
- [40] C. Pelletier & É. M.-F. Moreau, L'appropriation des technologies de l'Internet et ses facteurs critiques de succès : un défi de plus pour les PME ? *Revue internationale P.M.E : économie et gestion de la petite et moyenne entreprise*, vol. 21, 2, pp. 75-117, 2008.
- [41] P. Pestanes & B. Gautier, Essor des assistants vocaux intelligents : nouveau gadget pour votre salon ou fenêtre d'opportunité pour rebattre les cartes de l'économie du web ?, *Wavestone*, 2017.
- [42] R. Pinard, P. Potvin & R. Rousseau. Le choix d'une approche méthodologique mixte de recherche en éducation. *Recherches qualitatives*, Vol. 24, pp.58-82, 2004.
- [43] F. Portet, M. Vacher & S. Rossato, Les technologies de la parole et du TALN pour l'assistance à domicile des personnes âgées : un rapide tour d'horizon (Quick tour of NLP and speech technologies for ambient assisted living) [in French]. *JEP-TALN-RECITAL 2012, Atelier ILADI 2012: Interactions Langagières pour personnes Agées Dans les habitats Intelligents*, Jun 2012, Grenoble, France. pp.3-16, 2012.
- [44] Portolan, N., Nael, M., Renoullin, JL & Naudin, S. (1999). Will we speak to our TV remote control in the future? In: *Proceedings of the 17th international symposium on human factors in telecommunication*, HFT'99, Copenhagen.
- [45] S. Proulx, Trajectoires d'usages des technologies de communication : les formes d'appropriation d'une culture numérique comme enjeu d'une société du savoir. *Annales des télécommunications*, 57(3-4), pp. 180-189, 2002.
- [46] P. Rabardel. *Les hommes et les technologies, une approche cognitive des instruments contemporains*. Armand Colin, Paris.
- [47] Radiocentre, Getting vocal. How voice-activated devices are increasing radio listening & elevating audio branding, 2017. Full report, http://www.radioplayer.co.uk/sites/default/files/getting_vocal_-_final_report.pdf
- [48] D. Rasmus, Hey, Alexa: The Pros and Cons of My Amazon Echo Experience, 2016, *Iphonelife.com*, <https://www.iphonelife.com/content/hey-alexa-pros-and-cons-my-amazon-echo-experience>
- [49] J. Rodin, Aging and health: effects of the sense of control. *Science*, 233(4770): pp.1271- 1276, 1986.
- [50] B. Shackel, Human Factors for Informatics Usability. In B. Shackel & B. Richardson (Eds.), *Human Factors* (pp. 21-38). Cambridge: Cambridge University Press, 1991.
- [51] H. Soronen, M. Turunen & J. Hakulinen, Voice Commands in Home Environment - a Consumer Survey. In *Proceedings of Interspeech 2008*, pp. 2078-2081, 2008.
- [52] J. Theureau. *Appropriation 1, 2, 3 ou Appropriation, Incorporation & « Inculturation »*. Conférence journée Ergo-Idf Appropriation & Ergonomie, 16/06/11, Paris, 2011.
- [53] M. Turunen, A. Melto, J. Hella, T. Heimonen, J. Hakulinen, E. Mäkinen, T. Laivo, H. Soronen, User expectations and user experience with different modalities in a mobile phone controlled home entertainment system. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, New York, NY, USA, 2009.
- [54] K. Whinton, Voice Interaction UX: Brave New World...Same Old Story, 2016. *Nielsen Norman Group.com*, <https://www.nngroup.com/articles/voice-interaction-ux/>

AI techniques to analyse a social network on text, user and group level : application on Galaxy2

Guillaume Gadek^{1,2}, Alexandre Pauchet², Stéphan Brunessaux¹, Khaled Khelif¹, Bruno Grilhères¹

¹ Airbus, 78990 Elancourt, France

² Normandie Univ, INSA Rouen, LITIS, 76000 Rouen, France

guillaume.gadek@litislab.eu

Résumé

Les Réseaux Sociaux Numériques ont pris une place prépondérante dans l'espace informationnel, et sont souvent utilisés pour la publicité, le suivi de réputation, la propagande et même la manipulation, que ce soit par des individus, des entreprises ou des états.

Alors que la quantité d'information rend difficile son exploitation par des humains, des solutions sont accessibles pour aider les preneurs de décision, solutions provenant notamment de techniques d'IA pour obtenir du renseignement à partir des messages postés, pour qualifier les comportements des utilisateurs, et pour identifier la structure sociale émergente.

Dans cet article, nous illustrons l'exploitation de telles techniques sur un réseau social très particulier, nommé Galaxy2, caché sur le Dark Web.

Nous proposons une analyse de 1000 jours d'activité, extrayons des thématiques et découvrons des acteurs-clés. Nous procédons ensuite à une étape de profilage des comptes utilisateurs, basée sur leur comportement. Finalement, nous introduisons des scores d'influence et de cohésion pour les groupes d'utilisateurs, ce qui aide à les caractériser et à les évaluer.

Abstract

Online Social Networks have taken a huge place in the informational space, and are often used for advertising, e-reputation, propaganda, or even manipulation, either by individuals, companies or states.

As the quantity of information makes the human exploitation difficult, solutions to support the decision makers can only come from the use of AI techniques to extract intelligence from posted messages, to qualify the user behaviours, and to identify the social structure.

In this article, we illustrate how to exploit such techniques on a very peculiar social network, named Galaxy2, hidden in the Dark Web.

We propose an analysis of 1000 days of activity using NLP techniques to find the most interesting topics and

to discover key actors. We then proceed with a ML-based profiling of the user behaviours. Finally, we introduce influence and cohesion scores for groups of users, which help their characterisation and evaluation.

Keywords

Deep Web, Social network analysis, User profiling, Group discovery, Graph analysis

1 Introduction

Web 2.0 and the social media have introduced a new paradigm in the informational space. Its specificities as a broadcast media enable anyone to be a new source of information, entertainment, or propaganda, at will. It gives power to the consumer, signalling defects on products; it gives the opportunity to the citizen to directly chat with his/her representatives; overall, the social web enables the user to be *more* connected and informed.

However, the presence of threats on Twitter and other social platforms is attested : the jihadist group Daesh is said to benefit from the support of 100,000 accounts on Twitter [1]; Russia is suspected by the FBI of interferences in the 2016 US election; in almost every election, worldwide, the intensive use of deception has been widely observed during the last years. Automation enables any party to produce and diffuse content quite effectively, if we consider the appearance of malicious accounts in the classic media a victory.

The diversity of threats implies a multiplicity of the tools to tackle all of the aspects of the problem. The quantity of messages exchanged in an Online Social Network can be astounding, with almost 500 million new tweets each day. As is possible the automation of their publishing, it is also possible to automate the reading. Opinion mining and topic detection enable the analyst to grasp the trend on a mass of texts, providing a classification along various keywords and characteristics.

The user accounts can be classified according to various taxonomies. In this article we distinguish the profile (who the user claims to be), and the behaviour (how s/he performs actions). Unsupervised machine learning helps here to cluster the users along similar types, and enables the analyst to recognise similar situations.

The never ending list of “friendship” connections, as well as the list of interactions between user accounts, is fertile to grow graphs. This kind of graphs can be used to easily compute social influence scores [2], and to detect one of the most interesting features of a social network : people naturally gather in groups, where there is more continuous interaction. To detect these groups from graphs, we rely on community detection algorithms [3]; specific measures help the analyst to characterise the impact and the pertinence of such groups with regard to his preoccupations [4].

Galaxy2, on Tor, is a strange place. Created in 2015 to replace a predecessor social network (simply called *Galaxy*), it has been branded as the most frequented social network on Tor. In this article, we show the added value of a variety of tools to scan this very peculiar social network, analyse the topics of discussion during Galaxy2’s uptime, detect its key-actors, and discover and characterise its communities. Finally, using all this information, we perform a case study, showing the interest of dedicated tools for social network analysis.

This paper is organised as follows : Section 2 presents the website and the context of TOR, Section 3 presents the tools used, Section 4 shows the case study ; Section 5 proposes a discussion about our results and concludes this article.

2 Presentation of Galaxy2

2.1 The TOR network

Partly developed through DARPA funding in the 90s, Tor, *The Onion Router*, has been launched in 2002. Aiming to bring anonymity to the data flow, it relies on an encryption and routing protocol, named *onion*, to hide the content of the packets to the transiting servers.

Tor is used either to anonymously access the clear web, as well as to access the so-called *darknet* : some websites are only accessible through the *onion* protocol, protecting the host. Well-known examples include a Wikileaks portal, and a disrupted illegal marketplace, *the silk road*.

As anyone can use the Tor network and enjoy the increased privacy, it helps to breach censorship and can be used by journalists, political activists or anyone else. Unfortunately, Tor is also used to host illegal content, such as hacker forums, drug marketplaces, dark forums, porn and pedo-pornography ¹.

1. [https://www.fbi.gov/news/stories/playpen-creator-](https://www.fbi.gov/news/stories/playpen-creator-sentenced-to-30-years)

However, due to the perceived complexity of use and the strong incentive to link one’s Internet profile with his real life (shopping, all kinds of services...), the global usage of Tor is not very high. Moreover, some countries deploy firewalls, laws and rules to prohibit this kind of tools.

2.2 Galaxy2 : history and main features

Founded in 2015 after the disruption of a previous Tor-hosted social network (“Galaxy”), Galaxy2 is based on an open source framework named *elgg*, enabling to build small social websites. According to its anonymous founder, *Lameth*, “*The server broke down and your terrible host here (me, not the current host, mind you !)* hadn’t been keeping regular backups off the server.” The service is shut down since end of October 2017².

The main features of Galaxy2 include The Wire, a space for microblogging posts ; blogs, polls and pages ; image and file sharing. Because of the anonymity introduced by the Tor network, the users are less keen to use their real names, and do not expose personal data. Thus, only the private direct messages are kept private, and it was possible to consult all the posts, friendship connections, images and comments performed by the users over the almost three years of uptime. The collection of our corpus covers a time range from Galaxy2 creation, on 9th of January, 2015 to the 22nd of September, 2017. It includes all traces of activity logged and then publicly available, but does not consist of personal profiles nor photos, excluded of the collect.

3 Tools for social media analysis

3.1 Focused on text processing

The first axis of analysis deals with text processing. The tremendous amount of new messages emitted through the Online Social Networks pushed the development of machine text understanding and summary. Moreover, as users often express their feelings and opinions about products, facts and events, one of the most necessary tasks is opinion mining, which can briefly be explained as the automatic computation of a couple (*topic, sentiment*) matching a message [5]. The sentiment, or polarity, enables the evaluation of the support or reject of an entity by the author of a message. It can be either computed specifically with a topic, using a learning corpus, or generically using linguistic resources. Such a resource, SentiWordNet [6], associates polarities with sets of synonyms, sometimes considering contextual associations to modify the polarities according to the surrounding sentence. An ana-

sentenced-to-30-years

2. <https://socialmediaalternatives.org/archive/collections/show/10>

lyser, Vader [7], has been made available as an open source module, and compiles various resources and a classifier to compute sentiment scores, at document level.

Topic detection can be performed either through a supervised or an unsupervised method. An analyst can describe a topic using a few keywords, or rely on text clustering to let an algorithm cluster similar texts together, based on the semantics. Instead of matching pre-established categories, this approach adapts well to new corpora, enabling the analyst to grasp the topic distribution.

Often cited as a reference for text clustering, LDA (Latent Dirichlet Allocation) [8] can briefly be resumed as a dimensionality reduction applied on a matrix where each line consists of the tf*idf representation of a text of the training corpus. It is said adapted to short text topic classification and can be associated with a module of visualisation of the detected topics, to better depict their distribution [9].

3.2 User profiling

One of the main tasks of social network analysis, or social listening, aims to spot the key actors, i.e. the most important users of the network. The focus is commonly directed either towards the *influence* [10, 11], defined as the ability of a user to trigger actions from other users, or towards community detection [12, 13].

Influence is commonly defined as the ability to *engage*, i.e. to make the other users do something (typically, sending a message or clicking on a URL). A complete method has been exposed to build a social graph from Twitter follow links [2]. This enables the computation of a daily PageRank, as a score of influence. However, using the public free API, a very long time (counted in months) is needed to obtain the complete social graph. To evaluate influence, various indicators have been compared, among which the number of retweets, number of followers and PageRank computed on a graph of friendship links [10, 14].

Sometimes, the focus is not set on the most influential accounts, but on a general profiling to help the analyst understand “who” is a given account : what its typical behaviour is. Machine learning tools can answer this question using numeric features, through clustering or classification. When the data represent a graph, specific algorithms, such as RolX [15], enable a tagging of the nodes based on their graph features.

3.3 Group detection

Studies in psychology highlight the power of the “group” as a structure, and the variety of its impacts on individual behaviour [16, 17]. In computer science, a study on polarisation of the attitudes on online debates concludes similarly : groups can be considered as influencers, too [18].

Detection. To detect these groups, the main approach relies on community detection from friendship relation graphs. A community is a set of nodes, in a graph, more densely linked than a random graph. An intuitive way to discover groups is to look for communities in a graph built from the social network data.

However it has been pointed out that the social graph (“follow” graph) does not illustrate user interaction [19]. As a response, random walks community detection have been applied on a *graph of interactions* between Twitter users (retweets and mentions) [20].

To help choosing the best community detection method, a comparison of various state-of-the-art algorithms [21] on large networks (around 400k nodes) has been performed between FastGreedy [22], InfoMap [23] and Louvain/Blondel [3]. In this experiment, the Louvain method seems to be, by far, the fastest, for similar results in terms of quality.

Characterising the strength of the groups. Topological graph features such as the modularity value or the density of the detected groups can be considered to evaluate a community detection algorithm. A comprehensive review of scoring functions for community evaluation details the following measures [24].

Internal density (d) follows the intuition that communities are more dense than a random set of nodes. However, for large, non-complete communities, its value may be very low. On the other side, the smallest possible communities (pairs of nodes) have an internal density of 1.

Triad participation ratio (TPR) evaluates the number of nodes belonging to triads, or triangles, in community S . A value of 1 means that the group is strongly internally linked. Communities should include numerous triangles, showing interaction between the members.

Conductance (c), evaluates the quantity of edges linked with the other communities [25]. It illustrates the community behaviour, ranging from 0 (not linked towards the exterior) to 1 (strongly connected to other groups).

Modularity (Q) focuses on the number of internal edges (which link members of a community), compared to a “normal” quantity if the graph was random. Introduced by [26], a high modularity Q , close to 1, denotes a good partitioning at the graph level.

These topological metrics enable the analyst to check whether a community is sufficiently connected internally and/or with its boundaries : internal interaction and reach towards the exterior are measured.

Characterising the topical cohesion and relevance of the groups. To measure each of the obtained groups based on the topics expressed by the users, two topical metrics ξ and ρ , inspired from machine learning (ML) precision and recall, have been proposed [4].

These two ML inspired topical measures, which enable to better attribute a weight of the importance of a community on a given topic (ρ), and to evaluate the internal cohesion of a group (ξ), make the link between the graph and the semantic of the texts. Another measure ($\theta f.igf$) [4], similar to the natural language processing domain $tf * idf$, evaluates the specificity and the importance of a topic in a group.

The analysis of these scores yields interesting insights on a social and textual corpus. For marketing purposes, one can identify then target a customer group, helping to induce a positive opinion about a new product.

3.4 Discussion and contribution

Although much work has been done to tackle specific challenges in the domain of Online Social Networks, only a few systems and frameworks propose a complete analysis of a website. Most of the platforms focus on the graph and miss the texts, such as UCINET, Net-Draw, Pajek and ORA, presented in [27], or a work on graph characterisation [28].

As in a mirror, a set of efficient textual analysis tools is available, enabling specific platforms for news monitoring [29] or for disaster management and evaluation [30]. However they lack of user profiling features and social graph exploration and analysis.

In this paper, we propose a full-stack processing pipeline, integrating various open-source modules and providing our own model for specific aspects. We perform the textual analysis on the text, model the users and propose an influence ranking, and conclude with exploration through graph analysis and community detection and characterisation.

4 Case study

In this section, the Galaxy2 network is exposed through the explanation of the collected data type. Then, results are proposed to analyse the texts, the user accounts and the groups, which are an emerging social feature.

4.1 Types of actions logged

Fortunately for the privacy, not every action can be retrieved. In the collected data, actions are regrouped to better grasp their diversity : first are the comments. Users can add some words to react to the other’s actions : publication of photos, files, new pages or polls, and of course blog posts. A total of 4032 comments on blog posts are stored ; much less (375) on photos.

A second kind of actions gathers the upload of files and images : those are deemed to be shared, and frequently include pictures from the Anonymous. As Tor and Galaxy2 philosophies are not based on personal holidays pictures publication, the quantity of images is quite low (around 600 pictures, in total).

Creation of pages, groups and polls and their modification (new topics of discussion) are seen as an aspect of the network life. This step of page creation is to be linked with the cluster of votes, replies and publication on said pages/groups. Overall, 1944 blog posts have been emitted, and 548 groups created.

Blogging or “micro”-blogging features were the predominant ones, with 29,000 posts on The Wire, the message feed. Message size was not limited, though frequently short.

Finally, the last cluster includes some connection information : account creation notifications (19233 “joined the site”), link establishment between a user and a page, a group or another user (61027 friendship links).

4.2 Text : topic and sentiment repartition

LDA for topic detection. LDA is commonly used with a high number of topics ; a number of 200-500 is usually recommended as a first try³. This method clusters documents with similar $tf.idf$ representations, resulting in a predetermined number of emerging topics, seen simultaneously as sets of documents and as $tf.idf$ vectors. As a trade-off between computation time and representativeness of the topics, a number of 40 topics is chosen. Compared to other values, this choice results in a distribution of the documents over the topics, also limiting the number of too-close topics.

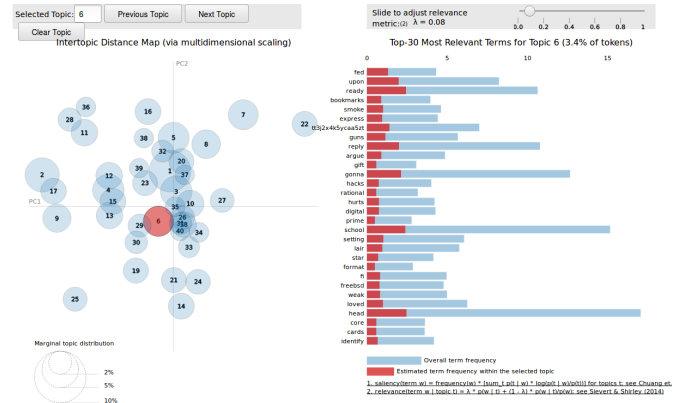


FIGURE 1 – Visualisation of the topics

Figure 1 shows a tool, proposed by [9], which enables to navigate along the 40 detected topics through the word frequencies in a given topic (in red), compared to word frequencies over the whole corpus (in blue), in the right panel. A principal component analysis (PCA) is performed to place the topics in a 2D-space, in the left of the diagram. The area of the circles is proportional to the quantity of documents clustered in the topics. Thanks to this tool, the analyst can visualise the topic distribution of, here, more than 30,000 documents.

3. <http://radimrehurek.com/gensim/models/ldamodel.html>

ments, while having the ability to see which words are more frequent or more relevant in each of the topics.

Sentiment exploitation. Each text is also analysed with regard to the *sentiment*, representing the overall polarity of a text, as perceived by a reader. The implementation relies on the *Vader* module [7]. Even if some of the sentiment predictions may be inaccurate, it results in a global view. For instance, Figure 2 proposes to compare the polarity of the texts containing some keywords : *onion*, *Tor*, *Snowden*, *Trump* and *Syria*⁴. Instead of an average value, and as one message is enough to set either a *max* or *min* sentiment value, we represent it as a boxplot, showing the statistical distribution. *Onion* is globally well perceived, with positive messages. Edward Snowden seems to enjoy a good reputation, while Galaxy2 members are more mitigated about Donald Trump. Finally, the polarity is more negative about Syria, because these messages are susceptible to mention the current conflict.

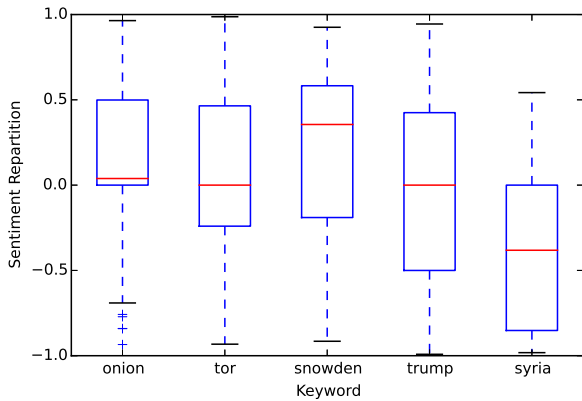


FIGURE 2 – Sentiment distribution for some keywords

The sentiment view can also be exploited to characterise a user activity. As an example, Figures 3 and 4 expose the sentiment repartition of the messages of two active authors : *XL33t* and *Fenris*. On both figures, an artefact, for a sentiment of 0.0, exists, representing the messages where no polarity could be extracted by the analyser (no known sentimental word). However, the repartition shows some signals. On the first figure, *XL33t* seems to mostly emit very positive messages. The second author, *Fenris*, covers the whole range of sentiment, avoiding to publish too positive messages. This view let the analyst grasp one aspect of a user behaviour, as well as to estimate his activity⁵.

4. No distinction has been made between uppercase and lowercase.

5. Thanks to the number of messages, indicated on the y-axis.

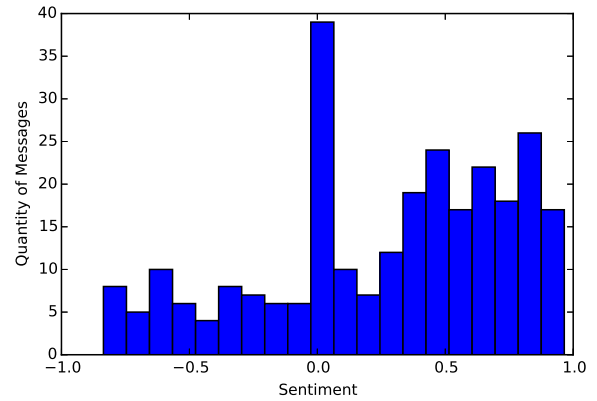


FIGURE 3 – Visualisation of the sentiment emitted by the user *XL33t*

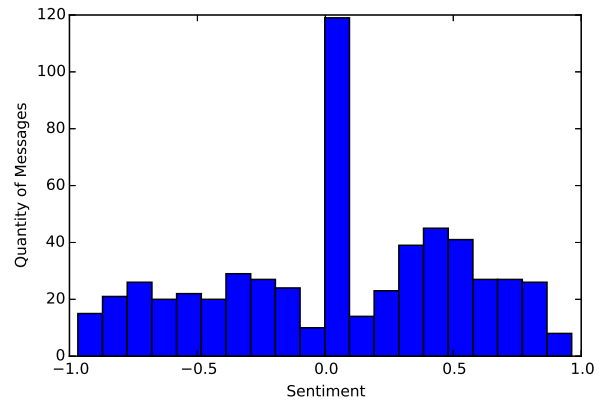


FIGURE 4 – Visualisation of the sentiment emitted by the user *Fenris*

4.3 Influential users and types of behaviour

Top-5 key users. Influence can be measured through a variety of prisms. In a first glance on a social network, one can look for the key actors, detecting a “top5”. Table 1 compares the most connected users (having the highest number of friends, noted *#Friends*); the most popular users (the most mentioned, noted *#Mentions*); the users of reference (a score based on the mentions graph, in the column *Sc(Mentions)*, explained in the following paragraph); and the most active ones (performing the highest number of actions on Galaxy2).

Inspired from the literature [10, 14], a score of influence is computed from G_M the graph of mentions. The intuition is the following : a mention gives some value, some social capital to the mentioned user ; this value is higher if the author of the mention already has some influence him/herself. The PageRank algorithm also follows this intuition and has been shown to

empirically match this representation of influence. In consequence, ScoreMentions is the ranking resulting from the Pagerank scores of each user-node in G_M .

TABLE 1 – Top5 of the users along three dimensions

#Friends	#Mentions	Sc(Me.)	#Actions
XsyntaX 8187	XsyntaX 514	Oxyy	XsyntaX 7974
prozac 3182	Oxyy 164	XsyntaX	kheper 1846
Spooky 3012	Lameth 121	Lameth	prozac 1783
xl33t 2140	Fenris 119	ChatTor	ChatTor 1583
kheper 1922	cpnemo 96	MahaKali	Spooky 1569

In a few words about Table 1 : *XsyntaX* outperforms every other account, with 7974 actions performed during the scope of the study ; the second, *kheper*, performed 1846 actions. *Lameth* is the founder of this social network ; as such he is often mentioned, either for thankful messages for hosting, or for administration purposes. *ChatTor* provides a chat service through Tor ; this account promotes the service and publishes news and updates.

Although some names are present on every influence score *Top5*, they do not all indicate the same kind of influence. *Lameth* is mentioned a lot, due to his status of host. However, he did not intend to be the friend of everyone else, or to be the most active. *xl33t* connected to many accounts, but is not mentioned as much as the other influential users. Influence and key-actors take many different forms, while only the symptoms can be measured.

Types of behaviour. The social network analyst cannot be expected to know who is every user on a platform. More precisely, the data of interest can be narrowed along a few aspects to grasp the essence of the behaviour.

Biography contains the user identity : username, ID, date of creation and other available data. *Style* explains how the user writes : length of messages and quantity of punctuation usually make a difference, and can be combined with the main topic of interest and global sentiment polarity. The *social* aspect details the number of friends, of mentions, and the resulting influence of the user, both from the graph of friendship and the graph of mentions. *Media* groups the features explaining the type of actions and objects posted (be it photos, texts, etc). Finally, the *temporal* aspect covers the global rhythm of publication of the user : average number of posts on TheWire per day, and average number of actions per day.

All these numerical features enable the repartition of the users along a few typical clusters, whose reduced number (commonly 4 to 6, depending of the data) enables the analyst to get used to this unsupervised labels.

Technically, this repartition is made in a few steps. First the data is cleaned, so as to avoid zero divisions and exponential distribution of the data : some features are converted to their log-values to reduce dis-

person, and normalised to follow a 0-mean, 1-standard deviation distribution. A PCA reduces the dimension of the problem, switching from 34 variables to 5 and still keeping most of the dataset variance. Finally, a k-means finds the groups ; a number of 4 groups is a good trade-off between dataset dispersion and cluster shape.

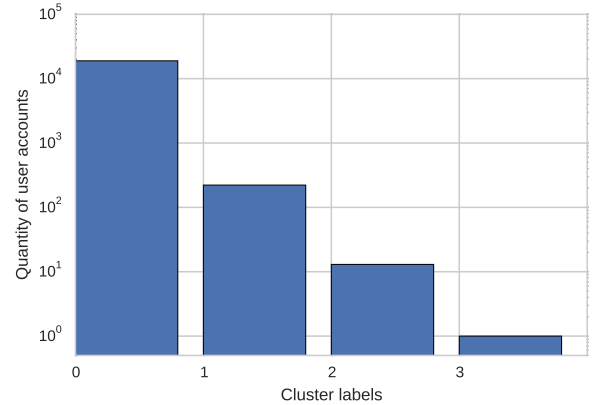


FIGURE 5 – Repartition of the user profiles

Figure 5 shows the repartition of the users along the 4 types of profiles. The first type (0) is the most common, including more than 90% of the profiles : most of the accounts are created, perform one or two actions, and then are forgotten. A second type, profile (1), characterises the regularly active users, even though they are not the central users of the network. Finally, profile (2) and the sole user in profile (3) are the central users, very active, those who produce most of the contents on the network. A similar repartition had already been remarked on networks such as Twitter [10].

To picture the shape of the profile clusters, Figure 6 represents each of the 19,177 Galaxy2 users along the first two dimensions of the PCA. These dimensions are linear combinations of the behavioural features, and consequently are quite abstract. However, positions near (0,0) are linked with very low levels of activities (the red cluster (0) being the most populated). In blue, cluster (1) shows the many different ways to be a standard active user. The green cluster (2) is very dispersed but matches high levels of publications, be it by posts, comments or creation of pages. Finally, the purple cluster (3) is constituted of only one account, *Spooky*, performing an average of 2.5 actions each day.

When the analyst reviews the posts, thanks to this step of profiling, he can quickly evaluate if the messages comes from well-established, influential users, or whether a given topic is dominated by usually inactive accounts, or show a unusual repartition.

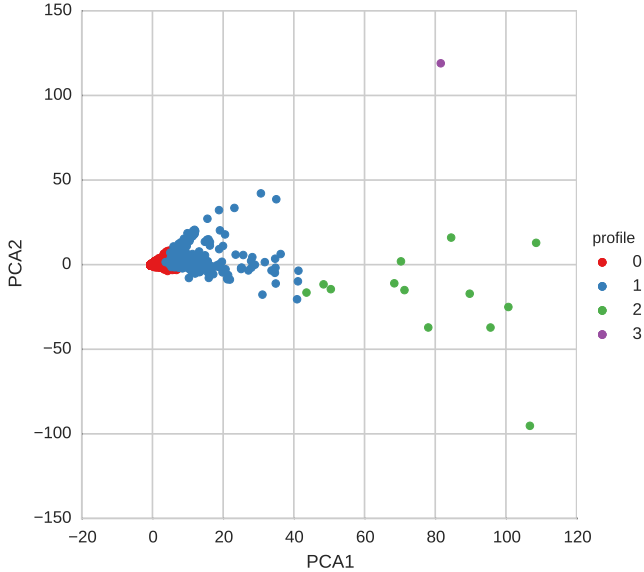


FIGURE 6 – Visualisation the clusters of users in a PCA space

4.4 Interactive groups, their cohesion and structure

About the user mentions. A common practice on social networks consists to mention people, using an “@” sign before their usernames in the text of the messages. At first this was only a social behaviour without any technical feature in the code of the social platform; nowadays most platforms recognise such *mentions*.

As Galaxy2 is based on the old-fashion social framework *elgg*, such a feature is not hard-coded, but is still used by some members. Following the intuition that this interaction (the mention) would bring us valuable information, a mention extraction module has parsed each of the publications in search for “@” signs followed by usernames.

From the set of mentions M associated with the authors A , a graph of mentions, $G_M = (V, E)$ is built. V is the set of the union of the authors and the mentioned users; edges $e \in E$ links the author of a message to the eventual mentioned user. This graph G_M is composed of 968 nodes, linked by 2342 edges⁶. The Louvain community detection [3] results on a set of 31 communities, amongst which only 11 contain more than 3 members.

Figure 7 illustrates one of these communities. Red nodes are part of the detected community; blue nodes are the only external contacts, the boundary of the community in the graph G_M . A central user, *Bishop*, is the target of a few mentions, notably from an auda-

6. there was a total of 5481 mentions : some accounts have been mentioned by a same user various times

cious account, *Nishikino_Maki*⁷. In a glance, we see that this community is centred on *Bishop*, who is not the most active in terms of community life (he does not mention the other members of the group).

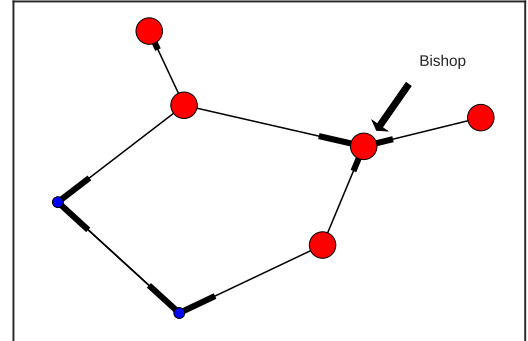


FIGURE 7 – Visualisation of a small community and its contacts through *mentions*

This community view helps the analyst to see the relations between a user and the rest of the network, and can guide his exploration along the interaction between accounts. However, as this graph of mentions is built from a social usage (using “@” to mention a user), the tool probably missed some differently written mentions and thus may be incomplete⁸.

The friendship graph. The most known feature of social networks is the friendship connection. On Galaxy2, one can claim his friendship with another user, which is considered reciprocal on Galaxy2. This action is part of the list of the collected activities, and enables to link users along these friendship relations. Let G_F be the graph of friendship, where nodes are user accounts. Edges represent a claim of friendship. From this graph, communities emerge, using the Louvain algorithm to reveal them. These communities may vary in number, based on the algorithm used. Here, 32 communities were obtained, amongst which 11 include more than 3 users. Figure 8 shows the distribution of the communities along two topological measures : x-axis gives the size of the group while y-axis represents the conductance, that is the proportion of boundary edges, linking the community to its environment, showing both its influence or its isolation.

On Figure 8, a group attracts the attention : with size=125 and conductance=0.74, this community is less tied to the rest of the network than it normally should. Its topical scores, based on [4], require some interpretation. $\xi = 0.064$ means that only 6,4% of its

7. He claimed to be “building a porn site cuz i can”

8. This graph G_M is *one of the* graphs of interaction, not the only one.

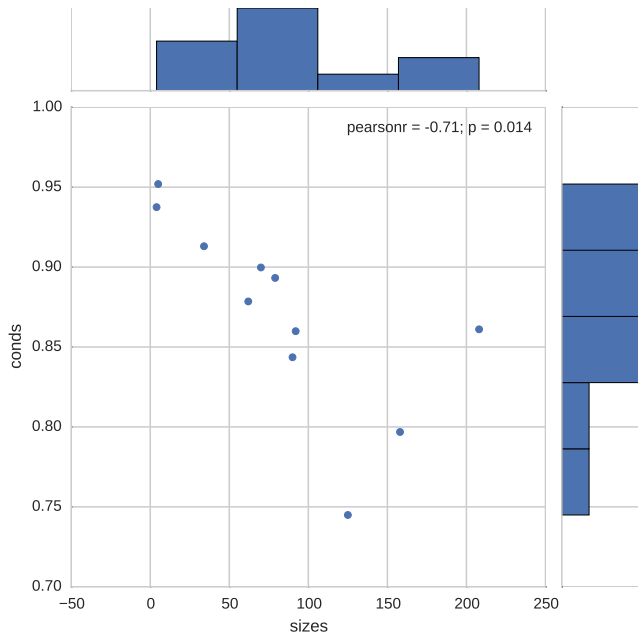


FIGURE 8 – Visualisation of the topological features of the communities

members, as a maximum, have been active about the same topic; however, compared with the duration covered on the corpus, this situation is not surprising. In particular, this “main topic” clusters mainly around two sets of texts at different dates. The first refers to a small pornographic discussion group; the second concerns the use of peer-to-peer mail through Tor. These topical measures may benefit from improvements on the topic detection part, notably covering the temporality of topics and their labelling. This would lead to a more usable topic modelling, and a more precise group characterisation for the analyst.

5 Conclusion

The emergence of modern social networks reveals the inherent complexity surrounding us. Although most of the analysis is focused on the most known Online Social Networks (OSN), such as Twitter or Facebook, other instances exist; Galaxy2 is much smaller than Twitter, but big enough to turn its analysis tedious due to the amount of exchanged messages.

We presented a whole system capable of analysing any type of OSN, showing its capabilities on a Tor-hosted social network. The functionalities include text analysis for topic and sentiment, user profiling and graph analysis, both on the friendship links as well as on the interaction between users. We showed that the tools, mostly developed for Twitter, can also be exploited for smaller networks. This approach has easily been adapted to the specific links and interactions of Galaxy2, as it could be for other platforms such as Reddit or

Facebook.

Future work will be oriented on real-time processing, to analyse a social network as it lives. We have the intuition that a stream of messages would better fit the requirements in scaling and response times. We also desire to publicly release the Galaxy2 corpus, as one of the first complete social network archives.

Références

- [1] G. Bertolin, *Digital Hydra : Security Implications of False Information Online*. NATO Strategic Communications Centre of Excellence, 2017.
- [2] P. Noordhuis, M. Heijkoop, and A. Lazovik, “Mining twitter in the cloud : A case study,” in *Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on*. IEEE, 2010, pp. 107–114.
- [3] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, “Fast unfolding of communities in large networks,” *Journal of statistical mechanics : theory and experiment*, vol. 2008, no. 10, p. P10008, 2008.
- [4] G. Gadek, A. Pauchet, N. Malandain, K. Khelif, L. Vercouter, and S. Brunessaux, “Measures for topical cohesion of user communities on twitter,” in *Proceedings of the International Conference on Web Intelligence*. ACM, 2017, pp. 211–218.
- [5] B. Pang, L. Lee, and S. Vaithyanathan, “Thumbs up? : sentiment classification using machine learning techniques,” in *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*. Association for Computational Linguistics, 2002, pp. 79–86.
- [6] S. Baccianella, A. Esuli, and F. Sebastiani, “Sentimentnet 3.0 : An enhanced lexical resource for sentiment analysis and opinion mining,” in *LREC*, vol. 10, 2010, pp. 2200–2204.
- [7] C. J. Hutto and E. Gilbert, “Vader : A parsimonious rule-based model for sentiment analysis of social media text,” in *Eighth International AAAI Conference on Weblogs and Social Media*, 2014.
- [8] D. M. Blei, A. Y. Ng, and M. I. Jordan, “Latent dirichlet allocation,” *Journal of machine Learning research*, vol. 3, no. Jan, pp. 993–1022, 2003.
- [9] C. Sievert and K. E. Shirley, “Ldavis : A method for visualizing and interpreting topics,” in *Proceedings of the workshop on interactive language learning, visualization, and interfaces*, 2014, pp. 63–70.
- [10] H. Kwak, C. Lee, H. Park, and S. Moon, “What is twitter, a social network or a news media?” in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 591–600.
- [11] A. Rao, N. Spasojevic, Z. Li, and T. DSouza, “Klout score : Measuring influence across multiple social networks,” in *Big Data (Big Data), 2015 IEEE International Conference on*. IEEE, 2015, pp. 2282–2289.
- [12] A. Java, X. Song, T. Finin, and B. Tseng, “Why we twitter : understanding microblogging usage and communities,” in *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*. ACM, 2007, pp. 56–65.

- [13] T. Aynaud and J.-L. Guillaume, "Static community detection algorithms for evolving networks," in *Modeling and optimization in mobile, ad hoc and wireless networks (WiOpt), 2010 proceedings of the 8th international symposium on*. IEEE, 2010, pp. 513–519.
- [14] C. Lee, H. Kwak, H. Park, and S. Moon, "Finding influentials based on the temporal order of information adoption in twitter," in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 1137–1138.
- [15] K. Henderson, B. Gallagher, T. Eliassi-Rad, H. Tong, S. Basu, L. Akoglu, D. Koutra, C. Faloutsos, and L. Li, "Rolx : structural role extraction & mining in large graphs," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2012, pp. 1231–1239.
- [16] S. Milgram, "Group pressure and action against a person." *The Journal of Abnormal and Social Psychology*, vol. 69, no. 2, p. 137, 1964.
- [17] G. L. Cohen, "Party over policy : The dominating impact of group influence on political beliefs." *Journal of personality and social psychology*, vol. 85, no. 5, p. 808, 2003.
- [18] S. Yardi and D. Boyd, "Dynamic debates : An analysis of group polarization over time on twitter," *Bulletin of Science, Technology & Society*, vol. 30, no. 5, pp. 316–327, 2010.
- [19] K. H. Lim and A. Datta, "An interaction-based approach to detecting highly interactive twitter communities using tweeting links," in *Web Intelligence*, vol. 14, no. 1. IOS Press, 2016, pp. 1–15.
- [20] Y. Tyshchuk, W. A. Wallace, H. Li, H. Ji, and S. E. Kase, "The nature of communications and emerging communities on twitter following the 2013 syria sarin gas attacks," in *Intelligence and Security Informatics Conference (IISIC), 2014 IEEE Joint*. IEEE, 2014, pp. 41–47.
- [21] R. Cazabet and F. Amblard, "Simulate to detect : a multi-agent system for community detection," in *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2011 IEEE/WIC/ACM International Conference on*, vol. 2. IEEE, 2011, pp. 402–408.
- [22] A. Clauset, M. E. Newman, and C. Moore, "Finding community structure in very large networks," *Physical review E*, vol. 70, no. 6, p. 066111, 2004.
- [23] M. Rosvall and C. Bergstrom, "Maps of information flow reveal community structure in complex networks," Technical report, Tech. Rep., 2007.
- [24] J. Yang and J. Leskovec, "Defining and evaluating network communities based on ground-truth," *Knowledge and Information Systems*, vol. 42, no. 1, pp. 181–213, 2015.
- [25] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 8, pp. 888–905, 2000.
- [26] M. E. Newman, "Modularity and community structure in networks," *Proceedings of the national academy of sciences*, vol. 103, no. 23, pp. 8577–8582, 2006.
- [27] S. F. Everton, *Disrupting dark networks*. Cambridge University Press, 2012, no. 34.
- [28] D. Ediger, K. Jiang, J. Riedy, D. A. Bader, and C. Corley, "Massive social network analysis : Mining twitter for social good," in *Parallel Processing (ICPP), 2010 39th International Conference on*. IEEE, 2010, pp. 583–593.
- [29] J. Sankaranarayanan, H. Samet, B. E. Teitler, M. D. Lieberman, and J. Sperling, "Twitterstand : news in tweets," in *Proceedings of the 17th ACM SIGSPATIAL international conference on advances in geographic information systems*. ACM, 2009, pp. 42–51.
- [30] F. Toriumi and S. Baba, "Real-time tweet classification in disaster situation," in *Proceedings of the 25th International Conference Companion on World Wide Web*. International World Wide Web Conferences Steering Committee, 2016, pp. 117–118.

Contribution à l'identification par logique floue d'un échangeur de chaleur à tube concentrique

Jacques KADIMA KAZAKU¹

Jimmy KALENGA KAUNDE KASONGO²

Richard NGENDA BANKA³

François NTAMBWE KADINDA⁴

^{1,2,3,4} Université de Lubumbashi/Polytechnique/Electromécanique

¹jacqueskazaku@gmail.com

^{2,3}{jimmykalenga,rickngenda}@yahoo.fr

⁴fntambwe@gmail.com

Résumé

Dans ce manuscrit nous traitons le problème lié à la modélisation d'un échangeur de chaleur pour des fins de surveillance. Ce système généralement décrit par des équations aux dérivées partielles obtenues à partir des lois de la physique, possède des paramètres mal connus et des paramètres variant conséquences des échanges thermiques dans l'échangeur. Ces observations traduisent la complexité du comportement dynamique de l'échangeur de chaleur. Par conséquent, l'obtention d'un modèle fiable de ce procédé est une tâche ardue. Pour contourner cette difficulté, une alternative est de recourir aux modèles flous de Takagi-Sugeno qui permettent une meilleure approximation des non linéarités. Les algorithmes d'identification par les méthodes des moindres carrés combinés aux techniques de classification des données entrées-sorties obtenus sur l'échangeur, ont fournis une bonne approximation du modèle global.

Mots Clef

Modélisation, identification floue, échangeur de chaleur.

1 Introduction

Cet article s'inscrit dans le domaine de supervision des systèmes. Il traite plus particulièrement de la modélisation d'un échangeur de chaleur en vue de sa surveillance.

Les lois de surveillances sont étroitement liées aux méthodes d'identification qui permettront d'obtenir le modèle de référence, et la conception des algorithmes de surveillance. De ce fait le modèle identifié doit être suffisamment représentatif et moins conservatif pour permettre la détection et le diagnostic des défauts.

Il est connu que le modèle fin d'un échangeur de chaleur thermique est composé d'équations aux dérivées partielles [7,17] qui en général sont difficile à manipuler, obtenues en partant des lois de la physique. Une des méthodes permettant de réduire le modèle sous

la forme d'équations différentielles ordinaires consiste à faire un découpage de l'échangeur en des volumes supposés homogènes (la température dans chaque volume est supposée uniforme) et d'appliquer le principe des bilans thermiques en considérant des échanges par conduction et par convection à travers la surface séparant les deux liquides traversant l'échangeur. Les modèles mathématiques obtenus en passant par ces simplifications sont moins satisfaisants pour des besoins de commande, observation, filtrage ou surveillance. Une façon de contourner la difficulté est d'établir un modèle de ce processus et de recourir au formalisme de Takagi-Sugeno (TS) [8]. Disons que ce formalisme est mieux adapté pour la construction de modèles non linéaires (NL). Les modèles TS peuvent être obtenus de plusieurs manières [12, 20,21]. Dans notre cas on se limitera à trouver le modèle mathématique de l'échangeur de chaleur à partir des données entrées-sorties (identification de type boîte-noire) en se basant sur des techniques de clustering [18] de Gustafson-Kessel (GK).

Pour modéliser le comportement dynamique de cet échangeur de chaleur en appliquant les techniques de clustering flou, le procédé à identifier peut-être représenté comme un système à plusieurs entrées et plusieurs sorties (MIMO), ce modèle MIMO ou globale peut être obtenu de deux manières [4,13]. Dans le but de réduire la complexité du modèle flou, le système MIMO à identifier sera décomposé en plusieurs sous-systèmes multi-entrée mono-sortie (MISO) de type TS. Ceci aura comme effet une réduction du nombre des fonctions d'appartenances et des règles.

Ce manuscrit est organisé comme suit : la section II donne une description du procédé et présente le problème lié à sa modélisation en utilisant des lois phénoménologiques. Une méthodologie d'identification du modèle flou de l'échangeur de chaleur est présentée à la section III, les résultats obtenus sont présentés à la section IV.

2 Description du procédé et problème lié à sa modélisation

2.1 Description du procédé

L'échangeur dont il est question est à tube concentrique dans une configuration contre-courant, assurant le refroidissement de l'eau chaude par l'eau froide. Une vue d'ensemble du processus est donnée à la figure 1.



Figure 1 Vue d'ensemble de l'échangeur

Cet échangeur est composé de deux circuits fermés (eau chaude et eau froide) :

1) Circuit d'eau chaude

Ce circuit comporte des éléments chauffants et de l'eau contenue tous dans un vase de 1.5 litres. Les éléments chauffants sont composés de deux résistances de 1KW chacune connectée en parallèle sur une tension de 115V ou en série sur une tension de 230V, commandées en température par un thermostat de 75°C comme seuil de température. Au démarrage l'eau est chauffée à une température T_1 . Elle est ensuite aspirée et refoulée vers l'échangeur par une pompe volumétrique entraînée par un moteur à courant continu. Après échange elle est évacuée vers le vase pour un nouveau cycle. Le débit d'eau chaude Q_1 à l'entrée de l'échangeur est commandé par une électrovanne.

2) Circuit d'eau froide

Ce circuit est composé d'un groupe frigorifique assurant le refroidissement de l'eau à une température de consigne. Après refroidissement l'eau est acheminée à l'entrée de l'échangeur via une électrovanne et y arrive à une température T_6 . Le débit d'eau Q_2 à l'entrée de l'échangeur dépend essentiellement de la position de l'électrovanne. Après échange, l'eau sort à une température T_5 et est évacuée pour à nouveau être refroidie.

Cet échangeur est équipé des capteurs fournissant des signaux convertis en tension normalisés dans différentes plages, transmis vers le dispositif d'acquisition de données. Les principales grandeurs mesurables ainsi que leurs plages de variation sont :

- Les températures, toutes dans la gamme [-5 5V], ce qui correspond à [0 133°C],
- Le débit d'eau chaude dans la gamme [0 5V] = [0 25l/min],
- Le débit d'eau froide dans la gamme [0 5V] = [0 5l/min].

2.2 Modélisation de l'échangeur

Le modèle mathématique global de dimension infinie du système est décrit par des équations aux dérivées partielles [7,17]. Si les phénomènes de diffusion sont négligés, ce qui est le cas le plus fréquent, alors le modèle de l'échangeur est linéaire de type convection-réaction.

L'écriture du bilan énergétique pour le tube interne et externe de l'échangeur de chaleur conduit respectivement aux équations aux dérivées partielles linéaires suivantes [24] :

$$\rho S_c c_c \frac{\partial T_c(x,t)}{\partial t} = -\rho S_c c_c v_c \frac{\partial T_c(x,t)}{\partial x} + \alpha_c D_c (T_f(t) - T_c(x,t)) \quad (1)$$

$$\rho S_f c_f \frac{\partial T_f(x,t)}{\partial t} = \rho S_f c_f v_f \frac{\partial T_f(x,t)}{\partial x} + \alpha_f D_f (T_c(t) - T_f(x,t)) \quad (2)$$

Où les indices f et c symbolise respectivement le froid et la chaleur,

Avec:

- D : périmètre d'échange thermique (m)
- S : aire de section traversée par chacun de flux (m^2)
- v : vitesse du flux (m/s)
- ρ : densité du flux
- c : chaleur spécifique ($Kcal/m^3 \cdot ^\circ C$)
- α : coefficient d'échange ($Kcal/m^2 \cdot s \cdot ^\circ C$)
- T : température ($^\circ C$)

Les équations (1, 2), obtenues en posant au préalable certaines hypothèses sont difficile à manipuler, et montrent que les échanges thermiques dans les échangeurs constituent des phénomènes à paramètres répartis. Aussi disons que, ces équations comportent des paramètres variant dans le temps (les vitesses d'écoulement et les températures) et des paramètres mal connus (les coefficients d'échange thermique). Toutes ces observations traduisent la complexité d'obtenir un modèle fiable du procédé. Dans le but de trouver un modèle beaucoup plus représentatif de cet échangeur, nous recourons dans la section suivante à la modélisation floue de type TS à partir des données d'entrées-sorties de ce système, qui approxime mieux les

3 Modélisation floue de l'échangeur de chaleur

Dans ce manuscrit, le modèle flou TS essaie de résoudre un problème complexe lié à la modélisation de cet échangeur en le décomposant en plusieurs sous problèmes simples. Il s'agit d'approximer ce procédé non linéaire par plusieurs modèles linéaires, et ce à partir des données prises en entrée et en sortie du procédé.

La méthodologie générale pour la construction du modèle flou TS passe par les étapes suivantes :

1. Validation du nombre de clusters
2. Génération des fonctions d'appartenance des antécédents
3. Obtention des paramètres des conséquences
4. Validation numérique du modèle final.

La théorie des ensembles flous offre un excellent outil pour représenter l'incertitude associée à la tâche de décomposition, en fournissant des transitions douces entre les sous-modèles linéaires [11,13,8] et afin d'intégrer divers types de connaissance dans un même cadre. A ce propos plusieurs algorithmes de classification existent [18,14,16], et admettent tous comme paramètre d'entrée le nombre de classe. Dans notre cas nous allons considérer la méthode de coalescence (clustering) floue de GUSTAFSON-KESSEL (GK) [18] basée sur la minimisation itérative d'une fonction objective, afin de générer automatiquement la décomposition à partir des données entrée-sortie du système. L'avantage de cet algorithme par rapport aux autres est sa capacité à détecter des classes possédant des formes et des orientations différentes dans un seul ensemble donné.

Disons que les méthodes de clustering nécessitent la sélection a priori du nombre de classes. La figure 2 montre la répartition de l'ensemble de données obtenue en appliquant l'algorithme de GK pour deux groupes ($c=2$) et dont le paramètre flou m est pris égal à 2.

En se référant à la description du procédé et dans le but de trouver son modèle flou, ce système peut être représenté comme un système multi-entrée multi-sortie (MIMO) non linéaire autorégressif avec entrée exogène (NARX) à quatre entrées ($Q_1(k), Q_2(k), T_1(k), T_6(k)$) et deux sorties ($T_3(k), T_5(k)$).

Ce modèle peut être représenté de deux manières : soit la fonction non linéaire est une fonction vectorielle, soit le système MIMO est décomposé en plusieurs systèmes MISO [13] de type TS mais qui n'ont pas le même type de structure au niveau de leurs règles. C'est cette dernière représentation qui va nous intéresser dans la suite de ce manuscrit car, cette approche réduit le nombre de fonctions d'appartenances et de règles [4].

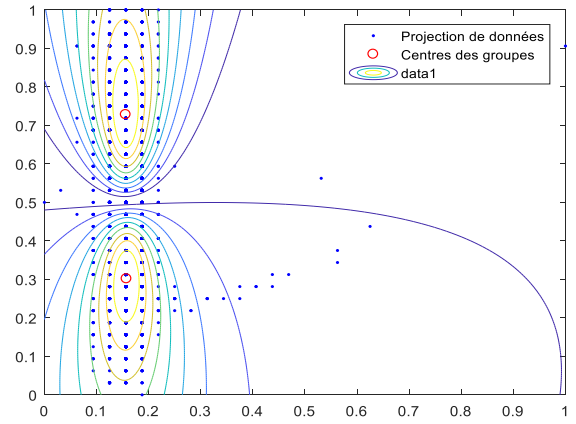


Figure 2 Résultat de l'algorithme de GK de regroupement de données

A partir de la connaissance préalable du procédé, nous avons choisi une dynamique pour les modèles flous TS telle que chaque sortie à l'instant $k + 1$ dépende des variables de sortie et d'entrée à l'instant k . Ce choix est dû d'une part, en supposant que le système est couplé et d'autre part, afin de maintenir un compromis entre une complexité réduite du modèle et une bonne capacité d'approximation.

Dans ces conditions, l'ordre du système, lié aux constantes n_y (ordre de la sortie), n_u (ordre de l'entrée) et n_d (retard pur), est exprimé alors par les structures suivantes :

$$n_y = 1, n_u = [1 \ 1 \ 1 \ 1], n_d = [1 \ 1 \ 1 \ 1]$$

Où les facteurs '1' indique la présence d'une dépendance fonctionnelle entre les variables [3,1,4].

La forme des règles pour la sortie $T_3(k)$ est donnée par l'équation suivante :

$$\text{Si } Q_1(k) \text{ est } A_{11} \text{ et } Q_2(k) \text{ est } A_{12} \text{ et } T_1(k) \text{ est } A_{13} \text{ et } T_6(k) \text{ est } A_{14} \text{ et } T_3(k) \text{ est } A_{15},$$

Alors

$$T_3(k + 1) = a_{11}Q_1(k) + a_{12}Q_2(k) + a_{13}T_1(k) + a_{14}T_6(k) + a_{15}T_3(k) + b_1$$

Celle en rapport avec la sortie $T_5(k)$ est :

$$\text{Si } Q_1(k) \text{ est } A_{11} \text{ et } Q_2(k) \text{ est } A_{12} \text{ et } T_1(k) \text{ est } A_{13} \text{ et } T_6(k) \text{ est } A_{14} \text{ et } T_5(k) \text{ est } A_{15},$$

Alors

$$T_5(k + 1) = a_{11}Q_1(k) + a_{12}Q_2(k) + a_{13}T_1(k) + a_{14}T_6(k) + a_{15}T_5(k) + b_1$$

Le modèle ainsi obtenu est linéaire.

Après obtention des sous-ensembles flous pour les antécédents, il faut estimer les fonctions affines pour chaque conclusion des règles. Les paramètres a_1 et b_1

partir de la structure géométrique des clusters ou par des techniques d'estimation. Deux approches existent [13] :

a) Critère global (GLS)

$$J_G = \sum_{i=1}^N (y(k) - y_{est}(k))^2$$

$y(k)$ et $y_{est}(k)$ sont respectivement les sorties du processus réel et du processus identifié. Les paramètres des conséquences a_i et b_i appartenant à la règle correspondant au i^{eme} cluster sont concaténés dans un seul vecteur de paramètres θ_i donnée par :

$$\theta = [a_1^T, b_1, a_2^T, b_2, \dots, a_k^T, b_k]^T$$

b) Critère local (c-WLS)

$$J_i = \sum_{k=1}^N \mu_{ik} (\hat{y}_{ik} - y_k)$$

En utilisant comme facteur de pondération des données, les degrés d'appartenance de la matrice de partition floue issus du processus de clustering. Cette approche conduit à une formulation de "c" problème indépendants de types moindres carrés pondéré. Les paramètres sont donnés respectivement par :

$$a_i = [\theta_1, \theta_2, \dots, \theta_p], \quad b_i = \theta_{p+1}$$

4 Présentation des résultats

Pour identifier le modèle flou de l'échangeur, nous lançons une séquence d'acquisition des données de 3400 échantillons sur chacune des variables manipulées, sur un temps d'échantillonnage d'une seconde. 2000 sont les données d'apprentissage pour la classification et l'identification des paramètres du modèle, et le reste pour tester la validité du modèle. Nous avons fait varier les deux variables de commande Q_1, Q_2 sur toute leurs plage de fonctionnement de manière à exciter toute la dynamique du système, et ceci pendant tout l'essai.

Pour estimer les paramètres des deux sorties T_3 et T_5 nous utilisons le critère d'optimisation local afin de privilégier la performance et l'interprétation locale des modèles.

Après identification, le modèle local de température du fluide chaud est :

$$T_3(k+1) = 0.4966T_3(k) + 0.772Q_1(k) + 0.0088Q_2(k) - 0.0198T_1(k) + 0.1821T_6(k) + 0.0309$$

Celui du fluide froid est :

$$T_5(k+1) = -0.2586T_5(k) + 0.6181Q_1(k) - 0.0253Q_2(k) + 0.0052T_1(k) + 0.0372T_6(k) + 0.356$$

On remarque que les modèles locaux obtenus pour chaque sortie sont du premier ordre. Les entrées sont un incrément de commande sur le débit eau chaude, le débit eau froide, la température eau chaude et la température eau froide. Ces modèles possèdent chacun un biais.

4.1 Validation graphique du modèle

La validation graphique consiste en un jugement visuel d'un ajustement graphique des données simulées sur les données observées.

Ainsi les figures 3 et 4 illustrent le comportement des deux modèles flous obtenus lors de l'apprentissage. On peut y observer que les dynamiques des sorties de l'échangeur de chaleur sont reproduites beaucoup mieux.

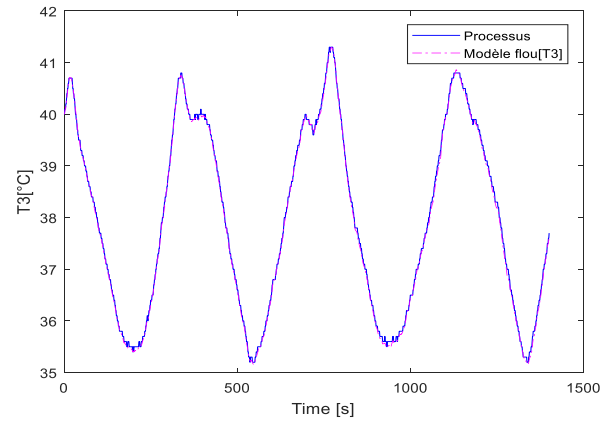


Figure 3 Validation hors ligne du modèle flou de T_3

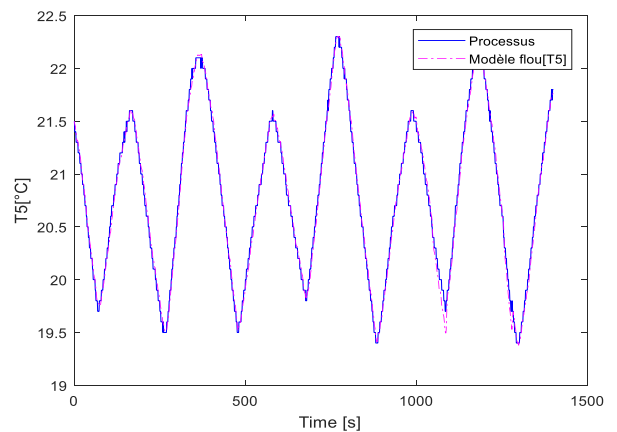


Figure 4 Validation hors ligne du modèle flou de T_5

Nous pouvons également faire un autre test de validation, mais cette fois en ligne, qui consiste à utiliser un autre jeu de données de signaux d'excitation différents que ceux employés pour l'identification.

Les figures 5 et 6 illustrent les évolutions des températures de sortie de l'échangeur de chaleur et du modèle flou obtenu.

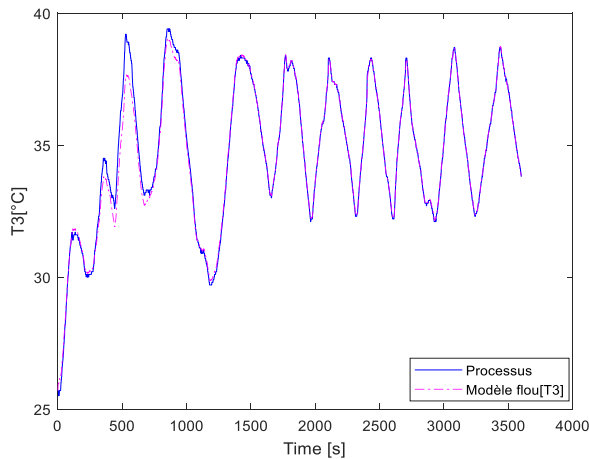


Figure 5 Validation en ligne du modèle flou T_3

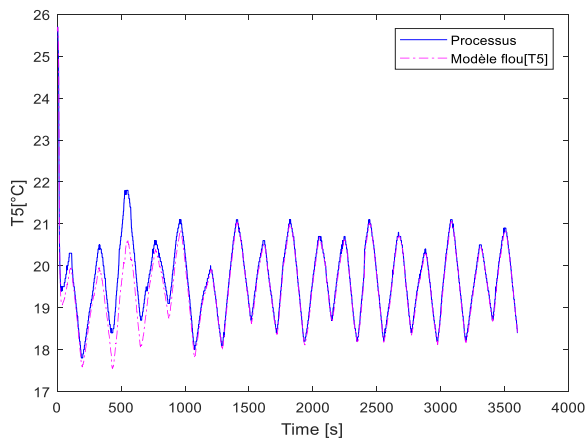


Figure 6 Validation en ligne du modèle flou T_5

On remarque sur ces deux dernières figures une divergence des grandeurs simulées en tout début, qui s'améliore avec au fur et à mesure que le système évolue. Ceci traduit la capacité d'adaptation en ligne de l'algorithme de GK.

