

Royaume du Maroc
Haut-Commissariat au Plan
Institut National de Statistique et d'Economie Appliquée

CENTRE D'ETUDES DOCTORALES

LABORATOIRE SI2M

FORMATION DOCTORALE

**« SYSTEMES D'INFORMATION, SYSTEMES
INTELLIGENTS ET MATHEMATIQUES
APPLIQUEES »**

SYSTEMES D'INFORMATION

Recommandation system for Implementing ubiquitous business processes in smart environment

Directeur de thèse : Rajaa SAIDI, Professeure Habilitée à l'INSEA

Mail : r.saidi@insea.ac.ma

Site Web : <https://sites.google.com/site/rajaasaidi/recherche>

Présentation

De nos jours, le développement et l'utilisation des systèmes d'information centrés processus métiers [1] font face à trois défis : l'évolution rapide des règles métiers, l'évolution technologique des dispositifs d'interaction qui intègrent parfois des terminaux mobiles et l'évolution des usages. En effet, une adaptation facile et rapide, d'un processus métier aux changements internes et externes, est ainsi un facteur clé améliorant son utilisabilité et sa performance.

Par ailleurs, introduite par Weiser [2], l'informatique ubiquitaire (Pervasive ou Omniprésente) [3] est considérée comme un paradigme de l'informatique personnelle ayant pour objectif de permettre aux utilisateurs de consulter des données n'importe où et n'importe quand, à travers leurs dispositifs d'accès. En effet, les environnements quotidiens de l'utilisateur sont composés des équipements hétérogènes, formant ainsi un système ouvert et dynamique, dans lequel les ressources disponibles, le contexte de l'utilisateur et ses activités, changent continuellement. Cela nécessite de construire une architecture type qui traite l'hétérogénéité des composants ubiquitaires et qui soit sensible au contexte d'utilisation [4,5].

Construire un environnement intelligent reste donc un sujet défi. L'idée ici est de capitaliser sur les avancées de l'intelligence artificielle pour doter ces systèmes d'une flexibilité traduite par deux caractéristiques à savoir : (1) la configuration automatique (ou semi-automatique) et (2) l'amélioration automatique (ou semi-automatique). La configuration automatique signifie que le système doit être capable de faire appel et recommander des services accessibles pour répondre à un besoin défini (modification du comportement du système, rémission en cas de panne, etc.). L'amélioration automatique signifie que ces systèmes doivent être capables de s'améliorer et d'optimiser leur comportement.

Objectifs

Les objectifs de cette thèse se résument comme suit :

- Etude des concepts de l'informatique ubiquitaire
- Proposition d'une architecture et d'une méthodologie pour développer un système de mise en œuvre des processus métiers ubiquitaires
- Validation des propositions par une plateforme et des études de cas (ex. Sécurisation de l'intervention des véhicules de secours, Localisation dans les transports en commun, Aide au stationnement en ville, Gestion des crises, etc.)

Mots clés : Processus Métiers, BPMN, Informatique Ubiquitaire, Composition de services, Prédiction et Recommandation de services

Références bibliographiques

[1] <http://www.bpmn.org>

[2] M. Weiser. The computer for 21st Century. Scientific American, USA, 1991.

[3] « Fundamentals of ubiquitous computing », <http://sociotech.pbworks.com/f/UbiquitousComputFundamen.pdf>

[4] Faieq Soufiane, Saidi Rajaa., Elghazi Hamid., Rahmani Moulay Driss, « A Conceptual Architecture for Cloud-Based Context-Aware Service Composition », The International Symposium on Ubiquitous Networking (Unet'16), LNEE, , Vol. 397, Casablanca, May 30 - June 1st, 2016

[5] Faieq Soufiane, Saidi Rajaa., Elghazi Hamid., Rahmani Moulay Driss, " C2IoT: A framework for Cloud-based Context-aware Internet of Things services for smart cities", 14th International Conference on Mobile Systems and Pervasive Computing (MobiSPC 2017), Leuven, Belgium, July 24-26, Procedia Computer Science 110, Elsevier 2017

Improving Business Processes using Blockchain technology

Pr. Maryam RADGUI

SI2M Laboratory, INSEA

Business processes rely on data from information systems outside their control, even data from other enterprises. This data must move between organizations, or even between enterprises, which is not only complex and expensive but often results in stale and inconsistent data. Lack of transparency and trust is the consequence.

In this sense, Blockchain is widely accepted as an answer to trust and transparency issues in business networks that span organizations for the following reasons :

- Participants in the network use a shared ledger to perform transactions on assets.
- Transactions are validated by participants through a consensus protocol.
- Smart contracts control transactions between participants, which therefore do not need to trust each other. Smart contracts ensure that contractual conditions are met, and obligations are enforced.
- Permissive blockchains ensure that all information and transactions on the blockchain are available only to network members with the right permission.

Indeed, In a solution that combine business process and Blockchain (cf. Fig1), the shared ledger provides the interface for business processes (a process queries asset information and performs transactions directly on the ledger). Stale or inconsistent data is a thing of the past! In addition, business events on the blockchain initiate or drive business processes, ensuring that the right organizations are involved and react to events in a timely manner.

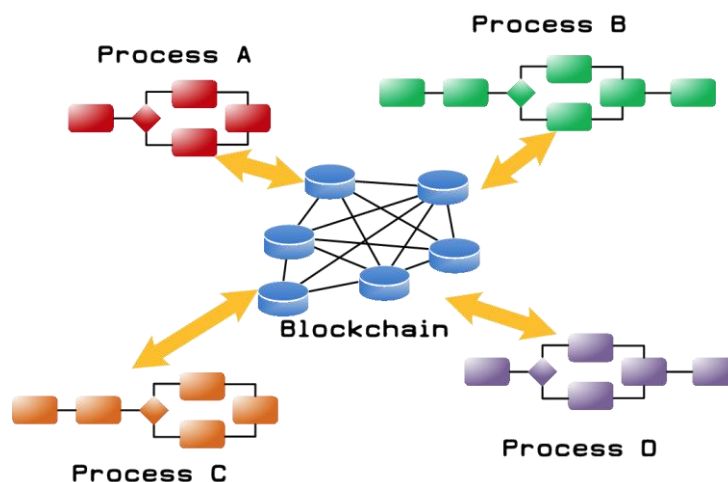


Fig 1 : Blockchain and BPM

Blockchain technology offers a sizable promise to rethink the way inter-organizational business processes are managed because of its potential to realize execution without a central party serving as a single point of trust (and failure).

In this sense, Blockchain technology provides basic building blocks to support the execution of collaborative

business processes involving mutually untrusted parties in a decentralized environment. Several research proposals have demonstrated the feasibility of designing blockchain-based collaborative business processes using a high-level notation, such as the Business Process Model and Notation (BPMN), and thereon automatically generating the code artifacts required to execute these processes on a blockchain platform. Hence, The combination of both business process management and blockchain helps to reach the next level of integration and automation of business processes. We can then significantly improve our business processes.

The objectif of this thesis is to stimulate research on this combination of business processes and Blockchain. It's then necessary to outline the challenges and opportunities of blockchain for Business Process

Management (BPM) in the existing works. The main focus is about detailing research directions for investigating the application of blockchain technology to BPM. This subject should deal with the use of blockchain for collaborative information systems. Conceptual, technical and application-oriented contributions must be also pursued within the scope of this thesis.

Keywords : Business Process Management (BPM), Blockchain, business process improvement, BPMN.

References:

1. Di Ciccio, Claudi, Cecconi, Alessio, Dumas, Marlon, "Blockchain Support for Collaborative Business Processes", *Informatik Spektrum*, Jun 2019"
2. Business Process Management : Blockchain and Central and Eastern Europe Forum, BPM 2019 Blockchain and CEE Forum, Vienna, Austria, September 1–6, 2019, Proceedings
3. Mendling, Jan & Weber, Ingo & Aalst, Wil M. P. & Brocke, « Blockchains for Business Process Management - Challenges and Opportunities », *ACM Transactions on Management Information Systems*, January 2018
4. JunyaoWang ShenlingWang JunqiGuo, A Summary of Research on Blockchain in the Field of Intellectual Property, *Procedia Computer Science*, Volume 147, 2019, Pages 191-197

An extension of BPMN for Social Business Process Management (SBPM)

Pr. Maryam RADGUI
SI2M Laboratory, INSEA

Several changes can affect the way how every organization interacts with customers and with partners to deliver goods and services. In this sense, Business Process Management (BPM) is used to enable the management of organisations's processes. In accordance with many management techniques, improvements can be made by analyzing, with varying degrees of granularity, how the processes are actually implemented.

An innovative way to improve business process management is to use social network analysis to observe real working relationships among users. This last technique allows the workflow manager to specifically examine how well users are tuned to their workflow. This gives us a way to reconstruct workflows or user practices.

In fact, social network is fundamentally affecting the way organizations are interacting, both internally and externally. However, when organizations adopt social media initiatives, they will not achieve its full potential unless the changes are integrated with both its operational processes and its approach to continuous improvement.

In this sense, Social Business Process Management (SBPM) (cf. Fig.1) is an approach to improve business process that aims to encourage a more collaborative and transparent approach to process improvement. A major benefit of social BPM is that it helps to break down the barrier between BPM

decision-makers and the users affected by their decisions. Social BPM takes advantage of social media tools such as real-time collaboration, communities and tagging to improve communication.

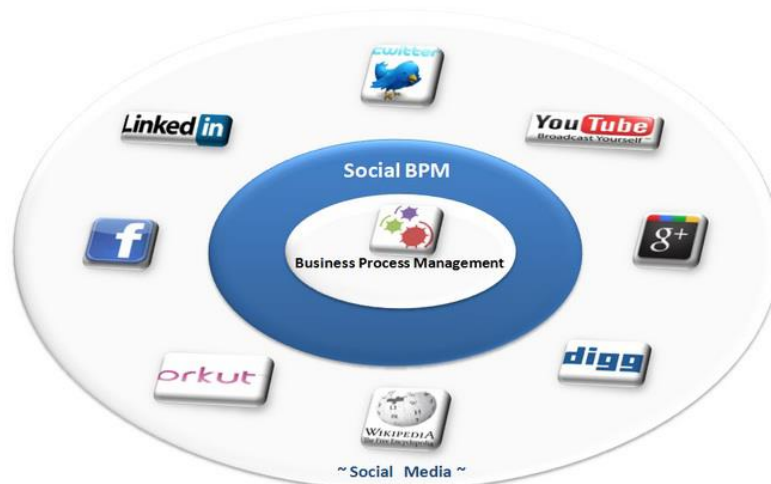


Fig.1 : Social BPM

This research field suffers from a lack of languages, notations and global approaches that can deal with business processes according to social network. Indeed, some research work tried to propose notation that support SBPM for example BPMN [3]. However, there is still a need for a global approach that cover all the aspects of SBPM.

Hence, in this thesis, the main focus will be on how Socially Oriented Process Modeling can be used effectively to enable business model renewal, service improvements, customer performance focus. The contributions should be about a proposition of an extension of Business Process Management and Notation (BPMN) that can support SBPM.

Objectives:

- Detailed study of BPM, BPMN and SBPM.
- State of the art about business process modeling approaches supporting social networks.
- Proposition of a DSL (Domain Specific Language) and / or an extension of BPMN which supports the SBPM.

Keywords: SBPM, BPM, BPMN, Social network, DSL.

References:

1. Eli Saad, The Social Process Model: Strategize At the Personal Level, Business and Economics Journal, 2015.
2. Nicolas Pflanzl and Gottfried Vossen, Human-Oriented Challenges of Social BPM: An Overview, Septembre 2013.
3. Marco Brambilla Weske, A BPMN Based Notation for Social BPM, BPMN Workshop, Luzern 2011.
4. Ademir Macedo Nascimento, Using Social Media for Business Process Improvement: A Systematic Review, New Advances in Information Systems and Technologies pp 47-58, 2016.
5. Peter Busch, Business process management, social network analysis and knowledge management: A triangulation of sorts?, January 2010.
6. Henrik von Scheel, Mona von Rosing and Zakaria Maamar, Social Media and Business Process Management, the complete business process handbook.

Improving user's services in smart environment using process mining techniques

Pr. Maryam RADGUI
SI2M Laboratory, INSEA

One of the main goals of a smart environment is to recognize the user's situation and to automatically adapt its behavior accordingly. Indeed, understanding what the user is doing in a Smart Environment is important not only for adapting the environment behavior, for example, by providing the most appropriate combination of services for the recognized situation, but also for identifying situations that could be problematic for the user.

In fact, prediction, proactivity and decision-making capabilities are important in helping users to achieve their goals in a smart environment. This assistance becomes fundamental when the smart home for example, is inhabited by elderly people or by persons with special needs. To this aim there is a huge amount of work in automatically learning models of the user's routines especially in the home environment. Most proposed approaches model routines of daily activities as process workflows. Manually building models of the user processes is a complex, costly and error-prone engineering task. Hence, the interest in automatically learning them from examples of actual procedures emerges. The ability to express and learn complex conditions on the involved tasks, are also desirable.

In this sense, process mining techniques could be an interesting way to express and learn user's behaviors and so predict his needs (cf. fig.1).

Indeed, Process mining is a relatively young research discipline that sits between computational intelligence and data mining on the one hand, and process modeling and analysis on the other hand. Process mining techniques aim to extract knowledge from event logs in order to discover, monitor and improve processes.

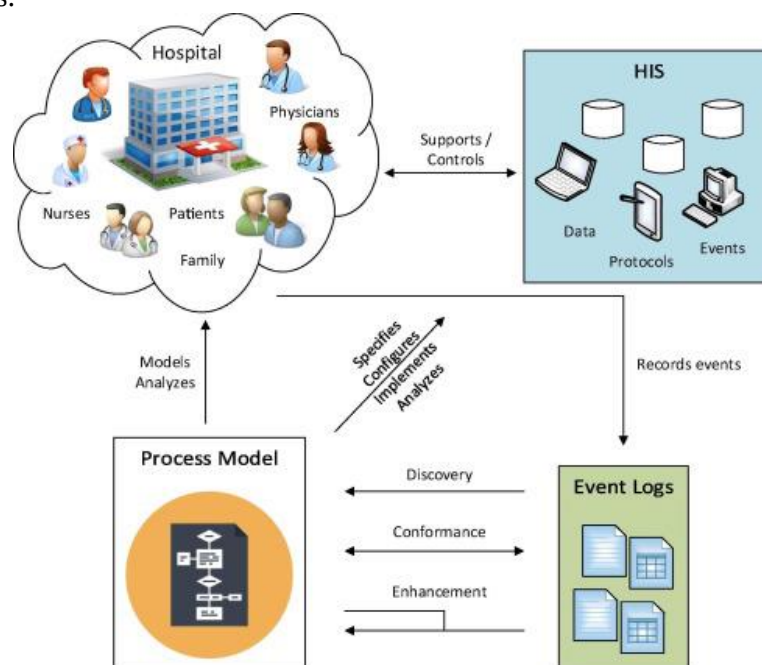


Fig 1. Process mining in healthcare

The main objective of this research will be applying process mining techniques to learn the user's daily routines, and so predicting his needs in smart environment. Then, focusing on a composition of appropriate services to adapt and to propose to the user according to his needs.

Keywords: Process mining, smart environment, business process, Micro Services, composition.

References:

1. Eric Rojas, Jorge Munoz Gama, Process mining in healthcare: A literature review, Journal of Biomedical Informatics Volume 61, Pages 224-236, June 2016.
2. Peyman Badakhshan, Using Process Mining for Process Analysis Improvement in Pre-Hospital Emergency, March 2018.
3. Bernis Tibeme ; Hossain Shahriar , Process Mining Algorithms for Clinical Workflow IEEE, April 2018.
4. Stefano Ferilli, Berardina De Carolis, Domenico Redavid, Logic-Based Incremental Process

Mining in Smart Environments, International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, pp 392-401, 2013.

5. Wil van der Aalst¹ and all, Process Mining Manifesto, BPM 2011 Workshops, Part I, LNBIP 99, pp. 169–194, 2012.

6. Vikramaditya R. Jakkula, Aaron S. Crandall, Diane J. Cook, Knowledge Discovery in Entity Based Smart Environment Resident Data Using Temporal Relation Based Data Mining, Data Mining Workshops, 2007.

7. Wil van der Aalst, Process Mining: Discovery, Conformance and Enhancement of Business Processes, Springer, 2011.

8. Jans, M., van der Werf, J.M., Lybaert, N., Vanhoof, K. (2011) A business process mining application for internal transaction fraud mitigation, Expert Systems with Applications, 38 (10), 13351–13359.