4.2 Validation numérique du modèle flou

Pour évaluer la qualité de l'approximation obtenue par les modèles flous TS, nous utilisons l'erreur quadratique moyenne (RMSE) qui est une mesure globale sur le nombre total de points de l'écart par rapport à la valeur attendue définie par l'expression [13,4] :

$$RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^N (y_k - \hat{y})^2}$$

Et la mesure VAF (Variance Accounted For) décrite par l'équation :

$$VAF = 100\% \left[1 - \frac{Var(y - \hat{y})}{Var(y)} \right]$$

Où $1 \leq k \leq N$ est le nombre de point considérés pour la modélisation, y est la sortie mesurée et \hat{y} est la sortie du modèle. Les valeurs idéales pour ces deux critères sont respectivement 0% et 100%.

Le tableau 1 compare les performances numériques de deux approximations. On remarque que les résultats obtenus en phase de validation hors ligne sont meilleurs par rapport à ceux obtenus en ligne.

Approximation	Validation hors ligne	Validation en ligne
$RMSE(T_3)$	0.0956	0,3055
$RMSE(T_5)$	0.0632	0,3165
$VAF(T_3)$	99,9428%	98,7928%
$VAF(T_5)$	99,3052%	91,6111%

Tableau 1 : Performance du modèle globale

Les différents tests de validation montrent l'aptitude du modèle flou global obtenu à reproduire les dynamiques des sorties de l'échangeur de chaleur, et ce malgré les bruits de mesures.

5 Conclusion

Dans ce manuscrit, une stratégie d'identification floue de type Takagi-Sugeno d'un échangeur de chaleur tubulaire a été présentée. Un modèle flou dynamique pour l'échangeur est développé sur des données entrée-sortie issues du procédé en appliquant les techniques de clustering flou, dans le but de contourner la complexité de modéliser le procédé par des lois phénoménologiques. En particulier, nous avons utilisés l'algorithme de Gustafson-Kessel qui nous a semblé le plus représentatif et le mieux adapté pour la tâche de modélisation. Le choix du multi modèle TS est fait dans le but d'obtenir un modèle mathématique fiable du processus. Le système étant MIMO avec quatre entrées (débit d'eau froide, débit eau chaude, température d'eau froide et la température d'eau chaude à l'entrée de l'échangeur et la) et deux variables de sorties (températures d'eau chaude et d'eau froide à la sortie de l'échangeur) a été décomposé en deux sous-systèmes MISO, représentés chacun par un modèle flou TS. Les tests de validation des modèles flous obtenus ont été faits dans l'espace de données d'apprentissage et de validation. Les performances numériques du modèle global obtenu selon les critères VAF et RMSE sont encourageant.

Remerciements

L'auteur remercie le responsable du laboratoire de régulation de la faculté polytechnique de l'université de Jubbumbashi et son assistant, pour lui avoir permis

d'accéder au laboratoire pour l'analyse et différents essais sur l'échangeur de chaleur.

Bibliographie

- [1] J.A.Roubos, S.Mollov, R.Babuska, H.B.Verbruggen. "Fuzzy model-based predictive control using Takagi-Sugeno models". *Elsevier, International Journal of Approximate Reasoning* 22, pp. 3-30, (1999)
- [2] M.Setnes, R.Babuska, H.B.Verbruggen. "Rule-based modeling: precision and transparency". *IEEE Transaction on Systems, Man and Cybernetics, C*, vol. 28, NO. 1, Feb. 1998.
- [3] H. Ouakka. "Contribution à l'Identification et la Commande Floue d'une Classe de Système Non Linéaires". Phd, *Université SIDI MOHAMED BEN ABDELLAH*. Juin 2009.
- [4] R.Babuska, J.A.Roulos, H.B.Verbruggen. "Identification of MIMO systems by input-output TS fuzzy". *IEEE Internationla Confernece on Fuzzy systems*, Anchorage, USA, vol. 1, pp. 657-662, 1998.
- [5] T.A.Johansen and R.Babuska. "Multi-objective identification of Takagi-Sugeno fuzzy models". *IEEE Transactions on Fuzzy Systems*, 11(6):847-860, 2003.
- [6] S.Mollov, R.Babuska, J.Abyoni and H.B. Verbruggen. "Effective Optimization for Fuzzy Model Predictive Control. *IEEE Transactions on Fuzzy Systems*, 12(5) :661-675, 2004.
- [7] H. Hacene. "Identification et surveillance de processus dynamiques complexes par logique floue, Application à un échangeur de chaleur". Phd, *UNIVERSITE M'HAMED BOUGARA DE BOUMERDES*. Juillet 2007.
- [8] T.Takagi and M.Sugeno. "Fuzzy identification of systems and its application to modeling and control". *IEEE Transaction on Systems, Man and Cybernetics*, vol. 15, NO. 1, pp. 116-132, 1985.
- [9] S.Mollov, T.J.Van den Boom, F. Cuesta, A. Ollero and R.Babuska. "Robust stability constraints for fuzzy model predictive control". *IEEE Transactions on Fuzzy Systems*, 10(1):50-64, 2002.
- [10] R.Sindelar and R. Babuska. "Input Selection for Nonlinear Regression Models". *IEEE Transactions on Fuzzy Systems*, 12(5):688-696, 2004.
- [11] R.Babuska. "Neuro-fuzzy methods for modeling and identification". In A. Abraham, L.C. Jain, and J. Kacprzyk, editors, *Recent Advances in Intelligent Paradigms ans Applications*, pages 161-186. Springer-Verlag, Heidelberg, 2002.
- [12] M. Jelali and A. Kroll. "Advances in Industrial Control". *Springer-Verlag*, London 2003.
- [13] R.Babuska. "Fuzzy Modeling for Control". *Kluwer Academic Publishers*, Boston, 1998.
- [14] D.Karaboga, B. Basturk, "A powerful and efficient algorithm for numerical function optimisation: artificial (ABC) algorithm", *Journal of Global optimization*, 39 (2007), 459-471.
- [15] H.B.Verbruggen, H.J. Zimmermann, R.Babuska. "Fuzzy Algorithms for Control". *Kluwer Academic Publishers*, Boston, USA, 1999.
- [16] D.Karaboga, B. Gorkemli, C. Ozturk, N. Karaboga. "A comprehensive survey: artificial bee colony (ABC) algorithm and applications", *Artificial Intelligence Review*, 42 (2014), 21-57.
- [17] S.E. Chouaba. "Contribution à l'estimation des modèles linéaires à paramètres variants à temps continu : Application à la modélisation des échangeurs de chaleur". Phd. *ECOLE NATIONALE DES INGENIEURS DE POITIERS*, Septembre 2012.
- [18] D.E. Gustafson, W.C. Kessel. "Fuzzy clustering with a fuzzy covariance matrix". In *Proceedings of the 18th IEEE Conference on Decision and Control*, pages 761-766, San Diego, CA, USA, 1979.
- [19] S.E.Chouaba, A. Chamroo, R. Ouvrard, T.Poinot. "A counter flow water to oil heat exchanger: MISO quasi linear parameter varying modeling and identification", *Simulation Modeling Praticce and Theory*, 23, 87-98, (2012).
- [20] J.H.Lilly. "Fuzzy control and identification". *John Wiley and Sons*, 2010.
- [21] J.A.Roubos, S.Mollov, R.Babuska, H.B.Verbruggen. "Fuzzy model-based predictive control using Takagi-Sugeno models". *Elsevier, International Journal of Approximate Reasoning* 22, pp. 3-30, (1999).
- [22] H.Hellendoorn, D. Driankov. "Fuzzy model identification". *Springer-Verlag*, Berlin, Heidelberg 1997.
- [23] J.Espinosa, J.Vandewalle and V. Wertz. "Fuzzy logic, identification and predictive control". *Springer-Verlag*, London, 2005.
- [24] A. Maldi. "Commande des systèmes décrits par des équations aux dérivées partielles hyperboliques". Phd, *UNIVERSITE MOULOUD MAMMERI*. 2008.

L'intelligence artificielle au service de la maintenance prévisionnelle

Juliette MATTIOLI¹

Pierre-Olivier ROBIC²

Thomas REYDELLET¹

¹ Thales Research & Technology, Campus Polytechnique 1 av. A. Fresnel, 91767 Palaiseau Cedex

² Thales Global Services, 19 av. M. Saulnier, 78140 Vélizy Villacoublay.

{juliette.mattioli, pierre-olivier.robic, thomas.reydellet}@thaligroup.com

Résumé

Avec la transformation numérique du Groupe Thales, l'intelligence artificielle (IA) qu'elle soit orientée par les modèles (IA symbolique) ou basée sur les données (IA connexioniste) devient un élément clé pour la maintenance prévisionnelle. L'objectif de cet article est de dresser un rapide panorama des différentes technologies d'IA mise en oeuvre dans Thales pour outiller les activités de maintenance prévisionnelle.

Mots Clef

Maintenance prévisionnelle, IA symbolique, IA connexioniste, diagnostic, pronostic, planification, ordonnancement.

Abstract

With the Thales Group digital transformation, model-driven AI (Artificial Intelligence) and data-driven AI are becoming a cornerstone to predictive maintenance. The aim of this paper is to give a brief overview of various AI techniques which are used to address predictive maintenance issues.

Keywords

Predictive maintenance, Symbolic AI, Connectionist AI, diagnostic, prognostic, planning, scheduling.

1 Maintenance prévisionnelle : un levier de compétitivité

Selon la définition de l'AFNOR [1], la maintenance couvre l'ensemble de toutes les actions durant le cycle de vie d'un système, destinées à le maintenir ou à le rétablir dans un état dans lequel il peut accomplir la fonction requise. Les activités induites sont aujourd'hui de deux types : corrective ou préventive. La maintenance corrective est programmée suite à une défaillance¹. Cependant, ces activités n'intègrent pas d'aspect préventif. Fondée sur l'adage "*mieux vaut prévenir que guérir*", la maintenance préventive repose sur un plan déterminé a priori. La fréquence des ces opérations s'appuie en général sur des options de garantie,

des instructions du fournisseur, la réglementation ou l'expérience. Mais elle intervient parfois trop tard et provoque une augmentation des activités de maintenance corrective, et cela parce que les conditions réelles et les usages ne sont pas pris en compte.

Avec la transformation numérique du Groupe Thales, une nouvelle forme de maintenance devient possible grâce aux technologies de l'intelligence artificielle (IA) : la maintenance prévisionnelle [2]. Cette dernière est réalisée suivant les prévisions extrapolées de l'analyse et de l'évaluation de paramètres significatifs de dégradation du système. Ce type de maintenance vise à prédire l'évolution de ces paramètres pour estimer le temps de fonctionnement restant avant défaillance (pronostic). Cette approche est souvent qualifiée de maintenance prédictive, car reposant sur les évolutions de paramètres surveillés dans le futur, et de maintenance proactive qui cherche à déterminer les causes à l'origine de ces dégradations et de ces défaillances.

Thales propose ainsi une offre complète de maintenance pour les systèmes critiques, offre de services outillée allant de la réparation et la fourniture de pièces de rechange à la maintenance prévisionnelle, préventive et corrective. En effet, tout système en fonctionnement qu'il soit critique ou non, manifeste des signes, visibles ou non, de dégradation annonçant une future défaillance. La détection d'anomalies ou de changements préfigurent alors à court terme d'une panne à venir. C'est pourquoi, comme nous le verrons dans la suite de ce papier, Thales propose des solutions innovantes reposant sur des approches à base d'IA, pour chacune des étapes de maintenance. En effet, l'IA que nous définissons comme "*des systèmes d'information et des applications pouvant percevoir un environnement, comprendre et décider en fonction du contexte et des usages, tout en apprenant au fil de l'eau*", qu'elle soit dirigée par les données (IA connexioniste) ou basée sur des modèles et la connaissance (IA symbolique) est devenue pour Thales, incontournable pour la surveillance de l'état de santé du système, la planification et la gestion des opérations de maintenance.

2 Diagnostic

L'idée de la maintenance prévisionnelle est de projeter l'état actuel d'un système dans le futur, afin d'en estimer sa

1. Une défaillance correspond à une cessation de l'aptitude d'une entité à accomplir une ou plusieurs fonctions requises.

durée de vie restante (i.e., durée de fonctionnement avant défaillance) appelée communément RUL (remaining useful life), ou de calculer sa probabilité pour qu'il fonctionne durant un certain temps ainsi que le risque d'existence ou d'apparition ultérieure d'un ou de plusieurs modes de défaillance. Plusieurs techniques d'IA permettent de diagnostiquer les fautes survenant dans un système qui provoquent des défaillances.

Certaines reposent sur la connaissance disponible du fonctionnement normal (comportement nominal) ou anormal (comportement en présence de faute) du système [3] pour établir des relations causales entre les causes et les effets observés ou les symptômes, les défaillances et les fautes. Ces méthodes exploitent la connaissance des experts sur le système, une connaissance acquise durant les opérations de maintenance ou/et d'un historique des dysfonctionnements du système. Elles peuvent être capturées par le biais par exemple d'**ontologies** [4] et/ou se basent sur l'existence d'un modèle formel du système qui impliquent une connaissance approfondie de son fonctionnement sous la forme de modèle mathématique et/ou logique. Ainsi pour certains systèmes critiques, Thales a mis en oeuvre des méthodes de diagnostic de défaillances [5] modélisations fonctionnelles et matérielles reposant sur des outils industriels comme l'AMDEC (Analyse des Modes de Défaillance, de leur Effet et de leur Criticité) des arbres de défaillances ou par le biais de modélisation physique consistant à comparer un modèle représentatif du bon fonctionnement du système avec les mesures observées [6]. Tout écart est alors synonyme d'un dysfonctionnement, c'est-à-dire la présence d'un ou plusieurs défauts. A partir de ces études, Thales a développé des **systèmes d'inférence** pour l'aide au diagnostic.

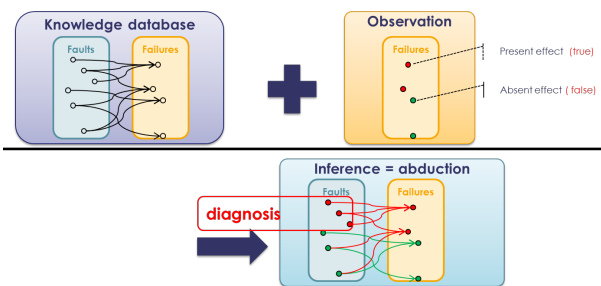


FIGURE 1 – Principe de raisonnement abductif pour l'aide au diagnostic

La propriété principale de cette approche est de pouvoir représenter et restituer les connaissances acquises par les experts. Rappelons qu'un tel système comporte une base de connaissances composée d'une base de règles qui modélise la connaissance et d'une base de faits qui contient les informations concernant le système observé et un moteur d'inférence capable de raisonner à partir des informations contenues dans la base de connaissances et de faire des déductions.

Quand cette connaissance n'est pas accessible, d'autres approches consistent à traiter le système comme étant une boîte noire. Les mesures (entrées-sorties) constituent souvent la plus forte et la plus sûre source d'informations pour comprendre les phénomènes de dégradation. L'objectif est d'associer ces observations à des états de fonctionnement connus du système. Ces approches à base d'IA basées sur les données, visent à apprendre (par les exemples) le phénomène de dégradation et à capturer les relations subtiles entre données, même si ces relations sont inconnues ou difficiles à décrire. Ainsi, Thales propose une combinaison d'approches [7] combinant des approches statistiques, l'**analyse en composante principale** pour identifier les interactions entre les différents composants du système et réduire la complexité, les **chaînes de Markov cachées** pour retrouver l'ensemble des états les plus probables à partir d'une séquence de sortie, les **réseaux bayésiens** pour permettre la corrélation d'événements et la détection de signaux faibles, les **arbres de décision** ou **réseaux de neurones** comme le **deep learning**. Ces approches exploitent les données de surveillance (issues de capteurs) ou d'interventions de maintenance, données traitées afin d'extraire des caractéristiques reflétant le comportement du système et sa dégradation. Celles-ci sont ensuite utilisées pour apprendre des modèles de prédiction des états du système, et calculer une estimation du RUL.

3 Prognostic

Le diagnostic traite de l'identification et de la quantification des dommages qui se sont produits, alors que le pronostic porte sur la prédiction non seulement des possibles états futurs dans lesquels un système peut évoluer mais aussi des dommages qui n'ont pas encore eu lieu. Sa fonction principale est de renforcer le choix des actions à entreprendre pour proposer des recommandations pour l'ordonancement des tâches de maintenance. Il permet d'aborder la maintenance prévisionnelle tout en réduisant ses coûts directs. Les méthodes pour le pronostic dépendent comme celles pour le diagnostic, de la connaissance et des données que l'on a sur le système à maintenir. Elles sont soit fondées sur un modèle physique, soit guidées par les données, soit basées sur l'expérience. Cette dernière s'appuie principalement sur des approches fiabilistes, sur des modélisations de processus stochastiques ou sur des systèmes experts.

Les approches à base de modèles, développées dans Thales, consistent à simuler le modèle de comportement du système pour caractériser la dynamique des paramètres reflétant la dégradation. Par exemple, le suivi de la dégradation et l'obtention de sa dynamique peuvent être modélisé par un filtre de Kalman, par un filtrage particulaire ou par des chaînes de Markov cachées [3].

Les méthodes guidées par les données sont, avec le renouveau de l'IA, en plines expansions. Elles exploitent des mesures représentant l'état actuel et/ou l'état passé. Elles utilisent des techniques d'**apprentissage** ou des techniques

d'**analyse de tendance** pour construire un modèle à partir de ces données pour ensuite extrapoler l'évolution du système. Cependant, elles supposent que la probabilité d'évolution est proche des conditions de fonctionnement utilisées lors de l'apprentissage du modèle. Les **réseaux de neurones adaptatifs**, les **machines à vecteurs de support** sont des approches qui ont révélées leur efficacité pour la prévision de séries temporelles. Leur intérêt réside dans leurs aptitudes à apprendre et à capturer les relations potentielles entre les données même si le comportement est complexe. Ces modèles sont appris par un ensemble d'échantillons (entrées/sorties) directement extrait de la séquence d'observations. Enfin, la **fusion d'informations** [8] permet la corrélation d'un indicateur de santé avec les conditions opératoires permet de déterminer un RUL en considérant à la fois l'état de santé courant du système et son utilisation. Cependant, ces approches sont difficiles à mettre en oeuvre pour les équipements complexes, leur réponse est souvent entachée d'incertitudes de modélisation ce qui induit un pronostic moins robuste. Pour avoir un modèle robuste permettant de prendre en compte les imperfections et les incertitudes, la **logique floue** et les **fonctions de croyances** sont très utilisées [9]. Combinée avec les méthodes citées ci-dessus, elles offrent un compromis raisonnable entre les modèles analytiques rigoureux, les approches reposant sur des connaissances métier et les approches guidées par les données.

4 Planification et Ordonnement

Dans toute opération de maintenance préventive, il s'agit d'établir un planning prévisionnel des tâches sous-jacentes incluant les ressources humaines et matérielles impliquées. Ainsi, les dates de début au plus tôt, les durées et les dépendances des tâches sont connues a priori. Concernant les tâches de maintenance corrective, leur durée est basée sur une estimation et dépend d'un bon diagnostic. Or, un tel planning est généralement figé et peu ouvert aux changements et donc à des actions correctives qui sont par essence non planifiées.

Pour construire un tel ordonnancement mixant maintenance préventive et corrective, Thales utilise une combinaison de la **programmation par contraintes** avec des approches multicritères [10, 11] pour affecter des tâches aux opérateurs sous contrainte de compétence, en minimisant la date de fin de l'ordonnement et les pénalités de retards [2]. Une approche outillée d'aide à la **décision multicritère** [12] a été mise en place (cf. figure 2), en particulier pour modéliser par le biais d'une fonction multicritère intégrant des dimensions de natures variées comme les arrêts de production, le retard dans l'exécution des tâches, les pénalités et d'autre part, des coûts de maintenance acceptables. De plus, avec la maintenance prévisionnelle, un réordonnement est à envisager, en tenant compte du RUL. L'évolution du planning devient alors dynamique suite aux dysfonctionnements du système de production [13]. Enfin, la prise en compte de la maintenance dès la conception per-

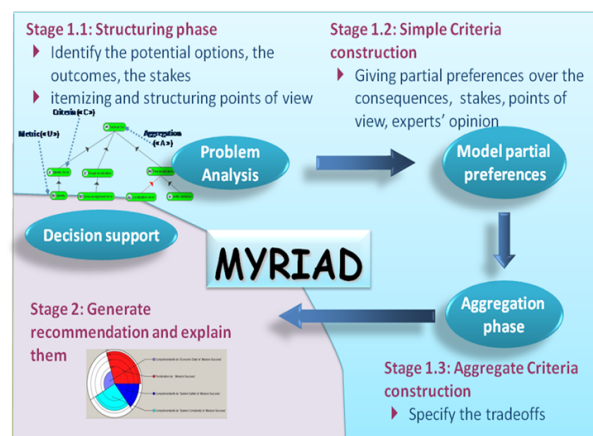


FIGURE 2 – Méthodologie Myriad développée par Thales pour construire une fonction multicritère

met de concevoir des outils d'aide à la décision pour une optimisation de la gestion du système en mode dégradé comme le montre [5].

5 Conclusion

Une politique de maintenance prévisionnelle efficace doit conduire à une réduction des pertes engendrées par des indisponibilités de production. Son efficacité est fortement liée à sa faculté de caractériser l'état de santé courant (diagnostic) et futur (pronostic) du système. Pour cela, il est nécessaire de pouvoir générer des données permettant de rendre compte de la dynamique de la dégradation du système, de modéliser ou d'apprendre ses modes de dégradation, d'estimer son état de santé et de prédire sa durée de vie résiduelle. Les techniques d'IA qu'elles soient symboliques ou connexionnistes deviennent incontournables pour les activités de maintenance de Thales. Cependant, comme le montre la figure 3, il s'agit d'une combinaison d'approches qui dépendent de la connaissance disponible du fonctionnement nominal et du type de fautes mais aussi des données d'observation de l'état de santé du système.

Références

- [1] Norme AFNOR. Terminologie de la maintenance. *NF-EN, 13306 :X60-319*, 2001.
- [2] Simon Fossier and Pierre-Olivier Robic. Maintenance of complex systems : From preventive to predictive. In *Live Maintenance (ICOLIM), 2017 12th International Conference on*, pages 1–6. IEEE, 2017.
- [3] Camille Baysse. *Analyse et optimisation de la fiabilité d'un équipement opto-électrique équipé de HUMS*. PhD thesis, Université Sciences et Technologies-Bordeaux I, 2013.
- [4] Luis Palacios, Gaëlle Lortal, Claire Laudy, Christian Sannino, Ludovic Simon, Giuseppe Fusco, Yue Ma, and Chantal Reynaud. Avionics maintenance ontology building for failure diagnosis support. In *Procee-*

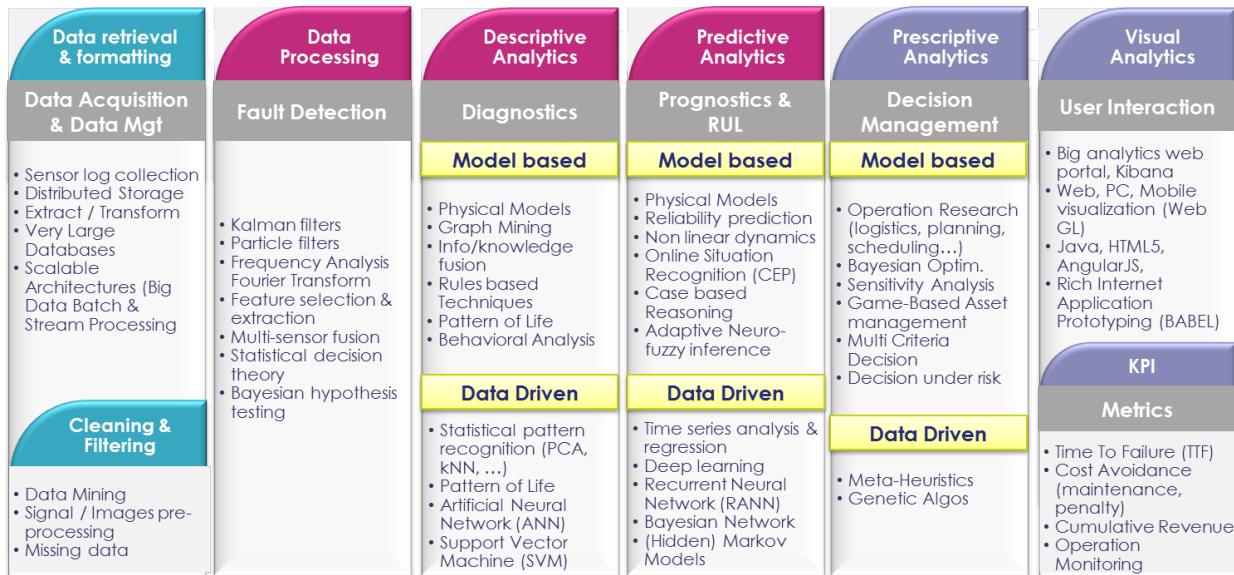


FIGURE 3 – Exemple d’approches mises en oeuvre dans Thales pour la maintenance

dings of the International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, pages 204–209. SCITEPRESS-Science and Technology Publications, Lda, 2016.

- [5] J. Sprauel, C. Sannino, and C. Seguin. Techniques d’aide à la décision appliquées à la maintenance d’un avion de type business jet decision aiding techniques applied to the maintenance of a business jet. In *Congrès Lambda Mu 19*, 2014.
- [6] Camille Baysse, Didier Bihannic, Anne Gégout-Petit, Michel Prenat, and Jérôme Saracco. Hidden markov model for the detection of a degraded operating mode of optronic equipment. *arXiv preprint arXiv :1212.2358*, 2012.
- [7] Michel Giraudeau, Emilie Ronfaut, and D Bihannic. Apport des hums sur les performances sdf et soutien des équipements électroniques embarqués. *7A-Maintenance-diagnostic/pronostic*, 2014.
- [8] Claire Laudy and Jean-Gabriel Ganascia. Introducing semantic knowledge in high-level fusion. In *Military Communications Conference, 2009. MILCOM 2009. IEEE*, pages 1–7. IEEE, 2009.
- [9] Simon Fossier, Claire Laudy, and Frédéric Pichon. Managing uncertainty in conceptual graph-based soft information fusion. In *Information Fusion (FUSION), 2013 16th International Conference on*, pages 930–937. IEEE, 2013.
- [10] Fabien Le Huédé, Michel Grabisch, Christophe Labreuche, and Pierre Savéant. Integration and propagation of a multi-criteria decision making model in constraint programming. *Journal of Heuristics*, 12(4-5) :329–346, 2006.
- [11] Fabien Le Huédé, Michel Grabisch, Christophe Labreuche, and Pierre Savéant. Mcs : A new algorithm

for multicriteria optimisation in constraint programming. *Annals of Operations Research*, 147(1) :143–174, 2006.

- [12] Michel Grabisch and Christophe Labreuche. A decade of application of the choquet and sugeno integrals in multi-criteria decision aid. *Annals of Operations Research*, 175(1) :247–286, 2010.
- [13] Christian Sannino, Jonathan Sprauel, and Christel Seguin. Modèles dysfonctionnels pour la gestion de la qualité de service des systèmes critiques. *3R-Recherche et Industrie 1*, 2016.

Towards automatic argumentation about voting rules

Michael Kirsten¹

Olivier Cailloux²

¹ Dept. of Informatics, Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany

² Université Paris-Dauphine, PSL Research University, CNRS, LAMSADE, 75016 Paris, France

kirsten@kit.edu, olivier.cailloux@dauphine.fr

Abstract

Voting rules aggregate the preferences of a group to make decisions. As multiple reasonable voting rules exist, the axiomatic approach has been proposed to exhibit both their merits and paradoxical behaviors. It consists in characterizing a voting rule by a set of understandable properties called axioms. It is however a difficult task to characterize a voting rule by such axioms, and even when a proof exists, it may be difficult to understand why a specific voting rule fails to satisfy a given axiom, especially for untrained users. In this article, we present an automatic method which determines whether a given rule satisfies a set of axioms. When the rule does not satisfy an axiom, the automatic prover generates comprehensible evidence of the violation in the form of a counter-example. It can be used by non-expert users to comprehend the violation and may serve to argue in favor of other rules which satisfy the axiom. Our method is based on the software analysis technique bounded model checking, which enables bounded verification for properties of software programs. It translates the program together with user-annotations into a reachability problem for those profiles and outcomes which adhere to our specification. The method can be applied to arbitrary voting rules; we demonstrate it on the case of the Borda axiomatization and compare the Borda rule to both the Black and the Copeland voting rules.

Keywords

Social choice theory, bounded model checking, argumentation theory, automated reasoning.

1 Goal and motivation

Voting rules serve to select winning alternatives from a set of a priori possible alternatives, given the preferences of a set of voters. Many voting rules that appear reasonable have been proposed in the literature, however each of them shows paradoxical behavior for at least some voting situations [Arrow, 2012]. Axiomatic analysis permits to distinguish the merits of different voting rules: it consists in determining axioms, properties of voting rules that seem a priori desirable for a voting rule to have, and sets of axioms which together characterize a voting rule. A set of axioms

characterizes a voting rule when there is a unique voting rule that satisfies all axioms in the set.

Our first interest in this article is, given a voting rule, to detect automatically which axioms it satisfies among a given set of axioms. Although it is not possible to do this efficiently in full generality, modern SAT- and SMT-solvers [Biere et al., 1999] enable solving this task efficiently when limiting the number of voters and alternatives. Indeed, as stated in the “small scope hypothesis” and also seen in related work on automated reasoning of voting rules, a high proportion of violations can often already be found for a comparatively small fraction of voters and alternatives [Jackson, 2006, Brandt et al., 2017]. With this prerequisite, the technique called *software bounded model checking (SBMC)* allows encoding the problem into a decidable set of logical equations, which can then be solved by a SAT- or SMT-solver. In this article we propose an automatic method that determines whether a rule satisfies an axiom within the specified bounds. We illustrate it by applying the approach to the case of Borda: we consider a set of four axioms that characterize the Borda rule, and given some rules different than the Borda rule, we find automatically which axioms from that set each rule fails to satisfy.

It is especially interesting to obtain short and readable proofs that a rule does not satisfy some axiom. This is our second objective in this article: our automatic prover outputs profiles, preferably small or satisfying some other property representing “simplicity” for a human eye, that exhibit a failure of a rule to satisfy an axiom. Such a short proof is interesting even if it is already known in the literature that the rule does not satisfy the axiom, as short proofs permit to help getting an intuitive understanding of voting rules. Furthermore, it can be used to talk about voting rules with non-expert users. Our illustration in the case of Borda permits to obtain a tool that “argues in favor” of Borda [Cailloux and Endriss, 2016] (assuming the axioms characterizing Borda of our choice are considered desirable properties of any voting rule): given any rule that is not the Borda rule, the prover finds a concrete profile which illustrates that the rule fails to comply with some of those axioms.

2 Concepts and notations

We consider an infinite set of voters \mathcal{N} and a fixed finite set of alternatives \mathcal{A} . An election involves a subset of voters $N \subseteq \mathcal{N}$ that each express preferences over all alternatives. Those preferences are elements of $\mathcal{L}(\mathcal{A})$, the linear orders over the alternatives: connected, transitive, asymmetric binary relations over \mathcal{A} . Given a voter $i \in \mathcal{N}$, we write as $\succ_i \in \mathcal{L}(\mathcal{A})$ the preference of voter i . Given a set of voters N , a profile $(\succ_i)_{i \in N} = R$ is an association of each voter to her preference. The set of all possible profiles is $\bigcup_{N \subseteq \mathcal{N}} \mathcal{L}(\mathcal{A})^N$.

A voting rule $f : \bigcup_{N \subseteq \mathcal{N}} \mathcal{L}(\mathcal{A})^N \rightarrow \mathcal{P}^*(\mathcal{A})$, where $\mathcal{P}^*(\mathcal{A})$ represents the non-empty subsets of \mathcal{A} , maps each possible profile $(\succ_i)_{i \in N} \in \bigcup_{N \subseteq \mathcal{N}} \mathcal{L}(\mathcal{A})^N$ to a non-empty set of winning alternatives $\emptyset \neq W \subseteq \mathcal{A}$ (sets of winners are used to account for tied winners).

An axiom is here simply defined as a property of a voting rule. One kind of axioms of interest (for their simplicity) are so-called functional properties [Beckert et al., 2016]. A functional property is one that can be represented as a relation $S \subseteq \bigcup_{N \subseteq \mathcal{N}} \mathcal{L}(\mathcal{A})^N \times \mathcal{P}^*(\mathcal{A})$: given any $R \in \bigcup_{N \subseteq \mathcal{N}} \mathcal{L}(\mathcal{A})^N$, $S(R) = \{A \mid (R, A) \in S\}$ contains all the admissible sets of winners given that profile. When $S(R) = \mathcal{P}^*(\mathcal{A})$, S expresses no constraint about that profile. A voting rule f satisfies the functional property S iff f picks sets of winners among the ones considered possible by S , namely, iff $f \subseteq S$ (where we view both f and S as relations). Whenever a rule f fails to satisfy a functional property, it can be proved in a very short and understandable way, by exhibiting a profile R on which f picks a wrong set of winners, $f(R) \notin S(R)$.

The property DOM is an example of a functional property.

DOM This property forbids Pareto-dominated alternatives to win. An alternative $a \in \mathcal{A}$ is Pareto-dominated in a profile $(\succ_i)_{i \in N}$ iff some alternative $a' \in \mathcal{A}$ is unanimously preferred to a in $(\succ_i)_{i \in N}$: $\forall i \in N, a' \succ_i a$. The property mandates that the rule selects winning alternatives among U_R , denoting the alternatives that are not Pareto-dominated in R : $\forall R \in \bigcup_{N \subseteq \mathcal{N}} \mathcal{L}(\mathcal{A})^N$, $S_{\text{DOM}}(R) = \mathcal{P}^*(U_R)$.

Another, more general, kind of axioms we are interested in, are k -relational properties, with $k \in \mathbb{N}$ [Beckert et al., 2016]. A k -relational property $S \subseteq (\bigcup_{N \subseteq \mathcal{N}} \mathcal{L}(\mathcal{A})^N \times \mathcal{P}^*(\mathcal{A}))^k$ contains all the admissible ways of associating sets of winners to k profiles. A voting rule f satisfies a k -relational property S iff $f^k \subseteq S$, viewing again f as a relation. For example, for $k = 2$, if S contains (R_1, W_1, R_2, W_3) and (R_1, W_1, R_2, W_4) , given a voting rule f such that $f(R_1) = W_1$, S declares as admissible that $f(R_2)$ be equal to W_3 or W_4 .

The property REINF exemplifies a 3-relational property.

REINF Reinforcement requires that elections uniting disjoint groups of voters elect the candidates chosen by both groups, if some such candidates exist: for each

$(\succ_i)_{i \in N_1}, (\succ_i)_{i \in N_2}$ with $N_1 \cap N_2 = \emptyset$ and $A \neq \emptyset$, defining A as $f((\succ_i)_{i \in N_1}) \cap f((\succ_i)_{i \in N_2})$, reinforcement imposes that $f((\succ_i)_{i \in N_1 \cup N_2}) = A$.

3 Checking properties automatically

Let us start by describing the general technique of checking for software properties, before we explain how we apply it to verify whether voting rules satisfy properties. In a nutshell, we translate the proving task to a reachability problem, which the SBMC tool encodes into an instance of the SAT problem.

3.1 Checking software properties

We assume we are given an algorithm A under the form of an imperative program (for example, written in the C language), that uses some parameter values taken among a set of possible values I . An entry $i \in I$ is a list of values, one value for each such parameter: it gives a value to everything that a run of A depends on, such as its input variables, or anything that is considered non-deterministic from the point of view of A . For this reason, those parameters are qualified as “non-deterministic”, to distinguish them from normal parameters used in a programming language to pass information around. (By contrast, some values can be “derived”, thus, computed in A from the non-deterministic parameter values, or declared as constants in A , and both values of non-deterministic parameters or derived values can then be used as normal parameters in the program.) We are also given a software property to be checked about A , in the form $C^{\text{ant}} \Rightarrow C^{\text{cons}}$, where *ant* and *cons* stand for antecedent and consequence respectively. Both C^{ant} and C^{cons} are sets of boolean statements. A boolean statement is a statement of A that evaluates to a boolean value, for example, a statement checking that some computed intermediate value is odd. An entry i is said to satisfy a set of boolean statements iff all boolean statements in the set evaluate to true during the execution of A using the non-deterministic parameter values i , and is said to fail the set of boolean statements otherwise. The property $C^{\text{ant}} \Rightarrow C^{\text{cons}}$ requires that for all possible entries $i \in I$, if i satisfies C^{ant} , then i satisfies C^{cons} . As an example, assume A computes, given i , two intermediate integer values v_1 and v_2 , and then returns a third value v_3 . The property to be checked could be: if v_1 is negative, then v_2 is positive and v_3 is odd. A solver that is asked to check a software property $C^{\text{ant}} \Rightarrow C^{\text{cons}}$ thus exhaustively searches for an entry i that satisfies C^{ant} but fails C^{cons} . The property is valid iff it is impossible to find such an entry. SBMC is a fully-automatic static program analysis technique used to verify whether such a software property is valid, given an algorithm and a property to be checked. It covers all possible inputs within a specified bound. It is static in the sense that programs are analyzed without executing them on concrete values. Instead, programs are symbolically executed and exhaustively checked for errors up to a certain bound, restricting the number of loop iterations to limit runs through the program to a bounded length.

This is done by unrolling the control flow graph of the program and translating it into a formula in a decidable logic that is satisfiable if and only if a program run exists which satisfies C^{ant} and fails C^{cons} . The variables in the formula are the non-deterministic parameters of A , and their possible values are taken from I .

This reduces the problem to a decidable satisfiability problem. Modern SAT or SMT-solving technology can then be used to verify whether such a program run exists, in which case an erroneous input has been found, and the run is presented to the user. If the solver cannot find such a program run, it may be either because the property is valid, or because it is invalid only for some run which exceeds the bound. In some cases, SBMC is able to infer statically which bound is sufficient to bring a definitive conclusion.

3.2 Checking voting rule properties

In our case, we are given a voting rule f implemented as a C function that returns winners given a profile, and a list of axioms. When the axiom is a functional property S , in order to check whether f satisfies S , we want to use SBMC to search for an input profile R such that $f(R) \notin S(R)$. To do this, we implement the requirements of S as a C function with annotations that indicate to the SBMC tool a set of boolean statements, constituting C^{ant} , that evaluate to true iff $f(R) \notin S(R)$. Other annotations serve to indicate non-deterministic parameters, including the input profile, and other non-deterministic parameters depending on what is necessary for our implementation to check whether a profile satisfies S . We use as C^{cons} a single boolean statement that is always false. This makes the solver search for an entry R (and possibly other parameter values) such that $f(R) \notin S(R)$.

For example, when searching for a violation of Pareto-dominance (DOM as defined in Section 2), it suffices to implement a C function F containing boolean statements that evaluate to true whenever some alternative a , a non-deterministic parameter of F , is Pareto-dominated by another alternative a' , another non-deterministic parameter of F , but wins. This way, the solver will search for an entry (R, a, a') corresponding to an input profile that contains a Pareto-dominated alternative that wins, thus, an input that violates S , and will report such an input, if found.

This approach can be extended to cope for k -relational properties, where $k > 1$. We do not detail this as this is similar in spirit, only requiring to duplicate the input parameters so that the solver searches for multiple profiles as entries, and writing the function checking for S so that it cares about the relationship between multiple sets of profiles and winners.

In order to apply SBMC, we need to set bounds for the number of loop iterations. We have implemented all tested axioms and voting rules (and more) in such a way that the number of loop iterations depend only on the number of voters and alternatives. In the experiments, we set those as constants. The SBMC tool can thus automatically infer the

right bounds in a syntactic pre-processing step. For more complex cases where it would not be possible to syntactically infer the bounds, e.g., when while-loops are used or the number of iterations involves some non-deterministic parameter, the bound could be specified manually.

Regarding computation time, finding an input that violates S , if such an input exists, scales much better than proving that no such input exists, as we are allowed to stop as soon as such an input is found. Moreover, our method generally scales better for smaller numbers of loop iterations, simply based on the smaller formula for which a model is searched. As we are generally looking for “small” profiles, this is most often the case. For very complex (i.e., having complex loop structures with a large number of iterations) voting rules and axioms, or for properties which are *not* violated, it may happen that the solver takes more computation time or resources than the user is willing to invest. Our method is generally well-suited for finding counter-examples to voting rule properties as most violations are exhibited for relatively small bounds. We substantiate these claims in the following section, where we exemplarily use our method to argue in favor of the Borda rule.

This approach can be generalized to search for profiles that illustrate the satisfaction of simultaneous axioms, or profiles that satisfy some axioms and not some others, etc. We have written a general program to search for such situations that cover about fifteen classical axioms¹.

For our experiments, we use CBMC 5.8 Clarke et al. [2004], an implementation of the SBMC approach for the C language, with the built-in solver based on the SAT-solver MiniSat 2.2.0 Eén and Sörensson [2003]. All experiments are performed on an Intel(R) Core(TM) i5-6500 CPU at 3.20 GHz with 4 cores and 16 GB of RAM.

3.3 A simple example

Listing 1 illustrates how we specify DOM, the “Pareto-dominance” functional property. The function is annotated with specifications in the form of assumptions (expressed by the function `__CPROVER_assume`), that represent the statements composing C^{ant} , as well as operations to model non-deterministic choice (indicated by the prefix `nondet_`) that represent what we called non-deterministic parameters in Section 3.1. These non-deterministic choice operators can be used anywhere inside the analyzed program and are also translated to the formula which is passed to the solver. The solver then searches for instantiations of these variables which lead to a violation of C^{cons} .

In the displayed function, the constants `N` and `M` denote the numbers of voters and alternatives respectively. The voters and alternatives are represented by integers from 0 to `N - 1` and from 0 to `M - 1`. The function accepts as input a profile `prof`, modeled as a two-dimensional array: `prof[i][r]` is the alternative that voter `i` associates to rank `r`, where

¹The runnable source code can be found at <https://github.com/mi-ki/voting-rule-argumentation>.

```

1 void dominance(uint prof[N][M], uint win[M]) {
2   uint bad = nondet_uint(), good = nondet_uint();
3   __CPROVER_assume (0 ≤ bad < M);
4   __CPROVER_assume (0 ≤ good < M);
5   __CPROVER_assume (bad ≠ good);
6   uint prefergtob[N];
7   for (uint i = 0; i < N; i++) {
8     uint rankg = M, rankb = M;
9     for (uint r = 0; r < M; r++) {
10      if (prof[i][r] == good) rankg = r;
11      if (prof[i][r] == bad) rankb = r;
12    }
13    prefergtob[i] = (rankg < rankb) ? 1 : 0;
14  }
15  for (uint i = 0; i < N; i++)
16    __CPROVER_assume (prefergtob[i]);
17  __CPROVER_assume (win[bad]);
18 }

```

Listing 1: Pareto-dominance specification for CBMC.

0 is the best, and $M-1$ the worst rank. The function also accepts as input a set of winning alternatives, modeled as an array `win` with one entry per alternative: `win[a]` equals 1 if the alternative `a` belongs to the set of winning alternatives, 0 otherwise. The non-deterministic parameters are `prof` (whose declaration is not shown here) and the two alternatives `bad` and `good` (Line 2). Lines 6 to 14 initialize the array `prefergtob` so that `prefergtob[i]` holds the value 1 if voter `i` prefers `good` to `bad`, 0 otherwise.

The function thus indicates to the solver that it must find two alternatives `bad` and `good`, thus integers in the suitable range (Lines 3 and 4) and different from each other (Line 5), such that everybody prefers `good` to `bad` (Lines 15 and 16), and yet the alternative `bad` is a winner of the election (Line 17). Any run which satisfies these specified statements is a valid counter-example which would prove that the profile and winners given as input violate Pareto-dominance.

We pursue the example with the setup given in Listing 2. Therein, we initialize a profile with symbolic non-deterministic values (Line 7) and restrict their ranges such that they are valid alternatives (Line 8). Furthermore, we use a helper array `used` with one entry per alternative, which is used to ensure that every ballot holds every alternative only once (Lines 5, 9 and 11). We assume a voting rule is given as a function `f` (Line 14) which gets a profile `prof` and the number of voters `N` as parameters, and returns the set of winning alternatives as output. For the experiments within this article, the implementations of `f` will follow directly from their definitions.

After calling the test methods for the properties to be checked (in this case only Pareto-dominance), we set a boolean statement that is always false (Line 16) as content of `Ccons` (recognized by the solver by the keyword `assert`), which indicate to the solver that any program run that reaches this point is a counter-example of interest to us.

```

1 int main(int argc, char *argv[]) {
2   uint prof[N][M], win[M];
3   for (uint i = 0; i < N; i++) {
4     uint used[M];
5     for (uint a = 0; a < M; a++) used[a] = 0;
6     for (uint r = 0; r < M; r++) {
7       a = nondet_uint();
8       __CPROVER_assume (0 ≤ a < M);
9       __CPROVER_assume (!used[a]);
10      prof[i][r] = a;
11      used[a] = 1;
12    }
13  }
14  win = f(prof, N);
15  dominance(prof, win);
16  assert (0);
17  return 0;
18 }

```

Listing 2: Setup for CBMC.

4 Definitions for the experiments

We propose here to consider the axioms that characterize the Borda voting rule (Def. 1), as defined in a variant of Young’s axiomatization [Cailloux and Endriss, 2016]. We also define a few axioms that are not satisfied by the Borda rule, and finally define two supplementary rules. All these concepts will be used in the experiments. (Readers knowledgeable in social choice theory can skim this section.)

Definition 1 (Borda). *The Borda rule, given a profile $(\succ_i)_{i \in N}$, associates to each alternative a and voter i the score $s(a, i)$ equal to the number of alternatives that a beats in \succ_i , and associates to each alternative a the score $s'(a) = \sum_{i \in N} s(a, i)$. The winners are the alternatives that have the maximal score: $f_{\text{Borda}}((\succ_i)_{i \in N}) = \arg \max_{a \in \mathcal{A}} s'(a)$.*

4.1 An axiomatization of the Borda rule

We first require a few definitions.

A profile $(\succ_i)_{i \in N}$ is *elementary* iff it has exactly two voters, and if the set of alternatives can be partitioned into disjoint subsets $A^{\text{top}}, A^{\text{bottom}} \subseteq \mathcal{A}$ with $A^{\text{top}} \cup A^{\text{bottom}} = \mathcal{A}$ and $A^{\text{top}} \neq \emptyset$ such that both voters have all alternatives in A^{top} preferred to all alternatives in A^{bottom} , and the voters have inverse preferences over A^{top} , and inverse preferences over A^{bottom} . Thus, denoting the voters by 1 and 2, and denoting by $\succ_i|_A$ the restriction of \succ_i to A , $\forall A \in \{A^{\text{top}}, A^{\text{bottom}}\}$: $\succ_1|_A = (\succ_2|_A)^{-1}$. Let $T(R)$ denote the “top alternatives” of an elementary profile, meaning, the set of alternatives corresponding to A^{top} . (This is legal as it is unique.)

Example 1. The profile R shown below, composed of the linear orders (a, b, c, d) and (b, a, d, c) , is an elementary profile corresponding to $A^{\text{top}} = \{a, b\}$, with \mathcal{A} equal to

$\{a, b, c, d\}$.

$$R = \begin{array}{cc} a & b \\ b & a \\ c & d \\ d & c \end{array}.$$

△

Given a linear order $\succ \in \mathcal{L}(\mathcal{A})$, with \mathcal{A} of size m , define the cycle corresponding to \succ as the set of m pairs consisting of the pairs of alternatives (a, b) such that b immediately succeeds to a in \succ (the rank of b is one more than the rank of a), union the pair of alternatives (z, a) where z and a are respectively the minimal and maximal elements of \succ . For example, the cycle corresponding to the linear order (a, b, c) is $\{(a, b), (b, c), (c, a)\}$. Observe that, given any $\succ \in \mathcal{L}(\mathcal{A})$, the cycle corresponding to \succ corresponds to exactly m linear orders in $\mathcal{L}(\mathcal{A})$. For example, the cycle $\{(a, b), (b, c), (c, a)\}$ also correspond to (b, c, a) and (c, a, b) .

We say that a profile $(\succ_i)_{i \in N}$ is cyclic iff it has exactly m voters and m different linear orders, and some cycle corresponds to all its linear orders (equivalently, the linear orders in $(\succ_i)_{i \in N}$ are all those that correspond to a given cycle).

Example 2. The profile R shown below is a cyclic profile corresponding to the cycle $\{(a, b), (b, c), (c, d), (d, a)\}$ with $\mathcal{A} = \{a, b, c, d\}$.

$$R = \begin{array}{cccc} a & b & c & d \\ b & c & d & a \\ c & d & a & b \\ d & a & b & c \end{array}.$$

△

Below is the axiomatization that we use for the Borda rule, composed of the axioms ELEM, CYCL, CANC and REINF. It is very similar but not identical to the axiomatization given by Young [1974]. The proof that these four axioms characterize the Borda rule is given in Cailloux and Endriss [2016].

ELEM This axiom mandates that the rule, when given any elementary profile, selects its top alternatives: $S_{\text{ELEM}}(R) = \{T(R)\}$ if R is an elementary profile and $S_{\text{ELEM}}(R) = \mathcal{P}^*(\mathcal{A})$ otherwise.

CYCL This axiom requires the rule to select all alternatives as tied winners when given any cyclic profile: $S_{\text{CYCL}}(R) = \{\mathcal{A}\}$ if R is a cyclic profile and $S_{\text{CYCL}}(R) = \mathcal{P}^*(\mathcal{A})$ otherwise.

CANC The cancellation axiom constrains the set of winners to be \mathcal{A} when all pairs of alternatives (a, b) are such that a is preferred to b for as many voters as b is to a in $(\succ_i)_{i \in N}$.

REINF As defined above.

4.2 Two axioms not satisfied by Borda

Borda notoriously fails (when $|\mathcal{A}| \geq 3$) to satisfy the following two functional properties COND and MAJ, as well as the derived property WMAJ.

Given a profile $R = (\succ_i)_{i \in N}$, we say that an alternative a obtains a strict majority against a' , denoted by $a M_R a'$, iff more than half of the voters prefer a to a' in $(\succ_i)_{i \in N}$: $a M_R a' \Leftrightarrow |\{i \mid a \succ_i a'\}| > |\{i \mid a' \succ_i a\}|$. An alternative is a Condorcet winner iff it obtains a strict majority against all other alternatives. Any Condorcet winner is unique.

COND This property mandates that if there is a Condorcet winner, it becomes the sole winner.

MAJ This property requires that, whenever some alternative is placed first by more than half of the voters in $(\succ_i)_{i \in N}$, it becomes the sole winner.

The Condorcet property is stronger than the majority property, i.e., $S_{\text{COND}} \subseteq S_{\text{MAJ}}$. Based on MAJ we define a weaker property, the weak majority property WMAJ, for which $S_{\text{MAJ}} \subseteq S_{\text{WMAJ}}$.

WMAJ This property requires that, whenever some alternative is placed first by more than half of the voters in $(\succ_i)_{i \in N}$, it becomes a (not necessarily unique) winner.

4.3 Two Condorcet compatible voting rules

To end this section, we define two famous voting rules that satisfy the Condorcet property.

Definition 2 (Black). *The Black [1958, p. 66] rule selects the Condorcet winner if there is one, otherwise, the Borda winners.*

Given a profile $R = (\succ_i)_{i \in N}$, let $M_R(a)$ denote the set of alternatives against which a obtains a strict majority, and $M_R^{-1}(a)$ the set of alternatives that obtain a strict majority against a .

Definition 3 (Copeland). *The Copeland [1951] rule (actually a close variant of a rule proposed by Ramon Llull in the 13th century [Colomer, 2013]), given a profile R , gives to each alternative the score $s(a) = |M_R(a)| - |M_R^{-1}(a)|$, and lets the alternatives with maximal score win.*

5 Experiments

In this section we first experiment with the Borda rule itself. In Section 5.1, we look at whether the solver is able to find out that the Borda rule satisfies Pareto-dominance, depending on the bounds we set on the numbers of alternatives and voters. We then check how long it takes to find an example that illustrates Pareto-dominance (one where some alternative is dominated and indeed not included among the winners). In Section 5.2, we search for counter-examples that illustrate that the Borda rule fails to satisfy the properties defined in Section 4.2.

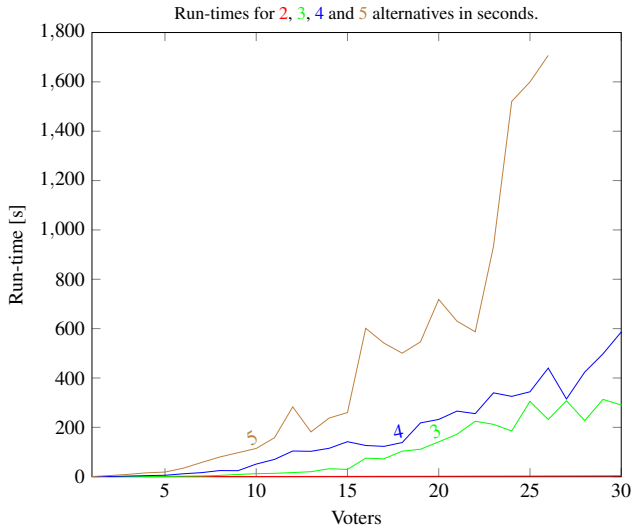


Figure 1: Run-times for the automatic verification of Pareto-dominance for the Borda voting rule.

Second, we experiment with the Borda axiomatization. We suppose we are given a rule f , different from Borda (Def. 1), as a C function: in our experiments we consider the Black rule and the Copeland rule. In this simple experimental setup, we thus know that f fails to satisfy at least one of the Borda axioms from Section 4.1. For each of these rules, we illustrate that we can find out automatically which axioms f fails to satisfy, and output a short proof of this, easy for a human to inspect. We also analyze in which situations we can prove that axioms are satisfied by f , for those that f satisfies.

5.1 Borda and Pareto-dominance

In Figures 1 and 2, we illustrate the sizes of situations, i.e., numbers of voters (on the x-axis) and alternatives (each having a different plot), that we are able to analyze. In this example, as well as in the following experiments, we only consider run-times below 30 minutes to be reasonable, and stop computations which require more time than 30 minutes. Run-times are given in seconds (on the y-axis) for up to 30 voters and 5 alternatives. Figure 1 shows the run-times for the verification that Pareto-dominance (Listing 1) holds for the Borda rule (the line for 2 alternatives is almost superposed to the x-axis). Experiments for 5 alternatives and more than 26 voters took more than 30 minutes and are thus not plotted.

Figure 2 shows the run-times for finding an example situation that illustrates that Borda satisfies DOM. We used the code from Listing 1 with the only difference that we negated the statement in Line 17. The function now indicates to the solver that it must find two different alternatives `bad` and `good`, such that everybody prefers `good` to `bad`, and the alternative `bad` is not a winner of the election.

In the first experiment, the solver proves that no profiles satisfy the provided conditions. We can see that the verifi-

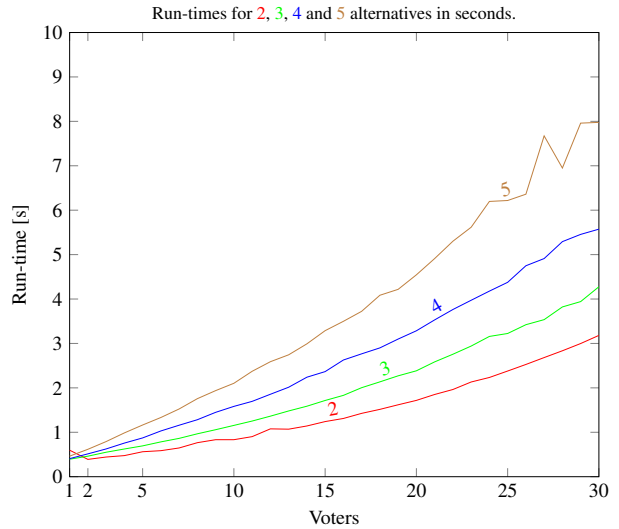


Figure 2: Run-times for finding an illustration of the dominance property for the Borda rule.

cation is feasible for small sizes of profiles within 15 minutes. Expectedly, in situations where some profile exists that satisfies the provided conditions, the run-times are significantly smaller. They stay well below 10 seconds for this example. Both experiments indicate that our method yields reasonable run-times for at least up to five alternatives and 25 voters.

5.2 Counter-examples to Borda

We illustrate here our approach by finding counter-examples which show that the Borda rule does not satisfy the properties defined in Section 4.2 (those properties are satisfied by both the Black and the Copeland voting rule, however).

We start with the stronger Condorcet property COND. The smallest example proving that Borda fails COND can be found in less than one second for three voters and three alternatives:

$$R = \begin{array}{ccc} c & c & b \\ b & b & a \\ a & a & c \end{array} .$$

For the profile R , the Borda rule elects the alternatives $\{a, c\}$ instead of the Condorcet winner c .

Whereas an isomorphic counter-example is found in less than one second for the failure of MAJ, we find two smallest examples (one regarding the number of alternatives, and another one regarding the number of voters) proving the failure of WMAJ.

For a minimal number of alternatives, we find the smallest proof for three alternatives and five voters in less than one second:

$$R = \begin{array}{ccccc} a & a & a & b & b \\ b & b & b & c & c \\ c & c & c & a & a \end{array} .$$

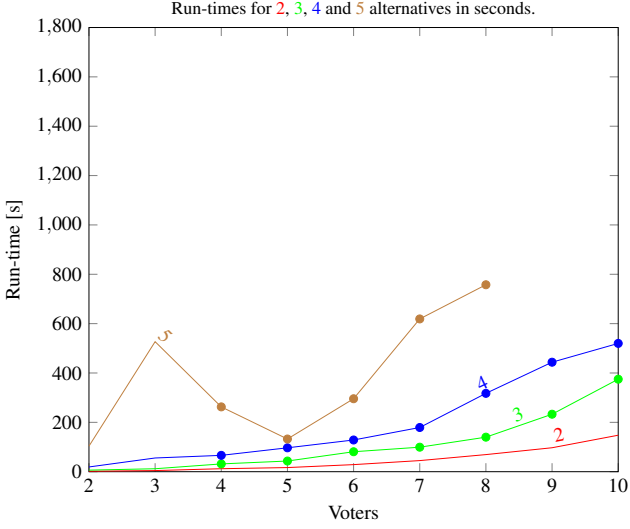


Figure 3: Run-times for the verification of REINF for the Copeland voting rule.

For the profile R , the Borda rule elects the alternative b instead of the majority winner a .

When searching a proof with a minimal number of voters, we find the smallest proof for four alternatives and three voters in less than one second:

$$R = \begin{array}{ccc} d & d & c \\ c & c & a \\ a & b & b \\ b & a & d \end{array}.$$

For the profile R , the Borda rule elects the alternative c instead of the majority winner d . All given examples can be easily inspected manually.

5.3 Automatic comparison of Borda with other voting rules

We here want to compare the Black rule and the Copeland rule with the axiomatization of the Borda rule shown before. We first encode all four axioms in a way similar to the dominance property shown in Listing 1. The axioms ELEM, CYCL and CANC are functional properties and can thus be encoded in the very same manner. As the axiom REINF is a 3-relational property, it needs further statements relating the profiles. The three profiles are initialized with symbolic non-deterministic values as seen in Listing 2. We consider the bound N as the number of voters of the joined profile prof , and fix the sizes of the two sub-profiles prof1 and prof2 using a non-deterministic value s to define prof1 's size as s , and the size of prof2 as $N - s$. We furthermore fix the profiles prof1 and prof2 with respect to prof using an array with non-deterministic values to model a mapping between the joined profile and its two sub-profiles.