9. <http://www.processmining.org/>

SYSTEMES INTELLIGENTS

Intelligence Artificielle pour la Traduction de la Langue Marocaine des Signes

Pr. Imade Benelallam et Pr. Walid Cherif

Vers un système d'interaction gestuelle : Langage des Signes
Marocain (LSM-Visuo-gestuelle)

Description du sujet de thèse

Le but de ce sujet de thèse est l'intégration sociale et éducative des personnes sourdes et malentendantes grâce à la vision artificielle et le traitement du langage naturel.

Aujourd'hui au Maroc, l'accès au savoir, à l'éducation et aux interactions sociales pour les personnes atteintes de handicap demeure négligeable malgré les nouvelles initiatives d'Etat. Au sein de cette population souvent marginalisée, les personnes atteintes de surdit  constituent une part importante : en effet, les sourds et malentendants repr sentent 15,8% du nombre total de personnes en situation de handicap et seuls 5% des 23 000 enfants sourds en  ge d' tre scolaris s sont pris en charge. Cette prise en charge demeure par ailleurs timide et peu adapt e aux besoins de ces enfants en situation de handicap, car les programmes d' ducation inclusifs n'ont pas encore  t  mis en place, les professeurs sont peu ou pas sp cialis s et sensibilis s aux besoins sp cifiques de ces  tudiants et en dehors des structures associatives d di es, l'int gration en soci t  de ces personnes au Maroc demeure complexe.

Cependant, les avanc es en mati re d'intelligence de vision artificielle, de traitement du langage naturel et de transcription de la voix au texte sont matures et en forte expansion. Ces technologies pourraient constituer un socle id al au d veloppement d'innovations d di es aux populations en handicap, plus particuli rement en ce qui concerne ce projet aux personnes sourdes et malentendantes. Il est en effet possible aujourd'hui de mettre en place des syst mes intelligents aptes   capter la voix, le texte et la vid o d'une part, et de retranscrire l'ensemble des informations acquises en langages des signes. D'autre part, l'op ration inverse est  galement r alisable, et donner acc s   la communication aux personnes sourdes en traduisant leurs signes en texte ou voix g n r s par intelligence artificielle est possible. Cela est par ailleurs un excellent moyen de standardiser la langue des signes marocaines et de cr er des ponts avec les autres langues des signes internationales.

Mots cl s

Langage des Signes Marocains, Visuo-gestuelle, Natural Language Processing, Image and Video processing, Text Mining, Voice Analysis, Intelligence Artificielle, Deep learning, Statistical Machine Learning, Interaction Homme-Machine.

Objectif et contexte

Ce projet a donc pour but de mettre en place une interface capable de traduire n'importe quelle source audio, textuelle ou vid o en langue des signes   fins p dagogiques : un professeur disposant un court en fran ais ou en arabe serait film  et son savoir retranscrit en temps r el sur un  cran proposant un avatar 3D parlant la langue des signes. Au-del  de ce premier cas d'usages, les m mes algorithmes traduiraient les contenus vid o des plateformes en ligne, le contenu textuel des ebooks et de fa on plus g n rale, tout contenu pr sent sur internet sous format audio, texte ou vid o. Inversement, pour faciliter l'int gration des personnes handicap es au sein de la soci t , la langue des signes capt e par une simple cam ra de smartphone produit du texte ou de l'audio   destination des destinataires avec qui la personne sourde cherche   communiquer.

M thode

La m thodologie propos e consistera en deux parties :

- Dans un premier temps, le candidat commencera par dresser l'état de l'art des travaux réalisés dans le cadre d'analyse de vidéos. Une collecte des représentations par signes devra être réalisée au cours de la première année. Cela permettra d'aboutir à un modèle qui divise le contenu d'une vidéo (extraite d'un smartphone ou d'un autre périphérique) en plusieurs frames de signes, et par la suite, les signes obtenus sont analysés, éventuellement corrigés avant de traduire leur contenu en texte. La conversion en audio devra aussi passer par le text to speech.
- La deuxième partie consiste en la réalisation inverse : à partir d'un texte, audio ou vidéo, le contenu devra être compris, corrigé en se basant sur la dataset disponible, traduit en signes et finalement ces signes seront fusionnés et convertis en animation sous forme d'avatar.

Résultats attendus

- Un modèle d'analyse de contenu vidéo.
- Un modèle d'analyse de contenu audio.
- Un modèle d'analyse de textes.
- Un module NLP pour la correction des textes obtenus.
- Un modèle\ des modèles Deep learning pour l'apprentissage et la prédiction.
- Un module d'animation pour l'exportation des signes générés en animation de l'avatar.

Précision sur l'encadrement

Directeur de thèse : Imade Benelallam (Directeur du Laboratoire SI2M de l'INSEA)

Co-directeurs de thèses : Walid Cherif (Membre du Laboratoire SI2M de l'INSEA)

L'équipe d'encadrement inclut également : : Bouyakhf El Houssine (Fondateur du LIMIARF, laboratoire d'Intelligence Artificielle de l'université Mohammed V)

Objectifs de valorisation des travaux de recherche du doctorant :

- Diffusion par communication dans les conférences nationales et internationales et publication dans les revues scientifiques internationales indexées et confidentialité, protection éventuelle par brevet du droit à la propriété intellectuelle, ...
- Valorisation sous forme de publications scientifiques

Collaborations envisagées

- Collaboration interne à l'INSEA et au LIMIARF: co-encadrement. Collaboration avec des Ecoles d'ingénieurs ainsi que la Start up AIOX Labs.

Profil et compétences recherchées

Connaissances solides en Traitement du langage Naturel, en text mining, en apprentissage automatique et en particulier en apprentissage profond (Deep Learning).

Avoir une expérience en analyse audio/ vidéo.

Avoir des notions en animation 3D pour la génération des signes.

Niveau de français et d'anglais requis : Compétent/Courant:

Vous pouvez utiliser la langue avec aisance et facilité en argumentant sur des sujets complexes.

Références :

Video

- [1] KUNDU, Malay, SRINIVASAN, Vikram, MIGDAL, Joshua, *et al.* *Method and apparatus for detecting suspicious activity using video analysis*. U.S. Patent Application No 10/318,818, 11 juin 2019.
- [2] Andersson, S. H., Cardinale, M., Whiteley, R., Popovic, N., Hansen, C., Lopez, F. S., ... & Myklebust, G. (2018). Video analysis of acute injuries and referee decisions during the 24th Men's Handball World Championship 2015 in Qatar. *Scandinavian journal of medicine & science in sports*, 28(7), 1837-1846.

Audio

- [3] MALLIK, Souvik, CHOWDHURY, Debjyoti, et CHTTOPADHYAY, Madhurima. Development and performance analysis of a low-cost MEMS microphone-based hearing aid with three different audio amplifiers. *Innovations in Systems and Software Engineering*, 2019, vol. 15, no 1, p. 17-25.
- [4] OWENS, Andrew et EFROS, Alexei A. Audio-visual scene analysis with self-supervised multisensory features. In : *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018. p. 631-648.

Deep learning / NLP

- [5] AL-AYYOUB, Mahmoud, NUSEIR, Aya, ALSMEARAT, Kholoud, *et al.* Deep learning for Arabic NLP: A survey. *Journal of computational science*, 2018, vol. 26, p. 522-531.
- [6] YOUNG, Tom, HAZARIKA, Devamanyu, PORIA, Soujanya, *et al.* Recent trends in deep learning based natural language processing. *IEEE Computational Intelligence Magazine*, 2018, vol. 13, no 3, p. 55-75.
- [7] DENG, Li et LIU, Yang (ed.). *Deep Learning in Natural Language Processing*. Springer, 2018.

Sign language

- [8] BRENTARI, Diane, FENLON, Jordan, et CORMIER, Kearsy. Sign language phonology. In : *Oxford Research Encyclopedia of Linguistics*. 2018.
- [9] MACDONALD, Kyle, LAMARR, Todd, CORINA, David, *et al.* Real-time lexical comprehension in young children learning American Sign Language. *Developmental science*, 2018, vol. 21, no 6, p. e12672.
- [10] CHENG, Qi et MAYBERRY, Rachel I. Acquiring a first language in adolescence: the case of basic word order in American Sign Language. *Journal of child language*, 2019, vol. 46, no 2, p. 214-240.