Having encoded all four axioms as either functional or k -relational properties in C functions, we can now compare

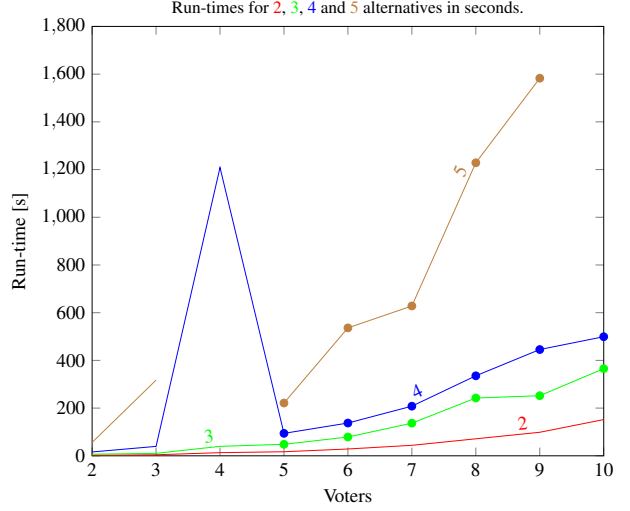


Figure 4: Run-times for the verification of REINF for the Black voting rule.

Table 1: Run-times for the verification that Black and Copeland rule each satisfy CANC.

m	n	verification time
≤ 3	≤ 8	< 3 minutes
4	≤ 5 and 7	< 2 minutes
4	6 and 8	> 30 minutes
5	≤ 3 and 5 and 7	< 1 minute
5	4 and 6 and 8	> 30 minutes

the Black rule and the Copeland rule with this axiomatization. We hereby rely on the “small scope hypothesis” and focus on small profiles with up to five alternatives and eight voters.

Our method successfully verifies that both the Black rule and the Copeland rule satisfy the axiom ELEM in less than 17 seconds (for “small profiles” as defined above). As this axiom must only be verified for two voters, only the number of alternatives must be bounded.

For the axiom CYCL, the verification for both voting rules can be done within 10 minutes (also addressing only “small profiles” as defined above). Here, the numbers of voters are fully determined by the numbers of alternatives (CYCL only addresses profiles with an equal number of alternatives and voters).

When checking the axiom CANC, for both rules, verification is achieved within a few minutes for the cases of eight voters or less when the number of alternatives is bounded to three; five voters or less when the number of alternatives is bounded to four; and three voters or less when the number of alternatives is bounded to five. As run-times for both rules are very similar, both are depicted in Table 1. Note that CANC trivially holds for odd numbers of n , and accordingly verification within the depicted range is achieved in less than 30 seconds.

For the axiom REINF, the run-times can be seen in Figures 3 and 4 for the Copeland rule and the Black rule, respectively. Round dots indicate that violating voting situations have been found for the respective numbers of voters and alternatives. Missing parts (for the case of 5 alternatives) indicate a timeout as defined in Section 5.1. Nonetheless, the time for finding counter-examples of up to ten voters and five alternatives is significantly lower than for verifying that none exists (see, e.g., the run-times for four and five alternatives in Figure 4, comparing run-times for four and five voters). The smallest counter-examples can be found for both voting rules within less than one minute. These simultaneously serve as proofs, which are both short and easy for a human to inspect.

The smallest example (in number of alternatives) proving that the Black voting rule fails REINF has been found in 48 seconds for three alternatives and five voters:

$$R_1 = \begin{array}{cc} a & c \\ c & a \\ b & b \end{array}, \quad R_2 = \begin{array}{ccc} c & b & a \\ b & a & c \\ a & c & b \end{array}.$$

The elected alternatives for the profiles R_1 and R_2 are $\{a, c\}$ and $\{a, b, c\}$ respectively. For the joined profile $R_1 \cup R_2$, the Black voting rule elects the Condorcet winner a instead of the set $\{a, c\}$ mandated by REINF.

The smallest example, both in number of alternatives and in number of voters, proving that the Copeland rule fails REINF has been found in 32 seconds for three alternatives and four voters:

$$R_1 = \begin{array}{cc} b & a \\ a & c \\ c & b \end{array}, \quad R_2 = \begin{array}{cc} a & b \\ b & a \\ c & c \end{array}.$$

The elected alternatives for the profiles R_1 and R_2 are a and $\{a, b\}$ respectively. For the joined profile $R_1 \cup R_2$, the Copeland rule elects the set $\{a, b\}$ instead of only alternative a , which would be required by the REINF property.

Both examples can be easily inspected manually.

6 Literature review

The automatic method presented within this article builds upon the techniques and concepts by Beckert et al. [2016]. It may be used to extend the argumentation framework by Cailloux and Endriss [2016].

Similarly to our approach, Brandt et al. [2017] apply automated reasoning based on SAT-solvers. They prove existing social choice theorems to obtain minimal bounds for the theorems to hold. More work on efficient applications of automated reasoning in social choice theory exists by Geist and Endriss [2011], who automatically detect many impossibility theorems. In another logic-based automatic approach, Xia [2013] incorporates axiomatic properties to a machine learning framework.

Besides fully automatic approaches, other approaches apply interactive techniques using tactical theorem provers in

order to yield unbounded results. This branch was initiated by Nipkow [2009], who provides a mechanized proof for the famous Arrow’s possibility theorem. Beckert et al. [2016] use optimized deductive program verification techniques to verify a number of simple voting rules with respect to classical axioms. Moreover, Dawson et al. [2015] apply interactive theorem proving to formalize and verify complex voting rules according to legal text. Another direction is pursued by Pattinson and Schürmann [2015], who encode complex voting rules in rules of higher-order logic, which serve as proof-carrying code for voting rules and may thus produce certified election results. Verification using tactical theorem provers may lead to very high confidence levels, but requires time-consuming, huge and difficult interactive proofs.

7 Conclusion and future work

In this article, we have presented and formally defined an automatic approach to argue for and against voting rules based on axiomatic social choice properties. The approach is fully automatic and based on the software analysis technique *bounded model checking*. Our case study on the voting rules Borda, Copeland and Black shows that illustrative proofs are obtained automatically in reasonable time for small numbers of voters and alternatives. This might appear surprising given the combinatorial structure of preference profiles and set-valued outcomes.

Based on the small-scope hypothesis, we argue that our approach can be used for arguing about voting rules, especially because it provides short and human-readable proofs when it detects that a voting rule fails to satisfy some axiom.

Ideas for future work include extensions to reasoning about rules that are incompletely specified. For example, we could automatically conceive an argument (in the form of an example profile and a set of winners as illustrated in the experiments) that would attack any rule that does not satisfy a given axiom. In this way, it would be possible to also argue against classes of rules, additionally to concrete rules as illustrated in this work. The approach could also be extended to automatically illustrate differences between two sets of axiomatic properties without the need for any explicit voting rule. A more ambitious extension would consist in automatically deriving proofs that a winner is a “right” winner by using several example profiles, as illustrated (using a non-automatic procedure) by Cailloux and Endriss [2016].

A longer-term objective is to mix the ideas proposed in this work and further optimization techniques [Beckert et al., 2016] with elicitation procedures in order to obtain a system that would permit to recommend a voting rule, possibly based on work in elicitation of scoring rules [Cailloux and Endriss, 2014] and on work that defines justified recommendations as those that resist counter-arguments [Cailloux and Meinard, 2017].

References

- K. J. Arrow. *Social choice and individual values*. Yale University Press, New Haven, 3rd edition, 2012.
- B. Beckert, T. Borner, M. Kirsten, T. Neuber, and M. Ulbrich. Automated verification for functional and relational properties of voting rules. In *Sixth International Workshop on Computational Social Choice (COMSOC 2016)*, 2016. URL <https://www.irit.fr/COMSOC-2016/proceedings/BeckertEtAlCOMSOC2016.pdf>.
- A. Biere, A. Cimatti, E. M. Clarke, and Y. Zhu. Symbolic model checking without BDDs. In R. Cleaveland, editor, *Proceedings of the 5th International Conference on Tools and Algorithms for Construction and Analysis of Systems (TACAS '99)*, volume 1579 of *Lecture Notes in Computer Science*, pages 193–207. Springer, 1999. doi:10.1007/3-540-49059-0_14.
- D. Black. *The Theory of Committees and Elections*, volume 36. Cambridge University Press, 1958. doi:10.1017/S0031819100058204.
- F. Brandt, C. Geist, and D. Peters. Optimal bounds for the no-show paradox via SAT solving. *Mathematical Social Sciences*, 90:18–27, 2017. doi:10.1016/j.mathsocsci.2016.09.003.
- O. Cailloux and U. Endriss. Eliciting a Suitable Voting Rule via Examples. In T. Schaub, G. Friedrich, and B. O’Sullivan, editors, *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI 2014)*, volume 263 of *Frontiers in Artificial Intelligence and Applications*, pages 183–188. IOS Press, 2014. doi:10.3233/978-1-61499-419-0-183.
- O. Cailloux and U. Endriss. Arguing about voting rules. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2016)*. IFAAMAS, 2016. URL <https://dl.acm.org/citation.cfm?id=2936968>.
- O. Cailloux and Y. Meinard. A formal framework for deliberated judgment, 2017. URL <https://arxiv.org/abs/1801.05644>.
- E. Clarke, D. Kroening, and F. Lerda. A tool for checking ANSI-C programs. In K. Jensen and A. Podelski, editors, *Proceedings of the 10th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2004)*, volume 2988 of *Lecture Notes in Computer Science*, pages 168–176. Springer, 2004.
- J. M. Colomer. Ramon Llull: from ‘Ars electionis’ to social choice theory. *Social Choice and Welfare*, 40(2): 317–328, 2013. doi:10.1007/s00355-011-0598-2.
- A. H. Copeland. A ‘reasonable’ social welfare function. seminar on applications of mathematics to social sciences. In *University of Michigan Seminar on Applications of Mathematics to the Social Sciences*, 1951.
- J. E. Dawson, R. Goré, and T. Meumann. Machine-checked reasoning about complex voting schemes using higher-order logic. In R. Haenni, R. E. Koenig, and D. Wikström, editors, *Proceedings of the 5th International Conference on E-Voting and Identity (VoteID 2015)*, volume 9269 of *Lecture Notes in Computer Science*, pages 142–158. Springer, 2015. doi:10.1007/978-3-319-22270-7_9.
- N. Eén and N. Sörensson. An extensible SAT-solver. In *International Conference on Theory and Applications of Satisfiability Testing (SAT 2003)*, *Selected Revised Papers*, pages 502–518, 2003.
- C. Geist and U. Endriss. Automated search for impossibility theorems in social choice theory: Ranking sets of objects. *Journal of Artificial Intelligence Research (JAIR)*, 40, 2011.
- D. Jackson. *Software Abstractions - Logic, Language, and Analysis*. MIT Press, 2006. URL <http://mitpress.mit.edu/catalog/item/default.asp?type=2&tid=10928>.
- T. Nipkow. Social choice theory in hol. *Journal of Automated Reasoning*, 43(3):289–304, 2009. doi:10.1007/s10817-009-9147-4.
- D. Pattinson and C. Schürmann. Vote counting as mathematical proof. In B. Pfahringer and J. Renz, editors, *Proceedings of the 28th Australasian Joint Conference on Advances in Artificial Intelligence (AI 2015)*, volume 9457 of *Lecture Notes in Computer Science*, pages 464–475. Springer, 2015. doi:10.1007/978-3-319-26350-2_41.
- L. Xia. Designing social choice mechanisms using machine learning. In M. L. Gini, O. Shehory, T. Ito, and C. M. Jonker, editors, *International conference on Autonomous Agents and Multi-Agent Systems (AAMAS '13)*, pages 471–474. IFAAMAS, 2013. URL <http://dl.acm.org/citation.cfm?id=2484995>.
- H. P. Young. An axiomatization of borda’s rule. *Journal of Economic Theory*, 9(1):43 – 52, 1974. doi:10.1016/0022-0531(74)90073-8.

Categorization of B2B Service Offers: Lessons learnt from the Silex Use case

Molka Tounsi Dhouib^{1 2}

Catherine Faron Zucker¹

Andrea Tettamanzi¹

¹ Université Côte d'Azur, Inria, CNRS, I3S, Sophia Antipolis, France

² Silex France

{dhouib, faron, tettamanzi}@i3s.unice.fr

Résumé

Dans le domaine de la recherche d'information et du traitement automatique du langage, la tâche de classification de textes est devenue une tâche cruciale. Dans cet article, nous partageons notre expérience de la classification de textes dans un contexte industriel et présentons une évaluation comparative de différents algorithmes de classification binaire et multi-label appliqués à des textes décrivant des offres de services, issus de la plateforme B2B SILEX pour la recommandation de prestataires de services. Nous montrons que dans certains cas pratiques comme celui que nous considérons, une représentation des données sous la forme de "bags of words" donne de meilleurs résultats de classification qu'une représentation réputée plus prometteuse par "word embeddings".

Mots Clef

Catégorisation de textes, Apprentissage automatique, Sac de mots, Plongement lexical

Abstract

In the domain of Information Retrieval and Natural Language Processing, text classification has become a crucial task. In this article, we share our experience of text categorization in an industrial context and we present a comparative evaluation of binary and multi-label classification algorithms applied to texts describing service offers, in the SILEX B2B platform. We show that for some use cases like the one we consider, a traditional representation of texts by "bags of words" gives better classification results than the promising representation by "word embeddings".

Keywords

Text categorization, Machine Learning, Bag of words, Word embedding

1 Introduction

The Silex France company offers a SaaS sourcing tool for identification of the service providers that are best suited to meet the service requests expressed by companies. The Silex platform is used by more than 5000 professionals to quickly identify and exchange with B2B service providers. Silex's ultimate aim is to provide a network of qualified

companies to service buyers combined with the best functionality to consult it.

In the framework of a collaborative project between Silex and the I3S research laboratory, we aim at introducing semantics into the B2B platform in order to enable automatic reasoning on service requests and offers and improve the recommendation of service providers.

As a first step, we are interested in automatically categorizing the textual description of companies, service requests and service offers. With data's exponential growth, text categorization has become a crucial issue in Natural Language Processing (NLP). It consists in classifying texts according to their contents, into one or more predefined categories [2]. In recent years, the number of machine learning (ML) techniques that automatically generate text categorization has increased considerably [3].

In this paper we report some experiments we conducted to answer the Silex use case, using supervised ML techniques to classify Silex textual data into predefined categories. Meanwhile, we addressed the following questions:

(1) *What is the best representation of Silex textual descriptions of service offers and requests to categorize it?*

(2) *What is the best ML algorithm to categorize Silex textual descriptions of service offers and requests?*

This paper is organized as follows: Section 2 presents state-of-the-art text categorization methods. Section 3 introduces our approach of text categorization to answer the SILEX use case. Section 4 reports and discusses the results of our experiments of categorization of NL descriptions of service requests and offers. Section 5 draws some conclusions and directions for future work.

2 Related Works

Text categorization is the task of classifying data into a predefined set of categories. In other words, given a set of categories and a set of textual documents, text categorization is the process of automatically finding the correct category for each document [4].

2.1 Feature Vector Models

In the text categorization process, the first step consists in preprocessing textual documents, to convert them into feature vectors, a representation that can be automatically interpreted by machine. This step includes tokenization,

stemming and removing of stop words before the creation of feature vectors.

Bag-Of-Words (BOW). is the most common feature vector model. In this model, the features are the frequencies of each word in the textual document. The feature space's dimension is the number of all different words in all documents [4]. The limitation of this model is that it ignores the semantic relations between words.

Word Embedding (WE). In order to overcome the weakness of BOW, a new model called Word Embedding (WE) has been successfully used in several NLP tasks. Word embeddings are projections in a continuous space of words that preserve the semantic and syntactic similarities between them [6]. There are many models that can produce a word embedding. In the following, we introduce the most important ones.

[7] proposes Word2Vec, a popular tool that produces word embeddings based on two models: *CBOW* and *Skip-gram*. *CBOW* is a Neural Network and log-linear model. It removes the non-linear hidden layer, and projects the contextual words on the same position. Word's prediction is obtained according to its past and future contexts. The process consists in computing the average of the contextual word vectors and running a log-linear classifier on the averaged vector.

Skip-gram is similar to *CBOW* and is also a Neural Network and log-linear model. Contrary to *CBOW*, *Skip-gram* predicts the contextual words given the current word [6].

Another model was introduced by [8] and is called *GloVE*. This model is based on global matrix factorization that calculates the co-occurrence of words in the corpus [6].

Finally, *FastText*¹ is a library for learning word representations and sentence classification. FastText published pre-trained word vectors for 294 languages, trained on Wikipedia. These vectors with dimension 300 were obtained using the skip-gram model [13].

2.2 Binary classification

After the construction of feature vectors, the second step of text categorization is the learning step. We present here some of the many learning algorithms for binary classification and multi-label classification.

Naive Bayes (NB) [9]. is a supervised, probabilistic algorithm based on Bayes theorem and the hypothesis of the independence between features. This algorithm is powerful due to the independence between features that ignores features' order, and consequently, the presence of a feature does not affect other features in classification tasks [14].

Support Vector Machine (SVM) [10]. is one of the most common and successful supervised classification algorithm used for text classification tasks [14]. This algorithm transforms training data into higher dimensions and searches for linear optimal separating hyperplane [1].

Neural Networks (NN). The NN is composed by many layers to perform text categorization. The perceptron is the simplest kind of a neural network and it has only two layers: the input nodes receive the feature values, the output nodes produce the categorization status values, and the link weights represent dependence relations. The NN text categorization process starts by loading feature weights into the input nodes; the categorization's final result is present in the output nodes after the propagation of the activation of the nodes forward through the network. The neural networks are trained by back propagation in order to minimize the error. In case of misclassification, the error is propagated back through the network and modifies the link weights [15].

2.3 Multi-label classification

Multi-label classification is an approach to classification problems that allows each data point to be assigned to more than one class at the same time. There are two multi-label classification methods: (i) problem transformation and (ii) algorithm adaptation [11]. Here we describe the most popular algorithms of each category.

Problem transformation. A first approach consists in transforming the multi-label problem into one or more single label classification problem. Among the methods adopting this approach, we can highlight:

- **The Binary Relevance method (BR)** is based on one-vs-all ensemble approach. BR transforms any multi-label problem into binary problems to predict the relevance of each label to a data point. All binary classifiers are then aggregated to form a set of relevant labels [11].
- **The label PowerSet method (LP)** considers the multi-label problem as a single multi-class classification problem. In order to create a transformed multi-class dataset, each combination of relevant labels is mapped to a class [11].
- **The Random K-label set (RAKEL)** aims at finding a better balance between BR and LP. RAKEL generates a series of label subsets and builds a label powerset model for each of them [11].

Algorithm adaptation. Algorithm adaptation extends specific algorithms to carry out multi-label classification.

- **Multi-label lazy algorithm (ML-KNN)** is one of the most famous adaptation algorithms. MLKNN uses the maximum a posteriori principle in addition to the K-nearest neighbour algorithm [11].
- **Multi-label decision tree (ML-DT)** is an extension of C4.5 decision tree algorithm. ML-DT allows the creation of multiple labels in the leaves, and chooses node splits based on a pre-defined multi-label entropy function [11].

¹<https://github.com/facebookresearch/fastText>

- **AdaBoost** is based on the addition of simple classifiers to a pool and the use of their weights to define the final classification [16].

3 Experimental methodology

To answer the Silex use case, we conducted some experiments to compare the performance of the above described state-of-the-art feature vector models and learning algorithms for the categorization of textual descriptions of service offers or requests.

3.1 Dataset

Silex distributes a service to purchasing departments that support the sourcing system within the company. Silex users can describe their company (company description), their needs (service request description) and the description of the offers provided by the company (service offer description) in Silex web application. All descriptions are written in French.

The dataset considered in our experiments comprises 3188 company descriptions, 580 service offers descriptions and 155 service requests descriptions.

Six main categories are defined by Silex to classify these descriptions. One description can be categorized in one or more classes. The categorization phase is done manually by experts in the current version of the Silex platform. Table 1 shows the distribution of all the descriptions into the different categories. We observe that the distribution of our dataset is unbalanced, for instance the *Informatique* (Information Technology) class has 915 more companies descriptions than the *Services_industriels* (Industrial Services) class.

Categories	Companies descriptions	service offer descriptions	service request descriptions
Informatique	1043	147	42
Finance	393	40	13
Services généraux	929	149	65
Marketing	926	160	18
Ressources humaines	297	57	7
Services industriels	128	32	19

Table 1: Distribution of the textual descriptions in the corpus.

3.2 Implementation

In our implementation, we focused first on two methods to construct our feature vectors:

- The BOW representation is achieved by applying a traditional process of tokenization, stop words removal and lemmatization. Then we use the TFIDF to construct the matrix. This matrix will be used in different algorithms.

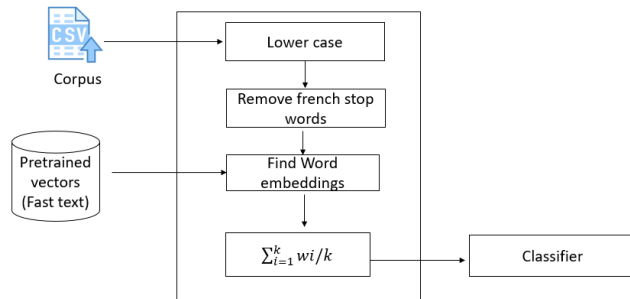


Figure 1: Building WE feature vectors representing the Silex textual descriptions from pre-trained word embeddings

- There are two methods to obtain Word embeddings representation: (i) build word embeddings from scratch using a model like word2vec or FastText. This approach needs a very large corpus to train these word embeddings; and (ii) use a pre-trained word embeddings in any desired language. For our case, we used the latter approach to obtain word embeddings because we do not have a large enough dataset. After studying the different pre-trained word embeddings, we decided to use the FastText model because it is the only model that provides French word embeddings pre-trained on a Wikipedia dump. This word embedding model ² contains 1,152,449 tokens, which are mapped to vectors in a space with 300 dimensions.

Our approach to produce the feature vectors representing the SILEX textual descriptions is inspired by [12] and [17]: As illustrated in Figure 1, the first step is to convert textual descriptions to lower case and remove all French stop words. Then the words are checked against the pretrained French vector. If the word exists in the dictionary, the vector representation of description is constructed by averaging the word embeddings vectors along each dimension for all the words in the description. If the dictionary does not contain the word, the zero vector is returned:

$$DescriptionWordEmbeddings = \sum_{i=1}^k \frac{w_i}{k}$$

where k is the number of words in a description and $w_i \in \mathbb{R}^n$ denotes the word embeddings vector of the i th word. The vector representation of each description has exactly the same dimensions as the word embeddings vector w_i [17]. The vector representations of each description are then used as features and input for each classifier.

²<https://fasttext.cc/docs/en/pretrained-vectors.html>

In order to test our feature vectors, we trained two methods of categorization: (i) three binary classification algorithms (NB, SVC, RNN) and (ii) three multi-label classification algorithms: two of them are problem transformation methods (Binary Relevance, Label PowerSet) and three of them are algorithm adaptation methods (Multi-label lazy algorithm (ML-KNN)). All these algorithms are implemented using scikit-learn³, and executed on a machine with 8 GB of memory and an Intel Core i7-7500U CPU.

To deal with the unbalanced data, there are two main state-of-the-art approaches to adjust the distribution of dataset: (i) over-sampling adds copies of instances from the under-represented class. The most common technique of over-sampling is called Synthetic Minority Over-sampling Technique (SMOTE), (ii) under-sampling deletes instances from the over-represented class. In our implementation, we tested first a very large parameter with a grid function in scikit-learn to identify the best parameters for each classifier and the best method to deal with unbalanced data. We obtained the best results with SMOTE method for SVM algorithm, and RandomOverSampler method for NB. Table 2 shows all parameters used in binary classification and Table 3 shows all parameters used in multi-label classification.

Classifier	Unbalanced data	Parameters
NB	RandomOverSampler	
SGDClassifier	SMOTE	penalty=l2 alpha=0.0001
LinearSVC	SMOTE	C=0.1 penalty='l2' loss='squared hinge'
SVC	SMOTE	kernel='rbf' C=1000.0 gamma=0.0001
RNN	balanced	hidden size=32 Activation('relu') Activation('sigmoid')

Table 2: Parameters used for binary classification

Classifier	multi-label methods	Parameters
LSVC	Binary Relevance	C=1, penalty='l2'
LSVC	Label Powerset	C=1, penalty='l2'
MLkNN	Algorithm adaptation	-

Table 3: Parameters used for multi-label classification

4 Experiments and discussion

4.1 Evaluation procedure

The experiments were conducted using the 5-fold cross-validation methodology. To evaluate our experiments, we calculated the classification duration and various evaluation measures inspired by [18] and described in the following.

³<http://scikit-learn.org/stable/>

Binary algorithms. To evaluate binary classification algorithms, we used the **F1 score** measure, defined as the harmonic mean between precision and recall:

$$F1score = \frac{1}{N} \sum_{i=1}^N \frac{2 \times |h(x_i) \cap y_i|}{|h(x_i) + y_i|}$$

Multi-label algorithms. To evaluate multi_labels algorithms we used the following metrics:

- **Accuracy** allows to compute the percentage of correctly predicted labels among all predicted and true labels. It is defined as follows:

$$Accuracy(h) = \frac{1}{N} \sum_{i=1}^N \left| \frac{h(x_i) \cap y_i}{h(x_i) \cup y_i} \right|$$

- **Micro-precision** is defined as the precision averaged over all the example/label pairs:

$$Microprecision = \frac{\sum_{j=1}^Q tp_j}{\sum_{j=1}^Q tp_j + \sum_{j=1}^Q fp_j}$$

where tp_j , fp_j are the number of true positives and false positive for label λ_j .

- **Micro-recall** is defined as recall averaged over all the example/label pairs:

$$Microrecall = \frac{\sum_{j=1}^Q tp_j}{\sum_{j=1}^Q tp_j + \sum_{j=1}^Q fn_j}$$

where fn_j is the number of false negatives for the label λ_j .

- **Micro-F1** is the harmonic mean between micro-precision and micro-recall and is defined as:

$$Micro - F1 = \frac{2 \times microprecision \times microrecall}{microprecision + microrecall}$$

- **Hamming loss** allows to evaluate how many times an example-label pair is misclassified. The performance is perfect when the value of this metric is equal to 0. Hamming loss is defined as

$$Hamming_loss(h) = \frac{1}{N} \sum_{i=1}^N \frac{1}{Q} |h(x_i) \Delta y_i|$$

where Δ is the symmetric difference between two sets, N is the number of examples and Q is the total number of possible class labels.

4.2 Results and Analysis

The results of binary classification algorithms reported in Table 4, Table 5 and Table 6 show that there is not a significant difference between the results with a BOW representation and the results with a word embeddings representation. For example, for category "Informatique", we obtain the same precision value that equals to 0.82. The recall value equals to 0.82 for BOW representation, and is a bit better compared to WE recall value that equals to 0.79. The classification processing time is much better for WE (0.48 seconds) than for BOW (1572 seconds). This is due to the BOW expansive phases of lemmatization and matrix building compared to the vector representation of word embeddings.

Table 4: Best classifier of binary classification of companies descriptions

Category	Method	P	R	F1	Time
Informatique	LSVC + WE	0.82	0.79	0.80	0.48
	SGD+BOW	0.82	0.82	0.82	1572
Finance	LSVC+WE	0.90	0.87	0.88	0.21
	NB+BOW	0.92	0.87	0.88	1793
Services généraux	LSVC+ WE	0.87	0.84	0.85	0.4
	NB + BOW	0.90	0.89	0.89	1925
Marketing	LSVC+WE	0.82	0.80	0.81	0.41
	SGD + BOW	0.85	0.85	0.85	1517
Ressources humaines	LSVC + WE	0.91	0.83	0.86	0.19
	NB + BOW	0.92	0.84	0.87	1904
Services industriels	LSVC+WE	0.94	0.89	0.91	0.05
	NB+BOW	0.97	0.94	0.95	1914

Table 5: Best classifier of binary classification of service offer descriptions

Category	Method	P	R	F1	Time
Informatique	RNN+WE	0.80	0.78	0.79	8.53
	RNN +BOW	0.78	0.78	0.78	332
Finance	SVC-RBF+ WE	0.82	0.87	0.84	0.083
	SGD + BOW	0.91	0.90	0.86	323
Services généraux	RNN+ WE	0.82	0.81	0.81	6.97
	RNN+BOW	0.79	0.79	0.79	348
Marketing	SGD+WE	0.80	0.77	0.78	0.028
	NB + BOW	0.90	0.90	0.90	459
Ressources humaines	RNN + WE	0.90	0.91	0.90	7.42
	NB + BOW	0.87	0.87	0.87	419
Services industriels	SVC-RBF+ WE	0.91	0.92	0.91	0.09
	SVC-RBF+BOW	0.96	0.97	0.96	347

We can draw similar conclusions when comparing multi_label classification algorithms, with a slight advantage to BOW compared to word embeddings in terms of accuracy and hamming_loss but not in term of time.

If we compare binary classification and multi_label classification algorithms, we can see that the results of the first algorithm are better than the second one. These results are as expected because first we don't have a very large corpus, and second when we use a binary classification we can play

Table 6: Best classifier of binary classification of service request descriptions

Category	Method	P	R	F1	Time
Informatique	NB +WE	0.82	0.77	0.79	0.006
	NB+ BOW	0.79	0.74	0.76	101.31
Finance	RNN +WE	0.97	0.48	0.62	2.44
	LSVC + BOW	0.94	0.94	0.92	84.82
Services généraux	NB+ WE	0.74	0.74	0.74	0.006
	NB + BOW	0.74	0.74	0.74	103.49
Marketing	RNN + WE	0.73	0.77	0.74	2.74
	NB +BOW	0.89	0.81	0.84	104.10
Ressources Humaines	NB+ WE	0.88	0.94	0.90	0.006
	BOW	-	-	-	-
Services industriels	RNN + WE	0.91	0.90	0.88	2.62
	LSVC+BOW	0.94	0.94	0.94	79.07

Table 7: Best classifier of multi_label classifications.

Classifier	Feature	Micro-F1	Accuracy	Hamming loss	Time
MLKNN	WE	0.66	0.53	0.12	7.00
	BOW	0.72	0.57	0.10	1945
BR	WE	0.61	0.35	0.17	6.54
	BOW	0.73	0.52	0.10	1850
Label Powerser	WE	0.54	0.39	0.19	9.44
	BOW	0.69	0.51	0.13	1945

on the classifier parameters for each category to enhance its performance, which is not the case for multi_label classification. The different binary classification algorithms perform more or less well in categorizing our dataset using the same representation. For example, only the LSVC algorithm categorized correctly this text as "informatique" *"Ascot accompagne les dirigeants et cadres de PME dans la définition, la maîtrise et le suivi de leur transformation numérique. Plus aucune organisation ne peut ignorer l'impact global sur son métier, ses outils et ses processus. Notre expérience, de plus de 20 ans, en Conseil, Assistance technique, Développement, Ingénierie, ... est entièrement dédié à nos clients. Sous contrat ou à la carte, nos prestations leur apportent valeur et sérénité."*

We obtained also different results when comparing the BOW and WE representations. For example, using the same LinearSVC algorithm, this text *"Agence WebMarketing 360 de nouvelle génération exploitant de nombreux leviers marketing pour optimiser vos campagnes digitales & booster votre chiffre d'affaires Kalipseo développe pour vous les stratégies digitales les plus pertinentes et s'emploie à promouvoir l'image de votre marque."* was wrongly categorized as "Informatique" when using the WE representation, but the categorization was right when using BOW representation.

As a conclusion, to best answer the Silex use case, we decided to use a binary classification algorithm with a word embedding representation. Table 8 shows the algorithms that are selected to classify each category of Silex data.

Table 8: Selected algorithms for Silex Company

Category	Description	Representation	Algorithm
Informatique	Companies	WE	NB
	Service offer	WE	RNN
	Service provider	WE	NB
Finance	Companies	WE	RNN
	Service offer	WE	SVC-RBF
	Service provider	WE	RNN
Services généraux	Companies	WE	NB
	Service offer	WE	RNN
	Service provider	WE	NB
Marketing	Companies	WE	RNN
	Service offer	WE	SGD
	Service provider	WE	RNN
Ressources Humaines	Companies	WE	NB
	Service offer	WE	RNN
	Service provider	WE	NB
Services industriels	Companies	WE	RNN
	Service offer	WE	SVC-RBF
	Service provider	WE	RNN

5 Conclusion

In this paper we reported the results of the experimental evaluations of various vector feature models and machine learning algorithms conducted to answer the real-world use case of the Silex company: how to categorize textual descriptions of service offers and requests with the ultimate goal of recommending service providers that better answer services requests. We compared the two main state-of-the-art feature vector models. A BOW representation is a bit better than a WE representation regarding the evaluation measures. This is due to the use of a generic pre-trained vector for WE descriptions, that is based on Wikipedia dump and that does not cover well our data set. On the other hand, a WE representation has the advantage of being less time consuming because of the additional lemmatization and matrix building phases specific to the BOW representation. Based on these results, we choose to use classifiers with word embeddings data for Silex company as shown in Table 8.

As future work, we aim to define a specific B2B pre-trained vector that best covers the Silex dataset instead of using a general reference like Wikipedia. We also aim to study a combined approach based on both word embeddings and knowledge engineering approaches. The objective is to improve the representation of texts by introducing semantics to increase the performance of algorithms. To achieve this, we are going to use our built ontologies to represent the sourcing domain, and we are going to enrich the original texts with additional information based on semantics relations between concepts such as generalization or specification.

6 Acknowledgement

This work is supported by the ANRT, CIFRE I3S-SILEX 2016/0947

References

- [1] Patra, Anuradha and Singh, Divaka A survey report on text classification with different term weighing methods and comparison between classification algorithms *International Journal of Computer Applications*, volume.75, pp.7,2013,Foundation of Computer Science.
- [2] Sadiq, Ahmed T and Abdullah, Sura Mahmood, Hybrid intelligent technique for text categorization *Advanced Computer Science Applications and Technologies (ACSAT), 2012 International Conference on*, pp.238-245,2012,IEEE.
- [3] Tan, Ah-Hwee and Lai, Fon-Lin, Text categorization, supervised learning, and domain knowledge integration *To appear, proceedings, KDD-2000 International Workshop on Text Mining, Boston*, Vol.20,2000.
- [4] Feldman, Ronen and Sanger, James, *The text mining handbook: advanced approaches in analyzing unstructured data*,2007,Cambridge university press.
- [5] Sriram, Bharath and Fuhry, Dave and Demir, Engin and Ferhatosmanoglu, Hakan and Demirbas, Murat, Short text classification in twitter to improve information filtering, *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*,pp.841–842, 2010, ACM
- [6] Ghannay, Sahar and Favre, Benoit and Esteve, Yannick and Camelin, Nathalie, *Word Embedding Evaluation and Combination*,LREC,2016
- [7] Mikolov, Tomas and Chen, Kai and Corrado, Greg and Dean, Jeffrey, *Efficient estimation of word representations in vector space*,arXiv preprint arXiv:1301.3781, 2013
- [8] Pennington, Jeffrey and Socher, Richard and Manning, Christopher, *Glove: Global vectors for word representation*,*Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*,pp.1532-1543,2014
- [9] Schneider, Karl-Michael, *Techniques for improving the performance of naive bayes for text classification*, *Computational Linguistics and Intelligent Text Processing*, pp.682-693, 2005,Springer
- [10] Cortes, Corinna and Vapnik, Vladimir, *Support-vector networks*, *Machine learning*, Vol.20,3,pp.273-297,1995,Springer
- [11] Pakrashi, Arjun and Greene, Derek and MacNamee, Brian, *Benchmarking Multi-label Classification Algorithms*,24th Irish Conference on Artificial Intelligence and Cognitive Science (AICS'16), Dublin, Ireland, 20-21 September 2016,2016, CEUR Workshop Proceedings

- [12] Bayot, Roy and Gonçalves, Teresa, Author Profiling using SVMs and Word Embedding Averages – Notebook for PAN at CLEF 2016, 2016, CEUR
- [13] Bojanowski, Piotr and Grave, Edouard and Joulin, Armand and Mikolov, Tomas, Enriching Word Vectors with Subword Information, arXiv preprint arXiv:1607.04606, 2016
- [14] Mohare, Punam Kishor Itkar, A, A Survey on Different Types of Approaches to Text Categorization, V5, International Journal of Emerging Trends & Technology in Computer Science
- [15] Bassil, Youssef, A Survey on Information Retrieval, Text Categorization, and Web Crawling, arXiv preprint arXiv:1212.2065, 2012
- [16] Benbouzid, Djalel and Busa-Fekete, Róbert and Casagrande, Norman and Collin, François-David and Kégl, Balázs, MultiBoost: a multi-purpose boosting package, V.13, pp.549-553, Journal of Machine Learning Research, 2012
- [17] Yang, Xiao and Macdonald, Craig and Ounis, Iadh, Using word embeddings in twitter election classification, arXiv preprint arXiv:1606.07006, 2016
- [18] El Kafrawy, Passent and Mausad, Amr and Esmail, Heba, Experimental comparison of methods for multi-label classification in different application domains, V.114, N.19, Foundation of Computer Science, International Journal of Computer Applications, 2015
- [19] Sokolova, Marina and Japkowicz, Nathalie and Szpakowicz, Stan, Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation, V.4304, pp.1015-1021, Australian conference on artificial intelligence, 2006

Bandits-Manchots Contextuels : Précision Globale Versus Individuelle

Nicolas Gutowski^{1,2}

Tassadit Amghar²

Olivier Camp¹

Fabien Chhel¹

¹ ESEO-TECH

10 boulevard Jean Jeanneteau, 49100 Angers, France

² LERIA, Université d'Angers (UBL)

2 boulevard Lavoisier, 49000 Angers, France

nicolas.gutowski@eseo.fr

Résumé

Dans la littérature, la plupart des travaux sur les bandits manchots sont évalués à l'aide d'une mesure de la précision globale. Concernant les bandits manchots contextuels, les approches existantes ont pour objectif d'atteindre une personnalisation individuelle. Ainsi, leur précision globale devrait refléter la précision individuelle pour chacun des utilisateurs. Afin de mesurer le niveau de personnalisation atteint par ces approches, nous avons défini une nouvelle évaluation comparant les précisions individuelles des recommandations faites à chaque utilisateur avec la précision globale. Sur la base de cette comparaison, démontrant des disparités entre la précision individuelle et la moyenne de précision globale, nous proposons *Sliding Window LinUCB (SW-LinUCB)*. *SW-LinUCB* est une combinaison de *LinUCB* (CMAB) et d'un mécanisme de diversification pénalisant les bras sélectionnés trop fréquemment. Notre approche, inspirée d'applications réelles, comme les systèmes de recommandation, ne nécessite pas uniquement d'atteindre une bonne précision globale mais doit aussi tenir compte de la précision individuelle. Nous expérimentons et discutons nos résultats sur plusieurs jeux de données réelles.

Mots Clef

Apprentissage par renforcement, LinUCB, Bandits manchots contextuels, Système de recommandation

Abstract

Most works on Multi-Armed Bandits (MAB) focus the evaluations of their methods on a global accuracy performance metric. In the case of Contextual Multi-Armed Bandit (CMAB), the existing algorithms claim to eventually provide full personalization, which might suggest that their global accuracy metric should reflect each user's individual accuracy. In order to verify this, we consider a novel approach of CMAB assessment focused on the evaluation of individual accuracy and compare it to global accuracy. Based on the results of this comparison highlighting some

users far from the average global accuracy, we propose *Sliding Window LinUCB (SW-LinUCB)*, a combination of the original *LinUCB* (CMAB) and a diversification mechanism penalizing arms which are pulled too frequently. It is motivated by the requirements of different real-world applications such as clinical trials or recommender systems, which must converge to a good global accuracy and should equally distribute it among individuals. We experiment and discuss the benefits and losses of the proposed method on several real-world datasets.

Keywords

Recommendation System, Reinforcement learning, LinUCB, Contextual Multi-armed Bandits

1 Introduction

De nos jours, les bandits manchots contextuels (Contextual Multi-Armed Bandit : CMAB) sont très largement considérés par de nombreuses applications se heurtant à des problèmes de décision séquentielle e.g., les systèmes de recommandation [1], ou encore les essais cliniques [2]. À chaque itération, les algorithmes d'apprentissage de CMAB ont pour objectif de choisir une action optimale (tirer le bras optimal) parmi un ensemble de possibilités, en tenant compte du contexte donné et des récompenses passées obtenues en regard de ces actions. Dans la majorité des cas, les récompenses retournées sont égales à 1 si l'algorithme réalise une bonne classification du contexte par rapport à l'action choisie, et à 0 sinon [1]. Ainsi, la plupart des études évaluent la performance de leurs méthodes, jusqu'à un horizon final, à travers une mesure de précision globale e.g., récompense moyenne, cumul des récompenses, ou nombre de regrets total [1, 3, 4]. Néanmoins, de telles métriques semblent inadéquates pour déterminer la précision à associer à chaque contexte [5]. Ainsi, si nous prenons l'exemple des systèmes de recommandation pour lesquels les utilisateurs peuvent être des visiteurs réguliers, ou des abonnés (exemple : applications mobiles), il semble

essentiel de tenir compte de leurs retours individuels en regard des recommandations qui leur sont faites.

De telles considérations nous ont amené à définir une méthode permettant de mesurer la précision individuelle, et à évaluer différents algorithmes existants au regard de notre nouvelle métrique. Le problème se pose alors comme étant double objectif puisqu'il convient de maximiser la précision individuelle tout en maintenant une bonne précision globale.

Ainsi, nous introduisons une mesure complémentaire de la performance qui est fondée sur l'observation de la fonction de distribution cumulative (Cumulative Distribution Function : CDF) de la précision individuelle. De plus, nous proposons l'adaptation d'une méthode de CMAB que nous nommons *SW-LinUCB*. Cette méthode est bâtie à partir de l'algorithme classique *LinUCB* combiné avec une fenêtre glissante. De ce fait, *SW-LinUCB* a pour objectif d'améliorer la précision individuelle en incorporant un mécanisme de diversification, tout en conservant une précision globale satisfaisante.

En nous appuyant sur des jeux de données d'applications réelles, nous étudions les performances de différentes méthodes existantes (MAB et CMAB) à travers notre mesure de CDF sur l'ensemble des contextes observés. Nous montrons que pour les deux méthodes — pour les CMAB, spécifiquement selon le niveau de parcimonie du vecteur de contexte — un fossé se creuse entre les utilisateurs pour lesquels nous obtenons un haut niveau de précision et ceux défavorisés par un très faible niveau de précision, malgré une précision globale satisfaisante. Nos expériences montrent que *SW-LinUCB* réussit à combler ce fossé tout en maintenant une précision globale qui ne décroît pas plus de 10% par rapport à celle d'origine. En outre, on constate que notre algorithme parvient à atténuer les effets négatifs d'un contexte insuffisamment renseigné.

La contribution de notre article est double : 1) nous proposons une nouvelle mesure de l'évaluation basée sur la précision individuelle ; 2) nous présentons *SW-LinUCB* qui est une adaptation de *LinUCB* dont l'objectif est de maximiser la précision individuelle tout en conservant une précision globale satisfaisante.

Cet article est organisé comme suit : la section 2 présente un état de l'art sur les problèmes de MAB et les différentes métriques utilisées. La section 3 introduit les travaux connexes sur les problématiques de CMAB, l'algorithme *LinUCB* et l'usage de fenêtres glissantes. La section 4 dresse notre problématique et la méthode que nous mettons en place. Dans la section 5 nous exprimons et discutons les résultats de nos expérimentations. Enfin, nous concluons et présentons les perspectives de notre travail dans la section 6.

2 Contexte

Le problème du Bandit-Manchot (Multi-Armed Bandit : MAB) est un sujet qui a suscité de nombreuses recherches depuis sa première formalisation en 1952 [6]. De nom-

breuses formulations ont pu être proposées : stochastiques [7, 8, 4], ou encore Bayésiennes [9]. Plus précisément, le défi pour tout problème de MAB consiste à construire une stratégie visant à tirer le bras optimal sans connaissance préalable de la rentabilité de chacun des bras disponibles. La résolution de ce problème consiste à trouver un compromis entre l'exploration de l'ensemble des bras pour en déduire leurs rentabilités et l'exploitation de ce qui a été inféré pour favoriser la sélection des bras optimaux. Une version étendue de ce problème prend en compte le contexte. Il s'agit du problème de Bandit-Manchot Contextuel (Contextual Multi-Armed Bandit : CMAB) [10, 3]. Ainsi, dans une approche CMAB, le défi visant à déterminer le bras optimal reste le même que pour un problème de MAB mais doit tenir compte du contexte des utilisateurs.

Dans la littérature, le critère le plus fréquemment observé pour mesurer la performance d'un algorithme de bandit reste la précision globale — i.e. le nombre de fois qu'une récompense positive a été obtenue en tirant les différents bras [7, 8, 1, 3, 4]. Néanmoins, en fonction du domaine dans lequel les bandits sont appliqués, l'évaluation de leur performance peut nécessiter de s'ouvrir à d'autres critères. En effet, comme c'est tout particulièrement le cas pour les systèmes de recommandations, il a été observé dans certaines études que les mesures de précision ne sont pas suffisamment adaptées et pourraient être préjudiciables et nuire à la satisfaction des utilisateurs [5]. De ce fait, même si des algorithmes de CMAB tels que *LinUCB* [1] ou encore *Contextual Thompson Sampling* [3] permettent à terme une personnalisation complète auprès de chaque utilisateur, une autre étude soutient en revanche que ceux-ci nécessitent un si grand nombre d'itérations pour atteindre cette personnalisation qu'ils risquent de causer la frustration des utilisateurs avant d'y parvenir [11].

De tels constats ont conduit des recherches sur les CMAB et les systèmes de recommandations vers deux directions : 1) Tenter de réduire le nombre d'itérations nécessaires pour atteindre la personnalisation pour chaque utilisateur [11, 12], 2) Prendre en considération d'autres critères d'évaluation de la performance comme : la qualité [13], la diversité et la nouveauté [14], la couverture et la sérénité [15], ou encore la satisfaction utilisateur [16].

La majorité des travaux tendent à montrer que la diversification serait l'un des points-clé pour améliorer la satisfaction utilisateur. Par exemple cela permettrait de mieux répondre aux besoins éphémères des utilisateurs [17], de les aider à découvrir de nouveaux éléments [13], ou d'éviter les recommandations redondantes [18].

À notre connaissance, aucune approche n'aborde spécifiquement le problème de la recherche d'un compromis entre précision individuelle et précision globale pour les CMAB à travers l'usage de techniques de diversification. Ceci constitue l'objectif principal de notre travail.

3 Travaux Antérieurs

Cette section présente les concepts clés sous-jacents à notre approche : le problème de CMAB, l'algorithme *LinUCB* qui le résout, et un principe de diversification s'appuyant sur une fenêtre glissante.

3.1 Bandits-Manchots Contextuels

Les approches contextuelles du problème de bandits-manchots (CMAB) [10] ont été très largement étudiées via des méthodes telles que *LinUCB* [1], Contextual Thompson Sampling *CTS* [3] ou encore Neural Bandit [19]. Ces méthodes résolvent le problème de CMAB en supposant une dépendance linéaire entre la récompense attendue d'une action et son contexte. Selon les travaux de Langford [10], le problème de CMAB peut être défini comme suit : Soit $A = \{a_1, \dots, a_k\}$ un ensemble donné de k bras indépendants. Soit $X \subseteq \mathbb{R}^d$ l'ensemble de vecteurs de contexte de dimension d caractérisant un utilisateur et son environnement e.g., $x \in X$ est un vecteur binaire codant les caractéristiques telles que : l'âge, le sexe, le métier, les préférences, les spécialités, la localisation ou encore les caractéristiques des bras eux-mêmes. Soit l'horizon $T \in \mathbb{N}^*$, à chaque itération $t \in [1, T]$, le contexte x_t incluant l'utilisateur, est pris en considération afin de permettre la sélection du bras optimal compte tenu des récompenses obtenues lors des itérations précédentes. Pour chaque itération t , soit $r_t = (r_{t,a_1}, \dots, r_{t,a_k})$ le vecteur de récompense où r_{t,a_i} correspond à la récompense obtenue après avoir sélectionné le bras a_i et $r_{t,a_i} \in \{0, 1\}$ dans notre cas où les récompenses sont tirées depuis des distributions de Bernoulli. Soit $\mathcal{D}_{x,r}$ la distribution conjointe entre les contextes x et les récompenses r , et soit $\theta_{t,a}$ le vecteur de coefficients inconnu (restant à déterminer) associé au bras a à l'itération t . Nous supposons que les récompenses attendues d'un bras a à l'itération t est une fonction linéaire du vecteur de contexte x_t de dimension d tel que $\mathbb{E}[r_{t,a}|x_t] = \hat{\theta}_a^\top x_t$ où $\hat{\theta}_a$ représente le vecteur de coefficients estimé associé au bras a . Ainsi, soit $\Pi : X \rightarrow A$ l'ensemble des politiques possibles où la politique optimale devant être déterminée est $\pi^* = \arg \max_{\pi \in \Pi} \mathbb{E}_{r,x}[r_{t,\pi(x)}]$. Alors, soit $\pi_t \in \Pi$ la politique empruntée par un algorithme de CMAB \mathcal{A} à l'itération t . Par conséquent, dans le cadre d'un environnement stationnaire où $\mathcal{D}_{x,r}$ ne varie pas, le pseudo-regret instantané à l'itération t peut alors être défini tel que $\rho_t(\mathcal{A}) = \mathbb{E}_{r,x}[r_{t,\pi^*(x_t)} - r_{t,\pi(x_t)}]$ et le pseudo-regret cumulé tel que $\rho(\mathcal{A}) = \sum_{t=1}^T \rho_t(\mathcal{A})$.

Les algorithmes tels que *LinUCB* [1] ou *Contextual Thompson Sampling (CTS)* [3] ont été modélisés et largement étudiés afin de résoudre ce problème de CMAB. Aussi, à la section suivante nous rappelons l'un des plus populaires d'entre eux : *LinUCB*.

3.2 LinUCB

Nous avons d'abord décidé de bâtir et d'expérimenter notre approche à partir de *LinUCB* [1] qui reste l'un des algo-

ritmes de CMAB les plus célèbres présentés dans la littérature.

LinUCB [1] est un algorithme contextuel à bornes supérieures de confiance qui renforce rapidement la sélection des bras optimaux en ajoutant un *bonus* (l'écart de la récompense) au gain total calculé. À chaque itération t , *LinUCB* sélectionne le bras $a \in A$ avec le gain calculé $p_{t,a}$ maximum parmi l'ensemble des bras disponibles. $p_{t,a}$ est construit à partir d'une combinaison linéaire du coefficient $\theta_{t,a}$ et du vecteur de caractéristiques x_t auxquels vient s'ajouter l'écart de récompense qui représente la valeur d'action optimiste du gain obtenu. Le vecteur de coefficient $\hat{\theta}_a$ est construit à partir de la matrice D_a de dimension $n \times d$ (n recommandations en correspondance de d caractéristiques), et $b_a \in \mathbb{R}^d$ représente le vecteur de réponse correspondant, dont les poids pour chaque dimension sont fonction des récompenses obtenues. Plus précisément, $\hat{\theta}_a = (D_a^\top D_a + I_d)^{-1} b_a$ où I_d représente la matrice identité de dimension $d \times d$. Par conséquent, à chaque itération t , *LinUCB* sélectionne le bras a_t tel que $a_t = \arg \max_{a \in A} p_{t,a}$ où $p_{t,a} = \hat{\theta}_a^\top x_t + \alpha \sqrt{x_t^\top (D_a^\top D_a + I_d)^{-1} x_t}$. Ainsi, $\hat{\theta}_a^\top x_t$ représente l'espérance de récompense et $\alpha \sqrt{x_t^\top (D_a^\top D_a + I_d)^{-1} x_t}$ l'écart de récompense où α est un paramètre pouvant être considéré comme un critère de robustesse face au bruit. De plus, selon [20], il y a une probabilité d'au moins $1 - \delta$ que $|\hat{\theta}_a^\top x_t - \mathbb{E}[r_{t,a}|x_t]| \leq \alpha \sqrt{x_t^\top (D_a^\top D_a + I_d)^{-1} x_t}$ avec $\alpha = 1 + \sqrt{\ln(2/\delta)}/2$. Si un ensemble de bras contient k bras, alors la borne supérieure du regret sera en $\tilde{O}(\sqrt{kdT})$.

Néanmoins, même si le regret total est ici bien identifié et que sa borne supérieure a été démontrée, il est encore nécessaire de surmonter les problématiques de faible précision individuelle possiblement induite par des environnements non stationnaires ou par des contextes trop pauvres en informations.

Dans la section suivante, nous rappelons plusieurs mécanismes de diversification reposant sur l'utilisation de fenêtres glissantes.

3.3 Mécanismes de diversification

Dans les systèmes de recommandation, la diversité est pertinente pour la satisfaction individuelle. La diversification peut aussi trouver son intérêt dans le cadre d'environnements non-stationnaires afin de permettre à l'algorithme de rester à jour et favoriser les observations les plus récentes. À l'aide d'une fenêtre glissante, des algorithmes tels que *SW-UCB* [21], ou encore *Windows Thompson Sampling with Restricted Context (Windows TSRC)* [12] permettent d'atténuer les effets résultant de la non-stationnarité. De plus, pour résoudre ces mêmes problèmes induits par la non-stationnarité et plus particulièrement dans le cadre d'une problématique de bandits de type *restless* [22], il existe une approche utilisant également une fenêtre glissante et dont l'objectif est de pénaliser les bras qui ont été

tirés trop souvent [23]. Cette approche intéressante a inspiré notre proposition.

4 Problématique et Méthodes

Dans cette section, nous posons notre problème, puis nous définissons notre nouvelle approche *SW-LinUCB*. Notre méthode est basée sur la combinaison de l’algorithme original *LinUCB* [1] et l’utilisation d’une fenêtre glissante inspirée de [23].

4.1 Énoncé du problème

Soit $\mathcal{U} = \{u_1, \dots, u_n\}$ l’ensemble des n agents disponibles dans un problème de bandits, et pouvant par exemple correspondre dans le cadre d’applications réelles, à des utilisateurs ou encore des patients. Inspiré par [24], nous supposons pour chaque bras $a \in A$ et étant donné $x \in X \subseteq \mathbb{R}^d$, que \mathcal{U} peut être partitionné en un nombre $m_a(x)$ de clusters $\mathcal{U}_{1,a}(x), \mathcal{U}_{2,a}(x), \dots, \mathcal{U}_{m_a(x),a}(x)$ d’utilisateurs partageant les mêmes comportements vis à vis des récompenses qu’ils octroient à chaque bras a . Faisant maintenant l’hypothèse de l’existence d’un vecteur de contexte optimal $x^* \in X^*$ qui posséderait toutes les caractéristiques pertinentes associées, avec une confiance de 100%, au bras optimal correspondant et cela pour chaque contexte disponible. Comme *LinUCB* suppose une dépendance linéaire entre la récompense attendue d’une action et son contexte tel que $\mathbb{E}[r_{t,a}|x_t] = \hat{\theta}_a^\top x_t$, alors lorsque x^* est fourni, *LinUCB* converge vers une précision de 100% et offre une personnalisation pour chaque individu. Cela signifie que toute précision individuelle convergera également vers 100%. Cependant, dans les situations réelles, x peut manquer d’informations et rester incomplet pour différentes raisons telles que : un manque d’information sur les caractéristiques des bras, une mauvaise modélisation du contexte c’est-à-dire un contexte spécifié de manière incomplète, des restrictions dues à des problématiques de confidentialité et de protection de la vie privée, un profil mal renseigné, des informations manquantes sur l’environnement de l’utilisateur (par exemple, une localisation temporairement indisponible). Dans les cas où $x \neq x^*$, les algorithmes de CMAB doivent faire face à des contraintes de parcimonies dans les données ou d’incomplétude sur les caractéristiques disponibles puisque les caractéristiques de x^* manquantes dans x ne peuvent pas être prises en compte. En effet, avec un vecteur de contexte insuffisamment décrit, les clusters associés à x^* ne seront pas pris en compte par *LinUCB*, qui peut finalement être incapable de tirer le bras optimal pour différentes situations. Les utilisateurs affectés par cette parcimonie vectorielle pourraient donc se retrouver insatisfaits de la sélection des bras qui leur est proposée par *LinUCB*. Cela entraîne une diminution de la précision globale mais également de la précision individuelle ciblant ces utilisateurs. Ces problématiques nous ont conduits à construire une nouvelle approche visant, à la fois, à garder une bonne précision globale et à atténuer la diminution de la précision individuelle. La sous-section suivante présente notre mé-

thode qui utilise un mécanisme de diversification afin de contrer le manque d’information contextuelle et favoriser la sérendipité.

4.2 Sliding Window LinUCB : SW-LinUCB

Notre Fenêtre Glissante : Notre nouvelle approche combine *LinUCB* et l’utilisation d’une fenêtre glissante permettant de pénaliser la sélection des bras optimaux (tirés plus fréquemment), afin de favoriser l’exploration des bras moins optimaux que nous appellerons ici l’ensemble des bras sous-optimaux. Les méthodes utilisant des fenêtres glissantes appliquent généralement un coefficient dit de *discount* pondérant les récompenses obtenues par leurs bras afin de favoriser les observations les plus récentes. Ainsi, il est possible de définir un coefficient de *discount* qui pondère les récompenses cumulées obtenues pour chaque bras tel que $\sum_{t=1}^T \gamma_t r_{t,a}$ [23]. Avec $\gamma_t = 1 - \frac{Occ_w(a,t)}{w}$ où w correspond à la taille de la fenêtre glissante et $Occ_w(a,t)$ représente le nombre de fois qu’un bras a a été sélectionné durant les t dernières itérations. $Occ_w(a,t) = \#_1(E_{t,a})$ où $E_{t,a} = \{0..(2^{(w+1)} - 1)\}$ représente les w dernières sélections d’un bras a donné e.g., pour une taille de fenêtre $w = 6$, $E_{t,a} = 101001$ signifie que a a été sélectionné aux itérations $t - 6$, $t - 4$ et $t - 1$. Néanmoins, même si il pourrait être intéressant de combiner une telle méthode à *LinUCB*, celle-ci reste un processus mettant en œuvre une mémoire à court-terme qui, dans notre cas, ne permettra pas de diversifier suffisamment. Dans notre cas, nous devons conserver un processus d’élimination des mauvaises solutions sur le long terme tout en diversifiant suffisamment parmi l’ensemble des bras sous-optimaux. Ainsi, nous proposons un nouvelle méthode de calcul du gain basé sur le $p_{t,a}$ originel tel que

$$p_{t,a}^w = \gamma_t \hat{\theta}_a^\top x_t + \alpha \sqrt{x_t^\top M_a^{-1} x_t} \quad (1)$$

où $M_a = (D_a^\top D_a + I_d)$. Ce calcul permet à la fois de garder la confiance (élimination à long terme) grâce au *bonus*, et de diversifier suffisamment parmi l’ensemble des bras sous-optimaux en pénalisant temporairement l’espérance calculée des récompenses pour les bras sélectionnés trop fréquemment.

L’algorithme SW-LinUCB : l’objectif de *SW-LinUCB* est de déterminer la politique π qui maximise les récompenses cumulées à l’horizon T tandis qu’une fenêtre glissante force la diversification parmi l’ensemble des bras sous-optimaux. Notre hypothèse est la suivante : en fonction du niveau de parcimonie du vecteur de contexte ($x \neq x^*$), diversifier parmi l’ensemble des bras sous-optimaux atténuera la perte de précision individuelle pour les utilisateurs pour lesquels la méthode d’origine obtient une très faible précision. Notre méthode est décrite dans l’algorithme 1.

5 Expérimentations et Résultats

Jeux de données : L’évaluation de notre proposition se base sur quatre jeux de données. Tout d’abord un jeu de

Algorithme 1 *Sliding Window LinUCB (SW-LinUCB)*

Require: L'ensemble des k bras $a \in A$ disponibles, $\alpha \in \mathbb{R}^+$, l'horizon T , et l'ensemble des n contextes fixes disponibles X

```
1:  $w \leftarrow k$ 
2: for  $t = 1$  to  $T$  do
3:   Considérer  $x_t \in X$  : un utilisateur et son contexte
4:   for all  $a \in A$  do
5:     if  $a$  n'a pas encore été sélectionné then
6:        $Occ_w(a, t) \leftarrow 0$ ;  $M_a \leftarrow I_d$ ;  $b_a \leftarrow 0_{d \times 1}$ 
7:     end if
8:      $\hat{\theta}_a \leftarrow M_a^{-1} b_a$ 
9:     if  $t > w$  then
10:      Calculer  $Occ_w(a, t) = \#_1(E_{t,a})$ 
11:     end if
12:      $p_{t,a}^w \leftarrow \left(1 - \frac{Occ_w(a,t)}{w}\right) \hat{\theta}_a^\top x_t + \alpha \sqrt{x_t^\top M_a^{-1} x_t}$ 
13:   end for
14:   Sélectionner le bras  $a_t = \arg \max_{a_t \in A} (p_{t,a})$  et observer la récompense  $r_t$  retournée par l'utilisateur
15:    $M_{a_t} \leftarrow M_{a_t} + x_t x_t^\top$ ;  $b_{a_t} \leftarrow b_{a_t} + r_t x_t$ 
16:    $\forall a \neq a_t$ , mettre à jour toutes les sous-séquences  $E_{t,a}$  en ajoutant un bit 0
17:   Mettre à jour la sous-séquence  $E_{t,a_t}$  en ajoutant un bit 1
18:   if  $t > w$  then
19:     Réaliser un décalage logique vers la gauche (Left Shift) de  $E_{t,a}$  et  $E_{t,a_t}$ 
20:   end if
21: end for
```

données a été artificiellement généré afin d'obtenir un x^* garantissant une équiprobabilité entre chacun des bras. Il servira de jeu de contrôle dans nos expérimentations. Enfin, nous avons utilisé trois autres jeux de données d'applications réelles : Recommendation System for Angers Smart City (RS-ASM)¹, Coverttype et Poker Hand². Chacun des jeux de données considéré est constitué d'un nombre d'instances, s'appuie sur un contexte d'une dimension donnée et propose un nombre défini de bras (voir Tableau 1).

Jeu	Instances	Dim	Bras	Source
Contrôle	1000	4	4	Generated
RS-ASM	2152	56	18	Kaggle
Coverttype	581 012	95	7	UCI
Poker Hand	1 025 010	11	9	UCI

Tableau 1 – Jeux de données

Mesure de Précision Globale : La précision globale est un critère de performance basé sur le total des récompenses positives cumulées à l'horizon T . De ce fait, pour obtenir la précision globale, nous calculons le gain c'est à dire le nombre total de récompenses positives $g(T)$ puis

¹<https://www.kaggle.com/>

²<http://archive.ics.uci.edu/ml/>

nous calculons enfin la précision (Accuracy : Acc , voir Tableau 2) tel que : $Acc(T) = \frac{g(T)}{T}$ où $g(T) = \sum_{t=1}^T r_t$ et $r_t = \{0, 1\}$.

Mesure de Diversité : La diversité de sélection parmi un ensemble fini et fixe de k bras peut être définie comme un critère de dispersion découlant du coefficient de variation (C_v) des bras sélectionnés. Il est par conséquent possible de calculer la diversité (Div , voir Tableau 2) comme suit : $Div(N) = 1 - \frac{c_v(N)}{\sqrt{k}}$ où $N = \{n_{a_1}, \dots, n_{a_k}\}$, et n_{a_i} correspond au nombre de fois qu'un bras $a_i \in A$ a été sélectionné. Ainsi, la dispersion de la sélection tend à son maximum quand $c_v(N) \rightarrow 0$ alors qu'elle tend vers son minimum quand $c_v(N) \rightarrow \sqrt{k}$ [25].

Mesure de Précision Individuelle : La précision individuelle par utilisateur $Acc_u(T)$ peut être définie comme étant $\forall u \in \mathcal{U}, Acc_u(T) = \frac{\sum_{t=1}^T r_{t,u}}{T_u}$ où T_u représente le nombre de fois qu'un utilisateur avec son contexte a été sélectionné à l'horizon T , et $r_{t,u}$ correspond à la récompense retournée par u à l'itération t . Les mesures de précision individuelle de chaque utilisateur peuvent être représentées par une fonction de distribution cumulative (CDF). La CDF nous permet ainsi d'observer la distribution de la précision individuelle sur l'ensemble des utilisateurs \mathcal{U} avec leur contexte à l'horizon T (voir Figure 1).

Comparaison des Algorithmes³ : Nous comparons notre algorithme *SW-LinUCB* avec les méthodes suivantes : *UCB* standard (MAB) [8], et *LinUCB* classique (CMAB) [1]. Notons que *LinUCB* et *SW-LinUCB* auront la même valeur du paramètre α calculé avec $\delta = 0.1$.

Protocole Expérimental : Pour chaque algorithme et pour chaque jeu de données, nous simulons 2 cas différents : 1) Avec le vecteur de contexte complet (vc), 2) Avec une partie tronquée à 25% du vecteur de contexte d'origine (vt). Ici, le terme tronqué représente la proportion (en pourcentage) des caractéristiques, sélectionnées aléatoirement, que nous décidons de perdre au début de l'expérience. Ainsi, pour chacun des différents cas et pour chaque algorithme, nous simulons 20 expériences de 10, 000, 000 d'itérations pour Poker Hand et Coverttype, et 100, 000 concernant RS-ASM et le jeu de données de contrôle. Comme le nombre d'instances de chaque jeu de données est plus ou moins important, nous devons mettre à l'échelle l'horizon T pour chacun d'entre eux afin d'obtenir une mesure suffisante de la précision individuelle.

De plus, pour simuler un flux de données d'utilisateurs se présentant pour recevoir une recommandation (voir ligne 3 de l'algorithme 1), nous sélectionnons séquentiellement et aléatoirement les contextes disponibles dans l'ensemble du jeu de données. Ensuite, nous déterminons les moyennes et écart-types de précision globale et de diversité de sélection des bras sur l'ensemble des 20 simulations. De plus, nous calculons la précision individuelle et déduisons sa CDF dont les données et la représentation sont représentées Figure 1 et Tableau 2. Enfin, nous réalisons un test

³Voir notre étude préliminaire <https://git.io/vxCcv>

		Mesures globales		Distribution de la Précision Individuelle				
		<i>Acc</i>	<i>Div</i>	10%	Q_1	<i>Med</i>	Q_3	90%
UCB	Contrôle	0.25 $\pm\epsilon$	10 $^{-3}$ $\pm\epsilon$	0.00 $\pm\epsilon$	0.00 $\pm\epsilon$	0.00 $\pm\epsilon$	0.25 $\pm\epsilon$	1.00 $\pm\epsilon$
	RS-ASM	0.52 ± 0.08	10 $^{-3}$ $\pm\epsilon$	0.00 $\pm\epsilon$	0.00 $\pm\epsilon$	0.80 ± 0.40	1.00 $\pm\epsilon$	1.00 $\pm\epsilon$
	Poker Hand	0.47 ± 0.04	10 $^{-3}$ $\pm\epsilon$	0.00 $\pm\epsilon$	0.00 $\pm\epsilon$	0.60 ± 0.49	1.00 $\pm\epsilon$	1.00 $\pm\epsilon$
	Coverttype	0.41 ± 0.05	10 $^{-3}$ $\pm\epsilon$	0.00 $\pm\epsilon$	0.00 $\pm\epsilon$	0.00 $\pm\epsilon$	0.90 ± 0.30	0.90 ± 0.30
LinUCB (<i>vc</i>)	Contrôle	1.0 $\pm\epsilon$	0.997 $\pm\epsilon$	1.00 $\pm\epsilon$	1.00 $\pm\epsilon$	1.00 $\pm\epsilon$	1.00 $\pm\epsilon$	1.00 $\pm\epsilon$
	RS-ASM	0.78 $\pm\epsilon$	0.86 $\pm\epsilon$	0.04 ± 0.02	0.77 ± 0.02	0.95 $\pm\epsilon$	0.99 $\pm\epsilon$	1.00 $\pm\epsilon$
	Poker Hand	0.53 $\pm\epsilon$	0.06 $\pm\epsilon$	0.00 $\pm\epsilon$	0.00 $\pm\epsilon$	0.90 ± 0.01	1.00 $\pm\epsilon$	1.00 $\pm\epsilon$
	Coverttype	0.72 $\pm\epsilon$	0.44 $\pm\epsilon$	0.00 $\pm\epsilon$	0.00 $\pm\epsilon$	1.00 $\pm\epsilon$	1.00 $\pm\epsilon$	1.00 $\pm\epsilon$
SW-LinUCB (<i>vc</i>)	Contrôle	0.991 $\pm\epsilon$	0.997 $\pm\epsilon$	0.98 $\pm\epsilon$	0.99 $\pm\epsilon$	0.99 $\pm\epsilon$	1.00 $\pm\epsilon$	1.00 $\pm\epsilon$
	RS-ASM	0.76 $\pm\epsilon$	0.88 $\pm\epsilon$	0.06 ± 0.02	0.68 ± 0.01	0.92 ± 0.01	0.98 $\pm\epsilon$	1.00 $\pm\epsilon$
	Poker Hand	0.48 $\pm\epsilon$	0.34 $\pm\epsilon$	0.00 $\pm\epsilon$	0.22 ± 0.02	0.50 $\pm\epsilon$	0.72 ± 0.02	0.87 ± 0.02
	Coverttype	0.69 $\pm\epsilon$	0.47 $\pm\epsilon$	0.00 $\pm\epsilon$	0.40 $\pm\epsilon$	0.89 $\pm\epsilon$	1.00 $\pm\epsilon$	1.00 $\pm\epsilon$
LinUCB (<i>vt</i>)	Contrôle	0.749 $\pm\epsilon$	0.88 ± 0.07	0.35 ± 0.09	0.52 ± 0.04	0.88 ± 0.04	1.00 $\pm\epsilon$	1.00 $\pm\epsilon$
	RS-ASM	0.629 $\pm\epsilon$	0.33 ± 0.01	0.01 $\pm\epsilon$	0.06 ± 0.01	0.92 ± 0.01	0.97 $\pm\epsilon$	0.99 $\pm\epsilon$
	Poker Hand	0.50 $\pm\epsilon$	0.01 $\pm\epsilon$	0.00 $\pm\epsilon$	0.00 $\pm\epsilon$	0.84 ± 0.04	1.00 $\pm\epsilon$	1.00 $\pm\epsilon$
	Coverttype	0.60 $\pm\epsilon$	0.35 $\pm\epsilon$	0.00 $\pm\epsilon$	0.00 $\pm\epsilon$	1.00 $\pm\epsilon$	1.00 $\pm\epsilon$	1.00 $\pm\epsilon$
SW-LinUCB (<i>vt</i>)	Contrôle	0.746 $\pm\epsilon$	0.96 ± 0.02	0.43 ± 0.02	0.50 $\pm\epsilon$	0.82 ± 0.01	0.99 $\pm\epsilon$	1.00 $\pm\epsilon$
	RS-ASM	0.567 $\pm\epsilon$	0.69 $\pm\epsilon$	0.08 ± 0.01	0.33 ± 0.01	0.61 ± 0.01	0.85 ± 0.01	0.95 ± 0.01
	Poker Hand	0.48 $\pm\epsilon$	0.33 $\pm\epsilon$	0.00 $\pm\epsilon$	0.26 ± 0.01	0.50 $\pm\epsilon$	0.67 $\pm\epsilon$	0.85 $\pm\epsilon$
	Coverttype	0.56 $\pm\epsilon$	0.41 $\pm\epsilon$	0.00 $\pm\epsilon$	0.25 $\pm\epsilon$	0.62 $\pm\epsilon$	0.88 $\pm\epsilon$	1.00 $\pm\epsilon$

Tableau 2 – Résultats sur plusieurs jeux de données avec vecteur complet (*vc*) et vecteur tronqué (*vt*) ($\epsilon = 0.0009$)

de *Kruskal-Wallis* pour vérifier l'inégalité des moyennes obtenues sur les critères observés sur l'ensemble des algorithmes, puis nous complétons ces tests par des comparaisons deux à deux en réalisant des tests de rang de *Wilcoxon* pour mettre en évidence la significativité statistique de ces inégalités.

5.1 Analyse Globale

Les analyses ci-dessous s'appuient sur les résultats présentés dans la Tableau 2 dont les CDFs sont illustrées Figure 1. **Tests Statistiques :** Un test de *Kruskal-Wallis* pour chaque expérience nous indique qu'il y a une différence significative entre les mesures de précision de chacun des 3 algorithmes ($p < 0.01$). De plus, le test des rangs signés de *Wilcoxon* met en évidence une différence significative entre chaque paire d'algorithmes ($p < 0.01$).

Diversité : En ce qui concerne les expériences sur le jeu de données de contrôle nous observons comme attendu que lorsque nous fournissons un vecteur optimal x^* les deux algorithmes de CMAB diversifient à 100%. Néanmoins, pour chaque jeu de données, lorsque nous perdons 25% de l'information du vecteur d'origine, alors la diversité décroît pour les deux algorithmes de CMAB. En effet, ils ne réussissent pas à trouver la bonne politique en regard de la règle de correspondance cachée entre récompenses et dimensions du contexte puisqu'une partie pertinente de ce contexte a été tronquée. En revanche, même si *LinUCB* ne diversifie pas autant que dans le cas où nous lui fournissons un vecteur plus complet, *SW-LinUCB* quant à lui offre dans ces mêmes conditions une meilleure diversifica-

tion que l'algorithme original. Enfin, comme prévu, *UCB* agit comme un algorithme glouton à bornes supérieures de confiance : il trouve le bras optimal et continue de le tirer tout au long des itérations ce qui résulte en une valeur de diversité proche de 0.

Précision Globale VS Individuelle : Comme attendu, *LinUCB* obtient la meilleure performance globale dans tous les cas et pour tout jeu de données. Sans surprise, la précision globale diminue lorsque nous tronquons le vecteur de contexte, mais il est important de noter que même avec le niveau d'éparsité choisi dans notre expérience, les algorithmes de CMABs restent encore meilleurs que l'algorithme de MAB représenté par *UCB*. Cependant, nous observons dans tous les cas (sauf quand $x = x^*$), que *LinUCB* crée un écart de précision individuelle très important entre les utilisateurs. D'autre part, sur l'ensemble des jeux de données et dans tous les cas (sauf quand $x = x^*$), *SW-LinUCB* perd en précision globale par rapport à *LinUCB* mais en revanche trouve, grâce à son mécanisme de diversification, un meilleur compromis en ce qui concerne la distribution de la précision individuelle. Enfin, pour tous les jeux de données, nous observons que plus l'incomplétude du vecteur de contexte est importante, plus un fossé se crée entre les différentes précisions individuelles d'où résultent distinctement une classe de précisions que l'on peut catégoriser de hautes et une classe de précisions dites basses. De la même manière, on remarque que plus x tend vers x^* , plus la distribution de la précision individuelle est uniformément répartie parmi les utilisateurs.

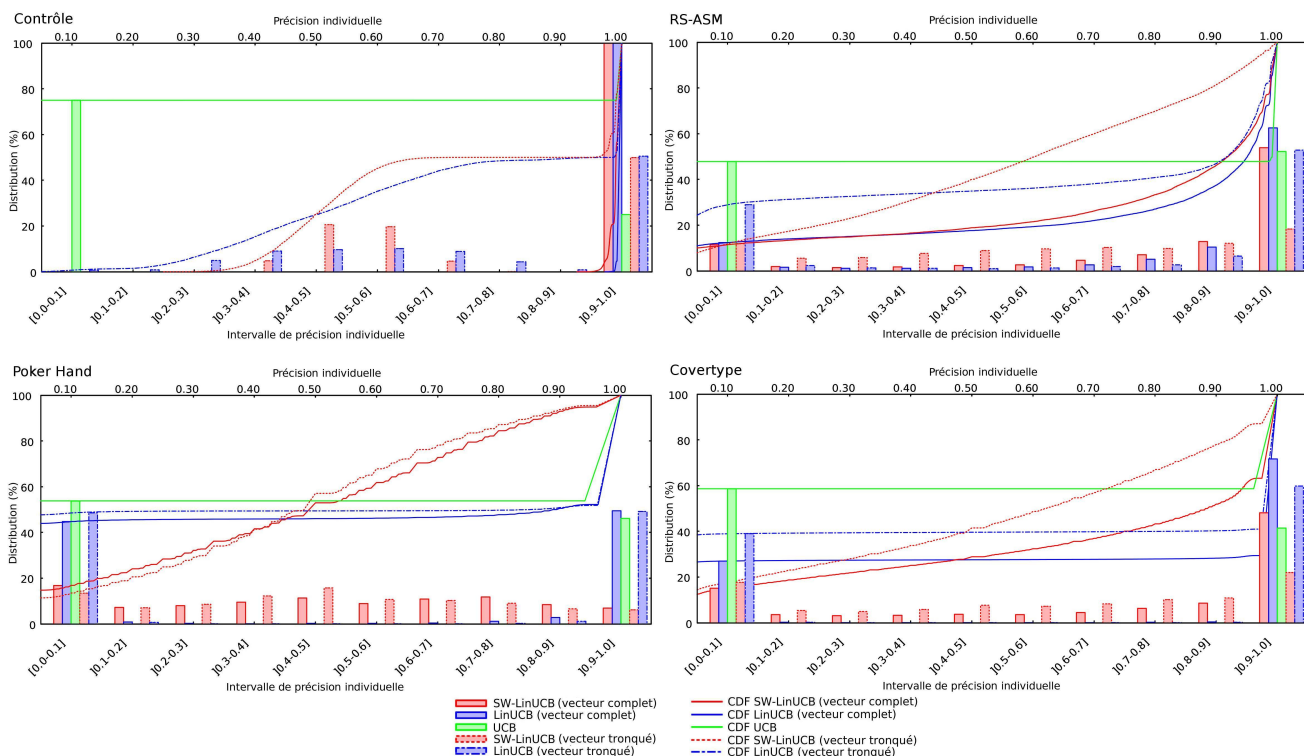


FIGURE 1 – Distribution de la précision individuelle pour chaque algorithme

5.2 Analyse Spécifique sur Coverttype

Diversité : On observe que *SW-LinUCB* diversifie plus ($vc : Div = 0.47$, $vt : Div = 0.41$) que *LinUCB* ($vc : Div = 0.44$, $vt : Div = 0.35$) alors que l'algorithme *UCB* continue de tirer le même bras tout au long des itérations ($Div = 10^{-3}$). De plus, nous remarquons que lorsque x tend vers x^* , les caractéristiques fournies en tant que dimension du vecteur de contexte permettent à *LinUCB* et *SW-LinUCB* de diversifier d'avantage.

Précision Globale VS Individuelle : On observe que *LinUCB* conserve une meilleure précision globale ($vc : Acc = 0.72$, $vt : Acc = 0.60$) que *SW-LinUCB* ($vc : Acc = 0.69$, $vt : Acc = 0.56$). De plus, il est important d'observer que le niveau d'incomplétude du vecteur de contexte n'est pas encore assez important pour permettre à notre algorithme de MAB *UCB* d'être plus précis (0.41). En outre, on observe Figure 1 et Tableau 2, que *SW-LinUCB* reste le meilleur en termes de distribution de la précision individuelle ($vc : Q_1 = 0.40$, $vt : Q_1 = 0.25$) que *LinUCB* ($vc : Q_1 = 0.00$, $vt : Q_1 = 0.00$). Ces derniers résultats montrent que notre mécanisme de diversification permet d'augmenter la précision individuelle de la classe dite basse avec la méthode d'origine. Enfin, la comparaison entre les résultats vc et vt montre que, pour les deux méthodes de CMAB, les précisions globales et individuelles sont toutes deux proportionnelles au niveau d'information et de complétude du vecteur contexte.

6 Conclusion et Perspectives

Dans cet article, nous proposons une nouvelle mesure pour les algorithmes de décision séquentielle visant à évaluer la distribution de la précision individuelle. Nous soutenons que dans certains cas pratiques, la mesure de précision globale n'est pas suffisante pour évaluer les algorithmes de CMAB et que la mesure de précision individuelle doit également être prise en compte. De plus, nous proposons une nouvelle approche adaptée de l'algorithme original *LinUCB* visant à la fois à améliorer la précision individuelle et à maintenir une bonne précision globale. Nous montrons qu'en privilégiant la diversité, notre algorithme *SW-LinUCB* offre un compromis entre précision globale et individuelle que nous pensons mieux adapté à un certain nombre d'applications du monde réel comme les systèmes de recommandations ou les essais cliniques.

Ainsi en perspectives, il semble pertinent de considérer les deux opportunités suivantes : 1) Mettre en place des techniques permettant la construction d'un vecteur plus précis notamment par l'observation approfondie du contexte de l'application concrète qui en découle afin de déterminer les dimensions manquantes et pertinentes ; 2) Concevoir un algorithme pour résoudre le problème multi-objectifs de la maximisation des trois critères de précision globale, de précision individuelle et de diversité. Nous pensons qu'une approche portfolio c'est à dire tirant parti des avantages de plusieurs algorithmes (notamment *LinUCB* et *SW-LinUCB*) pourrait être envisagée.

Références

- [1] L. Li, W. Chu, J. Langford, and R. E. Schapire, “A contextual-bandit approach to personalized news article recommendation,” in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 661–670.
- [2] S. S. Villar, J. Bowden, and J. Wason, “Multi-armed bandit models for the optimal design of clinical trials : benefits and challenges,” *Statistical science : a review journal of the Institute of Mathematical Statistics*, vol. 30, no. 2, p. 199, 2015.
- [3] S. Agrawal and N. Goyal, “Thompson sampling for contextual bandits with linear payoffs,” in *International Conference on Machine Learning*, 2013, pp. 127–135.
- [4] D. Bouneffouf and R. Feraud, “Multi-armed bandit problem with known trend,” *Neurocomputing*, vol. 205, pp. 16–21, 2016.
- [5] S. M. McNee, J. Riedl, and J. A. Konstan, “Being accurate is not enough : how accuracy metrics have hurt recommender systems,” in *CHI’06 extended abstracts on Human factors in computing systems*. ACM, 2006, pp. 1097–1101.
- [6] H. Robbins, “Some aspects of the sequential design of experiments,” *Bulletin of the American Mathematical Society*, pp. 527–535, 1952.
- [7] T. L. Lai and H. Robbins, “Asymptotically efficient adaptive allocation rules,” *Advances in applied mathematics*, vol. 6, no. 1, pp. 4–22, 1985.
- [8] P. Auer, “Using confidence bounds for exploitation-exploration trade-offs,” *Journal of Machine Learning Research*, vol. 3, no. Nov, pp. 397–422, 2002.
- [9] S. Agrawal and N. Goyal, “Analysis of Thompson sampling for the multi-armed bandit problem,” in *Conference on Learning Theory*, 2012, pp. 39–1.
- [10] J. Langford and T. Zhang, “The epoch-greedy algorithm for multi-armed bandits with side information,” in *Advances in neural information processing systems*, 2008, pp. 817–824.
- [11] L. Zhou and E. Brunskill, “Latent contextual bandits and their application to personalized recommendations for new users,” in *International Joint Conferences on Artificial Intelligence (IJCAI)*, 2016.
- [12] D. Bouneffouf, I. Rish, G. A. Cecchi, and R. Feraud, “Context attentive bandits : Contextual bandit with restricted context,” *International Joint Conferences on Artificial Intelligence (IJCAI)*, 2017.
- [13] S. Craw, B. Horsburgh, and S. Massie, “Music recommenders : user evaluation without real users ?” in *International Joint Conferences on Artificial Intelligence (IJCAI)*, 2015.
- [14] A. Lacerda, “Contextual bandits for multi-objective recommender systems,” in *Intelligent Systems (BRACIS), 2015 Brazilian Conference on*. IEEE, 2015, pp. 68–73.
- [15] M. Ge, C. Delgado-Battenfeld, and D. Jannach, “Beyond accuracy : evaluating recommender systems by coverage and serendipity,” in *Proceedings of the fourth ACM conference on Recommender systems*. ACM, 2010, pp. 257–260.
- [16] X. Wang, Y. Guo, and C. Xu, “Recommendation algorithms for optimizing hit rate, user satisfaction and website revenue,” in *International Joint Conferences on Artificial Intelligence (IJCAI)*, 2015, pp. 1820–1826.
- [17] A. Ashkan, B. Kveton, S. Berkovsky, and Z. Wen, “Optimal greedy diversity for recommendation,” in *International Joint Conferences on Artificial Intelligence (IJCAI)*, pp. 1742–1748, 2015.
- [18] L. Hu, L. Cao, S. Wang, G. Xu, J. Cao, and Z. Gu, “Diversifying personalized recommendation with user-session context,” in *International Joint Conferences on Artificial Intelligence (IJCAI)*, 2017.
- [19] R. Allesiardo, R. Féraud, and D. Bouneffouf, “A neural networks committee for the contextual bandit problem,” in *International Conference on Neural Information Processing*. Springer, 2014, pp. 374–381.
- [20] T. J. Walsh, I. Szita, C. Diuk, and M. L. Littman, “Exploring compact reinforcement-learning representations with linear regression,” in *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*. AUAI Press, 2009, pp. 591–598.
- [21] A. Garivier and E. Moulines, “On upper-confidence bound policies for switching bandit problems,” in *International Conference on Algorithmic Learning Theory*. Springer, 2011, pp. 174–188.
- [22] P. Whittle, “Restless bandits : Activity allocation in a changing world,” *Journal of applied probability*, vol. 25, no. A, pp. 287–298, 1988.
- [23] A. Goëffon, F. Lardeux, and F. Saubion, “Simulating non-stationary operators in search algorithms,” *Applied Soft Computing*, vol. 38, pp. 257–268, 2016.
- [24] S. Li, A. Karatzoglou, and C. Gentile, “Collaborative filtering bandits,” in *The 39th International ACM SIGIR Conference on Information Retrieval (SIGIR)*, 2016.
- [25] J. Katsnelson and S. Kotz, “On the upper limits of some measures of variability,” *Theoretical and Applied Climatology*, vol. 8, no. 1, pp. 103–107, 1957.

Apprentissage par Transfert reformulé via la Théorie des Catégories

L. Cordesses¹ S. Amar¹ O. Bentahar¹ A. Laurent¹ K. Poulet¹ T. Ehrmann¹ J. Page²

¹ Renault Innovation Silicon Valley, 1215 Bordeaux Drive, Sunnyvale, 94089 CA, USA

² Université Paris Diderot-CNRS, Laboratoire SPHERE, 5 rue Thomas Mann, 75205 Paris Cedex 13, France

{lionel.cordesses, kevin.poulet, thomas.ehrmann}@renault.com,
{aude.laurent, sarah.amar, omar.bentahar}@nissan-usa.com,
ju.page@hotmail.fr

Résumé

Les applications industrielles utilisant l'apprentissage automatique pour contrôler des systèmes embarqués doivent apprendre avec très peu d'essais. En effet, le nombre d'expériences possibles durant la phase d'apprentissage est souvent limité, par des contraintes de temps ou de budget. Cet objectif d'apprentissage rapide peut être atteint en utilisant du transfert de connaissances formalisé mathématiquement avec les outils de la théorie des catégories.

Nous présentons une approche originale de l'apprentissage automatique qui transmet ses connaissances à des situations nouvelles en s'appuyant sur les outils existants du domaine. Des résultats expérimentaux sont présentés sur un circuit de voitures miniatures.

Mots Clef

Apprentissage Automatique, Théorie des Catégories, Apprentissage par Transfert, Circuit de Voitures Miniatures.

Abstract

Industrial applications relying on Machine Learning (ML) to control robotic systems need to learn with few trials. The number of experiments in this context is limited by either the available time or budget. Such a goal can be attained by using knowledge transfer formalized mathematically with tools from category theory.

This leads to a new ML approach that transposes accumulated knowledge to new configurations, while still relying on existing ML tools. Illustrations of this innovative solution are presented using a slot car game.

Keywords

Machine Learning, Category Theory, Transfer Learning, Slotcar.

1 Introduction

La nécessité pour les algorithmes contrôlant des systèmes embarqués de devoir apprendre, avec peu de données, comment évoluer dans un environnement partiellement connu,

est de plus en plus importante. Beaucoup d'approches utilisent pour cela l'apprentissage par renforcement et, plus particulièrement, les réseaux de neurones qui ont prouvé leur efficacité sur des systèmes embarqués.

Néanmoins, ces approches nécessitent beaucoup de données pour l'apprentissage, qui peuvent ne pas être disponibles en fonction de contraintes de temps ou de budget.

Pour illustrer ce besoin d'apprentissage automatique utilisant peu de données, nous utilisons le secteur automobile. Plus particulièrement, les voitures à transmission automatique proposent habituellement plusieurs modes (sport, neige, etc.) qui configurent les réglages de la transmission. Néanmoins, adapter ces réglages aux préférences de chaque conducteur reste un défi. Par exemple, considérons un algorithme ajustant les préférences du conducteur pour le mode neige. Si 100 essais sont nécessaires pour l'apprentissage, le mode neige étant utilisé 1 fois par an, il faudra un siècle pour avoir des résultats. Les algorithmes d'apprentissage automatique nécessitant des milliers d'essais ne sont donc pas envisageables dans ce cas.

L'idée de départ de ce projet était de proposer une approche alternative sous forme d'un algorithme aussi performant que les méthodes de l'état de l'art, mais capable d'apprendre avec environ 1% du temps, des données et de la puissance de calcul que demandent ces méthodes.

Nous présentons d'abord les publications utilisant les circuits de voitures miniatures comme support expérimental, la plupart portant sur le contrôle de ces systèmes. Ensuite, nous présentons la théorie de notre approche et les deux types de problèmes sur lesquels nous travaillons : le cas bijectif où le système imite un savoir élément par élément, et le cas utilisant la théorie des catégories où le système transpose un savoir au niveau des *types* d'éléments. Ces deux cas utilisent le transfert de connaissances comme dans le Lifelong Machine Learning [4]. Nous décrivons enfin le support expérimental et les résultats obtenus.

2 Travaux Antérieurs

En raison de leur facilité d'approche et de leur faible coût, les circuits de voitures miniatures (voir figure 1) sont utili-

sés dans de nombreux domaines, de la théorie du contrôle à l'Intelligence Artificielle (IA). Nous présentons ici un aperçu des utilisations de ce type de circuits dans des travaux pertinents pour notre projet, puis nous élargissons à des problèmes liés à l'apprentissage avec peu d'exemples.



FIGURE 1 – Exemple de circuit de voitures miniatures

2.1 Outil pédagogique

Les circuits de voitures sont largement utilisés pour enseigner l'interfaçage et l'exploitation de capteurs [23] et d'actionneurs [5] avec un ordinateur. Les signaux réels, contrairement aux signaux simulés, ajoutent un degré de difficulté comme le mentionne [22] : « les moteurs des voitures génèrent une quantité importante de bruit électrique » ; cette publication propose une solution concrète pour limiter ce problème. De plus, un tel système n'est pas linéaire indépendant du temps (LTI) puisque le circuit et les voitures nécessitent tous les deux de l'entretien pour maintenir les performances constantes : graissage et nettoyage internes pour la voiture, et nettoyage pour la piste. L'article met également l'accent sur les incertitudes dans les mesures des signaux, en particulier pour le capteur utilisé pour indiquer la position de la voiture. La détection des défauts des capteurs est aussi détaillée dans [10] dans le cadre de la conception et du test d'IA pilotant la voiture.

Plus généralement, les circuits de voitures sont utilisés pour enseigner les micro-contrôleurs dans [31], la mécanique dans [14] et [11], les systèmes temps réel et l'IA dans [3].

2.2 Validation de théories du contrôle

Des codes barres imprimés sur les rails pour localiser la voiture et un micro-contrôleur embarqué pour ajuster la vitesse en conséquence peuvent être utilisés afin de contrôler de façon autonome la voiture [13]. Le contrôle de convois de véhicules, implanté sur un micro-contrôleur 32 bits, est validé dans [20] à l'aide de capteurs embarqués mesurant la vitesse et l'accélération. Le développement et la validation de nouveaux concepts de « Mobilité en tant que service » sont évalués sur un circuit de voitures dans [28].

Les circuits de voitures sont également mentionnés dans [24] pour l'important couple électrique du moteur comparé

au poids de la voiture. Cette caractéristique en fait un cas d'étude intéressant de lois de commande optimales limitant le couple pour protéger l'actionneur : le moteur.

2.3 Validation des algorithmes d'IA

La solution proposée par [15] contrôle la vitesse de la voiture grâce à un perceptron multicouche qui utilise un algorithme de traitement d'images pour localiser la position de la voiture. Le temps d'apprentissage de la stratégie de contrôle est de 45 minutes. La solution d'apprentissage par renforcement [17] repose sur un perceptron multicouche à convolution pour localiser la voiture, et sur un contrôle via quatre niveaux de tension. La phase d'entraînement du perceptron prend 12 heures, et celle de la stratégie de contrôle nécessite 30 minutes supplémentaires. Une autre méthode d'apprentissage par renforcement utilise un réseau de neurones pour contrôler la vitesse de la voiture : décrite dans [12], elle consiste à construire une représentation de l'état permettant à la voiture de rester sur la piste. Un algorithme évolutionniste est proposé dans [7] pour piloter la voiture grâce à une caméra neuromorphique.

Une solution plus rapide pour rendre autonome une telle voiture en quelques tours de circuit est décrite dans [27] et consiste à ajouter des accéléromètres et un micro-contrôleur embarqués pour cartographier le circuit (répartition des virages et des lignes droites). L'algorithme de contrôle est ensuite capable d'adapter la vitesse de la voiture grâce à une boucle à verrouillage de phase. Des algorithmes génétiques sont enfin comparés aux algorithmes d'apprentissage par renforcement dans [19] dans le but d'ajuster le régulateur PID (proportionnel, intégral, dérivé) embarqué pilotant la voiture.

2.4 Validation des algorithmes de prédictions

Les circuits de voitures sont également utilisés pour valider les algorithmes de prédictions basés sur des algorithmes d'apprentissage automatique comme les Modèles de Markov cachés dans [30] et [2], et les estimations de positions calculées grâce à des mesures inertielles utilisant des filtres de Bayes à processus Gaussien dans [16]. Cette dernière solution permet à la méthode de rejouer une stratégie apprise en observant un expert humain jouant sur 16 tours. Un autre estimateur exploitant des données inertielles pour piloter une voiture est proposé dans [21] et repose sur des modèles de mélanges gaussiens.

2.5 Apprentissage par imitation et analogies

En dehors du domaine du circuit de voitures, l'imitation d'un ou de plusieurs exemples est une solution au problème de l'apprentissage avec très peu de données. L'approche décrite dans [9] et utilisée dans un contexte de robotique manipulatrice consiste à transformer des exemples fournis par un humain en un programme compréhensible tant par l'humain que par la machine. Ce programme est réutilisé dans des cas similaires à ceux des exemples. La notion d'analogie est aussi exploitée en classification et dans le

contrôle d'agents [26], par exemple dans des jeux de stratégie en temps réel [29].

3 Contrôler une voiture électrique par apprentissage automatique

3.1 Objectifs

Bien que les méthodes présentées en partie 2 parviennent à contrôler avec succès une voiture électrique miniature sur un circuit inconnu, ces méthodes nécessitent un grand nombre de tours d'apprentissage, ou alors compensent une durée d'apprentissage réduite par l'utilisation de capteurs embarqués, tels que des accéléromètres. Notre objectif est d'aboutir à une méthode d'apprentissage automatique qui pourrait apprendre avec peu de ressources, en particulier sans l'ajout de capteurs supplémentaires, et qui parviendrait à contrôler la voiture en temps réel avec une fraction de la puissance de calcul auparavant requise. Ces conditions aboutissent à une approche d'apprentissage automatique qui est entraînée sur quelques tours, en n'utilisant comme données que les courants et tensions mesurés sur la piste, et qui fonctionne par exemple sur un microcontrôleur 8 bits d'Arduino Mega cadencé à 16 MHz.

À notre connaissance, de tels prérequis ne sont pas vérifiés par les méthodes existantes. C'est pourquoi nous nous sommes lancés dans la conception d'une nouvelle approche d'apprentissage automatique, en s'inspirant de concepts philosophiques, puis en la formalisant rigoureusement avec les outils mathématiques de la théorie des catégories.

3.2 Solution inspirée de concepts philosophiques

Notre approche travaille à l'échelle des entités, définies ci-dessous, et non à celle des échantillons. Travailler à cette échelle peut s'apparenter à travailler à l'échelle du morphème en linguistique, *i.e.* le plus petit élément significatif d'un langage, conformément à la définition de [6]. Cela comprend une analyse des signaux échantillonnés pour détecter les entités. Celles-ci sont nos objets du quotidien : les tronçons du circuit (lignes droites et courbes du circuit, zones de danger pour l'algorithme) et la voiture elle-même. Comme tout objet cognitif, elles disposent de certaines propriétés : cohérence relative, continuité de mouvement, permanence de l'existence.

Toutefois, le monde qui nous entoure n'est pas seulement constitué d'objets : il comporte aussi des sujets. En particulier, dans les problèmes considérés, il existe une représentation de ce que l'on appelle le « Moi », c'est-à-dire l'entité contrôlée par nos actions. Dans le cas spécifique du circuit de voitures, trouver ce « Moi » n'est pas nécessaire, puisque l'IA contrôle une seule entité : la voiture, donc le « Moi ». De plus, les modèles obtenus ne sont pas éternellement valides, puisque le « Moi » et l'environnement constitué des autres entités peuvent évoluer. Afin que notre approche continue de fonctionner dans un environnement en évolu-

tion, nous nous sommes inspirés de l'approche scientifique décrite dans [25] : « un système faisant partie de la science empirique doit pouvoir être réfuté par l'expérience ». Nous concevons notre logiciel de sorte qu'il puisse remettre en cause ses propres résultats, puis reconstruire ses modèles. Cette remise en cause se produit dès que l'erreur entre les mesures du système et les prédictions du modèle est supérieure à un seuil.

Enfin, en lien avec le concept heideggerien de *Dasein*, littéralement *être-là* [8], un humain confronté à la conscience de soi et de la mort percevra ce « Moi » comme plus qu'un simple système contrôlable, et le dotera ainsi d'un instinct de survie. Nous modélisons ce « Moi » à partir d'une philosophie similaire : au fur et à mesure qu'il expérimente le monde qui l'entoure, ce « Moi » classe les entités environnantes comme amies ou ennemies, selon les bénéfices ou pertes qu'elles occasionnent. Une stratégie basique est ensuite appliquée, à savoir transférer les comportements qui n'étaient pas dangereux lors de situations précédentes.

En résumé, notre algorithme prend en compte les récompenses qu'il tire de son environnement afin de classifier les entités, comme en apprentissage par renforcement. Ensuite, par analogie avec les méthodes scientifiques empiriques, cette classification et les modèles sont remis à jour lorsque les prédictions ne correspondent plus aux mesures. Enfin, le « Moi » est contrôlé dans le but de survivre, conformément à l'instinct de survie dont un joueur le dote. Nous présentons dans la section suivante comment formaliser mathématiquement une telle approche d'apprentissage automatique, en particulier dans le but de transposer des connaissances acquises à une situation inconnue.

4 Utilisation d'analogies pour le transfert de connaissances

L'un des outils les plus efficaces dont dispose l'humain pour évoluer dans une situation inconnue est sa capacité à faire des analogies entre cette nouvelle situation et ses expériences passées. Nous pressentons qu'une IA capable d'établir des analogies, et sachant comment conduire une voiture électrique miniature sur un circuit, sera capable de transposer ces capacités à une autre configuration, même si elle est de forme ou de taille différente.

Les situations où deux problèmes ont exactement le même nombre d'états et des structures isomorphes sont rares. Néanmoins, il existe des outils mathématiques pour identifier des structures non isomorphes tels que l'équivalence de catégories en théorie des catégories [18]. La théorie des catégories est un outil mathématique puissant apparu au milieu du XX^{ème} siècle en lien avec la topologie algébrique, et développé dans le but de transférer des concepts et des théorèmes entre différentes branches des mathématiques. Dans l'esprit de la théorie des ensembles, l'identification se réduit aux relations d'identité et aux cas bijectifs, tandis que la théorie des catégories apporte des descriptions plus riches des objets au travers de l'introduction de flèches (ou morphismes) entre objets, ce qui permet de nouvelles sortes

d'identifications. Comme la plupart des outils d'apprentissage automatique s'appuient sur la théorie des ensembles et non sur la théorie des catégories, nous illustrons ci-dessous la plus-value d'un tel cadre mathématique.

4.1 Analogies entre deux situations

Une catégorie \mathcal{C} est une collection d'objets et de morphismes (ou flèches) entre certains de ces objets, munie d'une composition de morphismes, et peut ainsi être assimilée à un graphe orienté. Si A et B sont des objets de \mathcal{C} , la flèche $a : A \rightarrow B$ est un isomorphisme si elle est inversible, i.e s'il existe une flèche $b : B \rightarrow A$, telle que $ba = Id_A$ et $ab = Id_B$. Dans ce cas, les objets A et B sont isomorphes. Les isomorphismes définissent une relation d'équivalence sur la classe des objets de \mathcal{C} . On note \mathcal{C}/\simeq son quotient. Aussi, si $F : \mathcal{C} \rightarrow \mathcal{C}'$ est ce qu'on appelle une équivalence de catégories, celle-ci induit une bijection $F' : (\mathcal{C}/\simeq) \rightarrow (\mathcal{C}'/\simeq)$ entre les classes d'objets isomorphes même si F n'est pas bijective. Dès lors, on n'identifie plus les objets (ou états) un à un entre deux situations, mais les types (ou classes d'isomorphismes) de ces états.

Ce procédé peut être utilisé dans le cas de problèmes observables. En effet, considérons deux ensembles d'états non vides \mathcal{C} et \mathcal{C}' sans autre hypothèse sur leur cardinalité. Supposons aussi l'existence de deux fonctions d'observations $f : \mathcal{C} \rightarrow O$ et $f' : \mathcal{C}' \rightarrow O'$. Quitte à restreindre O et O' , supposons que ces deux fonctions sont surjectives. Pour tout $o \in O$, on dit que les états $x \in \mathcal{C}$ dont l'observation associée est o (i.e $f(x) = o$) sont de type T_o . Cela définit une relation d'équivalence R_f sur l'ensemble $\mathcal{C} : \forall x, y \in \mathcal{C}, x R_f y$ si et seulement si (ssi) $f(x) = f(y)$. Transposé dans le champ lexical de la théorie des catégories, cela revient à placer une flèche inversible entre deux objets x et y de \mathcal{C} ssi $x R_f y$. \mathcal{C} devient alors une catégorie, où toutes les flèches sont inversibles et telle que \mathcal{C}/\simeq est exactement le quotient \mathcal{C}/R_f , quotient qui définit aussi l'ensemble des types d'états de \mathcal{C} . Puisque ces types ont été définis via les observations, la surjection $f : \mathcal{C} \rightarrow O$ induit une bijection $\tilde{f} : (\mathcal{C}/\simeq) \rightarrow O$ entre l'ensemble de types et l'ensemble d'observations. \tilde{f} est donc l'inverse de la fonction $T : O \rightarrow (\mathcal{C}/\simeq), o \mapsto T_o$ qui définit les types T_o . Le même raisonnement peut être reproduit à partir de \mathcal{C}' et de f' .

Enfin, supposons qu'il existe une bijection $G : O \rightarrow O'$ entre les ensembles d'observation et que O et O' sont donc de même cardinalité. Ainsi, on peut définir une autre bijection $F' = \tilde{f}'^{-1} \circ G \circ \tilde{f} : (\mathcal{C}/\simeq) \rightarrow (\mathcal{C}'/\simeq)$ entre les ensembles de types d'objets, bijection en réalité induite par l'équivalence de catégories $F : \mathcal{C} \rightarrow \mathcal{C}'$ définie comme suit : pour tout $x \in \mathcal{C}$, soit $o = f(x)$ et soit $x' \in f'^{-1}(G(o))$, définissons F telle que $F(x) = x'$. Si \mathcal{C} et \mathcal{C}' ont des cardinaux différents, F ne peut pas être bijective. Par construction, F envoie chaque état x vers un état x' du même type (modulo G). Cela permet de relier les catégories \mathcal{C} et \mathcal{C}' , de telle sorte que si l'on dispose d'une

stratégie applicable dans \mathcal{C}' , elle peut être transposée dans \mathcal{C} par la fonction F .

L'utilisation de la théorie des catégories permet de formaliser tout un éventail de situations. Nous choisissons de l'illustrer sur un circuit de voitures miniatures.

4.2 Analogies entre configurations d'un circuit de voitures

Nous introduisons les notations suivantes : soient N et N' les nombres de segments par configuration du circuit, et $\mathcal{C} = \{s, s \in [1, N]\}$, $\mathcal{C}' = \{s', s' \in [1, N']\}$ les ensembles de positions possibles de la voiture sur le circuit, qui ne sont utiles que pour le raisonnement théorique et ne sont pas utilisés lors des expériences. Soient $(u, i)_s$ (resp. $(u', i')_{s'}$) la tension et le courant mesurés lorsque la voiture passe sur la section s (resp. s'). Soit $1 \leq s_0 \leq N$ (resp. $1 \leq s'_0 \leq N'$) la position initiale de la voiture dans la configuration \mathcal{C} (resp. \mathcal{C}'). Nous notons k une ligne droite de \mathcal{C} et l un virage. De même, soient k' une ligne droite et l' un virage de \mathcal{C}' .

Dans la configuration \mathcal{C}' , le joueur peut agir sur $(u', i')'_{s'}$ en utilisant la manette de jeu, ce qui correspond à la stratégie π' définie par (1).

$$\pi'(s') = \begin{cases} (u', i')_{k'}, & \text{si } s' \text{ est une ligne droite} \\ (u', i')_{l'}, & \text{sinon} \end{cases} \quad (1)$$

Nous voulons identifier \mathcal{C} et \mathcal{C}' , pour pouvoir transposer la stratégie π' de \mathcal{C}' à \mathcal{C} . Les états de \mathcal{C} sont les positions s de la voiture sur le circuit. De même, les états de \mathcal{C}' sont les positions s' . Si $N = N'$ et $s_0 = s'_0$, nous pouvons définir une bijection entre \mathcal{C} et \mathcal{C}' et transposer π' de manière triviale. En revanche, si $N \neq N'$ ou $s_0 \neq s'_0$, il est impossible de définir une telle bijection.

Cependant, si nous transformons \mathcal{C} et \mathcal{C}' en catégories, en définissant des morphismes, nous serons capables de définir une équivalence de catégories $F : \mathcal{C} \rightarrow \mathcal{C}'$. Dès lors, nous définirons la stratégie π appliquée sur \mathcal{C} par $\pi = \pi' \circ F$. Pour définir ces morphismes, nous utilisons la fonction observable f définie sur les états s de \mathcal{C} comme suit : $f : \mathcal{C} \rightarrow \{1, 2\}$ avec $f = h \circ g$ où g et h sont telles que $g(s) = (u, i)_s$ et $h((u, i)_s) = 1$ si s est un virage, et 2 sinon. f' est définie de la même manière sur les s' de \mathcal{C}' , c'est-à-dire $f' : \mathcal{C}' \rightarrow \{1, 2\}$ avec $f' = h' \circ g'$ et g' et h' sont définies de la même manière que g et h , mais sur les états de \mathcal{C}' .

Nous définissons un isomorphisme entre deux états de \mathcal{C} ssi ils ont la même image par f , et un isomorphisme entre deux états de \mathcal{C}' ssi ils ont la même image par f' . Nous définissons ensuite $F : \mathcal{C} \rightarrow \mathcal{C}'$ par (2).

$$F(s) = \begin{cases} l', & \text{si } f(s) = 1 \\ k', & \text{si } f(s) = 2 \end{cases} \quad (2)$$

Il est facile de voir, si les définitions exactes sont connues, que (2) est une équivalence de catégories qui permet de transférer π de \mathcal{C}' à \mathcal{C} . F induit une bijection F' entre les

ensembles de classes (ou types de position – ici $\mathcal{C}, \mathcal{C}'$ sont les types de virage et $\mathcal{S}, \mathcal{S}'$ sont les types de lignes droites) : $F' : (\mathcal{C} / \simeq) \rightarrow (\mathcal{C}' / \simeq)$ où (\mathcal{C} / \simeq) est égal à $\{\mathcal{C}, \mathcal{S}\}$ et (\mathcal{C}' / \simeq) est égal à $\{\mathcal{C}', \mathcal{S}'\}$. On obtient finalement $F'(\mathcal{C}) = \mathcal{C}'$ et $F'(\mathcal{S}) = \mathcal{S}'$.

Cet exemple de catégorisation systématique et de généralisation prouve que nous ne travaillons pas à l'échelle des états, mais que nous considérons les types d'états.

5 Résultats Expérimentaux

5.1 Matériel

Le support expérimental est constitué d'un circuit MINI Challenge Set C1320 de la marque Scalextric dont le compteur de tours mécanique a été remplacé par un capteur à effet Hall omnipolaire à sortie binaire. Le courant est mesuré par une résistance shunt placée en série avec les rails d'alimentation. Le courant et la tension sont d'abord filtrés par un filtre RC passif du second ordre décrit en annexe, puis échantillonnés à $f_s = 100$ Hz. Le système ne contient aucun capteur supplémentaire. Les algorithmes sont écrits en C et exécutés en temps réel sur un Arduino Mega 2560 qui dispose de 8192 octets de mémoire vive (RAM). Nous définissons la période d'échantillonnage par $t_s = 1/f_s$, et l'instant associé à l'échantillon k est kt_s où $k \in \mathbb{N}$.

5.2 Implantation dans le cas bijectif

Le cas bijectif mentionné en 4.2 correspond au cas où $N = N'$ et $s_0 = s'_0$ et repose sur une procédure d'imitation en trois étapes. Un joueur humain commence par conduire la voiture sur n tours, avec $n = 3$ dans nos expériences. Les K valeurs de tensions $v(kt_s)$ et de courants $i(kt_s)$ ($0 \leq k < K$) échantillonnées sur le tour le plus rapide, ainsi que le temps du tour associé t_{best} , sont sauvegardés dans la RAM, pour être ensuite rejoués par l'IA. Une méthode d'optimisation (Newton) minimise la différence entre le temps de tour de l'IA et t_{best} en ajustant la commande par modulation en largeur d'impulsion (MLI) basée sur les échantillons enregistrés $v(kt_s)$.

5.3 Implantation dans le cas basé sur la théorie des catégories

Ce cas, inspiré de la théorie des catégories, repose sur deux modules : un module de récompense, ainsi qu'un module de prise de décision, tous deux décrits ci-dessous.

Comme en apprentissage par renforcement, notre approche s'appuie sur une récompense fournie par l'environnement. Celle-ci est le résultat de la combinaison de trois variables. La première variable est le temps du tour, mesuré directement par le compteur de tours. La seconde variable binaire reflète la présence de la voiture sur la piste, et la dernière variable binaire indique si la voiture se déplace ou pas. Ce module surveille constamment la voiture pour vérifier qu'elle n'est pas sortie de la piste à cause d'une vitesse trop élevée, ou qu'elle ne s'est pas arrêtée à cause de frottements trop importants. Ces deux détecteurs reposent sur une classification par k plus proches voisins (k-NN) en

utilisant les courants et tensions comme entrées. Le classificateur est entraîné sur 1301 échantillons avec $k = 3$. Rappelons que les données sont échantillonnées à une fréquence de 100 Hz : l'entraînement utilise donc 13 s de jeu. Une version réduite des données du classificateur est utilisée afin d'être exécutable en temps réel sur l'Arduino.

En utilisant ce module de récompense, l'IA peut piloter avec succès la voiture sur des circuits qu'elle découvre, sans rejouer ou manipuler des échantillons enregistrés lors d'une partie du joueur humain. La seule information conservée par l'algorithme est une vitesse de sécurité dont il sait qu'elle ne provoque ni sortie de piste, ni arrêt de la voiture. Comme la configuration du circuit change, il n'y a pas bijection entre les deux configurations : le cas bijectif ne peut pas s'appliquer. L'IA s'appuie sur des analogies et transpose les connaissances acquises dans une première configuration en utilisant l'équation (1), conformément au formalisme décrit en Partie 4. La fonction $h((u, i)_s)$ est évaluée par un k-NN entraîné sur 2327 échantillons (u, i) avec $k = 5$ et implémenté de la même façon que le premier classificateur.

En pratique, l'approche par analogies se déroule comme suit : la voiture démarre sur le circuit inconnu avec cette vitesse de sécurité importée de la première configuration. La fonction $h((u, i)_s)$ évaluée par le k-NN à partir des mesures de courant et de tension, détermine si la voiture est dans une configuration que nous appelons courbe ou ligne droite. L'IA choisit ensuite la meilleure commande utilisée au cours des expériences précédentes, en gardant l'objectif de minimiser le temps de tour en maintenant la voiture sur la piste. L'algorithme permet donc de généraliser les connaissances acquises au cours d'expériences passées et de les appliquer dans une configuration radicalement différente. En effet, les deux circuits schématisés en figure 2 sont de forme et de taille différentes, et une simple reproduction d'une stratégie ou d'une commande préalablement enregistrée conduirait rapidement à une sortie de route.

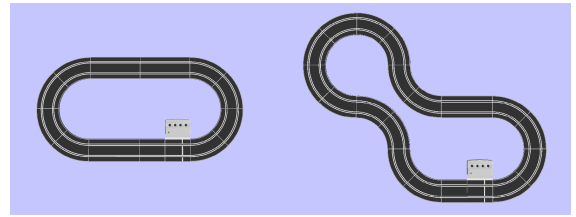


FIGURE 2 – Configuration des circuits : circuit 1 (à gauche), circuit 2 (à droite)

5.4 Résultats

Les expériences décrites dans cet article ont été menées sur deux configurations de complexité différente présentées en figure 2. La première est un anneau de 12 tronçons alors que la deuxième dispose de 18 tronçons. Les temps au tour sont décrits dans la table 1 pour le cas d'application de la théorie des catégories et dans le tableau 4 pour le cas bijectif. Ils sont donnés sous la forme (moyenne \pm écart-type),

TABLE 1 – Temps de tour en secondes pour le cas d’application de la théorie des catégories (moyenne \pm écart-type)

	HUMAIN (10 tours)	IA (jusqu’à 10 tours)	RÉGLAGES DE L’IA
CIRCUIT 1 (12 tronçons)			
Premier tour	2.99 \pm 0.46	3.12 \pm 0.09	MLI=39% de la vitesse maximale
Dernier tour	2.29 \pm 0.14	2.52 \pm 0.08	Analogies (adaptation de la vitesse)
CIRCUIT 2 (18 tronçons)			
Premier tour	4.30 \pm 1.16	3.66 \pm 0.03	MLI=39% de la vitesse maximale
Dernier tour	3.08 \pm 0.54	3.13 \pm 0.02	Analogies (adaptation de la vitesse)

sauf pour le meilleur tour humain qui est le meilleur temps obtenu parmi les 7 participants.

Impact de l’entretien. La voiture et le circuit sont régulièrement entretenus, comme stipulé en partie 2.1, afin que le système soit presque LTI et les résultats reproductibles. L’impact de cet entretien peut être estimé sur la figure 3. La non-invariance temporelle est aussi visible, puisque le temps au tour augmente dès que l’on dépasse la cinquantaine de tours parcourus, malgré une MLI constante.

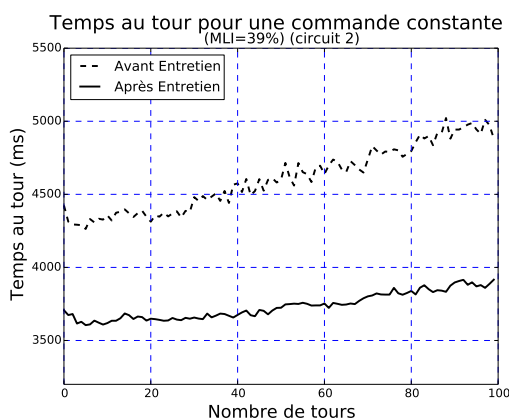


FIGURE 3 – Le système ne présente pas d’invariance temporelle

Évaluation des classificateurs. Les matrices de confusion des algorithmes de classification utilisés sont données dans le tableau 2 pour l’estimateur du signal d’arrêt de la voiture et dans le tableau 3 pour la fonction $h((u, i)_s)$, qui détermine la forme du tronçon sur lequel la voiture se trouve. Dans un souci de concision, l’estimateur de sortie de piste n’est pas analysé dans ce document, car il est similaire à celui du signal d’arrêt.

TABLE 2 – Matrice de confusion de l’estimateur "Voiture arrêtée" sur 641 échantillons (u, i)

		CLASSE ESTIMÉE	
		Arrêt	Mouvement
CLASSE RÉELLE	Arrêt	171	1
	Mouvement	4	465

TABLE 3 – Matrice de confusion de l’estimateur $h((u, i)_s)$ sur 287 échantillons (u, i)

		CLASSE ESTIMÉE	
		Droite	Courbe
CLASSE RÉELLE	Droite	51	5
	Courbe	1	230

TABLE 4 – Temps au tour en secondes pour le cas bijectif

	MEILLEUR HUMAIN	IA
CIRCUIT 1	2.11	2.12 \pm 0.05
CIRCUIT 2	2.67	2.65 \pm 0.02

Résultats dans le cas bijectif. Dans le cas bijectif, l’algorithme converge en moins de 5 tours de piste. Un exemple de cette convergence est représenté figure 4. Si le tour reproduit est le meilleur tour humain, cette approche aboutit au meilleur temps au tour. Néanmoins, cet algorithme ne peut être appliqué que si l’humain et l’algorithme pilotent sur des circuits identiques, contrairement à la solution qui s’appuie sur des analogies. En particulier, bien que le meilleur tour imité (bijectif) se fasse en 2.65 s sur le circuit 2 contre 3.13 s dans le cas où la voiture est pilotée par une approche basée sur l’établissement d’analogies, ce temps et ces vitesses ne sont viables que sur le circuit 2 et ne peuvent aboutir à une quelconque généralisation. Toutefois, ce temps est une borne inférieure du temps atteignable par un humain dans une configuration donnée.

Résultats dans le cas d’utilisation des analogies catégoriques. L’IA démarre à vitesse constante (tour 1 sur la figure 5), importée depuis la configuration précédente. À partir des mesures de courant et de tension, le k-NN détecte les tronçons du circuits qui ne présentent aucun danger. La MLI est ensuite augmentée dans le dernier de ces tronçons dès le deuxième tour. Pour aboutir à une amélioration moyenne du temps au tour due à cette accélération, le procédé est répété sur trois tours (tours 2, 3 and 4), car il n’y a pas d’invariance temporelle. Enfin, l’IA accélère dans le deuxième tronçon qui ne présente aucun danger sur les tours 5, 6 et 7 pour aboutir à un temps au tour final.

Dans les cas où le nouveau temps au tour est plus élevé

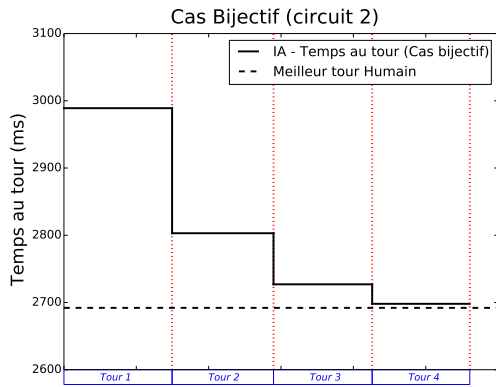


FIGURE 4 – Exemple de Cas Bijectif

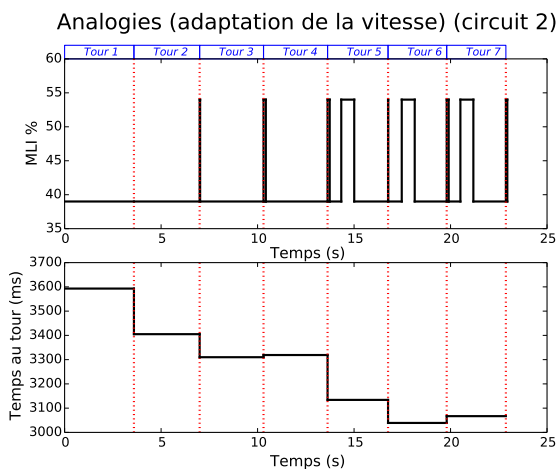


FIGURE 5 – Un exemple d'utilisation des analogies

que celui attendu, l'IA remet en cause ses connaissances, comme décrit en partie 3.2, en adoptant par exemple une vitesse plus élevée au travers de la MLI.

Cette IA, qui repose sur l'établissement d'analogies entre configurations et ne rejoue pas d'enregistrement d'une partie précédente, est capable d'améliorer les temps au tour en moins de 10 tours sur une piste inconnue. Sur le circuit le plus compliqué, l'IA atteint des temps similaires aux temps humains : 3.08 s pour les derniers tours des joueurs humains (*a fortiori* les meilleurs, car bénéficiant des stratégies mises en place lors des tours précédents) contre 3.13 s en moyenne pour l'algorithme sur le circuit 2.

Les améliorations futures de l'algorithme sur un circuit inconnu incluront l'optimisation des vitesses transposées par la fonction $h((u, i)_s)$. En effet, seule une vitesse de sécurité a été utilisée ici afin d'éviter les sorties de piste de la voiture pilotée par l'algorithme, alors que certains tours humains ont occasionné des sorties de piste et n'ont donc pas été comptabilisés.