Prédiction des complications prénatales par le Traitement du Langage Naturel

Pr. Imade Benelallam, Pr. Manar Jalal et Pr. Walid Cherif

Description du sujet de thèse

Dans cette thèse nous nous intéressons au développement d'un système prédictif des complications prénatales, chez les femmes enceintes, reposant sur des approches de Traitement du Langage Naturel, qui à ce jour fait office de référence dans la littérature en ce qui concerne l'analyse syntaxique, l'analyse sémantique, la lecture automatique des documents et l'extraction de l'information à partir des données archivées. Ce système permettra à la fois une prise en compte améliorée de l'incertitude inhérente au traitement des données en provenance de plusieurs sources de données (structurées et non structurées) et aussi l'optimisation automatique du modèle de prédiction grâce aux techniques d'apprentissage automatique, notamment, l'apprentissage profond et l'apprentissage par renforcement.

Les moteurs d'extraction de mesures seront développés sous la forme d'un ensemble d'expressions régulières et seront évalués par rapport à une vérité de terrain créée à partir d'une enquête rétrospective et ensuite une enquête prospective (cohorte) d'un échantillon de femmes enceintes, jusqu'à l'accouchement. La catégorisation automatisée de la temporalité des mesures sera définie comme un problème d'apprentissage automatique qui sera évalué par protocole de validation croisée. En plus de l'étude qui sera menée sur renforcement de la précision des modèles, la problématique de l'explicabilité des modèles développés sera examinée pour donner plus de transparence au processus d'aide à la décision.

Mots clés

Natural Language Processing, Deep learning, Reinforcement Learning, Diagnostic prénatal, modèles prédictifs, explicabilité.

Objectif et contexte

Le but de ce travail est de développer et d'évaluer deux moteurs de prédiction des complications prénatales chez les femmes enceintes. Ces deux moteurs seront basés sur les techniques de traitement du langage naturel qui extraient des constatations et des indicateurs mesurés à partir des rapports des surveillances et de radiologie inscrits par les praticiens, et catégorisent les mesures extraites selon leur "temporalité". Le caractère innovant et original de la thèse réside dans sa capacité d'explicabilité des décisions qui seront prises par la machine et permettent d'assister les praticiens dans la décision finale des diagnostics prénataux.

- Contexte sociétal, économique et/ou industriel :

Aujourd'hui, de nombreux problèmes de santé conduisent à des complications parfois dramatiques ou mortelles sans que nous disposions d'un outil d'aide à la décision qui permettrait au médecin et au patient de prédire et prévenir ces complications (exemple de la grossesse qui peut aboutir à une mortalité maternelle ou néonatale, le diabète qui peut conduire à la dialyse, l'hypertension artérielle qui peut conduire à l'AVC, ...)

Le projet sur lequel s'appuie cette thèse, vise dans sa globalité à travailler sur la Consultation Prénatale durant l'accouchement avec pour objectif de prédire les complications obstétricales et prénatales qui peuvent conduire à la morbidité et la mortalité maternelle et néonatale. Ce problème est actuellement parmi les grands défis de santé au Maroc et dans de nombreux pays et qui sont aussi inscrits dans les premières priorités des objectifs de développement durable de 2030 et appuyés par l'OMS et la communauté internationale.

- Contexte Scientifique :

L'apprentissage automatique est devenu aujourd'hui l'une des techniques les plus prometteuses en IA. Dès que les données sont disponibles dans un domaine particulier, la question de l'utilisation de ces techniques pour améliorer les processus de prise de décision se pose systématiquement. Deux familles d'algorithmes particulièrement efficaces aujourd'hui permettent d'effectuer un apprentissage profond (le « deep learning ») et un apprentissage par renforcement. De tels algorithmes simulent des architectures complexes, formées de couches de neurones artificiels, qui permettent d'implémenter des calculs distribués impossibles à programmer explicitement. L'objectif ici est d'utiliser ces dernières avancées en Machine Learning pour permettre la mise en place de modèles automatiques, à termes autonomes et qui permettent une identification précise de la pathologie, la prédiction des complications futures et leur degré d'explicabilité basé sur la corrélation de plusieurs indicateurs extraits par des techniques avancées du Traitement du langage Naturel.

Méthode

La méthodologie pour le développement d'algorithmes prédictifs basés sur les modèles d'apprentissage suivra les étapes ci-après:

- Dans un premier temps, le candidat commencera par dresser l'état de l'art des travaux réalisés dans le cadre de diagnostic prénatal, et d'autres types de diagnostics similaires, utilisant des approches d'apprentissage automatique combinées aux techniques de Traitement du Langage Naturel.
- En même temps durant cette première phase, le candidat approfondira ses connaissances dans le domaine de l'apprentissage automatique (supervisé, non supervisé, par renforcement) et des nouvelles approches de Traitement du Langage Naturel.

- Dans un second temps, il doit préparer la plateforme pour accueillir les documents numérisés et/non numérisés (fiche de surveillance, dossier d'accouchement) afin de construire une base de données labellisée qui servira à entraîner les modèles proposés ainsi qu'à les tester ;
- Par la suite le doctorant entamera les étapes de collecte et de prétraitement des données (dossiers médicaux) structurées et non-structurées : Content Sanitization, Sentence Tokenization, Word Tokenization, Stop Words Removal, Word Lemmatization, Part Of Speech (POS) Tagging, Named Entity Recognition (NER).
- L'étape précédente permettra entre autres la labellisation automatique des données d'entraînement des modèles de Machine Learning qui seront proposés.
- A ce niveau le doctorant procédera à l'extraction, la sélection et la classification des « features » qui serviront à l'alimentation des modèles prédictifs.
- De cette étape doivent sortir des propositions des modèles d'apprentissage profond et/ou par renforcement avec une architecture existante ou à développer.
- Les modèles doivent être ensuite entraînés sur un large échantillon et puis testés pour évaluation de leurs performances,
- Avant la validation des modèles développés, la question d'explicabilité des résultats de prédiction sera aussi étudiée pour identifier d'une manière précise les indicateurs et les combinaisons de paramètres qui ont poussé la machine à prendre une telle décision.
- Et enfin, la validation des modèles se fera sur un échantillon ne faisant pas partie de la base d'apprentissage et de test.
L'objectif étant de sortir avec le meilleur modèle prédictif basé apprentissage profond, apprentissage par renforcement ou système hybride. Un travail d'analyse rétroactive des « features » sélectionnées sera alors mis en place pour identifier et déterminer les meilleurs paramètres de diagnostic et de prédiction des complications obstétricales et néo-natales.

Résultats attendus

- Un modèle de labellisation automatique des données structurées et non-structurées des fiches et dossiers de suivi des femmes enceintes.
- Des modèles Deep learning/NLP, par Reinforcement learning/NLP ou hybrides, permettant de prédire les complications obstétricales et néonatales et d'une plateforme d'aide à la décision qui permet d'accompagner le médecin dans le diagnostic prénatale.
- Un POC (Proof Of Concept) permettant l'industrialisation de la plateforme d'aide aux diagnostics médicaux pour la détection et la prédiction des complications : cas de la femme enceinte.

Précision sur l'encadrement

Directeur de thèse : Imade Benelallam (Directeur du Laboratoire SI2M de l'INSEA)

Co-directeur de thèse : Manar Jalal (UM6SS) et Walid Cherif (Membre du Laboratoire SI2M de l'INSEA)

L'équipe d'encadrement inclut également : Bouyakhf El Houssine (Fondateur du LIMIARF, laboratoire d'Intelligence Artificielle de l'université Mohammed V)

Objectifs de valorisation des travaux de recherche du doctorant :

- Diffusion par communication dans les conférences nationales et internationales et publication dans les revues scientifiques internationales indexées et confidentialité, protection éventuelle par brevet du droit à la propriété intellectuelle, ...
- Valorisation sous forme de publications scientifiques

Collaborations envisagées

- Collaboration interne à l'INSEA et au LIMIARF: co-encadrement. Collaboration avec l'Université Mohammed VI des Sciences de la Santé et ses entités (Ecole Internationale de santé publique, Faculté de médecine, Ecole d'ingénieurs, faculté de pharmacie et faculté des sciences infirmières), l'Hôpital

Universitaire de l'UM6SS (Hôpital Cheikh Khalifa), la Start up AIOX Labs, le CHU de Casablanca et la Faculté de médecine de Casablanca.
Possible collaboration à l'international avec deux partenaires des Nations Unies l'OMS et le Fonds des Nations Unies pour la population (UNFPA).

Profil et compétences recherchées

Connaissances solides en Traitement du langage Naturel, en apprentissage automatique en général, et en particulier en apprentissage par Réseaux de Neurones, apprentissage profond (Deep Learning) et apprentissage par renforcement (Reinforcement Learning). Des expériences en modélisation et extraction de connaissance, en particulier en Text Mining, seront appréciées pour analyser les fiches de surveillance des femmes enceintes.

Niveau de français et d'anglais requis : Compétent/Courant:

Vous pouvez utiliser la langue avec aisance et facilité en argumentant sur des sujets complexes.

Références :

- [1] Sevenster M, Buurman J, Liu P, Peters JF, Chang PJ. Natural Language Processing Techniques for Extracting and Categorizing Finding Measurements in Narrative Radiology Reports. 2015 Sep 30;6(3):600-110. doi: 10.4338/ACI-2014-11-RA-0110. eCollection 2015.
- [2] Zech J, Forde J, Titano JJ, Kaji D, Costa A, Oermann EK, Detecting insertion, substitution, and deletion errors in radiology reports using neural sequence-to-sequence models. 2019 Jun;7(11):233. doi: 10.21037/atm.2018.08.11.
- [3] Tobi Olatunji and Al. Caveats in Generating Medical Imaging Labels from Radiology Reports with Natural Language Processing, Proceedings of Machine Learning Research – Under Review:1–4, 2019
- [4] Kreimeyer K and Al, Natural language processing systems for capturing and standardizing unstructured clinical information: A systematic review. 2017 Sep;73:14-29. doi: 10.1016/j.jbi.2017.07.012. Epub 2017 Jul 17

L'analyse des structures de communautés dans les réseaux sociaux pour la détection de fraude bancaire

Pr. Imade Benelallam Pr. Ahmed Doghmi et M. Ismail Boulahna (Head Data Scientist Salafin)

Description du sujet de thèse

La fraude bancaire continue à augmenter et à imposer des coûts énormes aux entreprises, exigeant le besoin de technologies et de méthodes de détection de la fraude efficaces. D'où le recours à l'application de l'analyse de réseau, une nouvelle frontière de l'analyse avancée des données qui inspecte l'aspect et le caractère social profond de la fraude dans les institutions financières.

Dans cette vision, ce projet de thèse s'intéresse à l'étude des méthodes de détection de communautés dans les réseaux sociaux et la contribution au défi grandissant de la lutte contre la fraude et de la criminalité financière via l'élaboration et l'application de divers algorithmes d'exploration de graphes sur des bases de données afin de fournir des informations approfondies pour détecter et prévenir les cas de fraude enchevêtrés et complexes.

Dans ce cadre, nous abordons ce défi par des nouvelles approches d'analyse prédictive orientée graphe (graph-based). Les bases de données orientées graphes utilisent des méthodes de connexion logiques entre les données (des nœuds ainsi que des arcs), mettant ainsi en avant les relations existantes entre les entités. Ce qui a pour effet de mieux modéliser des relations complexes entre les

clients d'un côté et les parties prenantes de l'autre côté et d'en accélérer ainsi le traitement par les méthodes d'extraction de connaissance et d'apprentissage automatique. D'autres techniques telles que la simulation multi-agents de scénarios de théorie des jeux peuvent aussi être appliquées pour modéliser les structures des contrôles, et les chaînes de transaction afin de détecter et de mieux comprendre le risque de fraude commerciale, un problème persistant qui provoque des pertes institutionnelles financières sans cesse.

Mots clés

Social network analysis, predictive modelling, random graph generator, fraud detection, game theory

Objectif et contexte

L'objectif de ce projet est de contribuer sur les méthodes de modélisation de réseaux sociaux appliquée à la prévention et la détection de la fraude au sein des institutions financières et permettant d'appréhender au mieux l'aspect tridimensionnel des données bancaires.

- Contexte sociétal, économique et/ou industriel :

L'impact de la fraude sur les organisations devient de plus en plus coûteux. Chaque année, les institutions financières se font imposer des coûts énormes en revenus à cause de la fraude bancaire. De plus, avec l'augmentation croissante de la mondialisation et les progrès inexorables des technologies de la communication, la fraude augmente chaque année en volume, en ampleur et en sophistication. Certes, la plupart des techniques existantes reposent uniquement sur les règles métier développées par les experts, qui nécessitent une grande quantité de saisie de l'utilisateur et doivent être mises à jour en permanence. Cependant, l'ampleur et la complexité croissantes des systèmes de fraude signifient que les fraudeurs sont de plus en plus capables de contourner et d'éviter de telles techniques. D'où l'intérêt à de des approches prometteuses et des développements clés dans l'utilisation de l'analyse de réseaux sociaux pour saisir, voire de prévoir, la fraude de manière plus efficace.

- Contexte Scientifique :

La pertinence de ce projet réside dans la possibilité d'automatiser plusieurs processus d'analyse des réseaux sociaux de fraudes à savoir les étapes de détection, de surveillance et d'intervention dans un cycle typique de fraude, tout en associant des données provenant de sources multiples, notamment des scores de crédit, des données transactionnelles d'entreprise et des médias sociaux. Les principales contributions dans la recherche sur la détection des fraudes au moyen d'algorithmes basés sur des graphes seront cités en références.

Méthode

La méthodologie pour le développement d'algorithmes prédictifs basés sur les modèles d'apprentissage suivra les étapes ci-après:

- Dresser l'état de l'art des travaux de recherche réalisés dans le cadre des modèles de prévention et de détection des fraudes via l'analyse des réseaux sociaux et d'autres techniques basées sur la simulation multi-agents de scénarios de théorie des jeux
- Etablir une approche d'extraction et de récolte de données supplémentaires sur les parties prenantes dans le crédit éventuellement les données dites de « third party data».
- Mettre en œuvre une approche de nettoyage et prétraitement des données provenant de plusieurs sources à caractère structuré et non structuré.
- En ce qui concerne la partie modélisation, l'approche de réseaux consiste en premier lieu à utiliser un concept très connu en sociologie 'Homophilic Networks' combiné à des tests statistiques, méthodes de simulation multi-agents, algorithmes de classification itérative.

Résultats attendus

- Établir une architecture de modèles de fraude en termes de mise en œuvre générale de haut niveau (présentant un intérêt pour la détection de calcul et l'apprentissage automatique) ainsi que de détails de niveau plus détaillé (intéressant pour les audits et les spécialistes des fraudes financières).
- Développer une plate-forme hybride permettant de superposer des détails concernant des cas de fraude en termes de domaines multiples (par exemple, transactions financières, communications, associations de personnes, propriété d'entreprise)
- Réaliser une plateforme as a service pour la visualisation interactive pour la découverte de modèles et une base pour la détection de modèle dans de grands ensembles de données sur la base de schémas pré-identifiés.
- Intégrer l'identification de nouveaux systèmes via l'analyse de modèles statistiques et l'apprentissage automatique.
- Fournir une mesure quantitative des schémas de fraude pouvant être utilisée plus généralement pour identifier la propension statistique (c'est-à-dire la longueur du chemin et le volume des transactions le long des chemins définis dans des graphes isolés constituant un facteur de risque mesurable pour la fraude).

Précision sur l'encadrement

Directeur de thèse :

Co-directeur de thèse :

L'équipe d'encadrement inclut également :

Objectifs de valorisation des travaux de recherche du doctorant :

- Diffusion par communication dans les conférences nationales et internationales et publication dans les journaux scientifiques internationaux indexés et confidentialité, protection éventuelle par brevet du droit à la propriété intellectuelle, ...
- Valorisation sous forme de publications scientifiques

Profil et compétences recherchées

- Connaissances solides en statistiques inferentielles, théorie de graphe, séries temporelles et machine learning en général, deep learning en particulier.
- Niveau de français requis : Compétent/Courant: Vous pouvez utiliser la langue avec aisance et facilité en argumentant sur des sujets complexes.

Références :

Social Network Analysis for Insurance Fraud Détection, Nuno C. Garcia, UNIVERSIDADE DA BEIRA INTERIOR, Engenharia

Classification for Fraud Detection with Social Network Analysis, Miguel Pironet San-Bento Almeida, Universidad tecnica de Lisboa

Implementing social network analysis for fraud prévention, CGI Group Inc., 2011

Fraud Analytics Using Descriptive, Predictive, and Social Network Techniques, A Guide to Data Science for Fraud Detection; Bart Baesens, Véronique Van Vlasselaer
Wouter Verbeke, wiley

APATE A Novel Approach for Automated Credit Card Transaction Fraud Detection Using Network-Based Extensions. Van Vlasselaer, V., Bravo, C., Caelen, O., Eliassi-Rad, T., Akoglu, L., Snoeck, M., Baesens, B. (2015b), *Decision Support Systems* 75 (2015): 38–48.