Pour résumer les travaux sur le circuit de voitures, le cadre théorique détaillé en section 4 permet au système d'être au niveau des meilleurs joueurs humains sur ce jeu, le tout en moins d'une minute, et sans ajout de capteurs embarqués. Ce cadre permet d'atteindre de tels résultats sur des circuits inconnus où la simple reproduction d'un tour humain

conduirait immédiatement à une sortie de piste.

6 Conclusion

Nous avons illustré l'utilisation de concepts élémentaires de la théorie des catégories dans un contexte d'apprentissage automatique pour un système embarqué : un circuit de voitures miniatures. Aucun capteur n'a été ajouté à la voiture. Les seules informations utilisées sont le courant et la tension du circuit, ainsi qu'un compteur de tours. Nous montrons comment transférer la connaissance acquise d'une situation à une autre lorsqu'il n'y a pas de bijection entre elles, en utilisant les analogies définies grâce à cet outil mathématique.

Les résultats expérimentaux prouvent qu'il est simple de mettre en œuvre un algorithme d'apprentissage automatique basé sur la théorie des catégories, y compris sur un calculateur à faible puissance de calcul. Cette théorie permet de transférer les connaissances à des situations jamais rencontrées auparavant, tout en utilisant des outils classiques et connus, ce qui permet à l'IA d'apprendre avec peu de données.

Annexe 1 : filtre en échelle RC

Bien que les filtres actifs soient omniprésents de nos jours, l'approche de « conception par les coûts », usuelle dans des industries de production en grandes séries, mène au choix d'un filtre anti-repliement passif en échelle utilisant uniquement des résisteurs (R) et des condensateurs (C). Notons que si les filtres passifs en échelles composés de solénoïdes et de condensateurs sont classiques, les versions sans solénoïdes sont assez peu usuelles, d'où les précisions apportées par cette annexe.

Le filtre a une fréquence de coupure f_c , et sa fonction de transfert est $G(s) = 1/(sRC + 1)^2$ où s est la variable de Laplace. Elle est approximée par un double réseau RC de Cauer décrit dans [1] dont les valeurs de R et C du premier réseau RC vérifient $f_c = 1/(2\pi RC)$. Les valeurs du second circuit RC, R/d et Cd avec $\{d \in \mathbb{R} : d > 0\}$, sont calculées pour vérifier une condition sur l'erreur maximale sur le gain $e(d)$ entre $G(s)$ et $G_a(s)$, la fonction de transfert du réseau de Cauer définie par $G_a(s) = 1/((sRC)^2 + s(d+2)RC + 1)$. La fonction $e(d)$ est définie par l'équation (3). Nous choisissons pour notre application $d = 0.1$, ce qui conduit à une erreur inférieure à 0.5 dB, sans répercussion sur les calculs qui en dépendent. Une application avec deux réseaux RC identiques donnerait $e(1) = 3.5$ dB, ce qui diminuerait la performance globale du système. La tension et le courant sont tous les deux filtrés grâce à de tels filtres RC en échelle.

$$e(d) = 20 \log_{10} \left(\frac{d+2}{2} \right) \quad (3)$$

References

- [1] N. Balabanian, *Network Synthesis*. Prentice-Hall Electrical Engineering Series, 1958.

- [2] B. Boots and G. J. Gordon, "An Online Spectral Learning Algorithm for Partially Observable Nonlinear Dynamical Systems," in *AAAI Conference on Artificial Intelligence*, August 2011, pp. 293–300.
- [3] M. Brejl and J. Necesany, "Student's contest: Self-driven slot car racing," in *IMCSIT 2008*, Oct 2008, pp. 589–592.
- [4] Z. Chen and B. Liu, *Lifelong Machine Learning*. Morgan & Claypool Publishers, 2016.
- [5] D. R. Cheriton, M. A. Malcolm, L. S. Melen, and G. R. Sager, "Thoth, a Portable Real-time Operating System," *Commun. ACM*, vol. 22, no. 2, pp. 105–115, Feb 1979.
- [6] F. de Saussure, *Cours de linguistique générale*. Payot, 1916.
- [7] T. Delbruck, M. Pfeiffer, R. Juston, G. Orchard, E. Muggler, A. Linares-Barranco, and M. W. Tilden, "Human vs. computer slot car racing using an event and frame-based DAVIS vision sensor," in *IEEE IS-CAS*, May 2015, pp. 2409–2412.
- [8] H. Dreyfus, *Being-in-the-world: A commentary of the Heidegger's Being and Time, Division I*. MIT Press, 1990.
- [9] A. Feniello, H. Dang, and S. Birchfield, "Program synthesis by examples for object repositioning tasks," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sept 2014, pp. 4428–4435.
- [10] J. A. Fulcher, "Fun and games and microcomputer interfacing (laboratory exercises)," *IEEE Micro*, vol. 11, no. 1, pp. 18–21, Feb 1991.
- [11] P. Gober, "Experiences with a Project to Design Autonomous Slotcars in a Mechatronics Master's Program," in *Proceedings of the WESE'15: Workshop on Embedded and Cyber-Physical Systems Education*. ACM, Oct 2015, pp. 5:1–5:3.
- [12] R. Jonschkowski and O. Brock, "State representation learning in robotics: Using prior knowledge about physical interaction," in *Proceedings of Robotics: Science and Systems*, July 2014.
- [13] S. Kane and J. B. Scott, "The slot car stig: Performance and consistency of a slot car driven by a heuristic algorithm in an embedded microcontroller," in *The 16th Electronics New Zealand Conference*, Dunedin, New Zealand, Nov 2009, pp. 177–180.
- [14] V. Kapila and S.-H. Lee, "Science and mechatronics-aided research for teachers," *IEEE Control Systems*, vol. 24, no. 5, pp. 24–30, Oct 2004.
- [15] T. C. Kietzmann and M. Riedmiller, "The Neuro Slot Car Racer: Reinforcement Learning in a Real World Setting," in *ICMLA '09*, Dec 2009, pp. 311–316.
- [16] J. Ko and D. Fox, "Learning GP-Bayes Filters via Gaussian process latent variable models," *Autonomous Robots*, vol. 30, no. 1, pp. 3–23, Jan 2011.
- [17] S. Lange, M. Riedmiller, and A. Voigtlander, "Autonomous reinforcement learning on raw visual input data in a real world application," in *IJCNN 2012*. IEEE, 2012, pp. 1–8.
- [18] S. Mac Lane, *Categories for the working mathematician*, 2nd ed. Springer-Verlag, 1998.
- [19] D. Martinec and M. Bundzel, "Evolutionary algorithms and reinforcement learning in experiments with slot cars," in *International Conference on Process Control*, June 2013, pp. 159–162.
- [20] D. Martinec, M. Sebek, and Z. Hurak, "Vehicular platooning experiments with racing slot cars," in *CCA 2012*, Oct 2012, pp. 166–171.
- [21] L. McCalman, S. O'Callaghan, and F. Ramos, "Multimodal estimation with kernel embeddings for learning motion models," in *ICRA 2013*, May 2013, pp. 2845–2852.
- [22] P. J. McKerrow, "Micro-computers, slotcars and education," University of Wollongong, Australia, Working Paper 82-7, 1982.
- [23] A. T. Mogill, "Calibration and use of a small motor as a slot car speedometer," *Physics Teacher*, vol. 13, no. 5, pp. 304–305, 1975.
- [24] J. B. Moore and B. D. O. Anderson, "Optimal linear control systems with input derivative constraints," *Electrical Engineers, Proceedings of the Institution of*, vol. 114, no. 12, pp. 1987–1990, Dec 1967.
- [25] K. Popper, *The Logic of Scientific Discovery*. Hutchinson & Co, London, 1959.
- [26] H. Prade and G. Richard, "Analogical proportions and analogical reasoning – an introduction," in *Case-Based Reasoning Research and Development*, D. W. Aha and J. Lieber, Eds. Springer, 2017, pp. 16–32.
- [27] L. Pusman and K. Kosturik, "Control algorithm based on phase locked loop," in *Telecommunications Forum (TELFOR)*, Nov 2013, pp. 605–607.
- [28] D. Richter, A. Grapentin, and A. Polze, "Mobility-as-a-service: A Distributed Real-Time Simulation with Carrera Slot-Cars," in *IEEE 18th International Symposium on Real-Time Distributed Computing*, Apr 2015, pp. 276–279.
- [29] G. Robertson, "Applying learning by observation and case-based reasoning to improve commercial RTS game AI," in *AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 2012.
- [30] L. Song, B. Boots, S. M. Siddiqi, G. J. Gordon, and A. J. Smola, "Hilbert Space Embeddings of Hidden Markov Models," in *ICML*, June 2010, pp. 991–998.
- [31] B. Sprunt, "A Novel Race Track Platform For Teaching Microcontroller System Design Concepts," in *Proc. of the American Society for Engineering Education Annual Conference*, June 2003.

FIGURE 6 – Algorithme utilisé.

Require: $bdc = [entites, modeles, strategies]$
Require: $strategie_par_defaut = MLI_constante$
Require: $seuil, N_{max}$

```

1: for  $N_{init} = 0$  to  $N_{max}$  do
2:    $entites.append(k\text{-}NN(u,i))$ 
3:    $action \leftarrow strategie\_par\_defaut$ 
4:    $applique\_MLI(action)$ 
5: end for
6:  $jeu\_similaire \leftarrow$  recherche jeu avec des  $entites$  similaires
7:  $modeles, strategies \leftarrow$  extrait( $bdc, jeu\_similaire$ )
8:  $bdc.append([entites, modeles, strategies])$ 
9:  $f\_refute\_savoir \leftarrow$  false
10: if  $modeles$  est vide then
11:    $modeles \leftarrow$  identification_parametrique( $entites$ )
12: end if
13: while experimentation en cours do
14:    $entites.append(k\text{-}NN(u,i))$ 
15:   if  $f\_refute\_savoir$  then
16:      $jeu\_similaire \leftarrow$  cherche jeu avec des  $entites$  similaires
17:      $modeles, strategies \leftarrow$  extrait( $bdc, jeu\_similaire$ )
18:      $f\_refute\_savoir \leftarrow$  false
19:      $bdc.append([entites, modeles, strategies])$ 
20:   else
21:      $action \leftarrow$  cherche_action( $entites, strategies$ )
22:      $predictions \leftarrow$  cherche_predictions( $entites, modeles$ )
23:      $erreur \leftarrow predictions - entites$ 
24:     if  $|erreur| > seuil$  then
25:        $f\_refute\_savoir \leftarrow$  true
26:     end if
27:   end if
28:    $applique\_MLI(action)$ 
29: end while
30: return  $bdc$ 

```

Annexe 2 : algorithme

L'algorithme utilisé sur le circuit de voitures est décrit en pseudo-code sur la figure 6. Il repose sur une base de connaissance bdc , une stratégie par défaut $strategie_par_defaut$, un seuil $seuil$ de rejet du modèle et un nombre d'itérations N_{max} pour le ou les premiers tours du circuit.

La base bdc est initialisée avec une stratégie obtenue en analysant une précédente course lorsqu'un humain pilotait le véhicule sur un autre circuit : une MLI par type de tronçon. La stratégie par défaut $strategie_par_defaut$ est une MLI constante, par exemple la plus petite MLI mesurée sur les courses précédentes et qui permet à la voiture d'avancer (39% dans nos expériences). La valeur de N_{max} correspond à 3 tours de circuit.

L'algorithme commence par acquérir la tension u et le courant i en ligne 2, puis assimile la paire (u,i) à un type d'entité (courbe ou droite) via le k-NN. Lors de la première utilisation, il n'existe aucun modèle dans la base bdc : un modèle affine donnant le temps au tour en fonction de la MLI est ensuite calculé en ligne 11. Si la base bdc contient

des informations issues d'une précédente course sur un circuit, alors la stratégie de la course la plus similaire à la configuration actuelle est transférée. Ici, la stratégie se résume à une MLI par type de tronçon.

Lorsque la voiture avance sur le circuit, le k-NN transforme à nouveau tension et courant en entités en ligne 14. Cette entité permet, à partir de la stratégie, de choisir une des deux actions suivantes : MLI pour le tronçon droit ou MLI pour le tronçon courbe. Une prédiction du temps au tour est calculée en ligne 22. Si l'erreur sur cette prédiction est supérieure à $seuil$ (50 ms), la variable f_refute_savoir devient vraie en ligne 25. À l'itération suivante en ligne 16, une nouvelle stratégie, voire un nouveau modèle seront calculés à partir de la base bdc puis appliqués et testés, et ainsi de suite jusqu'à la fin de l'expérience. La base de connaissance bdc est alors retournée et pourra être réutilisée sur un circuit de configuration différente.

Notons que l'algorithme est identique en phase d'apprentissage et en phase d'exploitation (test). Il continue de réfuter son savoir, si besoin est, même en phase d'exploitation.

Théorie des Catégories et Critère de Contrôlabilité pour Atari 2600

L. Cordesses¹ O. Bentahar¹ K. Poulet¹ A. Laurent¹ S. Amar¹ T. Ehrmann¹ J. Page²

¹ Renault Innovation Silicon Valley, 1215 Bordeaux Drive, Sunnyvale, 94089 CA, USA

² Université Paris Diderot-CNRS, Laboratoire SPHERE, 5 rue Thomas Mann, 75205 Paris Cedex 13, France

{lionel.cordesses, kevin.poulet, thomas.ehrmann}@renault.com,
{aude.laurent, sarah.amar, omar.bentahar}@nissan-usa.com,
ju.page@hotmail.fr

Résumé

L'objectif de l'apprentissage par renforcement est de contrôler un agent ou système dynamique dans le but de tirer la meilleure récompense de son environnement. Toutefois, dès lors que l'on prend en compte des contraintes réalistes telles qu'un nombre restreint d'expériences ou une puissance de calcul limitée, les approches classiques perdent de leur efficacité.

Nous introduisons des éléments de la théorie du contrôle pour diminuer le nombre d'essais requis. En parallèle, des outils mathématiques issus de la théorie des catégories nous permettent de formaliser une approche d'apprentissage automatique innovante reposant sur la formalisation d'analogies entre problèmes. Cette méthode suit une approche scientifique en remettant en cause – si besoin est – ses propres conclusions. Enfin, nous présentons une première application de la théorie des catégories sur les jeux Atari 2600, et illustrons les avantages d'utiliser des analogies pour transférer des connaissances entre jeux vidéo.

Mots Clef

Apprentissage Automatique, Théorie des Catégories, Atari 2600.

Abstract

Reinforcement Learning aims at controlling a dynamic system or agent to get the best reward from its environment. However, when adding real-world constraints such as a limited number of trials and limited computing power, classic and state-of-the-art approaches may perform sub-optimally.

We introduce mathematical concepts from control theory related to dynamic systems to alleviate the need for many trials, while formalizing a Machine Learning approach relying on mathematical analogies from category theory. We argue that our system is able to refute its own findings, as it is the case with any scientific approach. We then present what we believe is the first application of category theory to play Atari 2600 games, and we illustrate the use and benefits of analogies to transfer knowledge between games.

Keywords

Machine Learning, Category Theory, Atari 2600.

1 Introduction

Bien qu'elles soient excellentes lorsqu'il s'agit d'apporter des solutions, les méthodes à la pointe de l'apprentissage automatique n'atteignent pas d'aussi bons résultats dès lors que l'on ajoute des contraintes réalistes telles qu'un budget restreint ou des délais limités. Aussi, plus la recherche en Intelligence Artificielle (IA) progresse sur les problématiques du quotidien telles que la conduite automobile, plus la capacité à transférer des connaissances passées à des situations inconnues devient cruciale, puisque cela réduit considérablement le nombre d'essais requis. En effet, bien que les algorithmes les plus récents aient surpassé les humains sur des tâches spécifiques telles que le jeu de Go [26] ou les échecs [4], les humains restent supérieurs aux machines quand il s'agit d'évoluer et de prendre des décisions pertinentes dans un environnement inconnu. De manière générale, les humains savent généraliser les connaissances et expériences passées.

Un autre aspect prépondérant des algorithmes d'IA est leur robustesse. Certains algorithmes d'apprentissage par renforcement, par exemple, perdent leur efficacité dès que les données changent légèrement (un objet déplacé vers le haut de quelques pixels [10], ou une variation légère de la luminosité de l'image [14]). Même si entraîner de nouveau le système résout ce genre de problème, cela met en évidence le besoin d'algorithmes d'IA capables de remettre en cause ce qu'ils ont appris, en particulier si les observations contredisent le modèle établi.

Cela soulève aussi la question de l'échelle à laquelle le raisonnement doit se faire, puisque cette remise en cause nécessite une capacité de déduction. Par exemple, si les données d'entrée sont des images, certains considèrent que le pixel est une échelle trop réduite, et que se focaliser sur des amas de pixels est plus pertinent [25]. De la même façon, nous considérons qu'un cadre mathématique fondé sur la *théorie des catégories* apporte des outils rigoureux pour formaliser ces situations où l'utilisation de connaissances

passées donne un avantage. Ce cadre permet de détailler l'exploitation des analogies entre différentes situations, et définit aussi le processus de transfert.

Afin de réduire le nombre d'échantillons, et donc la quantité de données utilisées, nous proposons un concept d'entités. Nos travaux ont commencé dans le but d'obtenir des performances similaires à l'état de l'art, mais en n'utilisant qu'un centième des ressources et du temps habituellement consacrés. Nous validons cette approche sur les jeux Atari 2600. Nous utilisons aussi ces jeux pour valider la capacité à transférer des connaissances non seulement au sein d'un même jeu, mais aussi entre des jeux tactiquement ou visuellement similaires.

2 Travaux antérieurs

Les jeux proposent un éventail de problèmes qui peuvent être modélisés par un environnement fermé et fini. Les jeux vidéo ont eux la particularité de disposer d'un environnement qui évolue en temps réel. De plus, ils obéissent à un ensemble de règles et constituent donc une représentation d'un monde virtuel mais rationnel. Ce faisant, les jeux vidéo sont devenus une référence pour tester les algorithmes. En effet, l'essor de benchmarks vidéoludiques tels que Arcade Learning Environment (ALE) [2] a permis des progrès phénoménaux des méthodes d'apprentissage en temps réel, telles que l'apprentissage par renforcement [8][18]. Ce benchmark comprend une myriade de problèmes différents qui impliquent tous une prise de décision et ont pu être modélisés par de larges processus de décision Markoviens [19]. Aussi, la versatilité de ces problèmes permet de tester différentes méthodes. Les premiers travaux, à l'instar du Deep-Q-Network (DQN) décrit dans [20], utilisent des réseaux de neurones profonds sur des images pré-traitées de l'aire de jeu. Des algorithmes de recherche ou de planification classiques tels que la recherche arborescente Monte-Carlo [22] ou « Iterated Width Algorithm » [15] peuvent aussi être implémentés pour guider le système.

Néanmoins, entraîner un algorithme d'apprentissage profond est très chronophage et demande une quantité importante de données. Par exemple, les conditions expérimentales de [20] demandent 200 millions d'images de jeu. Des méthodes plus exotiques ont été développées afin de diminuer le nombre d'essais nécessaires, par exemple « Inverse Reinforcement Learning » [11] ou l'apprentissage par imitation [3]. Un autre aspect modulable en vue d'une diminution du nombre d'essais requis est l'architecture des réseaux utilisés et *a fortiori* la procédure d'entraînement choisie. Des architectures moins profondes [13] ou hybrides [12] ont été mises en oeuvre pour contourner ces problèmes.

L'humain et la machine comprennent et représentent le monde différemment puisque, contrairement à un algorithme, une personne est capable de transférer et généraliser ses connaissances, ce qui lui permet de s'adapter rapidement à un environnement inconnu. On peut considérer que, dans un premier temps, chaque nouvelle situation est

traitée en établissant des analogies entre problèmes. C'est pourquoi on peut prétendre que des algorithmes, qui ne peuvent transférer leurs connaissances de manière aussi intuitive, sont incapables de prendre des décisions pertinentes dès lors que l'environnement change. Toutefois, les techniques d'apprentissage par transfert cherchent à reproduire une caractéristique similaire et à abolir cette frontière, et ont récemment fait des progrès colossaux [1] [7] [16]. Un avantage évident de ces techniques demeure la possibilité de contourner les besoins en expériences en important des connaissances et raisonnements d'autres situations.

Cependant, transférer ses connaissances depuis un autre domaine peut ne pas suffire à atteindre les niveaux de versatilité et d'adaptation de l'être humain. En effet, la façon dont les gens utilisent leurs intuitions pour établir des liens de causalité pourrait être une raison expliquant cette aptitude à transférer des connaissances d'un domaine à un autre pourvu qu'ils partagent des causalités similaires. En plus de ce concept, les « Schema Networks » [10] s'appuient sur des entités plutôt que des images brutes, et font l'hypothèse qu'ils disposent d'un système capable de détecter et traquer ces entités sur l'écran. Dès lors, ils parviennent à réutiliser ce qu'ils ont appris pendant 100 000 frames d'entraînement sur le jeu Breakout à des variations de ce même jeu.

3 La Théorie des Catégories comme cadre

Bien que les méthodes présentées en Partie 2 parviennent à contrôler avec succès un agent dans les jeux Atari 2600, ces méthodes nécessitent un grand nombre d'images d'apprentissage. Notre objectif reste d'aboutir à une méthode d'apprentissage automatique qui pourrait apprendre avec peu de ressources, à savoir environ un centième des données et puissance de calcul habituellement requises par les meilleures solutions présentes dans la littérature.

À notre connaissance, de tels prérequis ne peuvent être validés à partir de méthodes existantes. C'est pourquoi nous nous sommes lancés dans la conception d'une nouvelle approche d'apprentissage automatique.

3.1 Construction des entités

Comparablement aux « Schemas Networks » [10], notre travail implique un concept « mésoscopique » d'entités. Ces entités sont en général nos objets du quotidien, ou bien dans le cas des jeux vidéo, tous les objets que l'on rencontre au cours de notre parcours dans cet environnement fermé. Par exemple, dans le jeu Breakout, on peut considérer que la raquette constitue une telle entité puisqu'elle se situe, en terme d'échelle, entre le « microscopique » pixel et la « macroscopique » image. Travailler à cette échelle peut s'apparenter à travailler à l'échelle du morphème en linguistique, *i.e.* le plus petit élément significatif d'un langage [6].

Par construction, à partir d'une analyse des signaux échan-

tillonnés, notre système est capable de détecter et traquer ces entités. Toutefois, le monde qui nous entoure n'est pas seulement constitué d'objets, mais aussi de sujets. Dans les problèmes que nous considérons, nous appelons « Moi » l'entité contrôlée par les actions du joueur. Nous supposons qu'une IA familière avec ces concepts sera plus à même de prendre des décisions pertinentes, et de transférer ses connaissances et raisonnements à tout un panel de situations qui ne partagent qu'une vague similarité.

Le « stade du miroir » est une étape clef du développement de l'identité de soi durant l'enfance. Elle correspond au moment où l'enfant commence à prendre conscience de son propre corps et à le distinguer de celui d'autrui. Nous concevons notre Intelligence Artificielle à partir de ce stade, en distinguant d'abord le « Moi » parmi toutes les entités présentes. Dès lors, les autres entités de notre environnement sont regroupées en deux grandes classes en fonction de leurs intentions, selon qu'elles sont plutôt hostiles (les « ennemis ») ou adjuvantes (les « amis »). Le parallèle est facilement établi avec la construction d'un modèle de récompense lors d'un algorithme classique d'apprentissage par renforcement : les amis apportent des récompenses positives, alors que les ennemis produisent des récompenses négatives.

3.2 Remise en cause des modèles

Les modèles obtenus ne sont pas éternellement valides puisque le « Moi », mais aussi l'environnement évoluent et peuvent changer. Ainsi, nous choisissons de nous orienter vers une approche scientifique telle que décrite par [23], c'est-à-dire qu'il doit être possible pour un système empirique d'être réfuté par le résultat de ses expériences, et concevons notre logiciel de sorte à ce qu'il puisse remettre en cause ses propres résultats, puis reconstruire ses modèles. Cette remise en cause se produit dès lors que l'erreur mesurée entre les observations du système et les prédictions du modèle dépasse un certain seuil.

3.3 Effet mémoire

Au fil des parties d'un même jeu, il est évident qu'un joueur humain retient les enseignements tirés des parties précédentes et sait toujours comment jouer. Notre logiciel mémorise les connaissances qu'il découvre (entités, modèles dynamiques, etc.) et les utilise lors des parties successives.

3.4 Utilisation d'analogies pour le transfert de connaissances entre problèmes

L'un des outils les plus efficaces à notre disposition pour évoluer dans une nouvelle situation est la capacité à faire des analogies entre cette situation et nos expériences passées. C'est pourquoi nous pressentons qu'une IA capable d'établir des analogies et sachant comment jouer au jeu Breakout sera capable de transposer ses capacités au jeu Pong, même si les aires de jeu et les règles ne sont pas identiques, à l'instar d'un joueur de tennis qui saurait comment manier une raquette de badminton même s'il n'a jamais pratiqué le sport.

Les situations où deux problèmes ont exactement le même nombre d'états et des structures isomorphes sont rares. Néanmoins, il existe des outils mathématiques pour identifier des structures non isomorphes tels que l'équivalence de catégories en théorie des catégories [17]¹. La théorie des catégories est un outil puissant et moderne qui est apparu au milieu du XX^{ème} siècle pour permettre le transfert de théorèmes entre différentes branches des mathématiques.

Dans les problèmes qui reposent sur la théorie des ensembles, l'identification se réduit aux relations d'identité et aux cas bijectifs, tandis que la théorie des catégories apporte des descriptions plus riches des objets au travers de l'introduction de flèches (ou morphismes) entre objets, ce qui permet de nouvelles sortes d'identifications. La plupart des outils d'apprentissage automatique s'appuyant sur la théorie des ensembles, et non sur la théorie des catégories, nous illustrons ci-dessous la plus-value d'un tel cadre mathématique.

Théorie des catégories. Une catégorie \mathcal{C} est une collection d'objets et de morphismes entre certains de ces objets, munie d'une composition de morphismes, et peut ainsi être assimilée à un graphe orienté. Si A et B sont des objets de \mathcal{C} , la flèche $a : A \rightarrow B$ est un isomorphisme si elle est inversible, i.e s'il existe une flèche $b : B \rightarrow A$, telle que $ba = Id_A$ et $ab = Id_B$. Dans ce cas, les objets A et B sont isomorphes. Les isomorphismes définissent une relation d'équivalence sur la classe des objets de \mathcal{C} . On note \mathcal{C}/\simeq son quotient. Aussi, si $F : \mathcal{C} \rightarrow \mathcal{C}'$ est ce qu'on appelle une équivalence de catégories, celle-ci induit une bijection $F' : (\mathcal{C}/\simeq) \rightarrow (\mathcal{C}'/\simeq)$ entre les classes d'objets isomorphes même si F elle-même n'est pas bijective. Dès lors, on n'identifie plus les objets (ou états) un à un entre deux situations, mais les types (ou classes d'isomorphisme) de ces états.

Ce procédé peut être utilisé dans le cas de problèmes observables. En effet, considérons deux ensembles d'états non vides \mathcal{C} et \mathcal{C}' sans autre hypothèse sur leur cardinalité. Supposons aussi l'existence de deux fonctions d'observations $f : \mathcal{C} \rightarrow O$ et $f' : \mathcal{C}' \rightarrow O'$. Quitte à restreindre O et O' , supposons que ces deux fonctions sont surjectives. Pour tout $o \in O$, on dit que les états $x \in \mathcal{C}$ dont l'observation associée est o (i.e $f(x) = o$) sont de type T_o . Cela définit une relation d'équivalence R_f sur l'ensemble $\mathcal{C} : \forall x, y \in \mathcal{C}, x R_f y$ si et seulement si (ssi) $f(x) = f(y)$. Transposé dans le champ lexical de la théorie des catégories, cela revient à placer une flèche inversible entre deux objets x et y de \mathcal{C} ssi $x R_f y$. \mathcal{C} devient alors une catégorie, où toutes les flèches sont inversibles et telle que \mathcal{C}/\simeq est exactement le quotient \mathcal{C}/R_f , quotient qui définit aussi l'ensemble des types d'états de \mathcal{C} . Puisque ces types ont été définis via les observations, la surjection $f : \mathcal{C} \rightarrow O$ induit une bijection $\tilde{f} : (\mathcal{C}/\simeq) \rightarrow O$ entre l'ensemble de types et l'ensemble d'observations. \tilde{f} est donc l'inverse de la fonc-

1. Une autre approche pourrait être l'utilisation de la théorie des modèles, plus spécifiquement du concept d'équivalence élémentaire [5] dans la perspective de doter l'IA de raisonnements logiques.

tion $T : O \rightarrow (\mathcal{C}/\simeq)$, $o \mapsto T_o$ qui définit les types T_o . Le même raisonnement peut être reproduit à partir de \mathcal{C}' et de f' .

Enfin, supposons qu'il existe une bijection $G : O \rightarrow O'$ entre les ensembles d'observation et que O et O' sont donc de même cardinalité. Ainsi, on peut définir une autre bijection $F' = \tilde{f}'^{-1} \circ G \circ \tilde{f} : (\mathcal{C}/\simeq) \rightarrow (\mathcal{C}'/\simeq)$ entre les ensembles de types d'objets, bijection en réalité induite par l'équivalence de catégories $F : \mathcal{C} \rightarrow \mathcal{C}'$ définie comme suit : pour tout $x \in \mathcal{C}$, soit $o = f(x)$ et soit $x' \in f'^{-1}(G(o))$, définissons F telle que $F(x) = x'$. Si \mathcal{C} et \mathcal{C}' ont des cardinaux différents, F ne peut être bijective. Par construction, F envoie chaque état x vers un état x' du même type (modulo G). Cela permet de relier les catégories \mathcal{C} et \mathcal{C}' , de telle sorte que si l'on dispose d'une stratégie applicable dans \mathcal{C} , elle peut être transposée dans \mathcal{C}' par la fonction F .

Une présentation plus détaillée et approfondie de la théorie des catégories peut être trouvée dans [17]. Un exemple de transfert de stratégie grâce à la théorie des catégories peut être trouvé en annexe de cet article.

Similarités entre paires d'entités mesurées grâce à la théorie des catégories. Les humains n'ont aucune difficulté à remarquer les similarités entre différentes situations. Nous montrons dans cette partie comment formaliser cette recherche de similarités, même lorsque les jeux sont complètement différents. Cela nous permettra de transférer des connaissances d'un jeu à un autre.

Nous avons identifié chaque situation par son espace d'états. D'autres paramètres auraient pu être pris en compte, comme le nombre d'entités et leurs interactions possibles. Considérons par exemple un jeu \mathcal{G} et son espace d'états \mathcal{S} . Si \mathcal{G} contient n entités visibles, \mathcal{S} pourrait être un sous-ensemble d'un certain produit $\prod_{i=1}^n \mathcal{S}_i$, où chaque \mathcal{S}_i serait l'espace d'états d'une entité E_i — ce serait un sous-ensemble strict car, en général, deux entités distinctes ne peuvent pas être au même endroit au même moment. Nous pouvons aussi associer à \mathcal{G} l'ensemble de ses entités $\mathcal{E}_1 = \{E_1, E_2, \dots, E_n\}$, l'ensemble des paires de ses entités $\mathcal{E}_2 = \{\{E_1, E_2\}; \{E_1, E_3\} \dots \{E_{n-1}, E_n\}\}$. Nous pourrions alors décrire l'état de \mathcal{G} avec l'information contenue dans ces deux ensembles \mathcal{E}_1 et \mathcal{E}_2 . L'état d'un autre jeu \mathcal{G}' serait décrit par l'ensemble de ses entités $\mathcal{E}'_1 = \{E'_1, E'_2, \dots, E'_{n'}\}$ et l'ensemble des paires de ses entités $\mathcal{E}'_2 = \{\{E'_1, E'_2\}; \{E'_1, E'_3\} \dots \{E'_{n'-1}, E'_{n'}\}\}$, n et n' pouvant être égaux ou différents.

Nous voulons comparer $(\mathcal{E}_1, \mathcal{E}_2)$ et $(\mathcal{E}'_1, \mathcal{E}'_2)$. Si $n = n'$, nous pouvons définir des bijections $\phi : \mathcal{E}_1 \rightarrow \mathcal{E}'_1$ et $\psi : \mathcal{E}_2 \rightarrow \mathcal{E}'_2$. Nous sommes alors dans le cas bijectif. Mais si $n \neq n'$ nous pouvons définir des fonctions $\phi : \mathcal{E}_1 \rightarrow \mathcal{E}'_1$ et $\psi : \mathcal{E}_2 \rightarrow \mathcal{E}'_2$ qui sont des équivalences de catégories. Pour ce faire, comme expliqué précédemment, nous définissons des flèches inversibles entre les éléments de $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}'_1$ et \mathcal{E}'_2 , transformant ces quatre ensembles en catégories. Ces flèches sont définies en utilisant des fonctions d'observation $f_i : \mathcal{E}_i \rightarrow \mathcal{O}_i$ ($i = 1, 2$)

et $f'_i : \mathcal{E}'_i \rightarrow \mathcal{O}'_i$ ($i = 1, 2$), qui définissent des relations d'équivalence $\mathcal{R}_{f_1}, \mathcal{R}_{f_2}, \mathcal{R}_{f'_1}$ et $\mathcal{R}_{f'_2}$ sur les ensembles $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}'_1$ et \mathcal{E}'_2 respectivement. Nous pouvons alors en déduire des bijections $\Phi : (\mathcal{E}_1/\mathcal{R}_{f_1}) \rightarrow (\mathcal{E}'_1/\mathcal{R}_{f'_1})$ et $\Psi : (\mathcal{E}_2/\mathcal{R}_{f_2}) \rightarrow (\mathcal{E}'_2/\mathcal{R}_{f'_2})$. Φ (resp. Ψ) peuvent être utilisées pour identifier les entités (resp. paires d'entités) entre des états de jeu différents.

Nous avons expérimenté le cas particulier où $f_1 : \mathcal{E}_1 \rightarrow \{0, 1\}$ et $f'_1 : \mathcal{E}'_1 \rightarrow \{0, 1\}$ valent 1 si l'entité est un « Moi » et 0 sinon ; et $f_2 : \mathcal{E}_2 \rightarrow \{0, 1\}$ et $f'_2 : \mathcal{E}'_2 \rightarrow \{0, 1\}$ valent 1 si les trajectoires des deux entités d'une paire s'intersectent, et 0 sinon. Ces fonctions d'observation nous permettent de transférer la connaissance de l'ami dans Breakout — la balle — pour identifier sans apprentissage l'ami dans Pong : la balle.

4 Implémentation et Validation

4.1 Détection et suivi d'entités par analyse d'images

ALE permet de valider cette approche sur des configurations complexes en traitant des signaux directement échantillonnés par l'émulateur. Cette validation de principe repose sur le fait que notre système est capable de détecter puis traquer les entités. Pour vérifier ce prérequis, nous concevons notre algorithme à partir de la librairie open-source OpenCV [21]. Puisque la meilleure description d'une entité correspond à un groupe de pixels identiques ou similaires, la première étape consiste en une détection de contours via un filtre de Sobel sur chacun des trois canaux de l'image en TSV, *i.e.* Teinte, Saturation, Valeur (deuxième image en partant de la gauche en figure 1), ce modèle étant privilégié car plus proche de la perception humaine des couleurs. Ensuite, des boîtes englobantes sont calculées et utilisées en guise d'entités (troisième image de la figure 1).

Pour suivre ces entités, nous les apparions grâce à leur description TSV à des objets identiques rencontrés dans les images passées. De plus, la librairie ORB décrite dans [24] permet aussi de regrouper certains objets similaires, ce qui résout partiellement le problème des objets déformables. Ensuite, dès lors que l'on dispose d'un objet dans l'image présente ainsi que de son « ancêtre » d'une image passée, nous en déduisons son vecteur vitesse, information qui est incluse au résultat de l'algorithme de traitement d'images et apparaît à droite en figure 1.

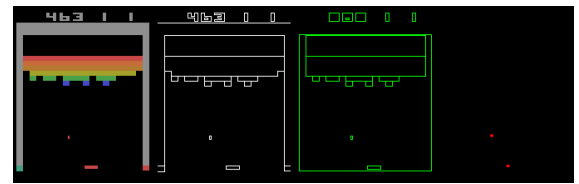


FIGURE 1 – Traitement d'images sur Breakout. De gauche à droite : Image de départ, Contours, Boîtes englobantes, Vitesses mesurées

4.2 Critère de contrôlabilité

Pour trouver l'entité qui est le « Moi » parmi le nombre important d'entités présentes (le nombre d'entités dans le cas des jeux Breakout et Space Invaders est représenté en figures 2 et 3 — l'histogramme serait très similaire pour le jeu Pong), nous utilisons une méthode d'identification de systèmes. La première étape est d'envoyer une séquence d'actions pseudo-aléatoires à ALE pour obtenir les entités affectées par ces signaux, ce qui permet ensuite de construire leur modèle dynamique.

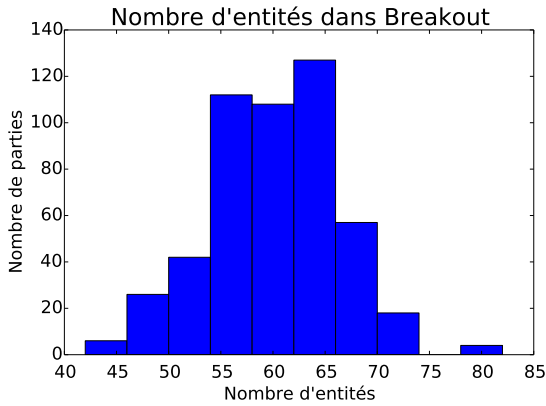


FIGURE 2 – Nombre d'entités détectées dans Breakout

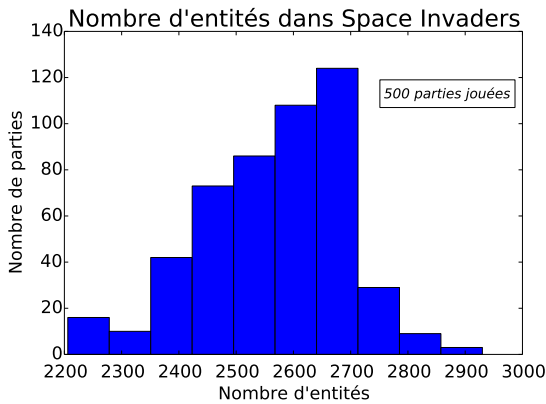


FIGURE 3 – Nombre d'entités détectées dans Space Invaders

Le mouvement de chacune de ces entités est modélisé par les équations d'état discrètes (1) qui correspondent dans notre cas à un système d'ordre $N_d = 2$, où x est le vecteur d'état, y la mesure, u la commande, A la matrice d'état, B la matrice de commande, C la matrice d'observation, D la matrice d'action directe, et k l'indice de l'échantillon.

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned} \quad (1)$$

Pour trouver le « Moi », nous nous restreignons aux entités dont l'erreur entre le modèle dynamique et les mesures est

faible (inférieure à cinq fois la hauteur de l'image). Ce faisant, nous utilisons des résultats de *théorie du contrôle*, à savoir le critère de contrôlabilité d'un système dynamique qui stipule qu'un système est complètement contrôlable ssi la matrice de contrôlabilité $C_{N_d} \triangleq [B|AB|\dots|A^{N_d-1}B]$ est de rang N_d [9].

Cependant, les approximations numériques dans les calculs successifs peuvent créer des faux positifs, tels que des matrices C_{N_d} de rang complet alors qu'un des vecteurs est presque nul (par exemple, la norme d'un des deux vecteurs peut être 10^{20} fois plus grande que celle de l'autre, et il devient absurde de parler de colinéarité pour des questions d'échelle). Pour éliminer ces faux positifs, nous passons par le calcul des déterminants de ces matrices de contrôlabilité. En effet, le déterminant étant un accès direct à l'angle entre les deux vecteurs à travers le produit vectoriel, il permet une évaluation plus fine de la colinéarité de nos vecteurs. Les déterminants les plus élevés sont favorisés. Le « Moi » est choisi en fonction de ce critère, et est le système (l'entité) avec le déterminant le plus élevé.

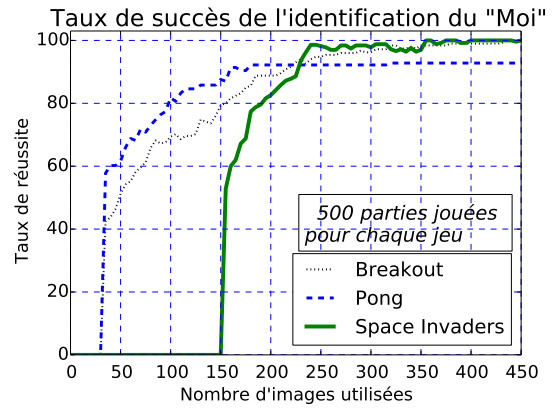


FIGURE 4 – Taux de réussite de l'identification dans trois jeux

Les résultats de cette procédure d'identification sont présentés en figure 4. Pour les trois jeux, notre algorithme cherche à identifier le « Moi » parmi des dizaines (Breakout et Pong) voire des milliers d'entités (Space Invaders) à partir de quelques centaines d'images (soit environ quelques secondes de jeu).

Étant donnée la robustesse de cette méthode et le nombre limité d'images requises, elle peut être appliquée à n'importe quel moment, y compris en cours de partie.

5 Résultats Expérimentaux

5.1 Procédure d'entraînement

L'entraînement de l'agent se fait en 2 étapes. Tout d'abord, il est nécessaire d'identifier le « Moi » parmi toutes les entités, grâce au procédé décrit en partie 4.2. Le système détermine aussi l'existence d'amis ou d'ennemis dans l'environnement en fonction des récompenses qu'ils apportent à l'agent. Dans cet exemple, la seule récompense positive est

l'augmentation du score, et les récompenses négatives sont la diminution du score et la perte de vies.

La stratégie de contrôle du « Moi » implémentée est très basique, et sera améliorée dans des travaux ultérieurs. Le « Moi » est contrôlé de manière à aller vers ses amis, et à fuir ses ennemis.

Les tests sont conduits avec les mêmes paramètres que [20] : l'IA joue des parties de 18 000 images (5 minutes à 60 images/seconde). Nous utilisons le DQN de [20] en guise de référence, car cette publication est l'une des publications sur les jeux Atari les plus citées. Ces travaux ont été reproduits à partir du code mis à disposition par les auteurs. Afin de permettre une comparaison avec notre approche dans des conditions similaires, le score atteint par le DQN suite à un apprentissage fait avec peu d'images d'entraînement est extrait.

5.2 Expériences

Remise en cause des modèles. Comme stipulé en partie 3.2, notre système a été conçu pour être capable de remettre en cause ses propres résultats. C'est pourquoi nous avons inclus une fonctionnalité permettant au système de remettre en cause son identification du « Moi » si celui-ci vient à changer. Par exemple, dans le jeu Breakout, la taille de la raquette diminue la première fois que la balle frappe le mur du haut. Dès lors, l'algorithme détecte que le système auparavant contrôlé est absent, et cherche ce nouveau « Moi » en temps réel par la procédure d'identification présentée en partie 4.2.

Nous évaluons les bénéfices de cette fonctionnalité en comparant le gain de score dû à cette procédure, c'est-à-dire que nous comparons les résultats dans le cas où cette ré-identification est possible et le cas où elle ne l'est pas. Chaque expérience est constituée d'une procédure d'entraînement à nombre d'images donné (entre 0 et 12 000) puis d'un test sur une partie de Breakout.

Effet mémoire. Comme décrit en section 3.3, l'algorithme mémorise les connaissances qu'il génère pour les réutiliser lors des parties successives. La procédure transfère tous les « Moi » rencontrés, ainsi que leurs modèles dynamiques, et transfère aussi la liste des amis et ennemis rencontrés au cours de parties précédentes. Chaque expérience inclut 5 parties consécutives et est conduite indépendamment sans, puis avec cette fonctionnalité.

Transfert de connaissances par analogies entre problèmes. L'humain est capable de reconnaître deux situations similaires en établissant des analogies. Ce faisant, lorsque confronté à une situation inconnue mais similaire à une expérience passée, il est capable d'importer une stratégie qu'il a appliqué auparavant (cf. section 3.4). C'est pourquoi notre système est conçu de sorte à être capable de se remémorer les connaissances acquises dans un premier jeu et de les transposer dans une nouvelle situation. Cette capacité est testée sur les jeux Breakout et Pong puisque dans ces deux jeux, la stratégie peut se réduire à déplacer le « Moi » vers la seule entité amie, à savoir la balle. L'objec-

tif est de transférer la stratégie de Breakout et de l'utiliser efficacement dans Pong sans avoir à entraîner complètement le système. Le système est entraîné à partir de 10 000 images de Breakout, puis 500 images de Pong sont utilisées pour identifier le « Moi » avant de tester l'agent sur Pong, en transférant la connaissance de l'ami à partir de Breakout en utilisant les fonctions d'observation décrites dans la Section 3.4.

5.3 Résultats

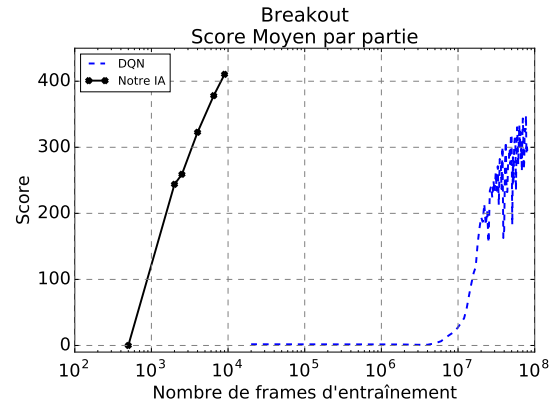


FIGURE 5 – Agent entraîné par DQN vs. notre système. Des scores similaires sont atteints 20 000 fois plus vite

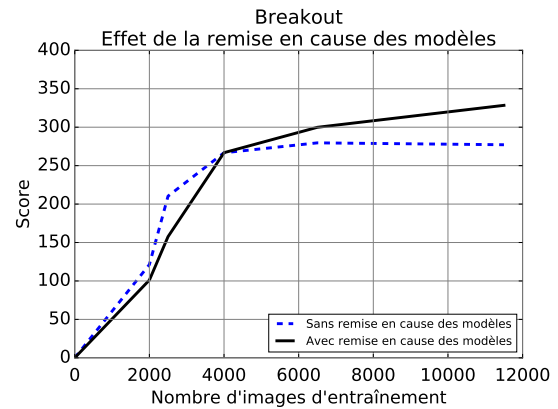


FIGURE 6 – Effet de la remise en cause des modèles. Chaque point (5 par courbe) est une moyenne de 40 expériences

Les résultats² sont présentés dans la table 1 et correspondent à 10 000 images d'entraînement (moins de 3 minutes de jeu). La figure 5 reflète le fait que notre système requiert 20 000 fois moins d'images pour atteindre des scores du même niveau que [20].

Remise en cause des modèles. La figure 6 montre les bénéfices apportés sur le jeu Breakout par une remise en

2. Les résultats du joueur humain et des parties aléatoires sont retranscrits depuis [20]. Le joueur s'entraîne préalablement sur chaque jeu pendant 2 heures (432 000 images).

TABLE 1 – Comparaison des scores obtenus après 10 000 images d’entraînement (moyenne \pm écart-type sur 8 parties)

JEU	ALÉATOIRE	JOUEUR HUMAIN	DQN	IA
Breakout	1.7	31.8	1.25 \pm 1.02	414.37 \pm 88.47
Pong	-20.7	9.3	-21.00 \pm 0.00	0.20 \pm 5.38

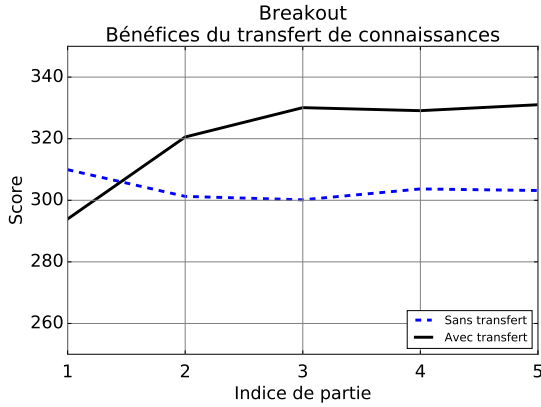


FIGURE 7 – Effet Mémoire. Chaque courbe est une moyenne de 112 expériences

cause des modèles telle que décrite en 5.2. À la fin de la procédure d’entraînement, un meilleur score est atteint si cette remise en cause est possible puisque les modèles sont invalidés dès lors que le « Moi » n’est plus correctement identifié. La procédure de réidentification permet de réactualiser le modèle et de fermer la boucle ouverte. En revanche, sans cette fonctionnalité, l’algorithme joue de façon aléatoire dès que le « Moi » disparaît, ce qui explique la courbe plate au delà de 5000 images d’entraînement. Enfin, pour les entraînements à nombre d’images très réduit, les résultats de l’identification sont trop aléatoires pour avoir une comparaison significative des algorithmes et les scores obtenus sont très similaires (de l’ordre d’une dizaine de points d’écart).

Effet mémoire. La figure 7 reflète l’évolution du score lorsque la connaissance que l’on a du « Moi » est transférée d’une partie à l’autre. Ce transfert de connaissances fait augmenter le score de façon évidente puisque, dès lors qu’une nouvelle forme du « Moi » a été découverte, le système conserve celle-ci en mémoire et reconnaît cette forme instantanément dès qu’elle réapparaît et l’utilise comme « Moi ». Dans l’autre cas, le score est à peu près constant, car des points peuvent être perdus pendant que la réidentification a lieu au cours de la partie.

Transfert de connaissances par analogies entre problèmes. Cette expérience est réalisée selon les paramètres décrits en partie 5.2. Quand l’IA importe la stratégie du jeu Breakout, les tests sur Pong conduisent aux résultats suivants : -2.53 ± 3.79 ce qui demeure très proche des

résultats obtenus lorsque notre système était préalablement entraîné sur Pong (0.20 ± 5.38). Cela prouve que notre cadre devient d’autant plus utile quand il s’agit de transposer des stratégies entre des problèmes analogues, bien que l’on ne dispose d’aucune connaissance du nouveau problème. En effet, cette expérience montre que l’IA est capable de jouer à ces deux jeux à partir de 10 000 images de Breakout et 500 images de Pong (10 500 images au total) aussi bien que si elle s’était entraînée séparément sur 10 000 images de chaque jeu (20 000 images au total).

6 Conclusion

Ces travaux introduisent un cadre basé sur des concepts élémentaires de la théorie des catégories qui permet de réduire de manière significative la quantité de données nécessaires pour entraîner et contrôler un agent sur Atari 2600. Pour ce faire, il a semblé pertinent d’échantillonner nos données à une échelle mésoscopique entre le pixel et l’image entière. Par ailleurs, il a fallu construire un modèle d’entités qui sont détectées puis suivies grâce à des outils usuels de traitement du signal. Nos expériences montrent qu’il est possible d’atteindre des résultats similaires à l’état de l’art avec seulement quelques minutes de jeu en guise d’entraînement. De plus, ce cadre peut aussi servir de base à des méthodes d’apprentissage par transfert. En effet, nos autres expériences ont mis en exergue la possibilité de transférer une stratégie d’un jeu à un jeu analogue, puis de commencer à jouer avec seulement quelques secondes de transition.

Les prochains travaux viseront à atteindre une meilleure compréhension de l’environnement par l’agent à partir des prédictions obtenues des modèles, ainsi qu’à développer un panel de stratégies applicables, puisqu’à ce jour seule une stratégie de survie basique a été utilisée.

Ces résultats théoriques, validés sur un environnement de simulation de jeu, pourront être appliqués dans l’industrie à des problèmes pour lesquels l’acquisition de données, et le temps ou la puissance de calcul requis sont des facteurs limitants, empêchant l’application des méthodes classiques en Intelligence Artificielle tels que les réseaux de neurones.

References

- [1] A. Barreto, W. Dabney, R. Munos, J. Hunt, T. Schaul, D. Silver, and H. van Hasselt, “Successor features for transfer in reinforcement learning,” in *Advances in Neural Information Processing Systems 30*, 2017, pp. 4058–4068.

- [2] M. G. Bellemare, Y. Naddaf, J. Veness, and M. Bowling, “The arcade learning environment: An evaluation platform for general agents,” *Journal of Artificial Intelligence Research*, vol. 47, pp. 253–279, June 2013.
- [3] M. Bogdanovic, D. Markovikj, M. Denil, and N. de Freitas, “Deep apprenticeship learning for playing video games,” in *AAAI Workshop on Learning for General Competency in Video Games*, 2015.
- [4] M. Campbell, A. J. Hoane, and F. Hsu, “Deep blue,” *Artificial Intelligence*, vol. 134, no. 1, pp. 57 – 83, 2002.
- [5] C. Chang and H. J. Keisler, *Model Theory*, 3rd ed. North Holland, 1990.
- [6] F. de Saussure, *Cours de linguistique générale*. Payot, 1916.
- [7] L. Fei-fei, “Knowledge transfer in learning to recognize visual object classes,” in *In: International Conference on Development and Learning (ICDL)*, 2006.
- [8] H. v. Hasselt, A. Guez, and D. Silver, “Deep reinforcement learning with double q-learning,” in *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, ser. AAAI’16. AAAI Press, 2016, pp. 2094–2100.
- [9] R. E. Kalman, “On the general theory of control systems,” in *Proceedings of the First International Congress on Automatic Control*. Butterworth, London, 1960, pp. 481–492.
- [10] K. Kansky, T. Silver, D. A. Mély, M. Eldawy, M. Lázaro-Gredilla, X. Lou, N. Dorfman, S. Sidor, D. S. Phoenix, and D. George, “Schema networks: Zero-shot transfer with a generative causal model of intuitive physics,” in *ICML 2017*, 8 2017, pp. 1809–1818.
- [11] G. Lee, M. Luo, F. Zambetta, and X. Li, “Learning a Super Mario controller from examples of human play,” in *2014 IEEE Congress on Evolutionary Computation (CEC)*, July 2014, pp. 1–8.
- [12] N. Levine, T. Zahavy, D. Mankowitz, A. Tamar, and S. Mannor, “Shallow updates for deep reinforcement learning,” in *Advances in Neural Information Processing Systems 30*, 2017, pp. 3138–3148.
- [13] Y. Liang, M. C. Machado, E. Talvitie, and M. Bowling, “State of the art control of Atari games using shallow reinforcement learning,” in *AAMAS ’16*, 2016, pp. 485–493.
- [14] Y. Lin, Z. Hong, Y. Liao, M. Shih, M. Liu, and M. Sun, “Tactics of adversarial attack on deep reinforcement learning agents,” in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, 2017, pp. 3756–3762.
- [15] N. Lipovetzky, M. Ramirez, and H. Geffner, “Classical planning with simulators: Results on the Atari video games,” in *IJCAI’15*, 2015, pp. 1610–1616.
- [16] Z. Luo, Y. Zou, J. Hoffman, and L. Fei-Fei, “Label efficient learning of transferable representations across domains and tasks,” in *Advances in Neural Information Processing Systems 30*, 2017, pp. 164–176.
- [17] S. Mac Lane, *Categories for the working mathematician*, 2nd ed. Springer-Verlag, 1998.
- [18] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu, “Asynchronous methods for deep reinforcement learning,” in *Proceedings of The 33rd ICML*, ser. Proceedings of Machine Learning Research, vol. 48. New York, New York, USA: PMLR, 20–22 Jun 2016, pp. 1928–1937.
- [19] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, “Playing Atari with deep reinforcement learning,” in *NIPS Deep Learning Workshop*, 2013.
- [20] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, “Human-level control through deep reinforcement learning,” *Nature*, vol. 518, no. 7540, pp. 529–533, February 2015.
- [21] OpenCV, “Open source computer vision library,” <https://opencv.org/>, 2017.
- [22] T. Pepels, M. H. M. Winands, and M. Lanctot, “Real-Time Monte-Carlo Tree Search in Ms. Pac-Man,” *IEEE Trans. on Computational Intelligence and AI in Games*, vol. 6, no. 3, pp. 245–257, Sept 2014.
- [23] K. Popper, *The Logic of Scientific Discovery*. Hutchinson & Co, London, 1959.
- [24] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, “Orb: An efficient alternative to sift or surf,” in *2011 International Conference on Computer Vision*, Nov 2011, pp. 2564–2571.
- [25] S. Sabour, N. Frosst, and G. Hinton, “Dynamic routing between capsules,” in *Advances in Neural Information Processing Systems 30*, 2017, pp. 3859–3869.
- [26] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel, and D. Hassabis, “Mastering the game of Go without human knowledge,” *Nature*, vol. 550, pp. 354–359, October 2017.

FIGURE 8 – Algorithme utilisé.

```

Require:  $bdc = [entites, modeles, strategies]$ 
Require:  $strategie\_par\_defaut = actions\_pseudo\_aleatoires$ 
Require:  $seuil, N_{max}$ 
1: for  $N_{init} = 0$  to  $N_{max}$  do
2:    $entites.append(boites\_englobantes)$ 
3:    $action \leftarrow strategie\_par\_defaut$ 
4:    $applique(action)$ 
5: end for
6:  $jeu\_similaire \leftarrow$  recherche jeu avec des  $entites$  similaires
7:  $modeles, strategies \leftarrow$   $extraire(bdc, jeu\_similaire)$ 
8:  $bdc.append([entites, modeles, strategies])$ 
9:  $f\_refute\_modeles \leftarrow$  false
10: if  $modeles$  est vide then
11:    $f\_refute\_modeles \leftarrow$  true
12: end if
13: while experimentation en cours do
14:    $entites.append(boites\_englobantes)$ 
15:   if  $f\_refute\_modeles$  then
16:      $action \leftarrow strategie\_par\_defaut$ 
17:      $modeles, strategies, f\_id \leftarrow$   $identification(entites)$ 
18:     if  $f\_id$  then
19:        $f\_refute\_modeles \leftarrow$  false
20:        $bdc.append([entites, modeles, strategies])$ 
21:     end if
22:   else
23:      $action \leftarrow$   $cherche\_action(entites, strategies)$ 
24:      $predictions \leftarrow$   $cherche\_predictions(entites, modeles)$ 
25:      $erreur \leftarrow predictions - entites$ 
26:     if  $|erreur| > seuil$  then
27:        $f\_refute\_modeles \leftarrow$  true
28:     end if
29:   end if
30:    $applique(action)$ 
31: end while
32: return  $bdc$ 

```

Annexe 1 : Algorithme

L'algorithme utilisé sur les jeux Atari 2600 est décrit en pseudo-code sur la figure 8. Il repose sur une base de connaissance bdc , une stratégie par défaut $strategie_par_defaut$, un seuil $seuil$ de rejet du modèle et un nombre d'itérations N_{max} correspondant à la durée de l'initialisation.

La base bdc , initialement vide, se remplit progressivement des stratégies et modèles obtenus au cours des parties jouées. La stratégie par défaut $strategie_par_defaut$ est une séquence pseudo-aléatoire d'actions envoyées en entrée du système.

L'algorithme commence par acquérir l'image brute provenant de l'émulateur ALE puis en déduit (en utilisant le procédé décrit en Section 4.1) des boîtes englobantes correspondant à des types d'entités en ligne 2. À partir de ces entités, l'algorithme détermine la situation qui présente les objets les plus similaires en ligne 6 avant d'extraire en ligne

7 les modèles et stratégies obtenus lors de cette situation. La base de connaissances bdc est ensuite complétée.

Lors de la partie, le système continue à acquérir des échantillons et à les transformer en entités. Si le modèle est valide (lignes 23 à 27), le système calcule à partir des entités présentes et stratégies disponibles la meilleure action à effectuer. Avec les entités et les modèles, une prédiction de l'état de l'environnement est aussi calculée. Si ces prédictions sont trop éloignées de l'état mesuré, alors la variable f_refute_savoir devient vraie. Dans ce cas, dès l'itération suivante, la procédure d'identification (lignes 15 à 20) se lance, et de nouveaux modèles et stratégies sont calculés puis appliqués dès l'itération suivante.

Notons que l'algorithme est identique en phase d'apprentissage et en phase d'exploitation (test). Il continue de réfuter son savoir, si besoin est, même en phase d'exploitation.