Learning and inference in massive social networks. Hill, S., Provost, F., & Volinsky, C. (2007, August). In: *The Fifth International Workshop on Mining and Learning with Graphs*. Florence, Italy.

Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. Geman, S., & Geman, D. (1984), *IEEE Transactions on Pattern Analysis and Machine Intelligence* (6): 721–741.

Link-based Classification. Lu, Q., & Getoor, L. (August 2003). In *ICML* 3:496–503.

Enhanced Hypertext Categorization Using Hyperlinks. Chakrabarti, S., Dom, B., & Indyk, P. (1998, June). In *ACM SIGMOD Record* 27 (2): 307–318. ACM.

Continuous

Fraud Monitoring and Detection via Advanced Analytics: State-of-the-Art Trends and Directions. Deloitte. 2014 ACFE European FRAUD CONFERENCE

ACFE 'Report to the Nations 2012 Global Fraud Study.

Machine learning and NLP-based approach for constraint acquisition problems

Pr. Imade BENELALLAM

Nowadays, the extent of the constraint programming paradigm as a friendly user platform remains limited. This level of restriction is due to the necessary mathematical modelling and programming knowledge to address many different types of NP-complete decision problems, such as planning, scheduling, and resource optimization problems.

Constraint acquisition algorithms [1][3][5], passive or active learners, are based on version space learning. The user provides solutions and non-solutions of his problem, and the learning algorithm builds the target constraint network, defining the set of variables, domains and constraints. The user needs to know beforehand solutions and non-solutions of his problem for passive learning and must classify correctly the given solutions for active learning. In addition to that, the constraint acquisition algorithms need to ask the user too many questions, which can be simple or more complex queries, to build an accurate target constraint network. The membership queries asked can be complete or partial assignments of the set of variables, and the user must classify them as solutions or non-solutions of his problem.

For this reason, we aim to combine natural language processing, through a conversational agent, with constraint acquisition, in order to make constraint programming paradigm accessible to a larger number of potential users. The aim of this work is to build an intelligent constraint acquisition system (figure 1), allowing a novice user to express a problem he wishes to solve in natural language. The system models the problem stated as a constraint network, solves it and offers possible solutions to the user through a solver.

This system would be based on an intelligent conversational agent, who allows, through an automatic learning process and interactive learning, to analyse and process the problem asked by the user and define the variables, domains and constraints of the problem. Once the target constraint network is formalized, it is provided to the solver to solve it.

A deep automatic learning algorithm [2][6] is used to identify the elements of the formal representation of constraint network. It is important that the interactions (questions / answers) with the user are optimized, and that the solver provides the user the best possible solutions.

To correctly formalize the problem, two requirements must be executed correctly: (a) Parsing the users question in natural language, (b) Correctly mapping the elements to the corresponding formal representation.

The first task is processed through natural language processing approaches. Recently, Deep Learning algorithms have made impressive improvements in NLP tasks [7] such as part of speech tagging, parsing and machine translation, with neural networks based on dense vector representations. Convolutional neural networks, recurrent neural networks and recursive neural networks are deep learning models, which applied to natural language processing, have good performance and results for the related tasks.

The approach we are adopting is to build a training database, through a variety of examples of users problems expressed in natural language, classified by domain, and train the neural model to correctly generate the parsing tree [8].

The mapping to the elements of the formal representation would be processed through a classification algorithm, which is a neural network model as well.

The formal representation of the problem is a constraint network, which is built from the elements identified from the user's utterance: The variables, domains and constraints. Once the consistency of the network verified, it is solved by the constraint solver.

The solutions provided to the user must be in an understandable format, not only assignments of variables, since assignments are in some cases not understood by the user.

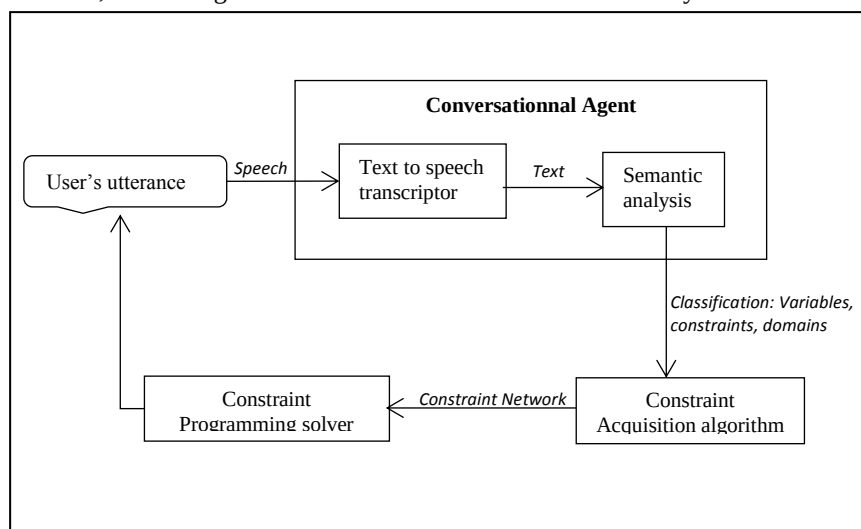


Fig.1: Targeted Acquisition system architecture

Keywords :

Constraint Acquisition, Constraint programming, Deep Learning, Constraint Solvers

References

- [1] Robin Arcangioli, Nadjib Lazaar, *Multiple Constraint Acquisition*, IJCAI workshop on Intelligent Personalization, 5 pages (2015)
- [2] Ronan Collober, Jason Weston, *A unified architecture for natural language processing: deep neural networks with multitask learning*, ICML '08 Proceedings of the 25th international conference on Machine learning, Pages 160-167 (2008)

- [3] C. Bessiere, R. Coletta, E. Hebrard, G. Katsirelos, N. Lazaar, N. Narodytska, C.G. Quimper, and T. Walsh, *Constraint Acquisition via Partial Queries*, IJCAI, 7 pages (2013)
- [4] Anton Dries, Angelika Kimmig, Jesse Davis, Vaishak Belle, Luc de Raedt, *Solving Probability Problems in Natural Language*, IJCAI, 7 pages (2017)
- [5] C. Bessiere, R. Coletta, A. Daoudi, N. Lazaar, Y. Mechqrane, E. Bouyakhf, *Boosting Constraint Acquisition via Generalization Queries*, ECAI, 6 pages (2014)
- [6] E. Cambria and B. White, *Jumping NLP curves: A review of natural language processing research*, IEEE Computational Intelligence Magazine, vol. 9, no. 2, pp. 48–57, 2014.
- [7] Y. Goldberg, *A primer on neural network models for natural language processing*, Journal of Artificial Intelligence Research, vol. 57, pp. 345–420 (2016).
- [8] R. Socher, C. C. Lin, C. Manning, and A. Y. Ng, *Parsing natural scenes and natural language with recursive neural networks*, Proceedings of the 28th international conference on machine learning (ICML11), pp. 129–136 (2011)

New approach to improve Writer identification system based on deep learning method

Pr. Adil Kabbaj et Pr Yaacoub Hannad

1. Domaine de recherche :

Ce sujet de recherche peut être classé sous plusieurs domaines globaux, notamment :

- Traitement d'image,
- Reconnaissance des formes (Pattern recognition),
- Biométrie.

2. Présentation du sujet:

La reconnaissance et la vérification de l'identité de l'auteur d'un texte manuscrit est devenu une nécessité dans le domaine bancaire et judiciaire. C'est dans ce contexte qu'on peut introduire l'importance de « l'identification hors ligne du scripteur à partir de l'image de son écriture manuscrite » qui est un sujet de recherche récent et actif.

L'identification du scripteur consiste à trouver la personne rédactrice d'un échantillon de texte en le comparant avec les caractéristiques de son écriture enregistrées dans une base de données. Cette identification est devenue possible grâce à la validation de l'hypothèse de l'individualité de l'écriture manuscrite à base duquel on peut extraire différents types de formes, indices et caractéristiques pour identifier la personne [Sirahi 2002].

3. Problématique:

Les systèmes d'identification hors ligne de scripteurs à base du texte Latin connaissent un taux d'identification largement supérieur à ceux appliquées sur le script Arabe, même si ces derniers utilisent presque les mêmes méthodes et démarches d'identification. Pour cela, nous nous intéressons dans ce sujet à proposer une nouvelle approche basée sur le deep learning pour réaliser un système d'identification de scripteurs performant pour différents scripts (surtout le script Arabe et le script Latin).

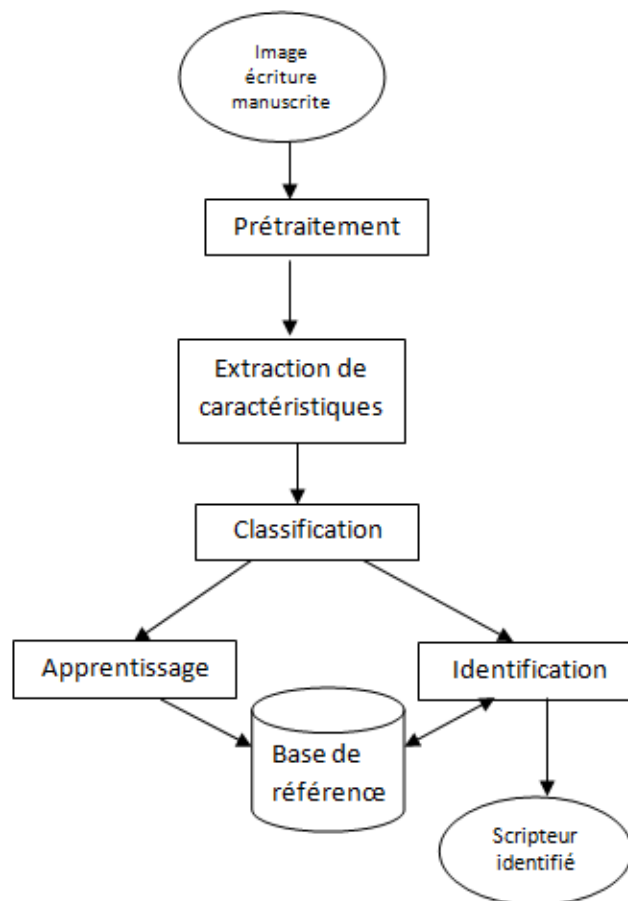


Figure 1 : Système d'identification hors ligne de scripteurs

4. Sujet proposé:

1^{er} volet: Une approche basée sur le Deep learning pour l'identification de scripteurs

L'étudiant chercheur est invité à proposer une nouvelle approche pour réaliser un système d'identification de scripteurs performant basé sur une approche Deep learning. Voir le travail récent de: [He 2019] *He, Sheng, and Lambert Schomaker. "Deep adaptive learning for writer identification based on single handwritten word images." Pattern Recognition 88 (2019): 64-74.*

L'approche proposée sera testée et validée en utilisant les bases de données les plus populaires dans ce domaine:

- La base de données IAM pour le script Latin.
- Les bases de données IFN/ENIT, KHATT et QUWI pour le script Arabe.

Etapes de travail:

- le chercheur doit établir un état d'art complet et analyser les différentes méthodes proposées: Points forts et points faibles de chaque méthode ...
- le chercheur doit également analyser et voir les nouveautés des différents systèmes biométriques des autres domaines: reconnaissance faciale, vérification de signature....
- le chercheur aura le but de proposer et réaliser une nouvelle approche pour l'identification de scripteurs: nouveau type de caractéristiques, nouvelle méthode d'identification basée sur une approche Deep learning...

2^{ème} volet: Système d'identification de scripteurs à base de Multi script

Actuellement, une nouvelle tendance dans le domaine d'identification de scripteurs vers l'utilisation de multi script, c.-à-d. l'apprentissage du système se fait en utilisant des manuscrits en anglais et le test se base sur des manuscrits d'une autre langue (par exemple l'arabe), et ces deux manuscrits sont rédigés par la même personne. voir le travail [Chawki multi script 2012] (Djeddi, C., Siddiqi, I., Souici-Meslati, L., & Ennaji, A. (2012, September). *Multi-script writer identification optimized with retrieval mechanism. In Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on* (pp. 509-514). IEEE).

Le but de ce volet du sujet est de réaliser un système d'identification de scripteurs puissant en multi script, et avoir des résultats concurrents par rapport à l'état d'art de ce domaine.

Pour ce faire, l'étudiant chercheur doit établir d'abord un état d'art sur les systèmes multi-scripts, ensuite il doit réaliser son propre système tout en proposant sa nouvelle approche pour ce sujet.

Références

- [Chawki 2008]*(Article en français)** D.Chawki, S. Labiba, "Une approche locale en mode indépendant du texte pour l'identification de scripteurs: Application à l'écriture arabe", CIFED'08, Rouen, France, pp. 151-156, October 2008.
- [He 2019]** He, Sheng, and Lambert Schomaker. "Deep adaptive learning for writer identification based on single handwritten word images." *Pattern Recognition* 88 (2019): 64-74.
- [Chawki multi script 2012]** Djeddi, C., Siddiqi, I., Souici-Meslati, L., & Ennaji, A. (2012, September). Multi-script writer identification optimized with retrieval mechanism. In *Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on* (pp. 509-514). IEEE
- [Yaacoub 2015]** Yaâcoub Hannad, Imran Siddiqi, and Mohamed El Youssfi El Kettani. Arabic Writer Identification Using Local Binary Patterns (LBP) of Handwritten Fragments. In *IbPRIA 2015, LNCS 9117*, (pp.237–244).Springer.
- [Yaacoub 2019]** Yaâcoub HANNAD, Imran SIDDIQI, Chawki DJEDDI and Mohamed El Youssfi EL KETTANI. « Improving Arabic Writer Identification using Score Level Fusion of Textural Descriptors », *IE-Biometrics* (2019), DOI: 10.1049/iet-bmt.2018.5009.
- [Srihari 2002]** Srihari, S. Cha, H. Arora, and S. Lee, "Individuality of Handwriting," *J. Forensic Sciences*, vol. 47, no. 4, pp. 1-17, July 2002.
- [Balucu 2007 IEEE]** M.Bulacu, L.Schomaker, Text-independent writer identification and verification using textural and allographic features, *IEEE Transactionson Pattern Analysis and Machine Intelligence* 29 (4)(2007)701–717.
- [Imran 2010]** Imran Siddiqi, NicoleVincent. Text independent writer recognition using redundant writing patterns with contour-based orientation and curvature features. *Pattern Recognition* 43(2010)3853–3865
- [Chawki 2010]** D. Chawki, S. Labiba, "A texture based approach for Arabic Writer Identification and Verification", *IEEE International Conference on Machine and Web Intelligence*, pp 115 – 120, October 2010.
- [Bertolini 2013]** D Bertolini, LS Oliveira, E Justino, R Sabourin, "Texture-based descriptors for writer identification and verification", *Expert Systems with Applications*, vol. 40, pp 2069-2080, 2013.
- [Balucu ICDAR 2007]** M. Bulacu, L. Schomaker, A. Brink, "Text-Independent Writer Identification and Verifica-tion on Offline Arabic Handwriting", *Proc. of 9th Int. Conference on Document Analysis and Recognition (ICDAR 2007)*, IEEE Computer Society, pp. 769-773, vol. II, 23 - 26 September, Curitiba, Brazil, 2007.
- [Schomaker 2004]** L. Schomaker, M. Bulacu, K. Franke, Automatic Writer Identification Using Fragmented Connected-Component Contours, *Proc. of 9th IWFHR, Japan, Los Alamitos IEEE Computer Society*, pp. 185-190
- [Khalifa 2015]** E Khalifa, S Al-maadeed, M A Tahira, A Bouridane, A Jamshed. Off-line Writer Identification Using an Ensemble of Grapheme Codebook Features. *Pattern Recognition Letters*, Volume 59, 1 July 2015, Pages 18–25
- [Nidhal 2014]** Mohamed Nidhal Abdi, Maher Khemakhem, A model based approach to offline text-independent arabic writer identification and verification, *Pattern Recognition*, Volume 48, Issue 5, May 2015, Pages 1890–1903

MATHEMATIQUES APPLIQUEES

**Problèmes d'optimisation à deux niveaux :
existence de solutions, dualité et conditions d'optimalité**

(Profs. Kadrani et Aboussoror)

Dans ce sujet on s'intéresse aux problèmes de l'optimisation à deux niveaux dans ses deux principales formulations forte et faible.

Soient $F, f : \mathbb{R}^p \times \mathbb{R}^q \rightarrow \mathbb{R}$ deux fonctions, X et Y deux sous-ensembles non vides de $\mathbb{R}^p$ et $\mathbb{R}^q$ respectivement.