Annexe 2 : Illustration du transfert de stratégie grâce à la théorie des catégories

Considérons un jeu \mathcal{C} (par exemple le jeu Breakout pour Atari 2600) où une raquette horizontale est située en x_0 sur un segment $[0, N]$ d'états initiaux possibles ($0 \leq x_0 \leq N$) et peut se déplacer vers la gauche ou la droite afin d'atteindre une balle devant arriver à la position $x_1 \in [0, N]$. La stratégie π est la suivante : si $x_1 < x_0$, se déplacer vers la gauche, si $x_1 = x_0$, ne pas bouger, mais si $x_1 > x_0$, alors se déplacer vers la droite. Désormais, considérons un autre jeu \mathcal{C}' (tel que Pong sur Atari 2600 par exemple) où cette fois-ci une raquette verticale se trouve en position y_0 sur le segment $[0, N']$ des positions atteignables ($0 \leq y_0 \leq N'$) et peut se déplacer de haut en bas pour toucher une balle qui atteindra la position $y_1 \in [0, N']$.

Nous souhaitons *identifier* \mathcal{C} et \mathcal{C}' afin de transposer la stratégie π de \mathcal{C} à \mathcal{C}' . Les états (objets) de \mathcal{C} sont les $N + 1$ positions x_1 de $[0, N]$. De la même façon, les états (objets) de \mathcal{C}' sont les $N' + 1$ positions y_1 de $[0, N']$. Si $N = N'$ et $x_0 = y_0$, alors il est facile de définir une bijection de \mathcal{C} à \mathcal{C}' puis de transposer π , mais si $N \neq N'$ ou $x_0 \neq y_0$ alors il est impossible de définir une telle bijection. Pour résoudre ce problème, nous transformons \mathcal{C} et \mathcal{C}' en catégories à travers la définition suivante de flèches, et ce faisant, nous pourrions définir une équivalence de catégories $F : \mathcal{C} \rightarrow \mathcal{C}'$. Pour définir ces flèches, nous utilisons les fonctions d'observation suivantes définies sur les états de \mathcal{C} et \mathcal{C}' : $f : \mathcal{C} \rightarrow \{0, 1, 2\}$ et $f' : \mathcal{C}' \rightarrow \{0, 1, 2\}$ telles que $f(x_1) = 0$ ssi $x_1 < x_0$, $f(x_1) = 1$ ssi $x_1 = x_0$ et $f(x_1) = 2$ ssi $x_1 > x_0$. De la même façon, $f'(y_1) = 0$ ssi $y_1 < y_0$, $f'(y_1) = 1$ ssi $y_1 = y_0$ et $f'(y_1) = 2$ ssi $y_1 > y_0$. Ensuite, nous définissons une flèche inversible entre deux états de \mathcal{C} ssi ils ont la même image par f , ainsi qu'une flèche inversible entre deux états de \mathcal{C}' ssi ils ont la même image par f' . Soit $F : \mathcal{C} \rightarrow \mathcal{C}'$, telle que si $x_1 < x_0$, alors $F(x_1) = 0$; si $x_0 = x_1$, alors $F(x_1) = y_0$; et si $x_1 > x_0$, alors $F(x_1) = N'$. Cette fonction F définit une équivalence de catégories et induit une bijection F' entre les ensembles de classes : $F' : (\mathcal{C} / \simeq) \rightarrow (\mathcal{C}' / \simeq)$. (\mathcal{C} / \simeq) correspond (grâce à la stratégie π) aux actions disponibles $\{Gauche, Attendre, Droite\}$ dans \mathcal{C} alors que (\mathcal{C}' / \simeq) correspond aux actions disponibles $\{Haut, Attendre, Bas\}$ dans \mathcal{C}' (en choisissant d'orienter l'axe des y du haut vers le bas). Dans ce cas-ci $F'(Gauche) = Haut$, $F'(Attendre) = Attendre$ et $F'(Droite) = Bas$.

En reproduisant ces raisonnements, la théorie des catégories permet de transférer rigoureusement des stratégies d'un problème à un problème analogue.

Planifier l'épreuve E5 à l'aide d'un solveur SAT

T. Delacroix¹

H. Sadighyan²

¹ IMT Atlantique - Dépt. LUSI, Brest, France

² Lycée Jacquard, Paris, France

thomas.delacroix@imt-atlantique.fr

Résumé

Cet article propose une modélisation à l'aide d'un problème SAT d'un problème spécifique de définition d'un planning des passages d'étudiants en examen (épreuve E5 en BTS-SN) ainsi qu'un lien vers une implémentation en Python prête à l'emploi par les enseignants de la filière concernée permettant de résoudre le problème. La démarche est présentée de manière détaillée afin de permettre une adaptation aisée à d'autres contextes.

Mots Clef

emploi du temps, SAT, PicoSAT, BTS.

Abstract

This article addresses a specific timetabling problem for student exam schedules (E5 exam for technical university students enrolled in a French BTS-SN curriculum) by modeling it thanks to a SAT problem and providing a link to a ready-to-use Python implementation to solve this problem. The detailed presentation of the process makes it easily transposable to other situations.

Keywords

timetabling, SAT, PicoSAT, BTS.

1 Introduction

La gestion des emplois du temps est, sans aucun doute, l'un des grands casse-têtes administratifs des établissements scolaires et universitaires. En particulier, la gestion d'un planning pour les passages à l'oral par groupes des étudiants peut donner du sacré fil à retordre aux personnels administratifs et aux enseignants. Ainsi, il est traditionnel qu'en classe préparatoire, on demande à l'enseignant de mathématiques de préparer le colloscope (i.e. le planning des oraux) de la classe car il faut au moins être mathématicien pour intégrer l'ensemble des contraintes logiques à la solution du problème ! À l'ère de l'intelligence artificielle, les enseignants seraient en droit de se demander pourquoi les scientifiques n'ont toujours pas trouvé de solution pour les débarrasser de ces problèmes.

Et pourtant, il s'agit tout simplement d'un gros malentendu. Car, bien au contraire, les chercheurs ont été par-

ticulièrement prolifiques dans le domaine de l'automatisation des emplois du temps (sûrement puisqu'ils sont confrontés à ce type de problèmes dans leur quotidien d'enseignants-chercheurs). À titre d'exemple, les états de l'art dans [17] et, plus récemment, dans [16] renvoient chacun à plus d'une centaine de publications dans ce domaine, dont les premières datent du début des années 60. Par ailleurs, la conférence internationale PATAT (Practice and Theory of Automated Timetabling <http://www.patatconference.org/>) réunit tous les deux ans de nombreux chercheurs autour de cette thématique. Alors, pourquoi tant de professeurs de mathématiques en classes préparatoires s'embêtent-ils tous les ans à refaire à la main leurs colloscopes ? Pourquoi les professeurs en BTS Systèmes numériques s'embêtent-ils tous les ans à refaire à la main les plannings des passages des étudiants pour les épreuves d'intervention sur système ?

Sans trop élaborer sur tous les obstacles à dépasser avant que le savoir d'une communauté scientifique soit mis en application au sein de la société de manière générale, il convient tout de même de faire quelques constats.

1. La diversité des contextes amène une diversité des solutions. Or ce n'est pas toujours évident de faire coller son propre problème à une méthodologie conçue pour en résoudre un autre.
2. La recherche scientifique se concentre sur des méthodes à la pointe, permettant de traiter les problèmes les plus complexes ou de les résoudre le plus rapidement possible. Mais celles-ci ne sont pas forcément les plus simples à implémenter ou accessibles à l'utilisateur.
3. L'utilisateur n'est pas forcément en capacité de transformer un algorithme en programme informatique, ni de se plonger dans un code de plusieurs milliers de lignes.

Ces différents constats doivent être pris en compte si l'on souhaite proposer un outil de résolution qui puisse être adopté localement, comme celui que l'on propose aux enseignants de BTS Systèmes numériques dans cet article. Le premier constat incite à opter pour une méthode de résolution suffisamment flexible, où l'utilisateur peut définir chacune des contraintes de son problème. Cela signifie en

pratique que les contraintes doivent tout de même être implémentées directement par l'utilisateur. Le second constat oriente vers une méthode qui soit conceptuellement compréhensible pour l'utilisateur. Enfin, le dernier constat entraîne le choix d'une méthode dont les algorithmes principaux ont déjà été implémentés dans les langages de programmation courants et sont librement disponibles. Ainsi, les algorithmes de résolution peuvent demeurer une boîte noire pour l'utilisateur qui se limite à implémenter les contraintes de son problème.

Ces considérations nous ont amené à choisir, dans le cadre d'un problème de planning des passages des étudiants en épreuve E5 d'une classe de BTS Systèmes numériques, une méthode de résolution s'appuyant sur un solveur SAT [4, 18] implémentée en Python3. Ce type de solveur est largement assez performant pour apporter des solutions acceptables aux problèmes posés ici si l'on se limite aux effectifs d'une seule classe. Quant au choix du langage de programmation, il a été fait en cohérence avec l'importance croissante que ce langage a pris dans les lycées depuis l'introduction d'un enseignement d'algorithmique au programme de seconde en 2009 [7] et celle du langage Python au programme des classes préparatoires scientifiques en 2013 [6] et qui a conduit de fait à une plus grande maîtrise du langage au sein du corps enseignant.

1.1 Travaux connexes

Comme cela a été mentionné en introduction, les problèmes d'emplois du temps scolaires et universitaires ont fait l'objet de nombreux travaux scientifiques mais le nombre de travaux abordant ces problèmes sous la forme de problèmes SAT est relativement limité. Dans [17], l'auteur fait mention d'un possible axe de recherche à venir sur l'utilisation de techniques de résolution de problèmes SAT, à condition de modéliser le problème du *school timetabling* par un problème SAT. Toutefois, près de 10 années plus tard, il n'est fait aucune mention de telles démarches dans [16] qui est un état de l'art pourtant très complet. En contraste, on y retrouve de nombreux travaux impliquant la résolution de problèmes CSP (Problèmes de satisfaction de contraintes), une approche proche mais bien distincte de la résolution de problèmes SAT [22].

Plus récemment, des articles ont, tout de même, abordé les problèmes de *timetabling* via le prisme des problèmes SAT [2] ou SMT [10], qui en est une généralisation. On peut également noter une approche SAT du *social golfer problem* qui n'est pas très éloigné des problèmes de gestion d'emplois du temps académiques [12, 21].

Enfin, il faut noter que le présent article se distingue de ces travaux en tant que l'on a fait le choix de ne pas inscrire la modélisation du problème considéré dans une classe de modélisations de problèmes pouvant être, à leur tour, modélisés par des problèmes SAT mais plutôt de passer directement à la modélisation en tant que problème SAT.

Plus largement, il existe différentes approches dont l'objectif général est apparenté à notre volonté de présenter

une méthodologie qui permette une utilisation accessible d'outils existants pour la résolution de problèmes de programmation sous contraintes. On peut citer dans ce sens le développement du langage MiniZinc et de solutions logicielles complètes associées [15, 20]. Notre approche se distingue d'une telle approche car elle ne nécessite ni l'apprentissage d'un nouveau langage spécifique, ni l'installation de logiciel spécifique, mais est limitée, par contre, aux problèmes SAT. On peut également remarquer qu'une modélisation du présent problème via le langage MiniZinc ne nous a pas semblé être une tâche aisée.

2 Modélisation du problème

2.1 Présentation du contexte

Les étudiants de BTS SN-IR et de BTS SN-EC (Systèmes numériques - Informatique et réseaux ; Électronique et communications) doivent réaliser chaque année une épreuve intitulée E5-Intervention sur un système numérique et d'information [5]. En première année, pour chacun des étudiants, l'épreuve dure 12h découpées en plusieurs séquences. Chacune de ces séquences consiste pour les étudiants en une intervention en équipe sur un système numérique au cours d'une séance d'examen. Une équipe est composée d'un chef d'équipe et de techniciens. Chaque étudiant doit être mis en situation de chef d'équipe pendant 4 heures et en situation de technicien pendant 8 heures. Lors des séances d'examen, plusieurs équipes travaillent en parallèle, chacune sur son système, sous la surveillance d'un enseignant qui évalue les étudiants. Il y a généralement entre 4 et 6 types de systèmes différents proposés et comme chaque système nécessite du matériel spécifique, il n'est possible de faire travailler qu'un nombre très réduit d'équipes sur des systèmes de même type en parallèle.

Par exemple, au lycée Jacquard de Paris (19^e arr.), les systèmes proposés sont DAMALISK (un réseau de téléphonie), WIFIBOT (un robot de surveillance), DMX (un système de gestion de projecteurs dans une salle de spectacle) et SIMRAD (un système de pilotage automatique de bateau). Le matériel disponible permet de faire tourner un système DAMALISK, un système SIMRAD, deux systèmes WIFIBOT et deux systèmes DMX en parallèle. Pour ces systèmes, une intervention requiert une équipe de trois étudiants (un chef d'équipe et deux techniciens) et se fait au cours d'une séance d'examen de 8 heures. Au cours de cette séance de 8 heures, chaque chef d'équipe travaille avec ses techniciens pendant 2 heures, ensuite les techniciens travaillent 4 heures en autonomie, puis le chef d'équipe travaille à nouveau avec les techniciens pendant 2 heures (ces dernières modalités peuvent varier d'un établissement à l'autre, mais cela n'a pas d'incidence sur le modèle proposé).

Ainsi, au cours d'une séance, un étudiant a : soit réalisé ses 4 heures en tant que chef d'équipe ; soit réalisé ses 8 heures en tant que technicien. Deux séances suffisent donc pour évaluer un étudiant. Toutefois, comme chaque intervention sur un système nécessite que deux techniciens y

consacrent 8 heures et que chaque étudiant doit passer au moins une fois en tant que chef d'équipe, il est demandé à chaque étudiant de passer deux fois en tant que technicien sur un système avec une évaluation sur un seul de ces passages. Ainsi, un étudiant doit travailler pendant trois séances, chacune correspondant à une équipe travaillant sur un système.

Les enseignants doivent donc proposer un planning pour des séances d'examen, en définissant les systèmes sur lesquels des équipes travailleront tout au long de cette séance et la composition de ces équipes, de manière à respecter les contraintes décrites précédemment. De plus, même si cela n'est pas imposé au niveau national, il paraît préférable qu'à chaque passage d'un étudiant, celui-ci travaille avec des étudiants différents et sur un système différent.

La tâche que le logiciel proposé dans la suite de l'article permet d'accomplir est la suivante :

Étant donné un planning des séances d'examen et des systèmes à traiter au cours de chacune de ces séances, renvoyer, s'il en existe une, une répartition des étudiants en équipes d'intervention sur ces systèmes de manière à ce que :

1. chaque étudiant intervient trois fois exactement ;
2. chaque équipe est composée de trois étudiants exactement ;
3. un étudiant ne peut intervenir dans plus d'une équipe par séance ;
4. pour chaque étudiant, chaque intervention correspond à un système différent ;
5. chaque étudiant ne peut collaborer qu'une seule fois avec un autre étudiant ;
6. chaque équipe possède exactement un chef d'équipe ;
7. chaque étudiant doit passer une fois chef d'équipe.

Le tableau 1 ci-contre donne un planning des systèmes par séances d'examen, et la répartition qui a été déterminée par le logiciel en fonction de ce planning. Chaque ligne correspond à un étudiant et chaque colonne à une séance. On indique qu'un étudiant est intervenu en tant que technicien ou en tant que chef d'équipe par les abréviations T et CE respectivement.

2.2 Résolution SAT

Le logiciel est construit autour d'une méthode de résolution d'un problème SAT. Sans rentrer dans les détails, en voici le principe.

Étant donnée une expression propositionnelle, il s'agit pour le problème de satisfiabilité (SAT) d'attribuer des valeurs de vérité aux variables propositionnelles de manière à ce que l'expression propositionnelle soit vraie, lorsque c'est possible, ou d'indiquer qu'il n'existe pas de telle solution sinon [13].

En pratique, les solveurs SAT sont généralement de type CNF-SAT (mis à part quelques rares exemples tel que

	S_1	S_2	S_3	S_4
DMX	2	2	1	1
WIFIBOT (WIF)	2	2	1	1
DAMALISK (DAM)	1	1	1	1
SIMRAD (SIM)	1	1	1	1

	S_1	S_2	S_3	S_4
E_1	T-WIF#2	T-SIM#2	CE-DMX#5	
E_2	CE-DAM#1	T-WIF#4		T-SIM#4
E_3	T-WIF#1	CE-DMX#4		T-DAM#4
E_4	CE-DMX#2	T-DAM#2		T-WIF#6
E_5	CE-WIF#2	T-DMX#4	T-SIM#3	
E_6	T-DMX#2	T-WIF#3		CE-SIM#4
E_7		T-DMX#4	CE-WIF#5	T-SIM#4
E_8		T-DMX#3	CE-SIM#3	T-DAM#4
E_9	CE-DMX#1	T-WIF#3	T-SIM#3	
E_{10}	T-DMX#2		T-WIF#5	CE-DAM#4
E_{11}	CE-SIM#1	T-DAM#2		T-DMX#6
E_{12}	T-SIM#1	CE-DMX#3		T-WIF#6
E_{13}	T-DAM#1	CE-WIF#3	T-DMX#5	
E_{14}	T-DMX#1	CE-DAM#2	T-WIF#5	
E_{15}	T-DAM#1	T-SIM#2		CE-WIF#6
E_{16}	T-SIM#1	CE-WIF#4	T-DMX#5	
E_{17}	T-WIF#1	CE-SIM#2		T-DMX#6
E_{18}	T-DMX#1	T-WIF#4	CE-DAM#3	
E_{19}	CE-WIF#1	T-DMX#3	T-DAM#3	
E_{20}	T-WIF#2		T-DAM#3	CE-DMX#6

TABLE 1 – Un planning des systèmes par séances et une répartition des étudiants admissible renvoyée par le logiciel

[14]), ce qui signifie que l'expression propositionnelle doit être renseignée sous sa forme normale conjonctive (CNF). En particulier, c'est le cas pour le solveur PicoSAT sur lequel s'appuie la librairie pycosat utilisée dans le logiciel [4, 18].

Les solveurs SAT sont aussi généralement équipés d'un *timeout* qui leur permet de stopper la recherche s'ils n'ont pas pu aboutir à une solution au problème dans un délai raisonnable. C'est le cas de PicoSAT qui permet de fixer une limite au nombre de propagations unitaires qui peuvent être réalisées lors de la recherche (le nombre de propagations unitaires réalisées étant quasi-linéaire au temps d'exécution).

2.3 Base de la modélisation SAT

On commence par poser les bases de la modélisation, puis on modélise chacune des contraintes présentées précédemment.

On remarque que, comme chaque étudiant doit intervenir sur exactement trois systèmes et qu'il y a exactement trois étudiants par équipe, le nombre d'équipes doit être égal au nombre d'étudiants. On note n ce nombre, E_1, \dots, E_n les étudiants et A_1, \dots, A_n les équipes.

Par ailleurs, chaque équipe travaille sur un système d'un certain type au cours d'une séance donnée. On note t le nombre de types de systèmes différents et s le nombre de séances sur lesquelles ils sont répartis. On fixe une partition $T_1 \sqcup \dots \sqcup T_t$ de $\llbracket 1, n \rrbracket$ de manière à ce que A_j est une équipe travaillant sur un système de type l si et seulement

si $j \in T_l$. De même, on fixe une partition $S_1 \sqcup \dots \sqcup S_s$ de $\llbracket 1, n \rrbracket$ de manière à ce que A_j intervient lors de la séance l si et seulement si $j \in S_l$. Pour faire cela, on part du planning des systèmes par séances que l'on représente sous la forme d'une matrice P de taille (t, s) et l'on attribue les indices de chacune des équipes en parcourant la matrice du haut vers le bas puis de droite à gauche. Dans le cas du planning des systèmes par séances donné en exemple précédemment, cela donne la matrice et les partitions ci-après.

$$\begin{array}{ll} T_1 = \{1, 2, 7, 8, 13, 17\} & S_1 = \{1, 2, 3, 4, 5, 6\} \\ T_2 = \{3, 4, 9, 10, 14, 18\} & S_2 = \{7, 8, 9, 10, 11, 12\} \\ T_3 = \{5, 11, 15, 19\} & S_3 = \{13, 14, 15, 16\} \\ T_4 = \{6, 12, 16, 20\} & S_4 = \{17, 18, 19, 20\} \end{array}$$

TABLE 2 – Partitions en types de systèmes et en séances définies pour l'exemple de la Table 1

Enfin, on modélise le fait qu'un étudiant E_i participe ou non à une équipe A_j par la variable propositionnelle $X_{i,j}$ et le fait qu'il en soit le chef d'équipe par la variable propositionnelle $Y_{i,j}$.

2.4 Modélisations des contraintes

On reprend ici une par une les contraintes listées à la section 2.1 en les exprimant en tant que contraintes de cardinalité lorsqu'il y a lieu. La contrainte de cardinalité $\#k(x_1, \dots, x_n)$ signifie que $N\#k$ où N est le nombre de variables évaluées à vrai parmi x_1, \dots, x_n , $\#$ est l'un des symboles $\leq, =, \geq$ et k un entier de $\llbracket 0, n \rrbracket$.

Contrainte (1) : trois équipes par étudiant.

$$\forall i \in \llbracket 1, n \rrbracket, = 3(X_{i,1}, X_{i,2}, \dots, X_{i,n}).$$

Contrainte (2) : trois étudiants par équipe.

$$\forall j \in \llbracket 1, n \rrbracket, = 3(X_{1,j}, X_{2,j}, \dots, X_{n,j}).$$

Contrainte (3) : non dédoublement des étudiants.

$$\forall i \in \llbracket 1, n \rrbracket, \forall k \in \llbracket 1, s \rrbracket, \forall \mathbf{j} \in \mathcal{C}_{S_k}^2, \leq 1(X_{i,j_1}, X_{i,j_2})$$

où $\mathcal{C}_{S_k}^2$ désigne l'ensemble des combinaisons de deux éléments de S_k .

Contrainte (4) : des systèmes différents.

$$\forall i \in \llbracket 1, n \rrbracket, \forall k \in \llbracket 1, t \rrbracket, \forall \mathbf{j} \in \mathcal{C}_{T_k}^2, \leq 1(X_{i,j_1}, X_{i,j_2}).$$

Contrainte (5) : des partenaires différents.

$$\forall i \in \mathcal{C}_n^2, \forall \mathbf{j} \in \mathcal{C}_n^2, \leq 3(X_{i_1,j_1}, X_{i_1,j_2}, X_{i_2,j_1}, X_{i_2,j_2}).$$

Contrainte (6) : un chef d'équipe par équipe.

$$\forall j \in \llbracket 1, n \rrbracket, = 1(Y_{1,j}, Y_{2,j}, \dots, Y_{n,j}).$$

Contrainte (7) : tout étudiant doit passer chef d'équipe.

$$\forall i \in \llbracket 1, n \rrbracket, = 1(Y_{i,1}, Y_{i,2}, \dots, Y_{i,n}).$$

Contrainte (8). On peut remarquer qu'un étudiant ne peut pas passer chef d'une équipe sans être dans cette équipe. On a donc une contrainte supplémentaire.

$$\forall i \in \llbracket 1, n \rrbracket, \forall j \in \llbracket 1, n \rrbracket, Y_{i,j} \implies X_{i,j}.$$

Contrainte assouplie (4'). Comme indiqué dans la présentation du contexte, les contraintes qui ont été considérées ne sont pas toutes imposées au niveau national et on peut essayer de les assouplir pour obtenir des résultats plus facilement. On peut considérer seulement qu'un étudiant ne doit pas intervenir plus de deux fois (au lieu d'une) sur un même système.

$$\forall i \in \llbracket 1, n \rrbracket, \forall k \in \llbracket 1, t \rrbracket, \forall \mathbf{j} \in \mathcal{C}_{T_k}^3, \leq 2(X_{i,j_1}, X_{i,j_2}, X_{i,j_3}).$$

2.5 Encodage CNF performant des contraintes de cardinalité

Hormi, la contrainte (8), toutes les contraintes ici correspondent à des contraintes de cardinalité de type $\#k(x_1, \dots, x_n)$. Ces contraintes peuvent être exprimées directement sous forme CNF via une approche dite naïve mais cela pose rapidement des problèmes de complexité dès que n augmente car le nombre de clauses de ces contraintes est prohibitif. Il existe dans la littérature de nombreux encodages plus performants de ce type de contraintes qui permettent d'obtenir une forme CNF équi-satisfiable comportant un nombre très réduit de clauses en rajoutant un nombre raisonnable de variables supplémentaires [1, 3, 11, 19]. Le logiciel proposé utilise de tels encodages en suivant le protocole défini dans [8].

2.6 Choix du planning des systèmes par séances

Jusqu'à présent dans cette section, il n'a été question que de la recherche d'une solution à un problème SAT défini par la matrice P du planning des systèmes par séances, cette matrice étant définie a priori par l'enseignant. La question du choix de ce planning est également une question valable et il peut être envisagé d'automatiser ce processus.

En effet, on commence par faire la remarque suivante : si le problème SAT défini par une matrice P admet une solution alors la matrice vérifie nécessairement les trois contraintes suivantes :

- la somme des coefficients de P vaut n ;
- pour chacune des colonnes de P , la somme des coefficients de cette colonne vaut au plus $n/3$ (sinon il faudrait qu'un étudiant soit présent dans deux équipes au cours de la même session, ce qui est proscrit) ;
- pour chacune des lignes de P , la somme des coefficients de cette ligne vaut au plus $n/3$ (sinon il faudrait qu'un étudiant intervienne sur deux systèmes du même type, ce qui est également proscrit).

Par ailleurs, on rappelle que, pour chaque type de système, il y a une borne sur le nombre de systèmes de ce type pouvant fonctionner simultanément. On note $R = (R_1, \dots, R_t)$ ces restrictions.

Enfin le nombre de systèmes est fixé et on peut fixer le nombre de séances possibles (ou une borne maximale si l'on préfère).

Ainsi, en ne considérant que les matrices qui vérifient l'ensemble de ces contraintes, on limite sérieusement les possibilités. Pour le cas présenté en exemple précédemment, c'est-à-dire 20 étudiants, travaillant pendant au plus 4 séances sur les 4 systèmes du lycée Jacquard avec les restrictions associées sur le fonctionnement en parallèle de ces systèmes (i.e. $R = (2, 2, 1, 1)$), il existe exactement 100 matrices satisfaisant ces contraintes. Mathématiquement, c'est l'ensemble $\mathcal{P}(n, s, t, R)$ des matrices $P \in \text{Mat}_{t,s}(\mathbb{N})$ telles que :

$$\left\{ \begin{array}{l} \sum_{i=1}^t \sum_{j=1}^s P_{i,j} = n, \\ \forall j \in \llbracket 1, s \rrbracket, \sum_{i=1}^t P_{i,j} \leq \frac{n}{3}, \\ \forall i \in \llbracket 1, t \rrbracket, \sum_{j=1}^s P_{i,j} \leq \frac{n}{3}, \\ \forall i \in \llbracket 1, t \rrbracket, \forall j \in \llbracket 1, s \rrbracket, P_{i,j} \leq R_i. \end{array} \right. \quad (1)$$

Un tel ensemble peut facilement être généré grâce à un algorithme récursif [9]. Comme cet ensemble est fini et calculable en un temps très raisonnable, il est envisageable de l'ordonner totalement par ordre de préférence selon un certain nombre de critères définis par les enseignants (comme regrouper les épreuves sur un minimum de séances ou proposer un minimum de types de systèmes différents). On peut ensuite parcourir cet ensemble ordonné de matrices, jusqu'à la première dont le problème SAT associé possède une solution.

Toutefois, dans la pratique cela n'est pas forcément la meilleure option et ceci pour deux raisons. D'abord, définir une fonction de préférence n'est pas immédiat et les critères de préférence sont tout à fait fluctuants d'un enseignant à l'autre. Ensuite, et c'est une découverte que nous avons réalisée en faisant différents tests sur le logiciel, pour toutes les valeurs testées, le problème SAT associé à chacune des matrices de $\mathcal{P}(n, s, t, R)$ admet une solution. Cette dernière remarque nous amène à poser la conjecture que les contraintes nécessaires énoncées dans cette section seraient en fait suffisantes pour que le problème SAT associée à une matrice P possède une solution. Si tel est le cas, il ne sert à rien d'ordonner totalement l'ensemble de ces matrices, il suffit simplement d'en choisir une (la meilleure).

On remarque par ailleurs que les considérations faites dans cette sous-section ne sont valables que si l'on considère les contraintes dans leur version première et ne tiennent plus si l'on considère la contrainte assouplie (4').

n	18	19	20	21	22	23	24
s_{min}	3	4	4	4	4	5	4
$ \mathcal{P} $	1	1120	100	208	16	11700	1

TABLE 3 – Seuils minimaux s_{min} et cardinaux des ensembles de matrices potentielles en fonction de n

3 Implémentation et résultats

3.1 Implémentation Python

L'implémentation proposée transpose directement ce qui a été décrit précédemment et permet de créer un fichier CSV contenant la solution si elle a été trouvée. À ce stade, la solution est calculée uniquement à partir d'une matrice P rentrée par l'utilisateur. Nous intégrerons également, à terme, une option qui permettra au logiciel de fournir lui-même une telle matrice. Le code source en Python3 est mis à disposition librement sous licence GNU [9].

3.2 Le cas d'une classe : $18 \leq n \leq 24$

Étant donné qu'une classe de BTS Systèmes numériques a un effectif situé entre 18 et 24 étudiants, nous avons déterminé pour chacun des effectifs possibles toutes les matrices vérifiant les conditions nécessaires définies au niveau de la sous-section 2.6 (en respectant les contraintes spécifiques sur les systèmes du lycée Jacquard de Paris) de taille (t, s) où s a été choisi comme étant le seuil minimum s_{min} à partir duquel il existe au moins une telle matrice.

Après vérification sur l'ensemble de ces matrices, chacune correspond à un problème SAT admettant une solution d'où notre conjecture que les conditions nécessaires de la sous-section 2.6 sont suffisantes. Nous n'avons toutefois pas cherché à démontrer mathématiquement cette conjecture. Les deux tableaux suivants permettent de synthétiser les résultats obtenus.

Nous avons également cherché à voir s'il était possible de réduire le nombre de séances en assouplissant les contraintes, puisqu'il peut être compliqué pour les enseignants d'organiser de nombreuses séances. Étant donné qu'il ne peut y avoir plus de 6 systèmes par séance, il est clair que l'effectif maximal pour une valeur fixe de s ne peut être supérieur à $6s$. On peut donc en déduire qu'un assouplissement des contraintes ne permettra pas de diminuer le nombre de séances, sauf éventuellement pour $n = 23$. Nous avons donc effectué des tests en remplaçant la contrainte (4) par sa version assouplie (4') et vérifié que, dans ce cas, la matrice P suivante permettait d'obtenir une solution au problème SAT assoupli.

$$P = \begin{bmatrix} 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

On précise également que le temps d'exécution est très satisfaisant par rapport à l'utilisation qui est attendue du logiciel. En effet, pour chacun des différents problèmes SAT

considérés, le temps d'obtention d'une solution est inférieur à la seconde, sachant que les calculs ont été réalisés sur un ordinateur portable bas de gamme au regard des standards actuels (processeur i5-2540M CPU @ 2.60GHz $\times 4$ et mémoire vive de 3.7Go).

3.3 Gérer la complexité pour des plus grandes valeurs de n

Une analyse de l'algorithme proposé permet d'approcher le nombre de variables de la modélisation par $14n^2$ et le nombre de clauses par $\frac{1}{4}n^4 - \frac{1}{2}n^3 + \frac{173}{4}n^2 - 20n$. Ces quantités restent raisonnables pour un solveur SAT opérant sur un ordinateur lambda si l'on considère les valeurs de n de la sous-section précédente. Cependant, si l'on considère des valeurs plus importantes comme $n = 80$, cela devient plus compliqué car la mémoire vive de l'ordinateur pourra saturer simplement sur le stockage des contraintes (dans l'implémentation Python, une clause est un tableau de jusqu'à 4 entiers de 28 bytes chacun, l'ensemble des clauses étant lui-même stocké dans un tableau). Même si, dans notre exemple particulier, il n'y a pas autant d'étudiants, ce nombre pourrait être largement dépassé si l'on souhaite organiser le déroulement d'épreuves semblables pour l'ensemble des étudiants d'une même académie.

Une solution consiste à découper le problème en sous-problèmes disjoints. Si l'on peut obtenir une solution à chacun de ces sous-problèmes, alors on peut les regrouper pour obtenir une solution au problème complet. Pour notre cas d'étude, il s'agit de considérer un premier groupe d'étudiant qui tourne sur un premier ensemble de séances, puis un deuxième groupe d'étudiant disjoint du premier qui tourne sur un nouvel ensemble de séances, également disjoint du premier, et ainsi de suite. Un tel découpage n'est pas toujours possible a priori. Toutefois, on peut montrer que, dans le cas précis qui est étudié ici (i.e. 4 systèmes différents et $R = (2, 2, 1, 1)$), c'est non seulement faisable mais que cela permet également d'obtenir une solution avec un nombre de séances minimal dès que $n \geq 18$.

Pour montrer cela, on commence par remarquer que le nombre de séances est minoré par $\lceil n/6 \rceil + 1$ si $n \equiv -1 \pmod 6$ ou par $\lceil n/6 \rceil$ sinon. En effet, les restrictions définies par R limitent à 6 le nombre d'équipes qui peuvent intervenir au cours d'une même séance et, dans le cas où $n \equiv -1 \pmod 6$, on doit nécessairement considérer une séance de plus de façon à ce que la somme des éléments d'une ligne de la matrice P ne dépasse $n/3$. Par ailleurs, nous avons fait tourner le logiciel proposé pour toutes les valeurs de n dans l'intervalle $\llbracket 18, 35 \rrbracket$ en considérant à chaque fois une matrice potentielle P dont le nombre de colonnes (i.e. le nombre de séances) est égal à ce minorant. Dans tous les cas, nous obtenons une solution au problème SAT, ce qui démontre que ce minorant est un minimum pour toutes ces valeurs de n . Il suffit maintenant, pour tout entier $n \geq 18$, de le décomposer sous la forme $n = 18p + r$ avec $r \in \llbracket 18, 35 \rrbracket$ et $p \in \mathbb{N}$. Si l'on considère une solution avec un minimum de séances pour 18 étudiants et une so-

lution avec un minimum de séances pour r étudiants, alors il suffit de recoller p fois la première solution, translatée à chaque fois sur un nouveau groupe de 18 étudiants et 18 équipes, puis une fois la deuxième solution, sur un dernier groupe de r étudiants et r équipes. On obtient alors bien une solution au problème pour n étudiants, avec un nombre de séances égal au minorant décrit précédemment donc, a fortiori, minimal. La complexité pour le solveur SAT est alors constante, il n'est utilisé que deux fois sur des problèmes de taille limitée. Quant au découpage et au recollement, ils représentent une part négligeable de l'utilisation des ressources informatiques pour les valeurs de n que l'on pourra être amené à considérer ici.

On rappelle que la démonstration qui précède ne porte que sur le cas où $R = (2, 2, 1, 1)$. Toutefois, la méthodologie du découpage, même si elle ne permet pas nécessairement d'obtenir une solution optimale, permettra généralement d'obtenir une solution alors même que les valeurs de n sont trop importantes pour une approche directe. Ce processus n'a pas, à ce stade, été implémenté dans le code Python mis à disposition mais devrait y être intégré prochainement.

4 Conclusion

L'objectif premier de cet article est de proposer un logiciel permettant de répondre à un besoin précis des enseignants en BTS Systèmes numériques, qui puisse être utilisé de manière immédiate par un enseignant ayant un usage basique de Python, et permettant de servir de base pour une adaptation à d'autres problèmes par un utilisateur plus expérimenté.

L'objectif second de cet article est d'illustrer, par cet exemple, la relative facilité qu'il peut y avoir à modéliser un problème réel par un problème SAT, à implémenter cette modélisation et à obtenir, ainsi, une solution rapidement à ce problème. Il est vrai que, pour le problème que nous avons considéré, de nombreuses approches sont possibles. Si la conjecture énoncée à la sous-section 2.6 était avérée et qu'il en existait une preuve constructive, alors une approche mathématique permettrait peut-être de développer un algorithme plus performant. Mais cette démarche est incertaine et moins généralisable que l'approche présente. En contraste, il était ici quasi certain, dès l'énonciation du problème, qu'une approche via une modélisation SAT permettrait d'y apporter une réponse en un temps raisonnable. Cette certitude est bien évidemment liée aux faibles dimensions du problème, mais un grand nombre de problèmes ont de faibles dimensions. Le développement de telles approches pourrait aider à la résolution de nombreux problèmes qui prennent bien plus de temps qu'ils ne devraient, à commencer par la gestion administrative de nos enseignements.

Les auteurs de cet article souhaitent remercier l'ensemble des étudiants en première année de BTS Systèmes numériques au lycée Jacquard de Paris qui ont servi de cobayes à cette expérience. Aucun n'étudiant n'a été maltraité durant ce processus.

Références

- [1] R. Asín, R. Nieuwenhuis, A. Oliveras, and E. Rodríguez-Carbonell. Cardinality networks : a theoretical and empirical study. *Constraints*, 16(2) :195–221, 2011.
- [2] R. Asín Achá and R. Nieuwenhuis. Curriculum-based course timetabling with sat and maxsat. *Annals of Operations Research*, pages 1–21, 2012.
- [3] O. Bailleux, Y. Boufkhad, and O. Roussel. New encodings of pseudo-boolean constraints into cnf. In *International Conference on Theory and Applications of Satisfiability Testing*, pages 181–194. Springer, 2009.
- [4] A. Biere. Picosat essentials. *Journal on Satisfiability, Boolean Modeling and Computation*, 4 :75–97, 2008.
- [5] BO. « systèmes numériques » option a « informatique et réseaux », option b « électronique et communications » : définition et conditions de délivrance. *Bulletin Officiel*, 47 du 19 décembre 2013, 2013.
- [6] BO. Programmes des classes préparatoires aux grandes écoles. *Bulletin Officiel*, spécial du 30 mai 2013, 2013.
- [7] DEGESCO. Ressources pour la classe de seconde - algorithmique -, 2009.
- [8] T. Delacroix. Choisir un encodage cnf de contraintes de cardinalité performant pour sat. In *CNIA, Conférence Nationale en Intelligence Artificielle*, 2018.
- [9] T. Delacroix. Sat pour e5. <https://github.com/Tdelacroix/SATpourE5>, 2018.
- [10] E. Demirovic and N. Musliu. Solving high school timetabling with satisfiability modulo theories. pages 142–166, 2014.
- [11] A. M. Frisch and P. A. Giannaros. Sat encodings of the at-most-k constraint. some old, some new, some fast, some slow. In *Proc. of the Tenth Int. Workshop of Constraint Modelling and Reformulation*, 2010.
- [12] I. Gent and I. Lynce. A sat encoding for the social golfer problem. In *IJCAI’05 workshop on modeling and solving problems with constraints*, 2005.
- [13] C. P. Gomes, H. Kautz, A. Sabharwal, and B. Selman. Satisfiability solvers. *Foundations of Artificial Intelligence*, 3 :89–134, 2008.
- [14] R. Muhammad and P. J. Stuckey. A stochastic non-cnf sat solver. In *Pacific Rim International Conference on Artificial Intelligence*, pages 120–129. Springer, 2006.
- [15] N. Nethercote, P. J. Stuckey, R. Becket, S. Brand, G. J. Duck, and G. Tack. Minizinc : Towards a standard cp modelling language. In *Proceedings of the 13th International Conference on Principles and Practice of Constraint Programming*, CP’07, pages 529–543, Berlin, Heidelberg, 2007. Springer-Verlag.
- [16] R. Qu, E. K. Burke, B. McCollum, L. T. Merlot, and S. Y. Lee. A survey of search methodologies and automated system development for examination timetabling. *Journal of scheduling*, 12(1) :55–89, 2009.
- [17] A. Schaerf. A survey of automated timetabling. *Artificial intelligence review*, 13(2) :87–127, 1999.
- [18] I. Schnell. pycosat : bindings to picosat (a sat solver). <https://github.com/ContinuumIO/pycosat>, 2016.
- [19] C. Sinz. Towards an optimal cnf encoding of boolean cardinality constraints. *CP*, 3709 :827–831, 2005.
- [20] P. J. Stuckey, T. Feydy, A. Schutt, G. Tack, and J. Fischer. The minizinc challenge 2008–2013. *AI Magazine*, 35(2) :55–60, 2014.
- [21] M. Triska and N. Musliu. An improved sat formulation for the social golfer problem. *Annals of Operations Research*, 194(1) :427–438, 2012.
- [22] T. Walsh. Sat v csp. *Principles and Practice of Constraint Programming—CP 2000*, pages 441–456, 2000.

Optimisation en temps réel de la gestion des convergences ferroviaires

Valentina Pozzoli¹

François Ramond¹

¹SNCF Innovation et Recherche
40, avenue des Terroirs de France
75611 Paris Cedex

Valentina.Pozzoli@sncf.fr
Francois.Ramond@sncf.fr

Résumé

L'exploitation ferroviaire en zone dense est une tâche complexe. La densité des circulations provoque une propagation très rapide des retards, même de petite amplitude. Afin de prendre des décisions d'exploitation rapides et optimales il est nécessaire de se doter d'outils d'aide à la décision performants.

Dans cet article, nous présentons un système d'aide à la gestion optimale en temps réel des zones de convergences ferroviaires, afin d'améliorer la ponctualité des voyageurs de la ligne. Nous détaillons les enjeux liés au projet, ainsi que les choix de modélisation et l'architecture du prototype. Le système a été testé avec succès sur un site pilote en Île-de-France : nous montrons les résultats de ces expérimentations et donnons les perspectives pour la suite du projet.

Mots Clef

Aide à la décision, Optimisation, Exploitation ferroviaire, Algorithmes en temps réel, Convergences ferroviaires

Abstract

Railway traffic management in mass transit system is a complex task. Trains run so frequently and so close to each other that delays propagate very fast. Operators have to take efficient decisions depending on a multitude of factors, that impact a large number of trains, and thus of passengers. Artificial intelligence methods come in help in the decision process.

In this paper, a real-time decision-support tool for the optimal management of junction areas is presented. It proposes the best order of trains to run through a junction, such that the passenger punctuality is maximized. Details are given on the modelling and the architecture of the tool, as well as on the goals and ambitions of the project. The results of promising experiments on a case-study on a junction in a Parisian area are presented.

Keywords

Decision-support, Optimization, Traffic management, Mass transit, Real-time algorithms, Railway junctions

1 Introduction

L'exploitation ferroviaire en zone dense, où la fréquence des trains est élevée et le nombre de voyageurs important, est un vrai défi dont l'enjeu principal est la qualité du service offert aux voyageurs. L'optimisation de l'exploitation ferroviaire relève de plusieurs facteurs avec des échéances temporelles différentes. En amont de la production se situent, par exemple, la conception de plans de transport robustes et résilients, la maintenance des voies et du matériel, l'affectation des conducteurs et la gestion des rames. En gestion opérationnelle, la mise en œuvre de mesures de régulation efficaces est nécessaire afin d'optimiser la production : changements de la desserte ou des voies utilisées, adaptation d'offre, etc. La gestion des zones de convergence figure parmi ces mesures de régulation.

Les zones de convergences sont des endroits du réseau ferroviaire où deux ou plusieurs voies se réunissent. L'exemple le plus simple, qui sera étudié au cours de cet article, est montré en Figure 1. La gestion des zones de convergence consiste à choisir l'ordre de passage des trains lors que deux trains de deux branches différentes s'apprêtent à rentrer dans la zone de convergence. Bien que la thématique d'optimisation de l'ordre de passage existe pour toutes zones de convergence, en zone dense le choix est particulièrement impactant : en effet, dans un réseau dense les retards se propagent facilement sur les trains suivants. Chaque décision de régulation peut donc avoir un impact sur un grand nombre de trains, et de conséquence sur un nombre important de voyageurs. Ces décisions dépendent de nombreux paramètres : par exemple, les retards des trains, leur desserte, le nombre de passagers dans les trains et sur les quais. En plus, les choix doivent être pris rapidement par le régulateur : pour ces raisons il est nécessaire de se doter d'outils d'aide à la décision performants qui puissent venir en support du travail des régulateurs.

En Ile-de-France, SNCF Mobilités, via son activité Transilien, et RATP sont les principaux opérateurs de transport ferroviaire. Ils opèrent par le compte d'Île-de-France Mobilité, qui définit également les objectifs de performance des opérateurs de transport. Transilien fait circuler 6200 trains chaque jour, qui transportent au quotidien 3 millions de voyageurs, nombre qui connaît une augmentation de 3% par an en moyenne (source Transilien). L'amélioration de la 'ponctualité des

voyageurs', premier élément pour la satisfaction des clients, est l'objectif principal de Transilien et le critère utilisé par Île-de-France Mobilité pour évaluer la performance d'une ligne.

Dans ce contexte, la gestion des zones de convergence, est particulièrement complexe et impactant. En effet, ces zones sont spécialement sollicitées car les trains des deux branches doivent y passer : la fréquence de passage peut aller jusqu'à 32 trains par heure. L'ordonnancement des trains dans une zone de convergence peut être modifié par les opérateurs. Bien que les horaires théoriques indiquent l'ordre de passage des trains à une zone de convergence, cet ordre peut être remis en cause par les aléas : la décision optimale doit considérer l'impact de ce choix sur toute la ligne d'un point de vue des voyageurs concernés.

La Direction de l'Innovation et de la Recherche de SNCF a développé pour Transilien un outil d'aide à la gestion des zones de convergence en temps réel. Cet outil préconise l'ordre de passage des trains à la convergence permettant de maximiser la performance globale de la ligne, définie par le critère de 'ponctualité des voyageurs'.

L'article est organisé de la façon suivante : la Section 2 est consacrée à une présentation succincte des travaux déjà existants sur le sujet ; la Section 3 détaille le critère d'optimisation et les enjeux du projet, la Section 4 décrit la modélisation du problème et l'architecture de l'outil ; la Section 5 présente les résultats de l'expérimentation menée dans un contexte opérationnel sur un site pilote ; enfin, des conclusions et perspectives sont exposées en Section 6.

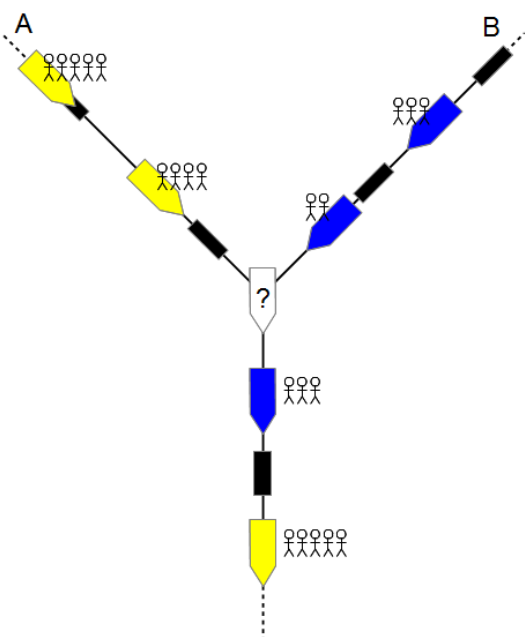


Figure 1 - Schéma d'une zone de convergence

2 Travaux antérieurs

La gestion du trafic en temps réel est un problème largement abordé dans la littérature. Une revue extensive des approches pour réduire les conséquences des incidents est présentée dans [1], en distinguant entre petits aléas et perturbations importantes. Les actions à mettre en place sont différentes dans les deux cas, et incluent la replanification d'itinéraires, la gestion des correspondances, la replanification des engins ou des conducteurs. Au niveau des adaptations du plan de transport en temps réel, les stratégies utilisées en zone dense sont la rétention d'un train en gare, l'ajustement de la vitesse, le changement de desserte, et l'ordre de passage en certains points du réseau. L'ordonnancement des trains à une convergence est un cas particulier de cette dernière stratégie de replanification.

L'article [2] résume les modèles existants dans la planification et replanification ferroviaire, dont les décisions d'ordonnancement à une convergence, ou de replanification d'itinéraires. Une autre revue des modèles de contrôle de trafic en temps réel est proposée par [3].

L'optimisation de l'exploitation du point de vue des voyageurs, et non pas que des retards des trains, est traitée dans plusieurs articles : par exemple dans [4], les horaires d'arrivée et de départ des métros sont choisis afin de minimiser les temps d'atteintes et de trajet, ainsi que la congestion; [5] proposent un algorithme de replanification qui minimise l'insatisfaction des voyageurs.

Plusieurs articles traitent spécifiquement le problème de gestion des convergences ferroviaires : [6] propose un modèle d'ordonnancement optimal des trains aux convergences afin de minimiser une somme pondérée des retards des trains aux gares desservies. Ce modèle est résolu grâce à un algorithme à évolution différentielle et, sur les instances de test, permet de réduire les retards par rapport à une stratégie du type premier entré, premier sorti. Un modèle de replanification en temps réel qui inclut l'optimisation de la vitesse des trains afin de minimiser les retards est proposé dans [7]. Une comparaison entre différentes stratégies de résolution des conflits aux convergences est décrite dans [8]: premier arrivé premier servi, priorité selon le type de train, et minimisation du nombre et de la durée des retards grâce à une heuristique. Les trois scénarios sont évalués grâce à des simulations.

Les travaux présentés dans cet article proposent un approche différente car centré sur une gestion optimale du point de vue de la performance globale de la ligne, en considérant le point de vue des voyageurs, et non pas que le retard des trains.

3 Critère d'optimisation

L'objectif de nos travaux est d'optimiser la gestion d'une zone de convergence afin de maximiser la performance globale de la ligne.

Aujourd'hui, en cas de conflit dans une zone de convergence, la décision d'ordonnement des trains est prise au cas par cas grâce à l'expertise humaine du régulateur, et dépend principalement des retards de deux trains en conflit, ainsi que de leurs missions. En cas de perturbation, le régulateur a une charge de travail élevée et doit prendre des décisions rapides qui ont un impact majeur sur toute la ligne. L'idée est de venir en aide aux régulateurs, en leur fournissant un outil qui puisse à la fois les décharger d'une partie de leur charge cognitive, leur donner des informations pour mieux analyser la situation, et leur proposer le meilleur ordonnancement du point de vue de la performance globale de la ligne.

Un des principaux indicateurs de la performance globale d'une ligne est appelé 'ponctualité voyageurs'. Ce critère mesure le nombre de voyageurs arrivés à leur gare de destination moins de 5 minutes après leur heure d'arrivée théorique. Il est utilisé pour évaluer la qualité de service de SNCF par Île-de-France Mobilités, et c'est le critère d'optimisation choisi dans l'outil d'aide à la gestion des zones de convergence.

Pour calculer la ponctualité voyageurs, des données statistiques sur le nombre de voyageurs pour chaque couple origine-destination par heure sont utilisées afin d'affecter les voyageurs aux différents trains, en se basant sur les horaires théoriques. Les données sur le nombre de voyageurs sont issues de comptages manuels des voyageurs effectués régulièrement sur les lignes concernées. Grâce à ces données et aux horaires théoriques des trains, il est possible de calculer l'heure théorique d'arrivée de chaque passager à la gare d'origine et à la gare de destination. Une fois les trains circulés, les horaires réels d'arrivée des voyageurs à destination sont calculés grâce aux horaires réels de passage des trains par les différentes gares. On considère que chaque voyageur arrive en gare à l'heure théorique prévu, et monte ensuite dans le premier train qui dessert sa destination. Par la suite, on calcule l'indicateur de ponctualité voyageurs comme le pourcentage de voyageurs arrivant à la gare de destination avec moins de 5 minutes de retard par rapport à l'heure théorique.

L'indicateur de ponctualité voyageurs a pour but de s'affranchir de la simple notion de retard du train pour représenter la perception du voyageur. En effet, en zone dense, où la fréquence des trains est élevée, l'heure théorique du train réellement emprunté a peu d'importance pour les voyageurs par rapport à d'autres paramètres tels que le temps d'atteinte en gare et le temps de trajet.

Le critère de ponctualité voyageurs diffère de celui du simple retard des trains pour plusieurs aspects :

- Le retard d'un train transportant beaucoup de voyageurs a plus d'impact sur le critère de

ponctualité voyageurs par rapport au même retard d'un train moins chargé.

- Il est possible d'avoir des voyageurs en retard dans un train à l'heure (par exemple, si un train est supprimé et que celui d'après et à l'heure, les voyageurs qui voulaient prendre le premier train seront en retard).
- Il est également possible d'avoir des voyageurs à l'heure dans un train en retard (notamment si un train roule à l'heure du train suivant, donc en retard, les voyageurs souhaitant prendre le deuxième train pourront monter dans le premier est être à l'heure).

Ces aspects rendent difficile pour les opérateurs une prise de décision en fonction du critère de ponctualité voyageurs, ce qui justifie le besoin d'un outil d'aide à la décision. En effet, selon le choix d'ordonnement les voyageurs pourront être affectés à des trains différents, et le choix peut impacter les retards des trains suivants (et donc l'heure d'arrivée à destination des voyageurs qui voudrait monter dans ces trains). La présence de missions directes et de missions omnibus, avec des temps de parcours différents, complexifie d'avantage la compréhension des impacts du choix sur la ponctualité voyageurs.

4 Modélisation du problème et architecture

La Direction de l'Innovation et de la Recherche a développé un outil visant à aider le choix d'ordonnement des trains du point de vue de la ponctualité voyageurs. L'outil calcule la ponctualité voyageurs associée aux différents scénarios d'ordonnement des trains en se basant sur un flux de géolocalisation des trains en temps réel. Les résultats sont restitués via une interface homme-machine (IHM). L'architecture du projet est synthétisée en Figure 2.

En entrée nous avons des données structurelles et des données temps réel. Les *données structurelles* incluent la modélisation du réseau et des paramètres de circulations tels que les espacements minimaux entre les trains, le plan de transport théorique et les données statistiques sur le nombre des voyageurs par heure par couple origine/destination. Certaines de ces données sont fixes pour un calcul donné, comme la modélisation du réseau, d'autre sont journalières, comme le plan de transport théorique. La modélisation du réseau inclut des informations sur les zones aux alentours de la convergence (voir Figure 3). En particulier, il contient une définition de la zone de décision, c'est-à-dire la zone en amont de la convergence où les trains doivent se trouver pour qu'il soit possible et pertinent d'effectuer un réordonnement.

Les *données temps réel* incluent les messages de géolocalisation des trains et les messages déclenchés par les actions de régulation telles que les adaptations de l'offre.

La localisation des trains se fait grâce à un système de balises placées sur les rails : le passage d'un train sur une balise déclenche un message de géolocalisation. Cela signifie que la connaissance de la position des trains n'est pas connue en continue, mais de façon discrète : une fois le train passé sur une balise, il n'est pas possible de connaître sa position ou sa vitesse jusqu'au passage par la balise suivante.

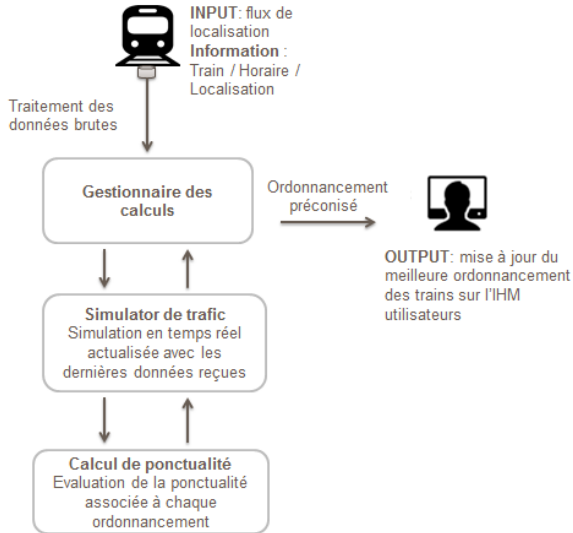


Figure 2 - Architecture du logiciel

Toutefois, une estimation de la position d'un train à un instant donné est possible en se basant sur l'horaire de passage à la dernière balise, sur les temps de trajet théorique et sur la position des autres trains sur le réseau.

À la réception de chaque nouveau message de localisation d'un train, l'outil met à jour la position des trains dans le réseau, identifie les différentes possibilités d'ordonnancement pour les trains concernés et simule l'impact de chaque possibilité sur les horaires des trains de l'instant présent jusqu'à la fin de la journée. Cela permet au module de ponctualité de

calculer la ponctualité voyageurs pour chaque scénario d'ordonnancement, et d'identifier le meilleur choix.

Tout d'abord le module de *gestion des calculs* reçoit les données de localisation et identifie les trains présents dans les différentes zones du réseau montrée en Figure 3. Les trains dans la 'zone de décision' sont potentiellement concernés par la décision d'ordonnancement. Bien que plusieurs trains de chaque branche puissent être dans la zone de décision à un instant donné, nous considérons que le réordonnancement n'est possible qu'entre les deux les plus proches de la convergence. Deux situations sont donc à examiner (en se référant à la Figure 3) : le train A passe avant le train B, et le train B passe devant le train A.

Le module de gestion des calculs transmet au *simulateur de trafic* l'information des trains concernés par le réordonnancement, ainsi que l'information sur le dernier message de localisation reçu. Le simulateur met à jour les positions des trains grâce au message de localisation, et une simulation est ensuite lancée pour chacun de scénarios d'ordre de passage. Ces simulations permettent d'évaluer l'impact du choix d'ordonnancement sur tous les trains de la ligne. La simulation est menée grâce à un simulateur de trafic développé à la Direction de l'Innovation et de la Recherche de la SNCF [9]. Le simulateur prend en entrée les horaires théoriques et les horaires réels de localisation des trains pour calculer les circulations futures des trains du début jusqu'à la fin de la journée.

Le *simulateur* retourne l'information sur l'estimation des horaires de passage à chaque gare pour tous les trains de la ligne. Le module de *calcul de la ponctualité* utilise cette information pour évaluer la ponctualité voyageurs telle que décrite dans la Section 3. Il retourne la valeur de l'indicateur de ponctualité voyageurs pour les deux scénarios d'ordonnancement (A passe devant B et B passe devant A).

Les informations sont par la suite transmises au module de gestion des calculs qui les élabore et les transmette à

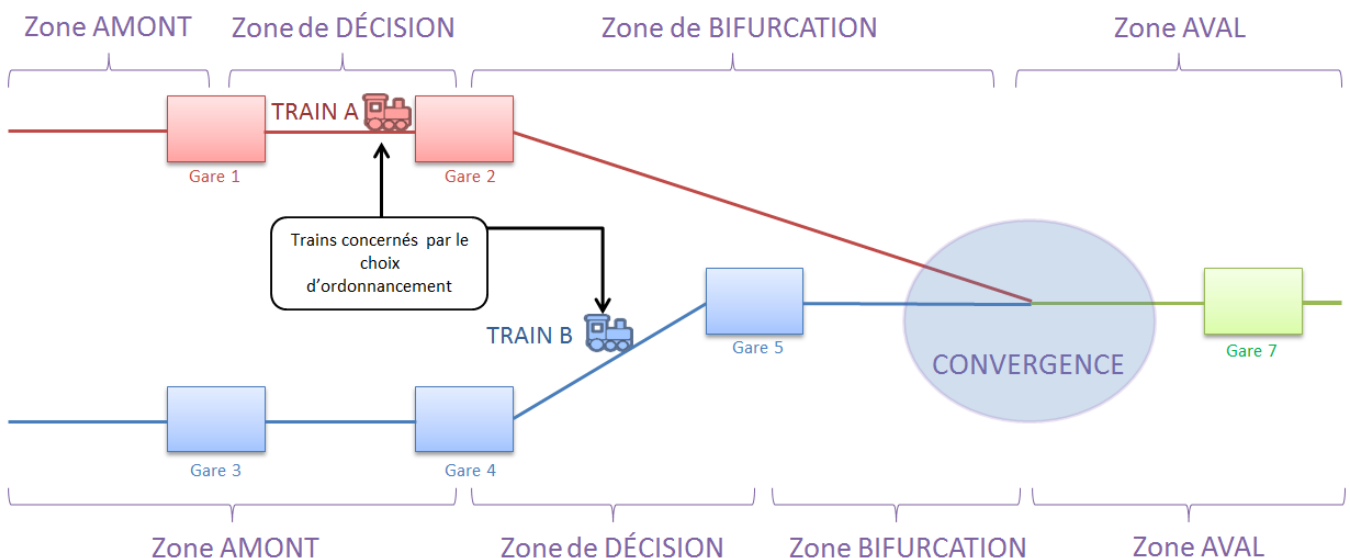


Figure 3 - Schéma des zones aux alentours de la convergence

l'IHM. L'interface a été élaborée en concertation avec des experts en design et en sciences cognitives, ainsi qu'avec les utilisateurs, afin de donner les indicateurs pertinents de façon simple et immédiate. Les enjeux principaux dans l'élaboration de l'interface ont été d'une part la nécessité d'expliquer le choix proposé par l'outil, et d'autre part de ne pas surcharger l'utilisateur avec trop d'informations. Cela est particulièrement crucial dans un contexte opérationnel où la charge de l'opérateur est élevée, notamment en cas de perturbations.

5 Cas d'étude et expérimentations

Afin d'éprouver la qualité des préconisations de l'outil pour l'amélioration de la ponctualité voyageurs, nous avons mené des expérimentations sur le périmètre de la ligne L et du RER A (voir Figure 4). L'outil propose en continu une préconisation d'ordre de passage à la convergence pour les deux trains dans la zone de prise de décision: le prochain train au départ de Poissy et le premier train en aval de la gare de Conflans Fin d'Oise. Ces trains sont marqués en rouge dans la Figure 4. Par rapport au schéma dans la Figure 3, il n'y a pas de zone amont pour les trains en provenance de Poissy, et le train est en zone de décision jusqu'à la détection de son départ de la gare d'origine.

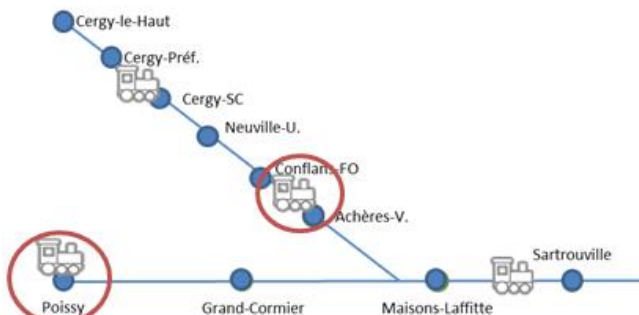


Figure 4 – Schéma de la zone de convergence étudiée

Les expérimentations se sont déroulées pendant trois semaines sur la pointe du matin, de 7:30 à 9:30, à la présence constante d'un membre de l'équipe projet. Dans une première phase, nous avons suivi le travail des régulateurs sans imposer la mise en place des préconisations de l'outil. Cela a permis d'évaluer la cohérence et l'apport de l'outil par rapport à la situation existante. Ensuite, une application systématique des préconisations a été demandée : cela a permis d'évaluer la qualité des préconisations de l'outil dans un contexte opérationnel.

Qualité des prévisions par rapport au réalisé

Nous nous sommes d'abord intéressés à l'étude de la qualité des prévisions des positions des trains par rapport au réalisé.

Pour cela, nous avons calculé les écarts entre les positions simulées et réalisées des trains sur des différentes plages horaires de 0 à 30 minutes à partir du moment de préconisation. La distribution de ces écarts, basée sur 1500 simulations, est montrée en Figure 5. Bien évidemment, nous ne pouvons à chaque fois que

confronter le scénario correspondant au réordonnement réalisé.

L'erreur moyenne sur la position de chaque train entre 0 et 5 minutes après la simulation est de 5 secondes. Cette erreur devient plus importante le plus on s'éloigne du moment de la simulation, tout en restant acceptable pour nos applications. Par exemple, il est inférieur à 40 secondes en moyenne pour le calcul des positions des trains entre 25 et 30 minutes après la simulation, bien que l'écart type soit plus important. L'augmentation de l'écart-type est due aux aléas ou incidents qui peuvent apparaître sur la ligne, et dont le simulateur ne peut pas avoir connaissance au moment du déclenchement de la simulation.

Nous pouvons noter que les sous-estimations des retards (partie droite du graphe) sont plus importantes que les surestimations. Cette différence est normale : le simulateur ne prévoit pas la possibilité qu'un train puisse rattraper son retard, car cela signifierait un temps de parcours plus court par rapport au temps de parcours théorique. Il y a tout de même des cas où le simulateur peut surestimer un retard : notamment si ce retard est dû à l'espacement avec un train qui sera au final supprimé.

La qualité de la simulation est très bonne par rapport à la précision requise pour nos usages, et nous conforte sur la qualité des préconisations données par l'outil.

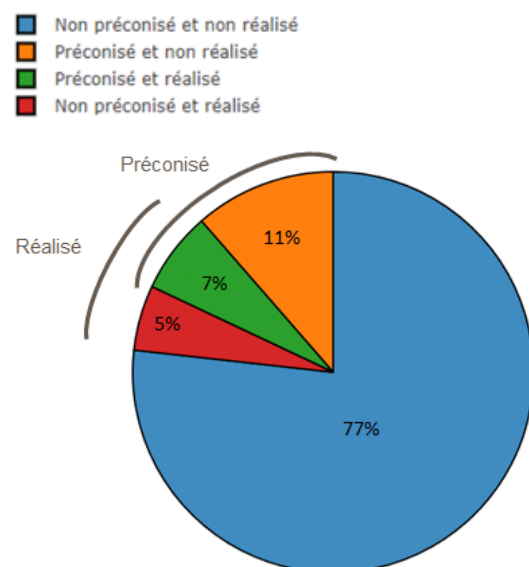


Figure 6 – Pourcentage de réordonnements préconisés vs réalisés

Réordonnements : effectués et préconisés

Nous avons par la suite analysé les réordonnements effectués par le régulateur par rapport aux réordonnements préconisés par l'outil. Les résultats sont montrés en Figure 6.

Nous retrouvons un accord entre la préconisation de l'outil et la décision de l'opérateur dans le 84% des cas (partie bleu et partie verte du graph). Cet accord montre que les préconisations de l'outil sont globalement cohérentes, bien qu'une marge

d'optimisation importante reste possible : 11% des réordonnements proposés ne sont pas effectués, et 5% des réordonnements effectués n'étaient pas préconisés par l'outil. La cause principale de cette discordance est à rechercher dans la complexité du critère de ponctualité voyageurs : en effet, la plupart de réordonnements non effectués (partie jaune) ne réduisait pas les retards des trains. Ce dernier critère est plus facile à estimer par l'opérateur par rapport au critère de ponctualité voyageurs. D'autres raisons de désaccord entre l'outil et la réalité incluent l'utilisation des voies supplémentaires afin de décharger la zone de convergences, ou la nécessité de faire passer un train en premier pour des sujets de réutilisation des rames ou de réaffectation des conducteurs. Certains de ces aspects pourront être intégrés par des versions futures de l'outil, bien que la disponibilité en temps réel des données nécessaires à l'implémentation de ces fonctionnalités reste pour l'instant un point bloquant.

Analyse de la phase d'application systématique des préconisations de l'outil

L'objectif est de comprendre si l'application des préconisations a amélioré la ponctualité voyageurs, comme prédit par l'outil. Pour cela, nous avons analysé dans le détail les situations où l'opérateur a suivi la préconisation de l'outil, mais où il aurait pris la décision opposée en dehors du cadre des expérimentations, sans l'outil à disposition. Un exemple est analysé en Figure 7, qui montre les Graphiques Espace-Temps (GET) pour les deux scénarios d'ordonnement des trains. Les GET affichent en abscisse les horaires et en ordonnée les positions sur la ligne. Les deux trains A et B viennent de deux branches différentes, mais nous affichons sur le graphique la position par rapport à la convergence.

La décision a des impacts sur les trains suivants, mais nous ne montrons par simplicité que les deux trains réordonnés.

Le scénario réalisé est celui du réordonnement : B passe devant A. La simulation prévoit les positions des trains avec une bonne précision par rapport au réalisé. À noter qu'à chaque fois nous manquons d'informations sur la qualité de la simulation par rapport au scénario non retenu.

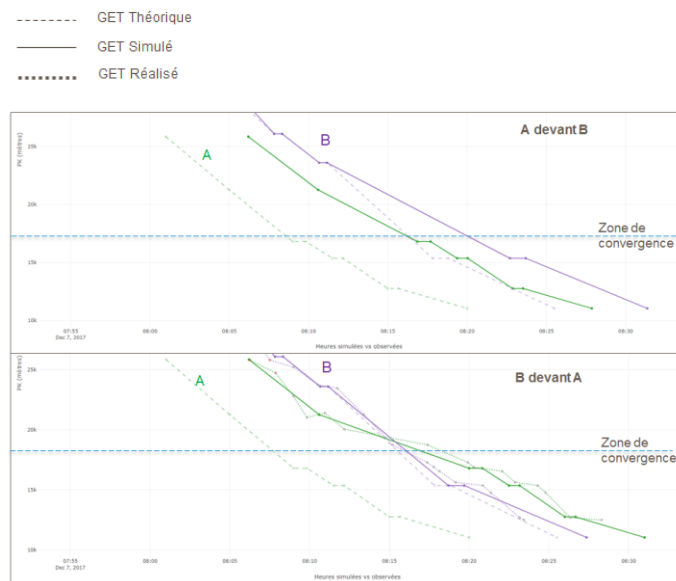


Figure 7 – GET pour le scénario A passe devant B et B passe devant A (scénario retenu)

Une analyse de chaque groupe de voyageurs impacté par la décision a confirmé l'amélioration de la ponctualité voyageurs par rapport au scénario non préconisé.

Estimation du gain en ponctualité voyageurs

Une estimation solide du gain en ponctualité voyageurs demanderait l'application systématique des préconisations de l'outil sur une période de plusieurs semaines, afin d'analyser si la ponctualité est

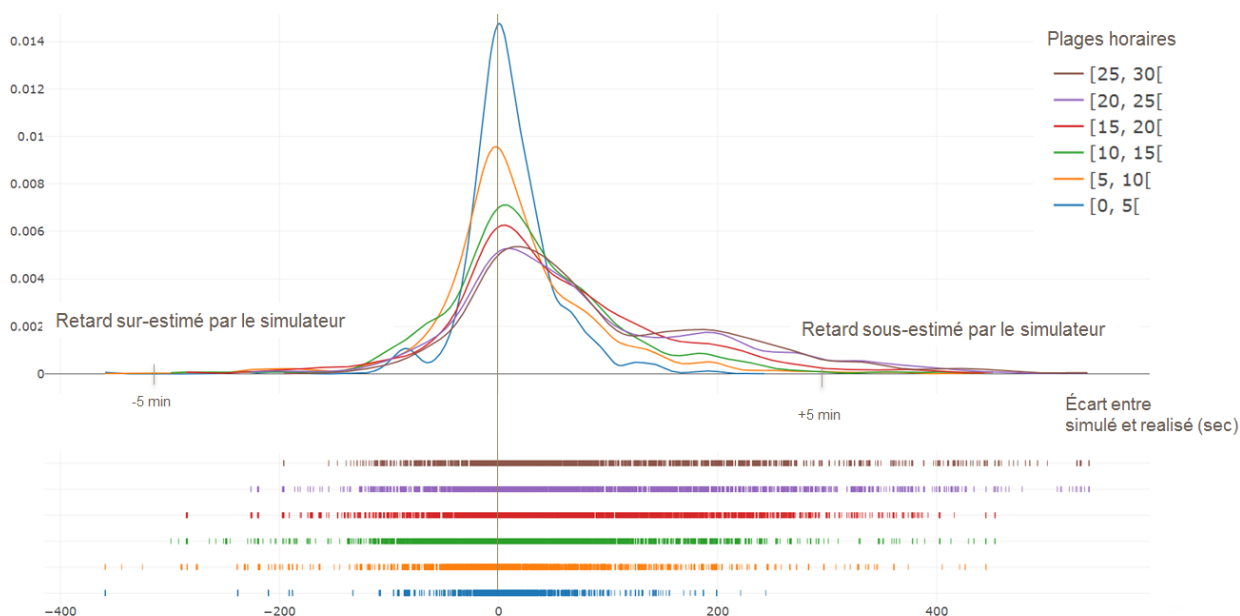


Figure 5 - Distribution des écarts entre réalisé et simulé sur des plages horaires de 5 minutes entre 0 et 30 minutes après la préconisation

significativement améliorée par rapport à une période de référence. Cela n'était pas possible dans le cadre de cette expérimentation. Nous nous sommes contentés d'effectuer une estimation moins fine en faisant une simple somme des gains en ponctualité générés par un suivi régulier de l'outil : grâce à une application systématique des préconisations, il serait possible d'éviter la perte en ponctualité liée aux réordonnements non préconisés et réalisés (partie rouge du graphe en Figure 6) et de profiter du gain générés par l'application des réordonnements préconisés mais non réalisés (partie jaune du graphe en Figure 6). L'estimation montre qu'un nombre important aujourd'hui en retard pourrait arriver à l'heure en suivant les ordonnancements préconisés par l'outil.

Retours des utilisateurs

Suite aux expérimentations, les utilisateurs ont complété un questionnaire de retour. Les questions visaient à comprendre le ressenti de leur travail quotidien, en particulier pour ce qui concerne la gestion de la convergence, en terme de difficulté et charge cognitive. Des avis ont été demandés sur les expérimentations et l'outil : confiance dans les préconisations, pertinence des informations présentée, etc. Les opérateurs ont donné un avis positif sur les préconisations de l'outil et les informations montrées au niveau de l'interface. La conduite du changement reste à travailler pour permettre l'utilisation de ce type d'algorithmes dans un contexte opérationnel. Malgré les nombreux échanges avec les utilisateurs pendant la conception du prototype, il reste difficile de changer des pratiques d'exploitations consolidées, notamment dans le cadre d'une expérimentation d'une durée de quelques semaines.

6 Conclusions et perspectives

Dans cet article, nous avons présenté un outil d'aide à la gestion des zones de convergence ferroviaire, et son application à un cas de test.

L'outil s'est montré pertinent sur plusieurs aspects : aider les décisions des opérateurs au quotidien, en leur fournissant des éléments objectifs pour une prise de décision rapide et optimale ; réduire la variabilité liée au facteur humain, en rendant l'exploitation plus homogène d'un jour à l'autre ; améliorer la ponctualité voyageurs, objectif primaire de Transilien.

Suite à ces résultats, le passage en industrialisation du prototype est prévu.

Du point de vue du projet de recherche, des améliorations sont envisageables pour traiter des cas plus complexes : par exemple, pour aider la gestion des lignes avec des convergences multiples, ou pour intégrer d'autres décisions en temps réel comme des adaptations de l'offre.

Bibliographie

- [1] V. Cacchiani, D. Huisman, M. Kidd, L. Kroon, P. Toth, L. Veelenturf, et J. Wagenaar (2014). An overview of recovery models and algorithms for real-time railway rescheduling. *Transportation Research Part B: Methodological*, 63, pages 15-37.
- [2] A. Toletti, M. Laumanns, P. Grossenbacher et U. Weidmann (2015). Meeting functional requirements for real-time railway traffic management with mathematical model. *Advanced Systems in Public Transport Conference*.
- [3] F. Corman and L. Meng (2013). A review of online dynamic models and algorithms for railway traffic control. *Intelligent Rail Transportation (ICIRT), 2013 IEEE International Conference*, pages 128-133.
- [4] C. J. Goodman et S. Murata (2001). Metro traffic regulation from the passenger perspective. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol 215, Issue 2, pages 137 – 147.
- [5] N. Tomii, Y. Tashiro, N. Tanabe, C. Hirai et K. Muraki (2005). Train rescheduling algorithm which minimizes passengers' dissatisfaction. Ali M., Esposito F., *Innovations in Applied Artificial Intelligence. IEA/AIE. Lecture Notes in Computer Science*, vol 3533. Springer, Berlin, Heidelberg
- [6] L. Chen, F. Schmid, M. Dasigi, B. Ning, C. Roberts, and T. Tang. (2010). Real-time train rescheduling in junction areas. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 224(6), pages 547-557.
- [7] A. Galapitge, A. R.Albrechta, P. Pudneyab, X. Vua et P. Zhoua (2017). Optimal real-time junction scheduling for trains with connected driver advice systems. *Journal of Rail Transport Planning & Management*.
- [8] S. Van Thielen, S. Burggraeve, and P. Vansteenwegen (2015). "Optimal train rescheduling after conflict detection". *Conference on Advanced Systems in Public Transport (CASPT)*
- [9] J. Damay, A. Boillot, (2014). "An efficient macroscopic railway system simulator". *20th Conference of the International Federation of Operational Research Societies (IFORS)*.

Remerciements

Ces travaux ont été financés par SNCF Transilien et ont été menés par la Direction de l'Innovation de la SNCF. Nous tenons à exprimer nos remerciements à toutes personnes impliquées dans le projet : le comité de pilotage, le Lab Mass Transit, la DSI de SNCF Mobilité, les opérateurs et les responsables du centre opérationnel où nous avons testé l'outil, les équipes de développements du logiciel, les fournisseurs des données de géolocalisation et de plans de transport.

Optimisation en temps réel de la gestion des convergences ferroviaires

Valentina Pozzoli¹

François Ramond¹

¹SNCF Innovation et Recherche
40, avenue des Terroirs de France
75611 Paris Cedex

Valentina.Pozzoli@sncf.fr
Francois.Ramond@sncf.fr

Résumé

L'exploitation ferroviaire en zone dense est une tâche complexe. La densité des circulations provoque une propagation très rapide des retards, même de petite amplitude. Afin de prendre des décisions d'exploitation rapides et optimales il est nécessaire de se doter d'outils d'aide à la décision performants.

Dans cet article, nous présentons un système d'aide à la gestion optimale en temps réel des zones de convergences ferroviaires, afin d'améliorer la ponctualité des voyageurs de la ligne. Nous détaillons les enjeux liés au projet, ainsi que les choix de modélisation et l'architecture du prototype. Le système a été testé avec succès sur un site pilote en Île-de-France : nous montrons les résultats de ces expérimentations et donnons les perspectives pour la suite du projet.

Mots Clef

Aide à la décision, Optimisation, Exploitation ferroviaire, Algorithmes en temps réel, Convergences ferroviaires

Abstract

Railway traffic management in mass transit system is a complex task. Trains run so frequently and so close to each other that delays propagate very fast. Operators have to take efficient decisions depending on a multitude of factors, that impact a large number of trains, and thus of passengers. Artificial intelligence methods come in help in the decision process.

In this paper, a real-time decision-support tool for the optimal management of junction areas is presented. It proposes the best order of trains to run through a junction, such that the passenger punctuality is maximized. Details are given on the modelling and the architecture of the tool, as well as on the goals and ambitions of the project. The results of promising experiments on a case-study on a junction in a Parisian area are presented.

Keywords

Decision-support, Optimization, Traffic management, Mass transit, Real-time algorithms, Railway junctions

1 Introduction

L'exploitation ferroviaire en zone dense, où la fréquence des trains est élevée et le nombre de voyageurs important, est un vrai défi dont l'enjeu principal est la qualité du service offert aux voyageurs. L'optimisation de l'exploitation ferroviaire relève de plusieurs facteurs avec des échéances temporelles différentes. En amont de la production se situent, par exemple, la conception de plans de transport robustes et résilients, la maintenance des voies et du matériel, l'affectation des conducteurs et la gestion des rames. En gestion opérationnelle, la mise en œuvre de mesures de régulation efficaces est nécessaire afin d'optimiser la production : changements de la desserte ou des voies utilisées, adaptation d'offre, etc. La gestion des zones de convergence figure parmi ces mesures de régulation.

Les zones de convergences sont des endroits du réseau ferroviaire où deux ou plusieurs voies se réunissent. L'exemple le plus simple, qui sera étudié au cours de cet article, est montré en Figure 1. La gestion des zones de convergence consiste à choisir l'ordre de passage des trains lors que deux trains de deux branches différentes s'apprêtent à rentrer dans la zone de convergence. Bien que la thématique d'optimisation de l'ordre de passage existe pour toutes zones de convergence, en zone dense le choix est particulièrement impactant : en effet, dans un réseau dense les retards se propagent facilement sur les trains suivants. Chaque décision de régulation peut donc avoir un impact sur un grand nombre de trains, et de conséquence sur un nombre important de voyageurs. Ces décisions dépendent de nombreux paramètres : par exemple, les retards des trains, leur desserte, le nombre de passagers dans les trains et sur les quais. En plus, les choix doivent être pris rapidement par le régulateur : pour ces raisons il est nécessaire de se doter d'outils d'aide à la décision performants qui puissent venir en support du travail des régulateurs.

En Ile-de-France, SNCF Mobilités, via son activité Transilien, et RATP sont les principaux opérateurs de transport ferroviaire. Ils opèrent par le compte d'Île-de-France Mobilité, qui définit également les objectifs de performance des opérateurs de transport. Transilien fait circuler 6200 trains chaque jour, qui transportent au quotidien 3 millions de voyageurs, nombre qui connaît une augmentation de 3% par an en moyenne (source Transilien). L'amélioration de la 'ponctualité des

voyageurs', premier élément pour la satisfaction des clients, est l'objectif principal de Transilien et le critère utilisé par Île-de-France Mobilité pour évaluer la performance d'une ligne.

Dans ce contexte, la gestion des zones de convergence, est particulièrement complexe et impactant. En effet, ces zones sont spécialement sollicitées car les trains des deux branches doivent y passer : la fréquence de passage peut aller jusqu'à 32 trains par heure. L'ordonnancement des trains dans une zone de convergence peut être modifié par les opérateurs. Bien que les horaires théoriques indiquent l'ordre de passage des trains à une zone de convergence, cet ordre peut être remis en cause par les aléas : la décision optimale doit considérer l'impact de ce choix sur toute la ligne d'un point de vue des voyageurs concernés.

La Direction de l'Innovation et de la Recherche de SNCF a développé pour Transilien un outil d'aide à la gestion des zones de convergence en temps réel. Cet outil préconise l'ordre de passage des trains à la convergence permettant de maximiser la performance globale de la ligne, définie par le critère de 'ponctualité des voyageurs'.

L'article est organisé de la façon suivante : la Section 2 est consacrée à une présentation succincte des travaux déjà existants sur le sujet ; la Section 3 détaille le critère d'optimisation et les enjeux du projet, la Section 4 décrit la modélisation du problème et l'architecture de l'outil ; la Section 5 présente les résultats de l'expérimentation menée dans un contexte opérationnel sur un site pilote ; enfin, des conclusions et perspectives sont exposées en Section 6.

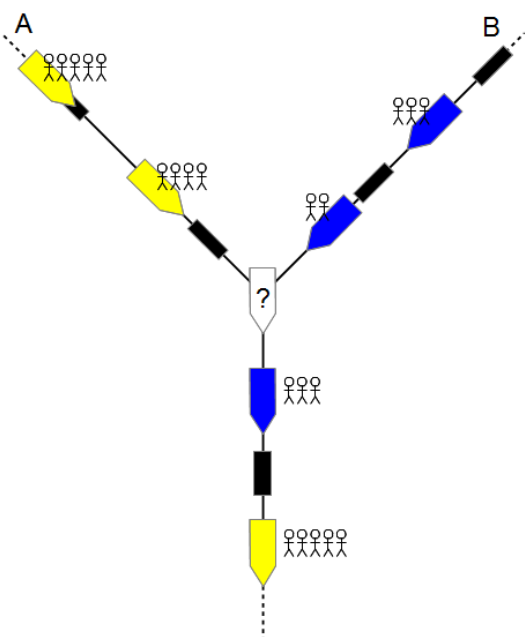


Figure 1 - Schéma d'une zone de convergence

2 Travaux antérieurs

La gestion du trafic en temps réel est un problème largement abordé dans la littérature. Une revue extensive des approches pour réduire les conséquences des incidents est présentée dans [1], en distinguant entre petits aléas et perturbations importantes. Les actions à mettre en place sont différentes dans les deux cas, et incluent la replanification d'itinéraires, la gestion des correspondances, la replanification des engins ou des conducteurs. Au niveau des adaptations du plan de transport en temps réel, les stratégies utilisées en zone dense sont la rétention d'un train en gare, l'ajustement de la vitesse, le changement de desserte, et l'ordre de passage en certains points du réseau. L'ordonnancement des trains à une convergence est un cas particulier de cette dernière stratégie de replanification.

L'article [2] résume les modèles existants dans la planification et replanification ferroviaire, dont les décisions d'ordonnancement à une convergence, ou de replanification d'itinéraires. Une autre revue des modèles de contrôle de trafic en temps réel est proposée par [3].

L'optimisation de l'exploitation du point de vue des voyageurs, et non pas que des retards des trains, est traitée dans plusieurs articles : par exemple dans [4], les horaires d'arrivée et de départ des métros sont choisis afin de minimiser les temps d'atteintes et de trajet, ainsi que la congestion; [5] proposent un algorithme de replanification qui minimise l'insatisfaction des voyageurs.

Plusieurs articles traitent spécifiquement le problème de gestion des convergences ferroviaires : [6] propose un modèle d'ordonnancement optimal des trains aux convergences afin de minimiser une somme pondérée des retards des trains aux gares desservies. Ce modèle est résolu grâce à un algorithme à évolution différentielle et, sur les instances de test, permet de réduire les retards par rapport à une stratégie du type premier entré, premier sorti. Un modèle de replanification en temps réel qui inclut l'optimisation de la vitesse des trains afin de minimiser les retards est proposé dans [7]. Une comparaison entre différentes stratégies de résolution des conflits aux convergences est décrite dans [8]: premier arrivé premier servi, priorité selon le type de train, et minimisation du nombre et de la durée des retards grâce à une heuristique. Les trois scénarios sont évalués grâce à des simulations.

Les travaux présentés dans cet article proposent un approche différente car centré sur une gestion optimale du point de vue de la performance globale de la ligne, en considérant le point de vue des voyageurs, et non pas que le retard des trains.

3 Critère d'optimisation

L'objectif de nos travaux est d'optimiser la gestion d'une zone de convergence afin de maximiser la performance globale de la ligne.

Aujourd'hui, en cas de conflit dans une zone de convergence, la décision d'ordonnancement des trains est prise au cas par cas grâce à l'expertise humaine du régulateur, et dépend principalement des retards de deux trains en conflit, ainsi que de leurs missions. En cas de perturbation, le régulateur a une charge de travail élevée et doit prendre des décisions rapides qui ont un impact majeur sur toute la ligne. L'idée est de venir en aide aux régulateurs, en leur fournissant un outil qui puisse à la fois les décharger d'une partie de leur charge cognitive, leur donner des informations pour mieux analyser la situation, et leur proposer le meilleur ordonnancement du point de vue de la performance globale de la ligne.

Un des principaux indicateurs de la performance globale d'une ligne est appelé 'ponctualité voyageurs'. Ce critère mesure le nombre de voyageurs arrivés à leur gare de destination moins de 5 minutes après leur heure d'arrivée théorique. Il est utilisé pour évaluer la qualité de service de SNCF par Île-de-France Mobilités, et c'est le critère d'optimisation choisi dans l'outil d'aide à la gestion des zones de convergence.

Pour calculer la ponctualité voyageurs, des données statistiques sur le nombre de voyageurs pour chaque couple origine-destination par heure sont utilisées afin d'affecter les voyageurs aux différents trains, en se basant sur les horaires théoriques. Les données sur le nombre de voyageurs sont issues de comptages manuels des voyageurs effectués régulièrement sur les lignes concernées. Grâce à ces données et aux horaires théoriques des trains, il est possible de calculer l'heure théorique d'arrivée de chaque passager à la gare d'origine et à la gare de destination. Une fois les trains circulés, les horaires réels d'arrivée des voyageurs à destination sont calculés grâce aux horaires réels de passage des trains par les différentes gares. On considère que chaque voyageur arrive en gare à l'heure théorique prévu, et monte ensuite dans le premier train qui dessert sa destination. Par la suite, on calcule l'indicateur de ponctualité voyageurs comme le pourcentage de voyageurs arrivant à la gare de destination avec moins de 5 minutes de retard par rapport à l'heure théorique.

L'indicateur de ponctualité voyageurs a pour but de s'affranchir de la simple notion de retard du train pour représenter la perception du voyageur. En effet, en zone dense, où la fréquence des trains est élevée, l'heure théorique du train réellement emprunté a peu d'importance pour les voyageurs par rapport à d'autres paramètres tels que le temps d'atteinte en gare et le temps de trajet.

Le critère de ponctualité voyageurs diffère de celui du simple retard des trains pour plusieurs aspects :

- Le retard d'un train transportant beaucoup de voyageurs a plus d'impact sur le critère de

ponctualité voyageurs par rapport au même retard d'un train moins chargé.

- Il est possible d'avoir des voyageurs en retard dans un train à l'heure (par exemple, si un train est supprimé et que celui d'après et à l'heure, les voyageurs qui voulaient prendre le premier train seront en retard).
- Il est également possible d'avoir des voyageurs à l'heure dans un train en retard (notamment si un train roule à l'heure du train suivant, donc en retard, les voyageurs souhaitant prendre le deuxième train pourront monter dans le premier est être à l'heure).

Ces aspects rendent difficile pour les opérateurs une prise de décision en fonction du critère de ponctualité voyageurs, ce qui justifie le besoin d'un outil d'aide à la décision. En effet, selon le choix d'ordonnancement les voyageurs pourront être affectés à des trains différents, et le choix peut impacter les retards des trains suivants (et donc l'heure d'arrivée à destination des voyageurs qui voudrait monter dans ces trains). La présence de missions directes et de missions omnibus, avec des temps de parcours différents, complexifie d'avantage la compréhension des impacts du choix sur la ponctualité voyageurs.

4 Modélisation du problème et architecture

La Direction de l'Innovation et de la Recherche a développé un outil visant à aider le choix d'ordonnancement des trains du point de vue de la ponctualité voyageurs. L'outil calcule la ponctualité voyageurs associée aux différents scénarios d'ordonnancement des trains en se basant sur un flux de géolocalisation des trains en temps réel. Les résultats sont restitués via une interface homme-machine (IHM). L'architecture du projet est synthétisée en Figure 2.

En entrée nous avons des données structurelles et des données temps réel. Les *données structurelles* incluent la modélisation du réseau et des paramètres de circulations tels que les espacements minimaux entre les trains, le plan de transport théorique et les données statistiques sur le nombre des voyageurs par heure par couple origine/destination. Certaines de ces données sont fixes pour un calcul donné, comme la modélisation du réseau, d'autre sont journalières, comme le plan de transport théorique. La modélisation du réseau inclut des informations sur les zones aux alentours de la convergence (voir Figure 3). En particulier, il contient une définition de la zone de décision, c'est-à-dire la zone en amont de la convergence où les trains doivent se trouver pour qu'il soit possible et pertinent d'effectuer un réordonnancement.

Les *données temps réel* incluent les messages de géolocalisation des trains et les messages déclenchés par les actions de régulation telles que les adaptations de l'offre.

La localisation des trains se fait grâce à un système de balises placées sur les rails : le passage d'un train sur une balise déclenche un message de géolocalisation. Cela signifie que la connaissance de la position des trains n'est pas connue en continue, mais de façon discrète : une fois le train passé sur une balise, il n'est pas possible de connaître sa position ou sa vitesse jusqu'au passage par la balise suivante.

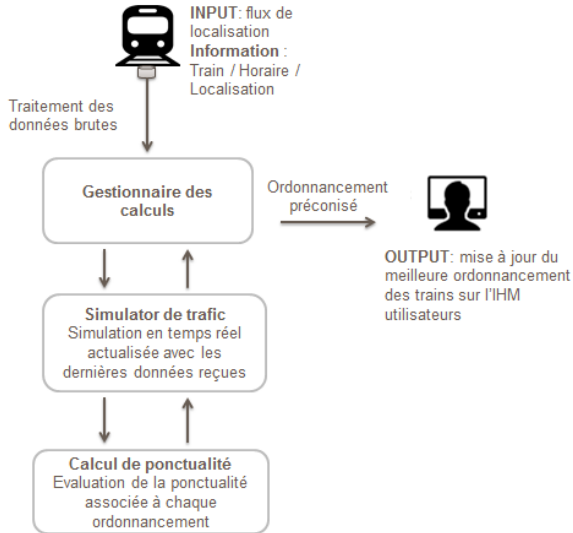


Figure 2 - Architecture du logiciel

Toutefois, une estimation de la position d'un train à un instant donné est possible en se basant sur l'horaire de passage à la dernière balise, sur les temps de trajet théorique et sur la position des autres trains sur le réseau.

À la réception de chaque nouveau message de localisation d'un train, l'outil met à jour la position des trains dans le réseau, identifie les différentes possibilités d'ordonnancement pour les trains concernés et simule l'impact de chaque possibilité sur les horaires des trains de l'instant présent jusqu'à la fin de la journée. Cela permet au module de ponctualité de

calculer la ponctualité voyageurs pour chaque scénario d'ordonnancement, et d'identifier le meilleur choix.

Tout d'abord le module de *gestion des calculs* reçoit les données de localisation et identifie les trains présents dans les différentes zones du réseau montrée en Figure 3. Les trains dans la 'zone de décision' sont potentiellement concernés par la décision d'ordonnancement. Bien que plusieurs trains de chaque branche puissent être dans la zone de décision à un instant donné, nous considérons que le réordonnancement n'est possible qu'entre les deux les plus proches de la convergence. Deux situations sont donc à examiner (en se référant à la Figure 3) : le train A passe avant le train B, et le train B passe devant le train A.

Le module de gestion des calculs transmet au *simulateur de trafic* l'information des trains concernés par le réordonnancement, ainsi que l'information sur le dernier message de localisation reçu. Le simulateur met à jour les positions des trains grâce au message de localisation, et une simulation est ensuite lancée pour chacun de scénarios d'ordre de passage. Ces simulations permettent d'évaluer l'impact du choix d'ordonnancement sur tous les trains de la ligne. La simulation est menée grâce à un simulateur de trafic développé à la Direction de l'Innovation et de la Recherche de la SNCF [9]. Le simulateur prend en entrée les horaires théoriques et les horaires réels de localisation des trains pour calculer les circulations futures des trains du début jusqu'à la fin de la journée.

Le *simulateur* retourne l'information sur l'estimation des horaires de passage à chaque gare pour tous les trains de la ligne. Le module de *calcul de la ponctualité* utilise cette information pour évaluer la ponctualité voyageurs telle que décrite dans la Section 3. Il retourne la valeur de l'indicateur de ponctualité voyageurs pour les deux scénarios d'ordonnancement (A passe devant B et B passe devant A).

Les informations sont par la suite transmises au module de gestion des calculs qui les élabore et les transmette à

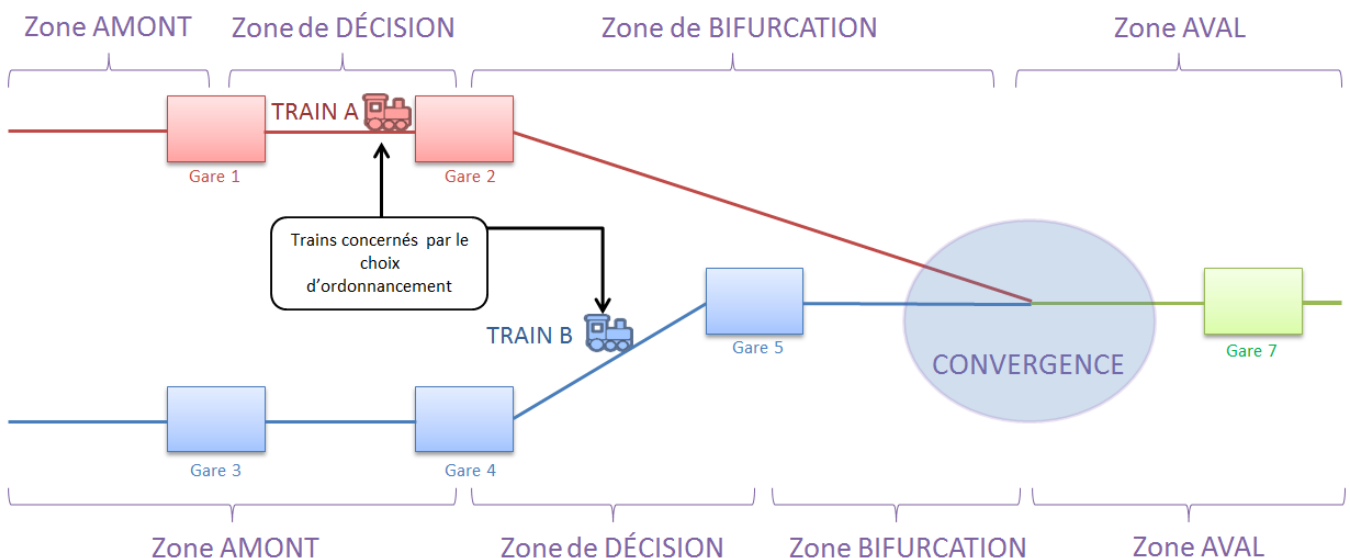


Figure 3 - Schéma des zones aux alentours de la convergence

l'IHM. L'interface a été élaborée en concertation avec des experts en design et en sciences cognitives, ainsi qu'avec les utilisateurs, afin de donner les indicateurs pertinents de façon simple et immédiate. Les enjeux principaux dans l'élaboration de l'interface ont été d'une part la nécessité d'expliquer le choix proposé par l'outil, et d'autre part de ne pas surcharger l'utilisateur avec trop d'informations. Cela est particulièrement crucial dans un contexte opérationnel où la charge de l'opérateur est élevée, notamment en cas de perturbations.

5 Cas d'étude et expérimentations

Afin d'éprouver la qualité des préconisations de l'outil pour l'amélioration de la ponctualité voyageurs, nous avons mené des expérimentations sur le périmètre de la ligne L et du RER A (voir Figure 4). L'outil propose en continu une préconisation d'ordre de passage à la convergence pour les deux trains dans la zone de décision: le prochain train au départ de Poissy et le premier train en aval de la gare de Conflans Fin d'Oise. Ces trains sont marqués en rouge dans la Figure 4. Par rapport au schéma dans la Figure 3, il n'y a pas de zone amont pour les trains en provenance de Poissy, et le train est en zone de décision jusqu'à la détection de son départ de la gare d'origine.

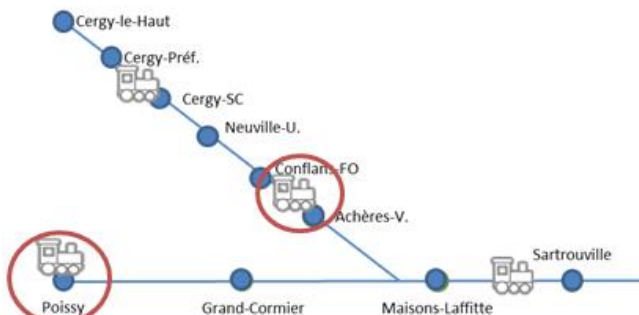


Figure 4 – Schéma de la zone de convergence étudiée

Les expérimentations se sont déroulées pendant trois semaines sur la pointe du matin, de 7:30 à 9:30, à la présence constante d'un membre de l'équipe projet. Dans une première phase, nous avons suivi le travail des régulateurs sans imposer la mise en place des préconisations de l'outil. Cela a permis d'évaluer la cohérence et l'apport de l'outil par rapport à la situation existante. Ensuite, une application systématique des préconisations a été demandée : cela a permis d'évaluer la qualité des préconisations de l'outil dans un contexte opérationnel.

Qualité des prévisions par rapport au réalisé

Nous nous sommes d'abord intéressés à l'étude de la qualité des prévisions des positions des trains par rapport au réalisé.

Pour cela, nous avons calculé les écarts entre les positions simulées et réalisées des trains sur des différentes plages horaires de 0 à 30 minutes à partir du moment de préconisation. La distribution de ces écarts, basée sur 1500 simulations, est montrée en Figure 5. Bien évidemment, nous ne pouvons à chaque fois que

confronter le scénario correspondant au réordonnement réalisé.

L'erreur moyenne sur la position de chaque train entre 0 et 5 minutes après la simulation est de 5 secondes. Cette erreur devient plus importante le plus on s'éloigne du moment de la simulation, tout en restant acceptable pour nos applications. Par exemple, il est inférieur à 40 secondes en moyenne pour le calcul des positions des trains entre 25 et 30 minutes après la simulation, bien que l'écart type soit plus important. L'augmentation de l'écart-type est due aux aléas ou incidents qui peuvent apparaître sur la ligne, et dont le simulateur ne peut pas avoir connaissance au moment du déclenchement de la simulation.

Nous pouvons noter que les sous-estimations des retards (partie droite du graphe) sont plus importantes que les surestimations. Cette différence est normale : le simulateur ne prévoit pas la possibilité qu'un train puisse rattraper son retard, car cela signifierait un temps de parcours plus court par rapport au temps de parcours théorique. Il y a tout de même des cas où le simulateur peut surestimer un retard : notamment si ce retard est dû à l'espacement avec un train qui sera au final supprimé.

La qualité de la simulation est très bonne par rapport à la précision requise pour nos usages, et nous conforte sur la qualité des préconisations données par l'outil.

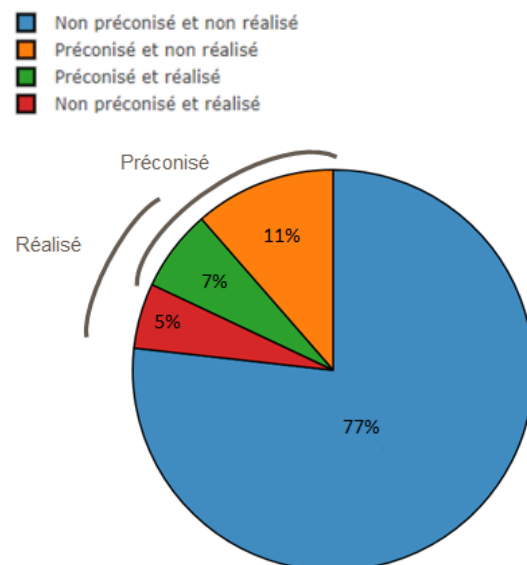


Figure 6 – Pourcentage de réordonnements préconisés vs réalisés

Réordonnements : effectués et préconisés

Nous avons par la suite analysé les réordonnements effectués par le régulateur par rapport aux réordonnements préconisés par l'outil. Les résultats sont montrés en Figure 6.

Nous retrouvons un accord entre la préconisation de l'outil et la décision de l'opérateur dans le 84% des cas (partie bleu et partie verte du graph). Cet accord montre que les préconisations de l'outil sont globalement cohérentes, bien qu'une marge

d'optimisation importante reste possible : 11% des réordonnements proposés ne sont pas effectués, et 5% des réordonnements effectués n'étaient pas préconisés par l'outil. La cause principale de cette discordance est à rechercher dans la complexité du critère de ponctualité voyageurs : en effet, la plupart de réordonnements non effectués (partie jaune) ne réduisait pas les retards des trains. Ce dernier critère est plus facile à estimer par l'opérateur par rapport au critère de ponctualité voyageurs. D'autres raisons de désaccord entre l'outil et la réalité incluent l'utilisation des voies supplémentaires afin de décharger la zone de convergences, ou la nécessité de faire passer un train en premier pour des sujets de réutilisation des rames ou de réaffectation des conducteurs. Certains de ces aspects pourront être intégrés par des versions futures de l'outil, bien que la disponibilité en temps réel des données nécessaires à l'implémentation de ces fonctionnalités reste pour l'instant un point bloquant.

Analyse de la phase d'application systématique des préconisations de l'outil

L'objectif est de comprendre si l'application des préconisations a amélioré la ponctualité voyageurs, comme prédit par l'outil. Pour cela, nous avons analysé dans le détail les situations où l'opérateur a suivi la préconisation de l'outil, mais où il aurait pris la décision opposée en dehors du cadre des expérimentations, sans l'outil à disposition. Un exemple est analysé en Figure 7, qui montre les Graphiques Espace-Temps (GET) pour les deux scénarios d'ordonnement des trains. Les GET affichent en abscisse les horaires et en ordonnée les positions sur la ligne. Les deux trains A et B viennent de deux branches différentes, mais nous affichons sur le graphique la position par rapport à la convergence.

La décision a des impacts sur les trains suivants, mais nous ne montrons par simplicité que les deux trains réordonnés.

Le scénario réalisé est celui du réordonnement : B passe devant A. La simulation prévoit les positions des trains avec une bonne précision par rapport au réalisé. À noter qu'à chaque fois nous manquons d'informations sur la qualité de la simulation par rapport au scénario non retenu.

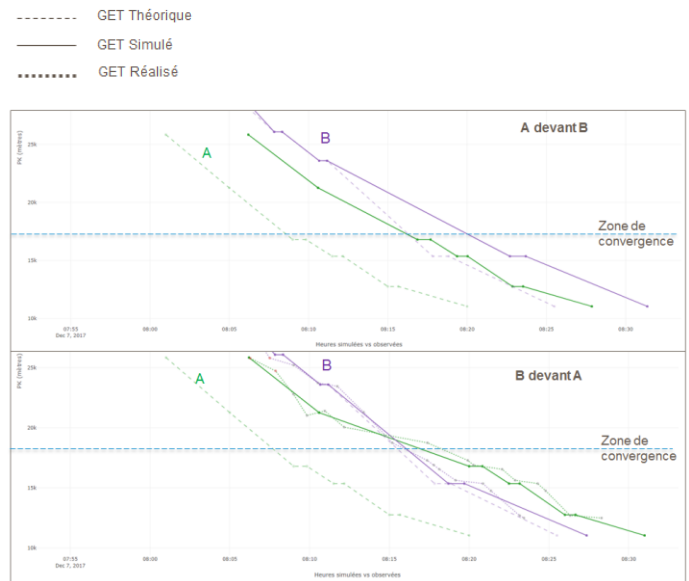


Figure 7 – GET pour le scénario A passe devant B et B passe devant A (scénario retenu)

Une analyse de chaque groupe de voyageurs impacté par la décision a confirmé l'amélioration de la ponctualité voyageurs par rapport au scénario non préconisé.

Estimation du gain en ponctualité voyageurs

Une estimation solide du gain en ponctualité voyageurs demanderait l'application systématique des préconisations de l'outil sur une période de plusieurs semaines, afin d'analyser si la ponctualité est

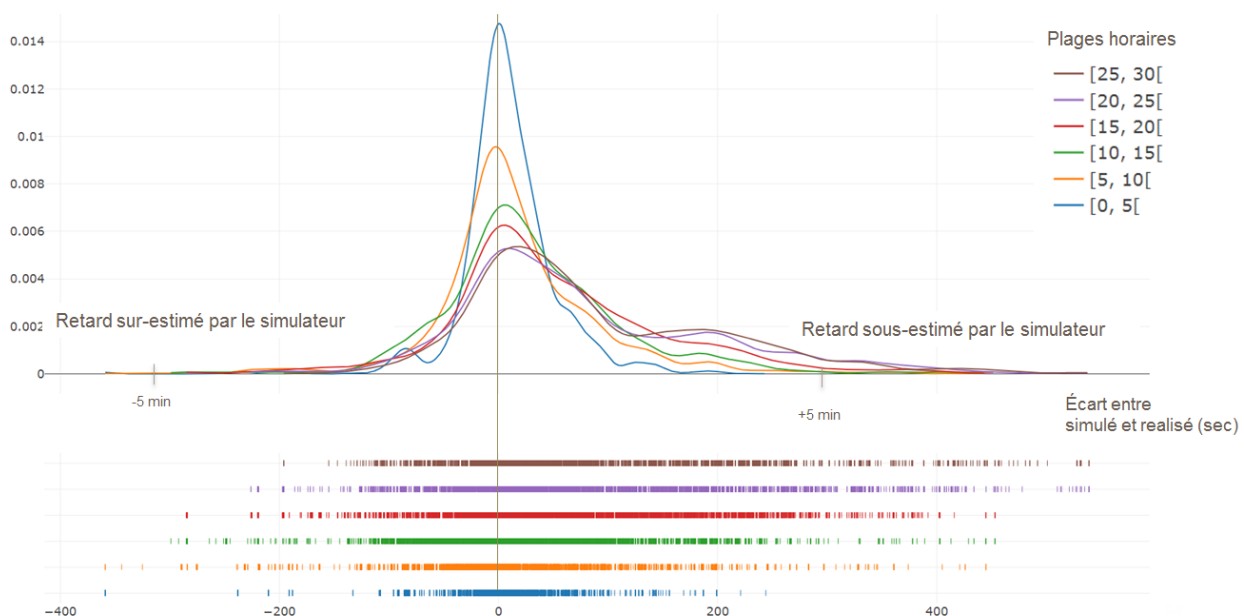


Figure 5 - Distribution des écarts entre réalisé et simulé sur des plages horaires de 5 minutes entre 0 et 30 minutes après la préconisation

significativement améliorée par rapport à une période de référence. Cela n'était pas possible dans le cadre de cette expérimentation. Nous nous sommes contentés d'effectuer une estimation moins fine en faisant une simple somme des gains en ponctualité générés par un suivi régulier de l'outil : grâce à une application systématique des préconisations, il serait possible d'éviter la perte en ponctualité liée aux réordonnements non préconisés et réalisés (partie rouge du graphe en Figure 6) et de profiter du gain générés par l'application des réordonnements préconisés mais non réalisés (partie jaune du graphe en Figure 6). L'estimation montre qu'un nombre important aujourd'hui en retard pourrait arriver à l'heure en suivant les ordonnancements préconisés par l'outil.

Retours des utilisateurs

Suite aux expérimentations, les utilisateurs ont complété un questionnaire de retour. Les questions visaient à comprendre le ressenti de leur travail quotidien, en particulier pour ce qui concerne la gestion de la convergence, en terme de difficulté et charge cognitive. Des avis ont été demandés sur les expérimentations et l'outil : confiance dans les préconisations, pertinence des informations présentée, etc. Les opérateurs ont donné un avis positif sur les préconisations de l'outil et les informations montrées au niveau de l'interface. La conduite du changement reste à travailler pour permettre l'utilisation de ce type d'algorithmes dans un contexte opérationnel. Malgré les nombreux échanges avec les utilisateurs pendant la conception du prototype, il reste difficile de changer des pratiques d'exploitations consolidées, notamment dans le cadre d'une expérimentation d'une durée de quelques semaines.

6 Conclusions et perspectives

Dans cet article, nous avons présenté un outil d'aide à la gestion des zones de convergence ferroviaire, et son application à un cas de test.

L'outil s'est montré pertinent sur plusieurs aspects : aider les décisions des opérateurs au quotidien, en leur fournissant des éléments objectifs pour une prise de décision rapide et optimale ; réduire la variabilité liée au facteur humain, en rendant l'exploitation plus homogène d'un jour à l'autre ; améliorer la ponctualité voyageurs, objectif primaire de Transilien.

Suite à ces résultats, le passage en industrialisation du prototype est prévu.

Du point de vue du projet de recherche, des améliorations sont envisageables pour traiter des cas plus complexes : par exemple, pour aider la gestion des lignes avec des convergences multiples, ou pour intégrer d'autres décisions en temps réel comme des adaptations de l'offre.

Bibliographie

- [1] V. Cacchiani, D. Huisman, M. Kidd, L. Kroon, P. Toth, L. Veelenturf, et J. Wagenaar (2014). An overview of recovery models and algorithms for real-time railway rescheduling. *Transportation Research Part B: Methodological*, 63, pages 15-37.
- [2] A. Toletti, M. Laumanns, P. Grossenbacher et U. Weidmann (2015). Meeting functional requirements for real-time railway traffic management with mathematical model. *Advanced Systems in Public Transport Conference*.
- [3] F. Corman and L. Meng (2013). A review of online dynamic models and algorithms for railway traffic control. *Intelligent Rail Transportation (ICIRT), 2013 IEEE International Conference*, pages 128-133.
- [4] C. J. Goodman et S. Murata (2001). Metro traffic regulation from the passenger perspective. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol 215, Issue 2, pages 137 – 147.
- [5] N. Tomii, Y. Tashiro, N. Tanabe, C. Hirai et K. Muraki (2005). Train rescheduling algorithm which minimizes passengers' dissatisfaction. Ali M., Esposito F., *Innovations in Applied Artificial Intelligence. IEA/AIE. Lecture Notes in Computer Science*, vol 3533. Springer, Berlin, Heidelberg
- [6] L. Chen, F. Schmid, M. Dasigi, B. Ning, C. Roberts, and T. Tang. (2010). Real-time train rescheduling in junction areas. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 224(6), pages 547-557.
- [7] A. Galapitige, A. R.Albrechta, P. Pudneyab, X. Vua et P. Zhoua (2017). Optimal real-time junction scheduling for trains with connected driver advice systems. *Journal of Rail Transport Planning & Management*.
- [8] S. Van Thielen, S. Burggraeve, and P. Vansteenwegen (2015). "Optimal train rescheduling after conflict detection". *Conference on Advanced Systems in Public Transport (CASPT)*
- [9] J. Damay, A. Boillot, (2014). "An efficient macroscopic railway system simulator". *20th Conference of the International Federation of Operational Research Societies (IFORS)*.

Remerciements

Ces travaux ont été financés par SNCF Transilien et ont été menés par la Direction de l'Innovation de la SNCF. Nous tenons à exprimer nos remerciements à toutes personnes impliquées dans le projet : le comité de pilotage, le Lab Mass Transit, la DSI de SNCF Mobilité, les opérateurs et les responsables du centre opérationnel où nous avons testé l'outil, les équipes de développements du logiciel, les fournisseurs des données de géolocalisation et de plans de transport.

Démonstrations et Posters

JOE : le robot-compagnon des enfants asthmatiques

Brage Anthony^{1,2}

Jean-Daubias Stéphanie¹

Loisel Elodie²

Basset Thierry²

¹ Université Lyon 1, LIRIS, UMR5205, F-69622, France

² LUDOCARE, 36 quai Fulchiron, Lyon, F-69005, France

elodie.loisel@ludocare.com

Résumé

JOE est un robot-compagnon connecté à destination des enfants asthmatiques visant à les aider à prendre leur traitement. "Autonomie de l'enfant" et "sérénité des parents" sont les maîtres mots des créateurs de ce robot : la start-up LUDOCARE [1]. Dans cette démonstration, nous nous proposons de présenter JOE et son fonctionnement en insistant sur les mécanismes de traçage et de personnalisation qui visent à rendre JOE plus intelligent. Nous détaillons comment s'effectue une prise de traitement pour l'enfant avec JOE, depuis l'alarme qui rappelle le traitement jusqu'à la personnalisation de la récompense donnée pour clôturer la prise. Nous expliquerons également comment sont exploitées les traces des interactions entre JOE et son jeune utilisateur.

Mots-clefs

Intelligence Artificielle, Asthme, Personnalisation, Récompenses, Analyse, Traces, Robot-Compagnon

Abstract

JOE is a connected companion-robot which aims to help young asthmatic children to take their medicines. "Child autonomy" and "parents serenity" are the key words of the creators of this small robot : LUDOCARE start-up [1]. In this demonstration, we will present how JOE works with a focus on the tracing system and the personalization that aim at making JOE smarter. Then we will explain all the steps of a medical treatment with JOE, from the alarm to the personalization of the reward given that ends the treatment, thanks to the trace-based analysis.

Keywords

Artificial Intelligence, Asthma, Personalization, Reward, Analysis, Traces, Companion-Robot

1 Introduction

L'asthme est une maladie qui touche un grand nombre de personnes, notamment des enfants : en moyenne un enfant sur dix est concerné. Des traitements efficaces existent, mais leur observance, c'est-à-dire le respect par le patient du traitement proposé par le médecin, n'est très souvent

pas bonne. Des études [2] [3] ont montré qu'en moyenne 60% des asthmatiques ne prennent pas régulièrement, ou prennent mal, voire pas du tout leur traitement de fond. En effet, nombreux sont les malades qui omettent volontairement ou involontairement leur prise de traitement sous prétexte que leur état actuel n'en montre pas le besoin. Or le traitement de fond doit justement être pris régulièrement et sans interruption pour être efficace. C'est de ce constat qu'est né JOE, un robot-compagnon qui vise à aider des enfants asthmatiques dès 3 ans à prendre rigoureusement leur traitement, de manière autonome. L'enjeu est double : améliorer l'observance des traitements de fond des enfants et rendre l'enfant plus autonome car les parents confient à JOE une partie de la gestion quotidienne de la maladie de leur(s) enfant(s). Dans cette démonstration, nous allons présenter JOE, en particulier le déroulement d'une prise de traitement, tout en expliquant son fonctionnement.



FIGURE 1 – JOE, le robot-compagnon

1.1 Les objets connectés pour la santé

Les objets connectés sont de plus en plus présents dans nos vies [4]. De par ses fonctionnalités, JOE entre dans la catégorie des objets connectés pour la santé. Un état de l'art détaillé dédié à ce type d'objets a été proposé par Islam et al [5]. Parmi les objets connectés pour la santé, certaines applications sur smartphone visent déjà à aider les personnes asthmatiques. Par exemple, ResApp [6] cherche à détecter et analyser la toux, elle pourrait être utilisée pendant le sommeil du malade pour prévenir une poten-

tielle crise. D'autres applications s'inscrivent spécifiquement dans la lutte contre l'asthme. Ainsi, WheezMe [7] se concentre principalement sur la communication entre l'enfant, ses parents, et les médecins pour permettre un suivi plus précis. L'application propose aussi d'autres fonctionnalités comme des vidéos de prise de médicaments ou des informations de sensibilisation sur l'asthme. Asthm'Activ [8] quant à elle est principalement destinée aux adultes et propose, en plus de la possibilité de programmer ses prises de médicaments, un suivi des symptômes et crises pour tenter ensuite de prévenir ces dernières. Enfin, Asthma Australia [9] est un bon exemple d'application informative. Principalement prévue pour être utilisée par les patients en milieu médical, elle fournit les informations nécessaires sur la maladie et les traitements associés. Aucune de ces applications ne vise directement à rendre l'enfant autonome. Les parents gèrent entièrement les traitements de l'enfant. Meyko [10] en revanche, comme JOE, est un compagnon connecté visant à rendre l'enfant plus autonome. Son objectif est de notifier l'enfant de l'heure de ses prises de médicaments, l'oubli étant la cause majeure de défaut d'observance. JOE va plus loin dans l'interaction avec l'enfant : il rappelle les moments des traitements ; il guide l'enfant dans les gestes à faire avec ses médicaments (les mésusages étant très fréquents) et il motive l'enfant sur le long terme, grâce à des mécanismes de récompenses personnalisés et évolutifs.

2 Architecture de JOE

JOE (cf. figure 1) ne possède pas les caractéristiques propres à un robot au sens habituel du terme. En effet, il ne se déplace pas et ne bouge pas les bras par exemple. Pourtant, "robot" est le terme que les enfants utilisent spontanément pour le qualifier dès leur première rencontre. Du point de vue de ses concepteurs, JOE est un "compagnon" dont l'objectif est d'aider l'enfant à prendre son traitement. Sa force réside dans son adaptation à l'enfant. Cette adaptation est rendue possible par l'analyse des traces laissées par son jeune utilisateur lors de ses interactions avec l'écran tactile qui incarne JOE. Associer le point de vue de ses utilisateurs et le point de vue de ses concepteurs, nous conduit à qualifier JOE de robot-compagnon.

2.1 L'écosystème de JOE

JOE est l'élément principal d'un environnement qui lie l'enfant, ses parents et la société LUDOCARE. Ce sont les parents qui, par le biais d'une application mobile, configurent le JOE de leur enfant. Ils spécifient les heures des prises de traitement et les médicaments nécessaires en accord avec l'ordonnance du médecin. L'enfant, quant à lui, interagit avec JOE lors des prises, puis obtient une récompense. Les interactions entre JOE et l'enfant sont tracées, qu'elles concernent ou non la prise de médicaments. Ces traces sont exploitées de trois façons : pour proposer des récompenses adaptées à l'enfant, pour faire un retour aux parents sur le déroulement de la prise, et pour établir des

statistiques (concernant la pertinence des récompenses et la bonne prise des traitements), utiles à LUDOCARE. C'est l'ensemble de cet écosystème qui garantit le bon fonctionnement de JOE.

2.2 JOE, un compagnon connecté

JOE est relié à Internet via la connexion wifi de la maison de son jeune utilisateur. Cette connexion est nécessaire pour que JOE puisse remplir certaines de ses fonctionnalités. Par exemple, pour connaître les prises de traitements configurées par les parents, JOE doit se connecter pour les récupérer auprès d'un serveur. Il utilisera également parfois cette connexion pour renouveler son catalogue de récompenses à proposer, ainsi que pour faire remonter les traces de l'activité de l'enfant et donc des informations aux parents.

3 Fonctionnement de JOE

Une prise de traitement avec JOE comporte plusieurs étapes que l'on retrouve dans la Figure 2 ci-dessous.

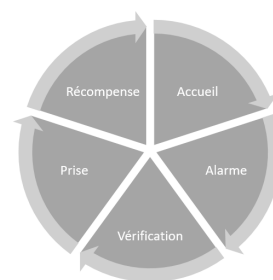


FIGURE 2 – Etapes des prises de traitement avec JOE

L'interface principale de JOE est l'écran qui représente ses yeux de manière animée et interactive. L'enfant peut librement interagir avec lui. Au moment programmé de la prise de médicaments, JOE déroule plusieurs étapes qui s'enchaînent interactivement au gré des actions de l'enfant : alarme, vérification des médicaments, prise du traitement et enfin récompense en cas de réussite, puis retour à l'écran principal jusqu'à la prochaine prise de traitement.

3.1 Alarme

L'alarme est le processus qui notifie l'enfant de l'heure de sa prise de traitement. À l'heure prévue configurée en amont par les parents dans l'application dédiée, JOE sonne pendant un délai prédéfini. L'enfant doit interagir avec JOE via l'écran tactile pendant ce délai pour accéder à l'étape suivante. Si ce n'est pas le cas, les parents seront notifiés qu'il n'a pas pris son traitement via leur application mobile.

3.2 Vérification des médicaments

La phase de vérification est l'étape où JOE mentionne tout ce dont l'enfant va avoir besoin pour cette prise de médicaments. Pour chaque médicament, l'enfant est invité à aller chercher le matériel, puis à confirmer ou non qu'il l'a

bien trouvé. Il est ainsi possible d'identifier facilement un problème et d'en informer les parents dans le cas où l'enfant n'a pas pu réunir tout ce dont il avait besoin pour cette prise. Si tout s'est bien passé, JOE passe à la phase de la prise de traitement.

3.3 Prise de traitement

Dans l'étape de prise de traitement, JOE propose un guide animé adapté aux spécificités de chaque traitement. JOE explique comment doit être pris le traitement, ainsi que les éventuelles manipulations à effectuer avant la prise (mettre le masque du nébuliseur par exemple). Dans le cas de traitements qui doivent être pris pendant une période donnée (par exemple respirer 3 minutes avec le nébuliseur), JOE joue le rôle de métronome pour rythmer la prise via l'animation de ses yeux, complétée par l'affichage d'un chronomètre et via la diffusion de sons. Dans le cas de traitements qui doivent être pris à plusieurs reprises (par exemple 4 bouffées d'aérosol), JOE affiche un compteur. À la fin de la prise de médicament, JOE demande à l'enfant s'il a bien réussi à prendre son traitement. S'il répond positivement, JOE le récompense, sinon les parents sont informés du problème rencontré. LUDOCARE a fait le choix d'une validation de la prise par auto-déclaration de l'enfant dans une démarche de responsabilisation de ce dernier sous la responsabilité des parents.

3.4 Récompense

JOE propose à l'enfant une récompense une fois qu'il a réussi à prendre tout son traitement correctement. Il peut s'agir d'une blague, d'une vidéo, d'une histoire, d'une chanson, ou encore d'une anecdote. L'enfant visualise la récompense attribuée automatiquement par le système, et, à la fin de la récompense, il a la possibilité de noter cette dernière sur une échelle de 1 à 5. Cette notation sera par la suite utilisée pour adapter au mieux les prochaines récompenses. Une fois cette étape terminée, JOE revient dans son état principal, l'écran d'accueil, signifiant ainsi la fin du processus de prise de traitement.

4 JOE, un compagnon intelligent

JOE est un tout jeune compagnon, aux fonctionnalités encore restreintes, mais de nombreuses possibilités vont venir s'ajouter progressivement à ses capacités actuelles afin de le rendre progressivement plus intelligent.

4.1 Spécification des traces

Comme mentionné plus haut, JOE utilisera certaines des traces issues des interactions de l'enfant avec lui pour mettre en place de la personnalisation. La prochaine version de JOE intégrera en particulier le mécanisme de traçage des interactions de l'enfant, mais également du fonctionnement de JOE, pour pouvoir connaître le contexte dans lequel l'enfant interagit. Ainsi, les traces produites comporteront des informations provenant de deux sources

différentes :

- L'enfant : toutes ses interactions (touché de l'écran à un seul ou à deux doigts, balayage), qu'elles conduisent ou non au déclenchement d'une réaction de JOE seront tracées. Les coordonnées des interactions sont également tracées, pour permettre une analyse des zones les plus actives par la suite.

- JOE : en plus des interactions avec l'enfant, chaque changement d'état de JOE sera tracé. Par exemple, dès qu'il est l'heure de prendre son traitement, JOE passe de sa phase de "veille" à sa phase "d'alerte".

Ces traces pseudonymisées seront utilisées en local, par JOE, pour améliorer l'adéquation des récompenses proposées à l'enfant. Puis par l'application parents pour fournir à ces derniers, le suivi du traitement de l'enfant. Et enfin, une fois anonymisées, elles seront utilisées par LUDOCARE à des fins statistiques.

4.2 Utilisation des traces dans JOE

L'exploitation des traces dans JOE a deux objectifs.

Dans un premier temps, l'analyse portera sur les traces liées aux récompenses que JOE a données à l'enfant, et aux notes que l'enfant leur a attribuées. Les informations résultantes permettront d'alimenter un système de recommandation qui sera le coeur de la personnalisation des récompenses : il permettra de donner des récompenses qui correspondront au mieux à ce que préfère l'enfant. Ce procédé devrait permettre d'engager davantage l'enfant vis-à-vis de JOE et de son traitement dans la mesure où le compagnon répondra mieux à ses attentes. Il s'agit ainsi d'une analyse ponctuelle qui intervient après chaque récompense que l'enfant a noté pour prévoir la prochaine.

À plus long terme, il sera possible d'utiliser les traces pour modifier le comportement de JOE. Concrètement, JOE pourrait réagir différemment si on se rend compte que l'enfant lui tape souvent dans l'oeil. Plutôt que de simplement déclencher une animation d'oeil qui se ferme, il pourrait déclencher une animation différente.

4.3 Utilisation des traces pour les parents

Une seconde analyse des traces, cette fois à destination des parents, a pour but d'extraire des informations concernant la prise des médicaments par l'enfant. L'idée est que les parents par le biais d'indicateurs (qui restent à définir), puissent savoir si l'enfant a bien pris son traitement (ou s'il il a oublié ou n'était pas là), s'il a bien réussi (ou dans le cas contraire, la raison de l'échec), les récompenses qu'il a bien aimées, ... L'important ici est de fournir un suivi efficace pour visualiser rapidement si quelque chose ne va pas. Par exemple, un calendrier regroupant l'ensemble des prises faites selon un code couleur permettrait de voir lorsqu'une prise n'a pas été réussie. Ce calendrier pourrait être utilisé comme information par les parents pour avoir une vue globale de la prise de traitement sur le long terme. Cette analyse aura lieu côté serveur et en continu pendant les prises de traitement pour que les parents aient des informations en temps réel. A terme, ces informations pour-

raient être communiquées au médecin.

4.4 Utilisation des traces pour LUDOCARE

L'exploitation des traces pour LUDOCARE est une analyse à plus large échelle. En effet, contrairement aux deux usages précédents qui se basent uniquement sur les traces d'un enfant particulier, on utilise ici l'ensemble des traces de tous les enfants.

L'un des objectifs de cette analyse est d'avoir un suivi global des JOEs. Par le biais d'indicateurs qui alimentent un tableau de bord (consultable localement et uniquement par LUDOCARE), il est ainsi possible de connaître le nombre d'échecs lors des prises, le nombre total de prises, le nombre total et moyen d'interactions, le temps moyen d'alarme, ... (cf. figure 3). Il s'agit d'une analyse plutôt statistique qui vise à fournir aux dirigeants de LUDOCARE des données directement exploitables concernant l'utilisation des JOEs, mais pas seulement, vu qu'elle permettrait également de détecter les mésusages des médicaments.

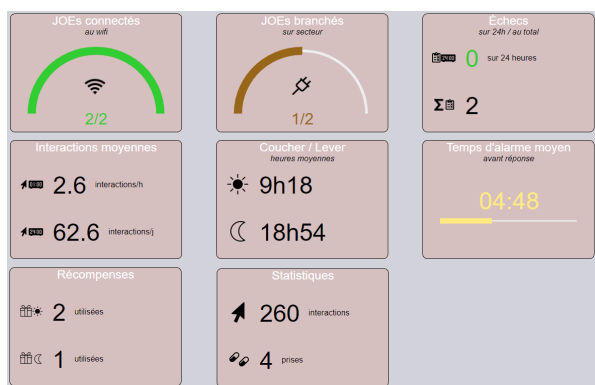


FIGURE 3 – Indicateurs du tableau de bord LUDOCARE

L'autre objectif de cette analyse est de permettre aux dirigeants de LUDOCARE d'effectuer des fouilles de données dans les traces par le biais d'une plate-forme d'analyse dédiée (ktBS¹ [11]). Le ktBS permet la modélisation et la transformation des traces en traces de plus haut niveau, exploitables pour identifier des nouveaux besoins ou confirmer que JOE répond bien aux attentes par exemple.

5 Conclusion

La solution connectée qu'est JOE aspire à améliorer l'observance des traitements de fond de l'asthme chez les jeunes enfants. Pour cela, l'intelligence artificielle va jouer un rôle croissant. L'un des enjeux est la mise en place d'une personnalisation toujours plus pertinente pour l'enfant, dans le but d'améliorer la motivation et l'engagement de l'enfant vis-à-vis de JOE et par conséquent vis-à-vis de son traitement.

Même si JOE n'en est qu'à une phase expérimentale très prometteuse, il est attendu avec impatience par nombre de familles. Et les nombreuses pistes de recherche envisagées laissent entrevoir un intérêt et une utilité grandissants.

1. kernel Trace-Based System

L'utilisation de procédés de gamification [12] pour renforcer l'engagement de l'enfant vis-à-vis de JOE est envisagée. Par exemple, une boutique pourrait permettre d'acheter des récompenses inédites grâce à une monnaie qu'il serait possible d'acquérir à chaque bonne prise de traitement. Notons que chez LUDOCARE, les familles jouent un rôle important dans le processus de conception et d'amélioration continue de JOE. Les expérimentations en situation réelle permettent d'obtenir des informations précieuses qui permettront d'améliorer JOE pour en faire le meilleur compagnon possible pour améliorer l'observance des traitements de fond des jeunes asthmatiques.

Références

- [1] LUDOCARE, Projet JOE.
<http://www.projetjoe.com/>
- [2] Chan, Y. F. Y., Wang, P., Rogers, L., Tignor, N., Zweig, M., Hershman, S. G., ... & Edgar, R. (2017). *The Asthma Mobile Health Study, a large-scale clinical observational study using ResearchKit*. *Nature biotechnology*, 35(4), 354
- [3] Delmas, M. C., & Fuhrman, C. (2010). *L'asthme en France : synthèse des données épidémiologiques descriptives*. *Revue des maladies respiratoires*, 27(2), 151-159.
- [4] Amy Nordrum (2016) *Popular Internet of Things Forecast of 50 Billion Devices by 2020 Is Outdated*
- [5] S. M. R. Islam and D. Kwak and M. H. Kabir and M. Hossain and K. S. Kwak *The Internet of Things for Health Care : A Comprehensive Survey*, IEEE Access, Vol. 3, pp 678-708, 2015
- [6] ResApp,
<https://www.resapphealth.com.au/>
- [7] WheezMe,
<https://www.wheezme.com/>
- [8] Asthm'Activ,
<https://play.google.com/store/apps/details?id=com.sophia.asthme>
- [9] Asthma Australia,
<https://itunes.apple.com/au/app/asthma-australia-asthma-app/id1057041779?mt=8>
- [10] Meyko,
<http://www.hellomeyko.com/fr/>
- [11] Kernel Trace-Based System
<https://projet.liris.cnrs.fr/sbt-dev/tbs/doku.php/tools:ktbs>
- [12] Nah F.FH., Zeng Q., Telaprolu V.R., Ayyappa A.P., Eschenbrenner B. (2014) *Gamification of Education : A Review of Literature*. In : Nah F.FH. (eds) *HCI in Business*. HCIB 2014. Lecture Notes in Computer Science, vol 8527. Springer, Cham

SmartCRF, un prototype pour visualiser, rechercher et annoter les informations d'un dossier patient dans I2B2

Cossin Sébastien^{1,2}, Jouhet Vianney^{1,2}, Lebrun Luc², Niamkey Aymerick²,
Mougin Fleur¹, Lambert Mathieu³, Diallo Gayo¹, Thiessard Frantz^{1,2}

¹ EQUIPE DE RECHERCHE EN INFORMATIQUE APPLIQUÉE À LA SANTÉ (ERIAS), INSERM BPH U1219, Université de Bordeaux, F-33000 Bordeaux, France
sebastien.cossin@u-bordeaux.fr

² Informatique et Archivistique Médicales (IAM), Service d'Information Médicale, Pôle de Santé Publique, CHU de Bordeaux, France

³ Service de médecine post-urgences, CHU de Bordeaux.

Résumé : Dans cet article, nous présentons SmartCRF, un prototype développé dans le cadre de l'entrepôt de données du CHU de Bordeaux permettant de visualiser, rechercher et annoter des informations d'un dossier patient informatisé. Cet outil permet d'aider les chercheurs à recueillir plus rapidement les données dont ils ont besoin pour une étude clinique. Les annotations réalisées pourront servir à entraîner des algorithmes d'apprentissage automatique pour faciliter la réutilisation secondaire des données.

Mots-clés : Electronic Health Records, User-Computer Interface, Information Storage and Retrieval, Semantics, Natural Language Processing, Data Warehousing, Data Curation

1 Contexte

L'entrepôt de données biomédicales du CHU de Bordeaux a été mis en place pour permettre la réutilisation secondaire des données du système d'information hospitalier (SIH) à des fins de recherche. Il intègre de nombreuses sources d'information d'un dossier patient informatisé (Programme de Médicalisation des Systèmes d'Information (PMSI), examens de biologie, prescriptions médicamenteuses, formulaires de saisie, compte-rendus de consultation et d'hospitalisation...).

Dans le cadre d'études rétrospectives menées sur ces données, l'une des étapes essentielle et préalable à l'analyse statistique est de recueillir et de structurer les informations pour chaque patient inclus. Cette collecte de données est chronophage et complexe en l'absence d'interfaces et d'outils dédiés (Pan & Cimino, 2014).

Des outils d'indexation sémantique, de recherche et de visualisation peuvent aider les médecins et les chercheurs à trouver rapidement des informations dans un dossier patient informatisé (DPI) (Thiessard *et al.*, 2012).

Notre objectif a été de réaliser un prototype fonctionnel permettant de visualiser l'ensemble des données d'un DPI, mais également de rechercher et de collecter des données pour une étude.

2 Fonctionnement

L'entrepôt du CHU de Bordeaux utilise la solution open source I2B2 (Murphy *et al.*, 2006). I2B2 possède un modèle de base de données relationnelles en étoile. Dans ce modèle, chaque donnée est reliée à un patient, à une venue à l'hôpital (consultation ou hospitalisation) et à une source de données (PMSI, biologie...). Chaque donnée possède une date (quand la donnée a été enregistrée) et un type (textuel ou structuré).

Le prototype a été développé avec le paquet Shiny du langage de programmation R. Ce dernier permet de réaliser des interfaces web interactives avec des bibliothèques JavaScript.

Pour visualiser un dossier patient, les données sont d'abord chargées en mémoire côté serveur (RShiny) en interrogeant I2B2. Quatre fonctionnalités, présentées ci-dessous, permettent d'interagir avec les données : une timeline, un moteur de recherche, un nuage de mots et un module d'annotation des éléments d'intérêt.

2.1 Timeline

Les données d'un patient sont affichées sur une timeline interactive permettant d'obtenir une vue d'ensemble des données d'un patient (figure 1).

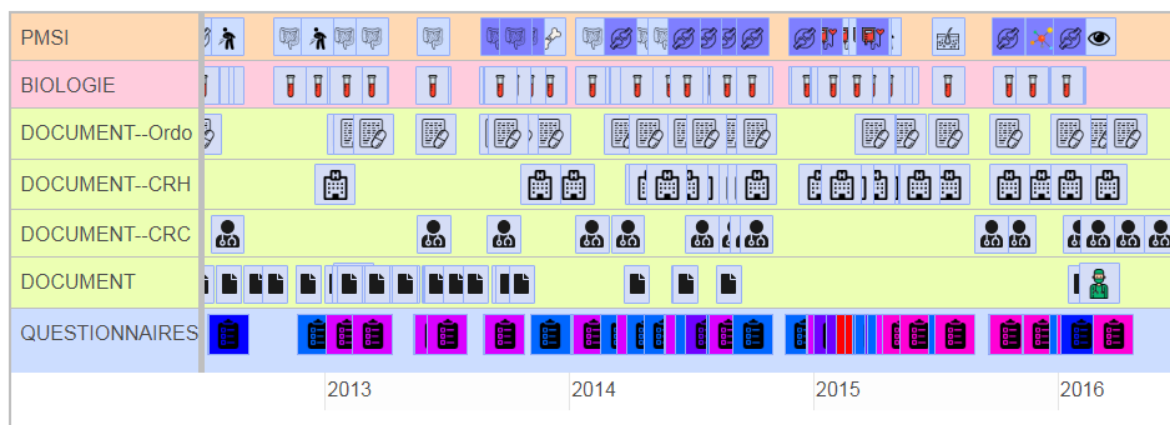


FIGURE 1 – Timeline interactive permettant de visualiser l'ensemble des données d'un patient. Chaque icône est cliquable afin d'afficher les données associées à l'événement sélectionné.

La timeline contient plusieurs groupes (PMSI, biologie...) qui correspondent à différentes sources de données. Chaque icône de la timeline contient une à plusieurs données du patient. Par exemple, tous les résultats biologiques d'une prise de sang sont regroupés sous une même icône. Chaque icône est cliquable permettant d'afficher le détail de son contenu.

L'utilisateur peut ajouter dynamiquement des sous-groupes à la timeline. Il peut par exemple ajouter le sous-groupe "plaquettes" au groupe "Biologie" pour visualiser rapidement les résultats des plaquettes au cours du temps.

2.2 Moteur de recherche

Les syntagmes nominaux sont extraits par un programme Java. Ils correspondent aux principales informations d'un document textuel (Kathait *et al.*, 2017). Ceux-ci sont normalisés (transformation du texte en minuscule, retrait des accents et des caractères spéciaux) et lemmatisés avec TreeTagger (Schmid, 1997) avant d'être indexés avec ElasticSearch™. L'auto-complétion du moteur de recherche utilise ces syntagmes nominaux pour guider l'utilisateur dans sa recherche (figure 2). Les résultats d'une recherche sont affichés par ordre antéchronologique. Seule la phrase contenant le terme recherché est affichée et l'utilisateur peut accéder au document dans son intégralité s'il le souhaite.

2.3 Nuage de mots

Certains concepts comme les symptômes, les maladies et les médicaments sont détectés dans le texte libre en utilisant des terminologies issues de l'UMLS (Bodenreider, 2004). Les concepts détectés sont affichés à l'utilisateur dans un nuage de mots cliquable (figure 3). Un clic conduit à placer le terme dans la barre de recherche précédente et à trouver les documents le mentionnant.

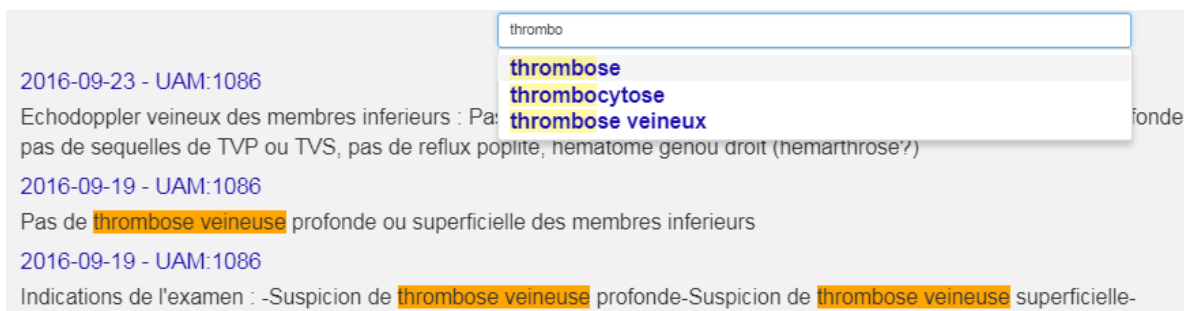


FIGURE 2 – Moteur de recherche avec autocomplétion (syntagmes nominaux normalisés et lemmatisés) permettant de rechercher des informations dans les données en texte libre d'un dossier patient informatisé. Seule la phrase contenant le terme recherché est affichée et l'utilisateur peut accéder au document dans son intégralité. UAM : service hospitalier



FIGURE 3 – Nuage de mots permettant de visualiser les symptômes et les maladies détectés dans les documents en texte libre. Un clic sur un terme permet de lancer une recherche.

2.4 Annotation

L'utilisateur peut surligner les informations pertinentes dans le dossier pour tracer et enregistrer les éléments lui ayant permis de structurer l'information. Par exemple, il enregistre le statut tabagique "non fumeur" et justifie son choix en sélectionnant le passage "pas de tabagisme" dans le dossier (figure 4). Ces annotations pourront être utilisées par la machine pour apprendre à trouver l'information du statut tabagique dans d'autres dossiers.

3 Discussion

Le prototype est en phase de tests avec des utilisateurs.

Nous réfléchissons à la manière d'implémenter un système de recommandation visant à présenter les éléments d'information susceptibles d'intéresser le chercheur.

Par exemple, si le chercheur a besoin de collecter l'information sur le statut tabagique, le

Variable : Statut tabagique patient numero : 123456789

Choix :

Fumeur actif	Fumeur sévère	✓ Non Fumeur
--------------	---------------	--------------

Annotations :

Pas de tabagisme

📁 Sauvegarder

FIGURE 4 – L'utilisateur classe le patient comme "Non Fumeur" et enregistre les éléments du dossier lui ayant permis de faire son choix ("Pas de tabagisme").

nuage de mots devrait afficher les termes en lien avec ce concept (tabac, tabagisme, cigarette, paquet-année, nicotine...) trouvés dans le dossier et les éléments de la timeline concernés pourraient être mis en exergue.

Le système de recommandation devra apprendre à rechercher et à visualiser les informations pertinentes en fonction des annotations déjà réalisées. Le chercheur sera enclin à annoter les termes pertinents à sa recherche pour que la machine l'aide dans ses recherches futures, créant ainsi une interaction homme-machine vertueuse.

Un jeu de données annotées dans le cadre d'une étude pourra servir à entraîner des algorithmes de classification, comme par exemple prédire le statut tabagique d'un patient. Les annotations par un expert du domaine étant habituellement coûteuses à obtenir, le recueil de données dans le cadre d'études cliniques offre une formidable opportunité pour entraîner les algorithmes de machine learning. Ces algorithmes entraînés pourront être utilisés pour faciliter la réutilisation secondaire des données de l'entrepôt de données du CHU de Bordeaux.

Références

- BODENREIDER O. (2004). The Unified Medical Language System (UMLS) : integrating biomedical terminology. *Nucleic Acids Research*, **32**(Database issue), D267–270.
- KATHAIT S. S., TIWARI S., VARSHNEY A., SHARMA A., MESTROVIC A., GELBUKH A. F., RENDON E., GARCÍA-HERNÁNDEZ, DA CRUZ R. P., ARNULFO R., MONTIEL R., LEDENEVA Y., GUTWIN C., NEVILL-MANNING C. G., FRANK E., PAYNTER G. W., HASAN K. S., NG V., CARAGEA C., GOLLAPALLI S. D., ALHAKIM A., DUTTA B., GIUNCHIGLIA F., LOPER E., KLEIN E., GODEA A. & BULGAROV F. A. (2017). Unsupervised key-phrase extraction using noun phrases. *International Journal of Computer Applications*, **162**.
- MURPHY S. N., MENDIS M. E., BERKOWITZ D. A., KOHANE I. & CHUEH H. C. (2006). Integration of clinical and genetic data in the i2b2 architecture. *AMIA ... Annual Symposium proceedings. AMIA Symposium*, p. 1040.
- PAN X. & CIMINO J. J. (2014). Locating Relevant Patient Information in Electronic Health Record Data Using Representations of Clinical Concepts and Database Structures. *AMIA Annual Symposium Proceedings*, **2014**, 969–975.
- SCHMID H. (1997). Probabilistic part-of-speech tagging using decision trees. In D. JONES & H. SOMERS, Eds., *New Methods in Language Processing*, Studies in Computational Linguistics, p. 154–164. London, GB : UCL Press.
- THIESSARD F., MOUGIN F., DIALLO G., JOUHET V., COSSIN S., GARCELON N., CAMPILLO B., JOUINI W., GROSJEAN J., MASSARI P., GRIFFON N., DUPUCH M., TAYALATI F., DUGAS E., BALVET A., GRABAR N., PEREIRA S., FRANDJI B., DARMONI S. & CUGGIA M. (2012). RAVEL : retrieval and visualization in EElectronic health records. *Studies in Health Technology and Informatics*, **180**, 194–198.

Humanités numériques : un cas pratique avec l'identification de hiéroglyphes

S. Madeleine¹

S. Maillot¹

S. Thomas¹

M. Servajean^{1 2}

¹ Université Paul Valéry Montpellier, Route de Mende, Montpellier 34199, France
{sarah.madeleine, sylvain.maillot, sara.thomas}@etu.univ-montp3.fr

² LIRMM, 860 rue de St Priest, Montpellier 34095, France
maximilien.servajean@lirmm.fr

Résumé

Dans le cadre du projet VEgA et des humanités numériques en général, nous nous intéressons à l'utilisation des méthodes de reconnaissance d'image par réseaux de neurones à convolutions et, plus particulièrement, à la classification de hiéroglyphes. Nous présentons ici le modèle de classification utilisé, l'interface graphique permettant son utilisation par le plus grand nombre ainsi que l'analyse des résultats expérimentaux. Bien que disposant de très peu d'images (environ 44 exemples par classe), notre modèle obtient un score supérieur à 90% en validation.

Mots Clef

Reconnaissance d'image, apprentissage automatique, réseaux de neurones à convolution, ResNet-18, hiéroglyphes égyptiens, humanités digitales.

Abstract

As part of the VEgA project, we are interested in the use of image recognition methods using convolutional neural networks in the context of digital humanities. We show in particular how classification can be performed on hieroglyphs. We present our classification model and the graphical user interface and we analyse our experimental evaluation. Although we only have few training examples (around 44 hieroglyphs per class), our model achieves more than 90% accuracy on the validation set.

Keywords

Image recognition, automatic learning, convolutional neural networks, ResNet-18, Egyptian hieroglyphs, digital humanities.

1 Introduction

Ces dernières années ont vu les performances de diverses tâches d'apprentissage statistique, notamment en classification d'images, exploser et ce, principalement grâce au renouveau de l'apprentissage profond [10]. Cette opportunité s'est initialement cantonnée aux sciences dites dures,

où la disponibilité des données nécessaires à l'apprentissage d'un modèle profond est plus immédiate.

De leur côté, les humanités numériques (ou *digital humanities*) [1] se sont tout d'abord manifestées par une révolution des usages avec des plateformes comme LexArt¹, dédiée au lexique artistique de 1600 à 1750, ou VEgA² (Figure 1), premier dictionnaire numérique de l'égyptien ancien. L'impact de tels projets se mesure autant dans la centralisation de l'information et que dans leur mise à disposition au travers d'interfaces innovantes.

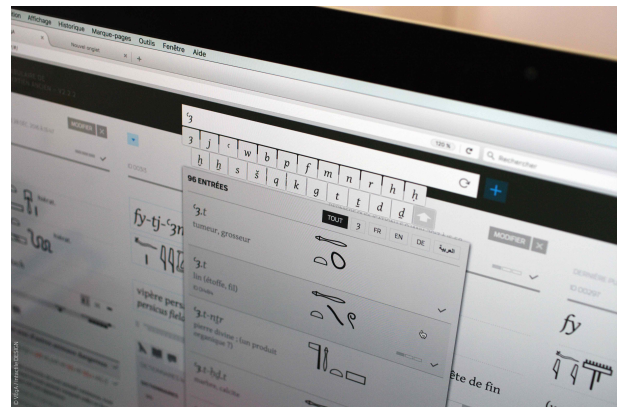


FIGURE 1 – Interface du dictionnaire VEgA.

Cependant, les techniques d'apprentissages statistiques (e.g. classification, régression) ne sont encore que très peu répandues dans ces disciplines qui, pourtant, pourraient en bénéficier [7]. En Egyptologie, objet d'étude de cette contribution, les hiéroglyphes peuvent ainsi être identifiés (i.e. classification) à partir d'images.

La classification consiste, à partir d'une donnée d'entrée x_i , à en prédire la classe (ou le label) y_i . Il s'agit d'une stratégie supervisée dans le sens où nous disposons d'un jeu d'apprentissage $D = \{(x_1, y_1), \dots, (x_N, y_N)\}$. Pour ce travail, x_i représente une image de hiéroglyphe et y_i la détermination (le nom) de ce dernier. Le cas particulier des

1. <http://lexart.fr>

2. <http://vega-vocabulaire-egyptien-ancien.fr>

images fait face à deux problèmes importants : la malédiction de la dimension et la notion de séparabilité linéaire. La grande dimension favorise le sur-apprentissage [8] et la non séparabilité linéaire [11] empêche tout classifieur linéaire de s'adapter correctement aux données.

Les réseaux de neurones profonds [11] jouent sur ces deux tableaux. Concernant la dimensionnalité, la succession des représentations des données (les couches du réseaux) ferait office de goulot d'étranglement de l'information [16] et l'algorithme d'optimisation stochastique SGD jouerait comme un régularisateur [2] limitant par là même les possibilités de sur-apprentissage. En outre, la succession d'activations non-linéaires permet de "tordre" l'espace d'entrée et de projeter les données dans un espace de dimension réduite où chaque classe devient séparable linéairement des autres. En omettant la dernière couche $g(\mathbf{x})$ de classification linéaire du modèle, l'ensemble du réseau peut ainsi être vu comme permettant l'apprentissage de *features* simples et séparables linéairement des données d'entrées. En d'autres termes, un modèle profond peut donc être vu comme la composée d'une fonction d'apprentissage de *features* f et d'une fonction de classification linéaire $g : (g \circ f)(\mathbf{x})$.

Malheureusement, de tels modèles nécessitent une très grande quantité de données, incomparable avec ce qui est réellement disponible dans le monde des humanités numériques. Ce travail montre qu'il est malgré tout possible, en se concentrant sur des techniques de l'état de l'art, d'obtenir d'excellentes performances (supérieures à 90% de précision en validation). Pour cela, il est par exemple nécessaire de s'appuyer sur la notion de *transfert learning*. En effet, la fonction de coût d'un réseau de neurones admet un grand nombre de minimum locaux. Ainsi, le point de départ de l'apprentissage a un impact sur le point de convergence. Le *transfert learning* s'appuie sur cette idée. En entraînant le modèle sur un jeu de données tiers conséquent, il apprend des *features* visuelles pertinentes et utiles pour la seconde tâche où la disponibilité des données est plus problématique. L'augmentation des données permet également de compenser cette limite en déformant aléatoirement les données d'entrées.

Afin de rendre utilisable notre modèle par le plus grand nombre, nous présentons également l'interface graphique qui l'accompagne. Celle-ci permet à la fois de lancer l'apprentissage d'un modèle lorsque de nouvelles données deviennent disponibles ; et d'identifier des hiéroglyphes par glisser/déposer.

La suite de ce papier se structure comme suit. La section 2 présente les travaux antérieurs. Notre modèle est introduit dans la section 3 et nous montrons ses résultats expérimentaux dans la section 4. Enfin nous concluons ce travail de recherche dans la section 5.

2 Travaux antérieurs

Les articles [4], [7], [9] et [12] font état des travaux de recherche dans le domaine de la reconnaissance des

symboles hiéroglyphiques par apprentissage automatique. Dans [9], Kriege et al. ont travaillé sur la reconnaissance de symboles cunéiformes à l'aide de deux méthodes basées sur la mesure de similarité de deux graphes : la méthode du plus proche voisin et celle d'un CNN. La première étant très coûteuse en termes de temps de calcul et nécessitant un volume important de données pour la phase d'apprentissage ; les auteurs se sont tournés vers un CNN. Bien que la démarche d'utiliser un CNN soit identique à la nôtre, leur modèle se base sur un a priori concernant la forme des symboles, ce qui n'est pas notre cas où nous nous appuyons sur les images brutes sans pré-traitement.

Dans [4], Duque-Domingo et al. proposent une méthode de décryptage des cartouches hiéroglyphiques basée sur les techniques d'analyse d'image et vision par ordinateur. Ils identifient notamment les hiéroglyphes présents dans une image grâce à la comparaison des distances de Chamfer et d'Hausdorff d'une image connue (i.e. une image-étalon) à celles de l'image étudiée. Bien que cette méthode s'affranchisse des biais induits par la luminosité de l'image ou par la texture du support du cartouche ; elle nécessite d'avoir à disposition autant d'images-étalons qu'il existe de symboles dans la langue étudiée, en incluant leur diversité. À l'inverse, notre modèle réalise la majorité de son apprentissage sur le jeu de données ImageNet. Cela nous permet de finaliser son apprentissage (fine-tuning) sur un jeu de données, spécifique aux symboles étudiés, restreint et non-exhaustif.

Dans [7] Hu et al. présente la méthode de l'histogramme du contexte de l'orientation de la forme pour identifier les symboles hiéroglyphiques. Cette méthode nécessite un important travail de traitement d'image en amont de l'identification par des épigraphistes pour faire disparaître le bruit. Tandis que notre modèle tire profit du bruit des images pour augmenter ses performances.

De nombreuses architectures ont été proposées et ont obtenu des résultats significatifs en classification d'images [3, 6, 10, 13, 14, 15, 17]. Il ressort que les réseaux de neurones à convolution avec des connexions résiduelles sont plus simples à entraîner et obtiennent de meilleurs résultats en terme d'apprentissage que les réseaux sans. Dans [13], les auteurs démontrent notamment que la profondeur du réseau a une influence sur sa précision. De même que dans [17] il est démontré que la méthode dite "*weighted-pooling*" obtient de meilleurs résultats que les méthodes de pooling traditionnelles, telles que *max-pooling*, *average pooling*.

Nous nous sommes servis de ces travaux pour adapter le modèle ResNet-18 aux contraintes inhérentes à la reconnaissance des hiéroglyphes égyptiens. Nous montrons la faisabilité des modèles CNN à partir du moment où un

minimum d'images sont disponibles. Nous montrons les limites de notre méthode lorsque la variabilité d'un signe est trop grande.

3 Proposition & démonstration

3.1 Jeu de données

Le jeu de données provient de deux sources : des signes du temple de Karnak³ et d'autres de la pyramide d'Ounas (M. Franken), représentant un total de 4193 symboles répartis en 88 catégories différentes. Cela représente la quantité extrêmement faible de 44 images par classe.

Chaque signe est labellisé à la main selon la liste de Gardiner. Les images de Karnak n'ont pas subi de pré-traitement. Les images du jeu de données de Morris Franken [5] sont recoupées en 50x75. Les pixels vides sont remplis avec une texture rappelant l'arrière-plan réel. Pour augmenter le jeu d'entraînement, les images sont recadrées aléatoirement et transformées par un flip horizontal.

3.2 Architecture

Implémentation PyTorch de ResNet-18 en utilisant des poids pré-entraînés sur le *dataset* ImageNet. Le modèle est constitué de 20 couches : 1 convolution 7x7, 16 convolution 3x3 et 3 couches connectées. La dernière couche donne en sortie une des 88 classes de signes possibles.

3.3 Optimisation

Plusieurs tests d'optimiseur ont été réalisés. Notre choix s'est porté sur Adam avec un *learning rate* initialisé à 0.0001 et divisé par 10 toutes les 7 epochs.

3.4 Activation

Chaque couche contient une opération de normalisation et d'activation. Cette dernière (*ReLU*) introduit une non-linéarité permettant de répondre plus efficacement à notre tâche de reconnaissance et classification d'image. La normalisation (*Batch Normalization*) augmente la stabilité et la vitesse d'apprentissage du modèle en normalisant les sorties des couches cachées à chaque itération.

Le modèle utilise la technique de *DropOut*, initialisée à 0.5, qui consiste en la désactivation temporaire et aléatoire de neurones pendant l'entraînement, puis en leur réactivation lors de la phase de test. Cela simule un apprentissage sur un ensemble de modèles différents. L'objectif est de prévenir les liens d'interdépendance qui pourraient se créer entre les neurones et ainsi éviter l'*overfitting*.

3.5 Entraînement

Le modèle est entraîné sur GPU (GTX 1080 ti) sur 100 epochs.

La fonction d'entropie croisée est utilisée pour évaluer la précision du modèle. La convergence est rapide et dépasse les 90% de précision.

3. <http://sith.huma-num.fr/karnak>

3.6 IHM

Afin de rendre possible l'utilisation de notre travail par le plus grand nombre, nous avons développé une interface graphique à l'aide de la librairie Python PYQT5.

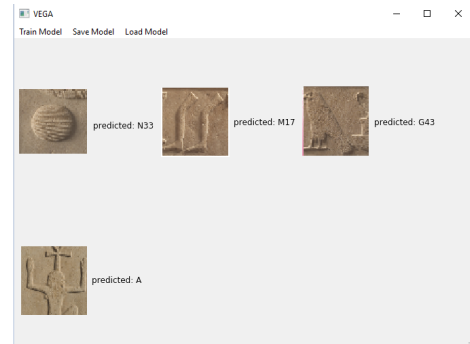


FIGURE 2 – Maquette de l'interface graphique.

Elle est composée d'un bandeau de navigation et d'une zone centrale. Trois onglets composent la barre de navigation : *Train model*, *Save model* et *Load model*. Le premier onglet permet de calculer un nouveau modèle (e.g. quand de nouvelles images sont disponibles et viennent enrichir le jeu de données d'apprentissage). À partir du second onglet l'utilisateur peut sauvegarder le modèle pour une réutilisation future. Enfin le chargement d'un modèle déjà existant se fait par le troisième et dernier onglet.

La zone centrale quant à elle, initialement vide, permet d'utiliser la classification. Elle s'utilise en y glissant des images contenant des hiéroglyphes inconnus pour ensuite afficher le résultat.

L'application permet de charger un modèle n'ayant pas nécessairement de lien avec les hiéroglyphes.

4 Évaluation expérimentale

En nous appuyant sur l'état de la recherche dans le domaine de la reconnaissance d'image par apprentissage automatique et des résultats de nos tests, nous sommes partis d'un modèle ResNet-18. Nous avons ensuite choisi les éléments qui correspondaient au mieux à nos besoins pour composer notre modèle final :

- Couches de convolution : 1 taille 7x7, 16 taille 3x3 ;
- Couche de pooling : Max pool et Average pool ;
- Couche de correction : Batchnorm, ReLU ;
- Couche de perte : Cross entropy ;
- Couche connectée : linéaire ;
- Optimiseur : Adam.

Notre modèle est entraîné pendant 100 epochs. Il converge rapidement (i.e. vingtième epoch) vers une valeur limite supérieure à 90% de précision sur le jeu de validation.

La précision de notre modèle est meilleure en phase de validation que lors de celle d'apprentissage. Nous expliquons cela par les transformations infligées au jeu

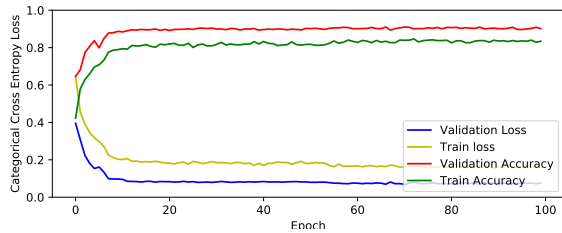
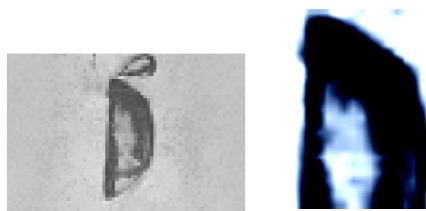


FIGURE 3 – Performances du modèle.

de données d'apprentissage, relativement importantes et complexifiant ainsi la tâche de classification. À l'inverse, le jeu de validation ne subit aucune transformation.



(a) Sans transformation. (b) Avec transformation.

FIGURE 4 – Image d'un signe V31.

5 Conclusion

Notre travail de recherche évalue les possibilités d'utilisation de modèles profonds pour la reconnaissance et la classification automatiquement des hiéroglyphes. Le déploiement d'un modèle résiduel simple à 20 couches apporte des résultats concluants. Il est cependant important de noter que l'efficacité des modèles profonds est largement impactée par la volumétrie et la variété (que l'on retrouverait en test) du jeu de données d'apprentissage.

Notre modèle converge très rapidement. Un modèle plus profond apporterait de meilleurs résultats de classification mais demanderait plus de temps d'entraînement. L'IHM est simple et pourrait à l'avenir proposer d'autres fonctionnalités.

Références

- [1] David M Berry and Anders Fagerjord. *Digital humanities : knowledge and critique in a digital age*. John Wiley & Sons, 2017.
- [2] Pratik Chaudhari and Stefano Soatto. Stochastic gradient descent performs variational inference, converges to limit cycles for deep networks. *CoRR*, abs/1710.11029, 2017.
- [3] François Chollet. Xception : Deep learning with depthwise separable convolutions. *arXiv preprint*, 2016.
- [4] Jaime Duque-Domingo, Pedro Javier Herrera, Enrique Valero, and Carlos Cerrada. Deciphering egyptian hieroglyphs : towards a new strategy for navigation in museums. *Sensors*, 17(3) :589, 2017.
- [5] Morris Franken and Jan C. van Gemert. Automatic egyptian hieroglyph recognition by retrieving images as texts. In *Proceedings of the 21st ACM International Conference on Multimedia*, MM '13, pages 765–768, New York, NY, USA, 2013. ACM.
- [6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [7] Rui Hu, Carlos Pallán Gayol, Jean-Marc Odobez, and Daniel Gatica-Perez. Analyzing and visualizing ancient maya hieroglyphics using shape : From computer vision to digital humanities. *Digital Scholarship in the Humanities*, 32(suppl_2) :ii179–ii194, 2017.
- [8] Eamonn Keogh and Abdullah Mueen. Curse of dimensionality. In *Encyclopedia of Machine Learning and Data Mining*, pages 314–315. Springer, 2017.
- [9] Nils M Kriege, Matthias Fey, Denis Fisseler, Petra Mutzel, and Frank Weichert. Recognizing cuneiform signs using graph based methods. *arXiv preprint arXiv :1802.05908*, 2018.
- [10] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [11] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553) :436, 2015.
- [12] Edgar Roman-Rangel and Stephane Marchand-Maillet. Assessing deep learning architectures for visualizing maya hieroglyphs. In *Mexican Conference on Pattern Recognition*, pages 137–146. Springer, 2017.
- [13] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv :1409.1556*, 2014.
- [14] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alexander A Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. In *AAAI*, volume 4, page 12, 2017.
- [15] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich, et al. Going deeper with convolutions. *Cvpr*, 2015.
- [16] Naftali Tishby and Noga Zaslavsky. Deep learning and the information bottleneck principle. *CoRR*, abs/1503.02406, 2015.
- [17] Xiaoning Zhu, Qingyue Meng, Bojian Ding, Lize Gu, and Yixian Yang. Weighted pooling for image recognition of deep convolutional neural networks. *Cluster Computing*, pages 1–13, 2018.

Humanités numériques : un cas pratique avec l'identification de hiéroglyphes

S. Madeleine¹

S. Maillot¹

S. Thomas¹

M. Servajean^{1 2}

¹ Université Paul Valéry Montpellier, Route de Mende, Montpellier 34199, France
{sarah.madeleine, sylvain.maillot, sara.thomas}@etu.univ-montp3.fr

² LIRMM, 860 rue de St Priest, Montpellier 34095, France
maximilien.servajean@lirmm.fr

Résumé

Dans le cadre du projet VEgA et des humanités numériques en général, nous nous intéressons à l'utilisation des méthodes de reconnaissance d'image par réseaux de neurones à convolutions et, plus particulièrement, à la classification de hiéroglyphes. Nous présentons ici le modèle de classification utilisé, l'interface graphique permettant son utilisation par le plus grand nombre ainsi que l'analyse des résultats expérimentaux. Bien que disposant de très peu d'images (environ 44 exemples par classe), notre modèle obtient un score supérieur à 90% en validation.

Mots Clef

Reconnaissance d'image, apprentissage automatique, réseaux de neurones à convolution, ResNet-18, hiéroglyphes égyptiens, humanités digitales.

Abstract

As part of the VEgA project, we are interested in the use of image recognition methods using convolutional neural networks in the context of digital humanities. We show in particular how classification can be performed on hieroglyphs. We present our classification model and the graphical user interface and we analyse our experimental evaluation. Although we only have few training examples (around 44 hieroglyphs per class), our model achieves more than 90% accuracy on the validation set.

Keywords

Image recognition, automatic learning, convolutional neural networks, ResNet-18, Egyptian hieroglyphs, digital humanities.

1 Introduction

Ces dernières années ont vu les performances de diverses tâches d'apprentissage statistique, notamment en classification d'images, exploser et ce, principalement grâce au renouveau de l'apprentissage profond [10]. Cette opportunité s'est initialement cantonnée aux sciences dites dures,

où la disponibilité des données nécessaires à l'apprentissage d'un modèle profond est plus immédiate.

De leur côté, les humanités numériques (ou *digital humanities*) [1] se sont tout d'abord manifestées par une révolution des usages avec des plateformes comme LexArt¹, dédiée au lexique artistique de 1600 à 1750, ou VEgA² (Figure 1), premier dictionnaire numérique de l'égyptien ancien. L'impact de tels projets se mesure autant dans la centralisation de l'information et que dans leur mise à disposition au travers d'interfaces innovantes.

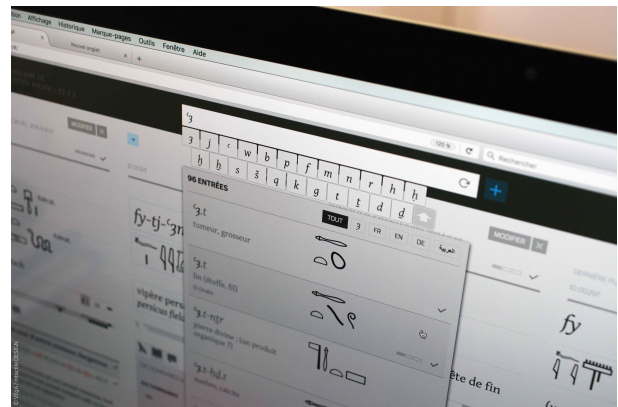


FIGURE 1 – Interface du dictionnaire VEgA.

Cependant, les techniques d'apprentissages statistiques (e.g. classification, régression) ne sont encore que très peu répandues dans ces disciplines qui, pourtant, pourraient en bénéficier [7]. En Egyptologie, objet d'étude de cette contribution, les hiéroglyphes peuvent ainsi être identifiés (i.e. classification) à partir d'images.

La classification consiste, à partir d'une donnée d'entrée x_i , à en prédire la classe (ou le label) y_i . Il s'agit d'une stratégie supervisée dans le sens où nous disposons d'un jeu d'apprentissage $D = \{(x_1, y_1), \dots, (x_N, y_N)\}$. Pour ce travail, x_i représente une image de hiéroglyphe et y_i la détermination (le nom) de ce dernier. Le cas particulier des

1. <http://lexart.fr>

2. <http://vega-vocabulaire-egyptien-ancien.fr>

images fait face à deux problèmes importants : la malédiction de la dimension et la notion de séparabilité linéaire. La grande dimension favorise le sur-apprentissage [8] et la non séparabilité linéaire [11] empêche tout classifieur linéaire de s'adapter correctement aux données.

Les réseaux de neurones profonds [11] jouent sur ces deux tableaux. Concernant la dimensionnalité, la succession des représentations des données (les couches du réseaux) ferait office de goulot d'étranglement de l'information [16] et l'algorithme d'optimisation stochastique SGD jouerait comme un régularisateur [2] limitant par là même les possibilités de sur-apprentissage. En outre, la succession d'activations non-linéaires permet de "tordre" l'espace d'entrée et de projeter les données dans un espace de dimension réduite où chaque classe devient séparable linéairement des autres. En omettant la dernière couche $g(\mathbf{x})$ de classification linéaire du modèle, l'ensemble du réseau peut ainsi être vu comme permettant l'apprentissage de *features* simples et séparables linéairement des données d'entrées. En d'autres termes, un modèle profond peut donc être vu comme la composée d'une fonction d'apprentissage de *features* f et d'une fonction de classification linéaire $g : (g \circ f)(\mathbf{x})$.

Malheureusement, de tels modèles nécessitent une très grande quantité de données, incomparable avec ce qui est réellement disponible dans le monde des humanités numériques. Ce travail montre qu'il est malgré tout possible, en se concentrant sur des techniques de l'état de l'art, d'obtenir d'excellentes performances (supérieures à 90% de précision en validation). Pour cela, il est par exemple nécessaire de s'appuyer sur la notion de *transfert learning*. En effet, la fonction de coût d'un réseau de neurones admet un grand nombre de minimum locaux. Ainsi, le point de départ de l'apprentissage a un impact sur le point de convergence. Le *transfert learning* s'appuie sur cette idée. En entraînant le modèle sur un jeu de données tiers conséquent, il apprend des *features* visuelles pertinentes et utiles pour la seconde tâche où la disponibilité des données est plus problématique. L'augmentation des données permet également de compenser cette limite en déformant aléatoirement les données d'entrées.

Afin de rendre utilisable notre modèle par le plus grand nombre, nous présentons également l'interface graphique qui l'accompagne. Celle-ci permet à la fois de lancer l'apprentissage d'un modèle lorsque de nouvelles données deviennent disponibles ; et d'identifier des hiéroglyphes par glisser/déposer.

La suite de ce papier se structure comme suit. La section 2 présente les travaux antérieurs. Notre modèle est introduit dans la section 3 et nous montrons ses résultats expérimentaux dans la section 4. Enfin nous concluons ce travail de recherche dans la section 5.

2 Travaux antérieurs

Les articles [4], [7], [9] et [12] font état des travaux de recherche dans le domaine de la reconnaissance des

symboles hiéroglyphiques par apprentissage automatique. Dans [9], Kriege et al. ont travaillé sur la reconnaissance de symboles cunéiformes à l'aide de deux méthodes basées sur la mesure de similarité de deux graphes : la méthode du plus proche voisin et celle d'un CNN. La première étant très coûteuse en termes de temps de calcul et nécessitant un volume important de données pour la phase d'apprentissage ; les auteurs se sont tournés vers un CNN. Bien que la démarche d'utiliser un CNN soit identique à la nôtre, leur modèle se base sur un a priori concernant la forme des symboles, ce qui n'est pas notre cas où nous nous appuyons sur les images brutes sans pré-traitement.

Dans [4], Duque-Domingo et al. proposent une méthode de décryptage des cartouches hiéroglyphiques basée sur les techniques d'analyse d'image et vision par ordinateur. Ils identifient notamment les hiéroglyphes présents dans une image grâce à la comparaison des distances de Chamfer et d'Hausdorff d'une image connue (i.e. une image-étalon) à celles de l'image étudiée. Bien que cette méthode s'affranchisse des biais induits par la luminosité de l'image ou par la texture du support du cartouche ; elle nécessite d'avoir à disposition autant d'images-étalons qu'il existe de symboles dans la langue étudiée, en incluant leur diversité. À l'inverse, notre modèle réalise la majorité de son apprentissage sur le jeu de données ImageNet. Cela nous permet de finaliser son apprentissage (fine-tuning) sur un jeu de données, spécifique aux symboles étudiés, restreint et non-exhaustif.

Dans [7] Hu et al. présente la méthode de l'histogramme du contexte de l'orientation de la forme pour identifier les symboles hiéroglyphiques. Cette méthode nécessite un important travail de traitement d'image en amont de l'identification par des épigraphistes pour faire disparaître le bruit. Tandis que notre modèle tire profit du bruit des images pour augmenter ses performances.

De nombreuses architectures ont été proposées et ont obtenu des résultats significatifs en classification d'images [3, 6, 10, 13, 14, 15, 17]. Il ressort que les réseaux de neurones à convolution avec des connexions résiduelles sont plus simples à entraîner et obtiennent de meilleurs résultats en terme d'apprentissage que les réseaux sans. Dans [13], les auteurs démontrent notamment que la profondeur du réseau a une influence sur sa précision. De même que dans [17] il est démontré que la méthode dite "*weighted-pooling*" obtient de meilleurs résultats que les méthodes de pooling traditionnelles, telles que *max-pooling*, *average pooling*.

Nous nous sommes servis de ces travaux pour adapter le modèle ResNet-18 aux contraintes inhérentes à la reconnaissance des hiéroglyphes égyptiens. Nous montrons la faisabilité des modèles CNN à partir du moment où un

minimum d'images sont disponibles. Nous montrons les limites de notre méthode lorsque la variabilité d'un signe est trop grande.

3 Proposition & démonstration

3.1 Jeu de données

Le jeu de données provient de deux sources : des signes du temple de Karnak³ et d'autres de la pyramide d'Ounas (M. Franken), représentant un total de 4193 symboles répartis en 88 catégories différentes. Cela représente la quantité extrêmement faible de 44 images par classe.

Chaque signe est labellisé à la main selon la liste de Gardiner. Les images de Karnak n'ont pas subi de pré-traitement. Les images du jeu de données de Morris Franken [5] sont recoupées en 50x75. Les pixels vides sont remplis avec une texture rappelant l'arrière-plan réel. Pour augmenter le jeu d'entraînement, les images sont recadrées aléatoirement et transformées par un flip horizontal.

3.2 Architecture

Implémentation PyTorch de ResNet-18 en utilisant des poids pré-entraînés sur le *dataset* ImageNet. Le modèle est constitué de 20 couches : 1 convolution 7x7, 16 convolution 3x3 et 3 couches connectées. La dernière couche donne en sortie une des 88 classes de signes possibles.

3.3 Optimisation

Plusieurs tests d'optimiseur ont été réalisés. Notre choix s'est porté sur Adam avec un *learning rate* initialisé à 0.0001 et divisé par 10 toutes les 7 epochs.

3.4 Activation

Chaque couche contient une opération de normalisation et d'activation. Cette dernière (*ReLU*) introduit une non-linéarité permettant de répondre plus efficacement à notre tâche de reconnaissance et classification d'image. La normalisation (*Batch Normalization*) augmente la stabilité et la vitesse d'apprentissage du modèle en normalisant les sorties des couches cachées à chaque itération.

Le modèle utilise la technique de *DropOut*, initialisée à 0.5, qui consiste en la désactivation temporaire et aléatoire de neurones pendant l'entraînement, puis en leur réactivation lors de la phase de test. Cela simule un apprentissage sur un ensemble de modèles différents. L'objectif est de prévenir les liens d'interdépendance qui pourraient se créer entre les neurones et ainsi éviter l'*overfitting*.

3.5 Entraînement

Le modèle est entraîné sur GPU (GTX 1080 ti) sur 100 epochs.

La fonction d'entropie croisée est utilisée pour évaluer la précision du modèle. La convergence est rapide et dépasse les 90% de précision.

3. <http://sith.huma-num.fr/karnak>

3.6 IHM

Afin de rendre possible l'utilisation de notre travail par le plus grand nombre, nous avons développé une interface graphique à l'aide de la librairie Python PYQT5.

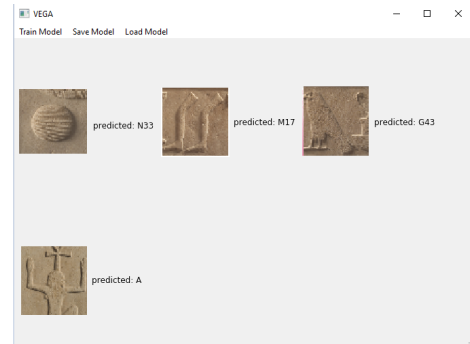


FIGURE 2 – Maquette de l'interface graphique.

Elle est composée d'un bandeau de navigation et d'une zone centrale. Trois onglets composent la barre de navigation : *Train model*, *Save model* et *Load model*. Le premier onglet permet de calculer un nouveau modèle (e.g. quand de nouvelles images sont disponibles et viennent enrichir le jeu de données d'apprentissage). À partir du second onglet l'utilisateur peut sauvegarder le modèle pour une réutilisation future. Enfin le chargement d'un modèle déjà existant se fait par le troisième et dernier onglet.

La zone centrale quant à elle, initialement vide, permet d'utiliser la classification. Elle s'utilise en y glissant des images contenant des hiéroglyphes inconnus pour ensuite afficher le résultat.

L'application permet de charger un modèle n'ayant pas nécessairement de lien avec les hiéroglyphes.

4 Évaluation expérimentale

En nous appuyant sur l'état de la recherche dans le domaine de la reconnaissance d'image par apprentissage automatique et des résultats de nos tests, nous sommes partis d'un modèle ResNet-18. Nous avons ensuite choisi les éléments qui correspondaient au mieux à nos besoins pour composer notre modèle final :

- Couches de convolution : 1 taille 7x7, 16 taille 3x3 ;
- Couche de pooling : Max pool et Average pool ;
- Couche de correction : Batchnorm, ReLU ;
- Couche de perte : Cross entropy ;
- Couche connectée : linéaire ;
- Optimiseur : Adam.

Notre modèle est entraîné pendant 100 epochs. Il converge rapidement (i.e. vingtième epoch) vers une valeur limite supérieure à 90% de précision sur le jeu de validation.

La précision de notre modèle est meilleure en phase de validation que lors de celle d'apprentissage. Nous expliquons cela par les transformations infligées au jeu

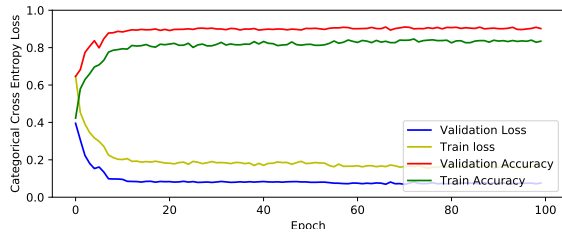
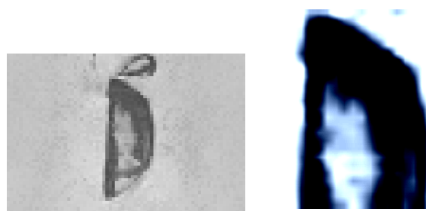


FIGURE 3 – Performances du modèle.

de données d'apprentissage, relativement importantes et complexifiant ainsi la tâche de classification. À l'inverse, le jeu de validation ne subit aucune transformation.



(a) Sans transformation. (b) Avec transformation.

FIGURE 4 – Image d'un signe V31.

5 Conclusion

Notre travail de recherche évalue les possibilités d'utilisation de modèles profonds pour la reconnaissance et la classification automatiquement des hiéroglyphes. Le déploiement d'un modèle résiduel simple à 20 couches apporte des résultats concluants. Il est cependant important de noter que l'efficacité des modèles profonds est largement impactée par la volumétrie et la variété (que l'on retrouverait en test) du jeu de données d'apprentissage.

Notre modèle converge très rapidement. Un modèle plus profond apporterait de meilleurs résultats de classification mais demanderait plus de temps d'entraînement. L'IHM est simple et pourrait à l'avenir proposer d'autres fonctionnalités.

Références

- [1] David M Berry and Anders Fagerjord. *Digital humanities : knowledge and critique in a digital age*. John Wiley & Sons, 2017.
- [2] Pratik Chaudhari and Stefano Soatto. Stochastic gradient descent performs variational inference, converges to limit cycles for deep networks. *CoRR*, abs/1710.11029, 2017.
- [3] François Chollet. Xception : Deep learning with depthwise separable convolutions. *arXiv preprint*, 2016.
- [4] Jaime Duque-Domingo, Pedro Javier Herrera, Enrique Valero, and Carlos Cerrada. Deciphering egyptian hieroglyphs : towards a new strategy for navigation in museums. *Sensors*, 17(3) :589, 2017.
- [5] Morris Franken and Jan C. van Gemert. Automatic egyptian hieroglyph recognition by retrieving images as texts. In *Proceedings of the 21st ACM International Conference on Multimedia*, MM '13, pages 765–768, New York, NY, USA, 2013. ACM.
- [6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [7] Rui Hu, Carlos Pallán Gayol, Jean-Marc Odobez, and Daniel Gatica-Perez. Analyzing and visualizing ancient maya hieroglyphics using shape : From computer vision to digital humanities. *Digital Scholarship in the Humanities*, 32(suppl_2) :ii179–ii194, 2017.
- [8] Eamonn Keogh and Abdullah Mueen. Curse of dimensionality. In *Encyclopedia of Machine Learning and Data Mining*, pages 314–315. Springer, 2017.
- [9] Nils M Kriege, Matthias Fey, Denis Fisseler, Petra Mutzel, and Frank Weichert. Recognizing cuneiform signs using graph based methods. *arXiv preprint arXiv :1802.05908*, 2018.
- [10] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [11] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553) :436, 2015.
- [12] Edgar Roman-Rangel and Stephane Marchand-Maillet. Assessing deep learning architectures for visualizing maya hieroglyphs. In *Mexican Conference on Pattern Recognition*, pages 137–146. Springer, 2017.
- [13] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv :1409.1556*, 2014.
- [14] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alexander A Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. In *AAAI*, volume 4, page 12, 2017.
- [15] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich, et al. Going deeper with convolutions. *Cvpr*, 2015.
- [16] Naftali Tishby and Noga Zaslavsky. Deep learning and the information bottleneck principle. *CoRR*, abs/1503.02406, 2015.
- [17] Xiaoning Zhu, Qingyue Meng, Bojian Ding, Lize Gu, and Yixian Yang. Weighted pooling for image recognition of deep convolutional neural networks. *Cluster Computing*, pages 1–13, 2018.