- Formulation faible (ou pessimiste) : Le problème d'optimisation à deux niveaux considéré se formule dans ce cas comme suit:

$$(\mathcal{P}_p) \quad \min_{x \in X} \sup_{y \in \mathcal{M}(x)} F(x, y)$$

où $\mathcal{M}(x)$ est l'ensemble des solutions du problème

$$\mathcal{P}(x) \quad \min_{y \in Y} f(x, y).$$

- Formulation forte (ou optimiste) : Le problème d'optimisation à deux niveaux considéré se formule dans ce cas comme suit

$$(\mathcal{P}_o) \quad \min_{x \in X} \inf_{y \in \mathcal{M}(x)} F(x, y).$$

Les problèmes $(\mathcal{P}_p)$ et $(\mathcal{P}_o)$ sont appelés problèmes de premier niveau, et pour un x donné dans X , le problème $\mathcal{P}(x)$ est appelé problème de second niveau. Les problèmes d'optimisation à deux niveaux sont bien connus par leurs difficultés théorique et numérique. L'objet de cette thèse est de définir des dualités et de donner des conditions d'optimalité pour ces deux problèmes. Pour les problèmes à formulation faible qui présentent des difficultés dans leur étude, nous allons établir les conditions suffisantes qui assurent l'existence de solutions et essayer de les étendre à d'autres espaces topologiques.

Bibliographie

- [1] A. Aboussoror, S. Adly and F. E. Saissi, *An extended Fenchel-Lagrange duality approach and optimality conditions for strong bilevel programming problems*, SIAM Journal on Optimization, Vol. 27(2), 1230-1255, 2017.
- [2] A. Aboussoror, *Weak bilevel programming problems: Existence of solutions*, Adv. Math. Res. 1, 83-92, Hauppauge, New York, 2002.
- [3] K. Shimizu Aiyoshi, Y. Ishizuka and J. F. Bard, *Non differentiable and two-level mathematical programming*, Kluwer Academic Publishers, Boston, 1997.

Bilinear Stabilization of Delayed Systems

(Profs. Kadrani et El harraki)

Let X be a complex Hilbert space with norm and inner product denoted respectively by $\|\cdot\|_X$ and $\langle \cdot, \cdot \rangle$. Let A be a linear unbounded self-adjoint and strictly positive operator in X . Let $V = D(A^{1/2})$ be the domain of $A^{1/2}$. We denote by $(D(A^{1/2}))'$ the dual space of $(D(A^{1/2}))$ with respect to the inner product in H . Suppose that $B_i \in \mathcal{L}(X, D(A^{1/2})')$. The systems that we consider are described as follows:

$$\begin{aligned} y_{tt} + Ay + u(t)B_1y + B_2y_t(t - \tau) &= 0 \\ y(x, 0) = y_0(x), \quad y_t(x, 0) = y_1(x), \quad B_2^*y_t(-s) &= f_0(-s), \quad s \in (0, \tau), \end{aligned} \quad (1)$$

where the scalar function $v(t)$ denotes the control. We aim to establish an implication between the stabilization of this system and the observability of the corresponding undamped system. This will generalize an earlier result by Haraux [3] who studied the following system:

$$\begin{aligned} y_{tt} + Ay + By &= 0 \\ y(x, 0) = y_0(x), \quad y_t(x, 0) &= y_1(x), \end{aligned} \quad (2)$$

with $B \in \mathcal{L}(X)$. The unbounded version of (2) was studied by [1] where the considered second order evolution equations with unbounded feedbacks. Under a regularity assumption we show that observability properties for the undamped problem they proved decay estimates for the damped problem. For the delayed case we refer [6,7] however this result considers a stabilization by a linear feedback. For the problem (1), we consider that will be of a great interest to treat the bilinear control in the presence of the delay.

Key words: Semilinear Systems with delay, Feedback stabilization, Observability inequality, Exponential decay, Semigroup.

References

- [1] K. Ammari, S. Nicaise, *Stabilization of elastic systems by collocated feedback*, Lecture Notes in Mathematics, Vol. 2124, Springer-Verlag, Berlin, (2015).
- [2] H. Brezis, *Analyse fonctionnelle, Théorie et applications*, Masson, Paris, (1983).
- [3] A. Haraux, *Une remarque sur la stabilisation de certains systèmes du deuxième ordre en temps*, Port.Math.46 245–258, (1989).

- [4] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Springer-Verlag, New-York, (1983).
- [5] L. Tebou, *Equivalence between observability and stabilization for a class of second order semilinear evolution equation*, Discrete Contin. Dyn. Syst. 744-752, (2009).
- [6] S. Nicaise, C. Pighotti, *Exponential stability of abstract evolution equations with time delay*, J. Evol. Equ. 15(2015), 107-129.
- [7] E.M. Ait Ben Hassi, K. Ammari, S. Boulite, L. Maniar, *Feedback stabilization of a class of evolution equations with delay*, J. Evol. Equ. 1 (2009), 103-121.

Phd students in games theory

Ahmed Doghmi

October 25, 2019

*National Institute of Statistics and Applied Economics, Madinat Al Irfane, Rabat Institutes,
10100 Rabat, Morocco.*

Game theory does not rest only limited to take the rules of the game modeling interaction as given, but it was considerably developed to study the consequences of different rules. This was due to its openness to other theories, especially its connection to *social choice theory*. This has led to the development of the important area in the literature of games theory: *mechanism design theory*, in particular its part called *implementation theory*. Conversely to the classical economic theory which studies *social outcomes* generated by the institutions, this theory identifies in the first a *social goal* or *desired outcomes*, and in the second, the institutions under which the *predicted outcomes* of the individual strategic interactions follow. The outcomes can take many forms following the context. For instance in some political situations, they represent the candidate elected in a society where each member has a preference ranking over the candidates, and in certain economic environments they constitute the production and allocation of public and private goods and services in a community where each agent has *preferences* over different alternatives. The institutions are the set of rules, procedures, or game forms for determining outcomes, and so they called *mechanisms*. They, once again, deponed on the setting. As examples; in national or communal political elections, mechanisms are electoral procedures, which are usually prescribed ahead of time, sometimes even by the country's constitution; in the case of the public goods provision, the mechanisms are the choices of the methods by which the levels of provision and financing are determined; in the situation of the house-building, mechanisms are contracts between the home buyers and the builders, which establish the rights and responsibilities of each, etc...The study of the relationship between institutions and outcomes can be viewed as a *design*, and in this form, is called *the theory of mechanism design*. This designation is often used when there is a *unique equilibrium* as the relevant solution concept that captures the strategic interactions of agents, and a *direct mechanism*, which are used as a common tool. When there are multiple equilibria which have desirable properties, and *specific mechanisms*, which are *indirects*, to keep track of all equilibria, we are talking about a specific part of the theory of mechanism design, termed *implementation theory*.

Given a social goal, which is represented by an objective function called *social choice correspondence* (SCC) or *rule* that selects the desired outcomes, if a planner (Representant of State, a Local collectivity, a Region, Manager of a firm,...) knew in advance the preferences of agents, they will have no problem for achieving this goal. However, the difficulty lies in the

asymmetry of information between planner and agents. Implementation theory characterizes when the planner can design a mechanism whose predicted outcomes, that can be produced by the individual strategic interactions, coincide with the desired outcomes, according to that social objective. To achieve this goal, some conditions should be imposed on the individual behaviors. Thus, Maskin (1977,1999) was the first who discovered an important connection between the class of implementable SCC in Nash equilibria and a property called *Maskin monotonicity*. This condition stipulates that if an outcome a is optimal according to an SCC in a state or profile R and if the alternatives ranked below a for all agents remain ranked below it (in large sense) in going from R to a new profile R' , then the outcome a must be socially chosen in R' . Maskin proved that any Nash implementable SCC must satisfy Maskin monotonicity. However, this condition alone can not guarantee the implementability of an SCC. Hence, Maskin provided an additional condition called *no-veto power*. This property requires that if an outcome is top-ranked for all agents except possibly one, it must be chosen. Maskin showed that any SCC satisfying monotonicity and no-veto power can be implemented in Nash equilibria with at least three agents.

In this context, we would like to study the problem of implementation in two different solution concepts: 1) Dominant strategy equilibrium and 2) Undominated strategy equilibrium. For the first solution concept, Yao and Yi (2007) provided a complete characterization of a SCC social choice rules which can be implemented in dominant strategies when the preferences of agents are strict. Hence, we would like to relax this result by allowing indifferences between some alternatives (like single-peaked preferences, single-dipped preferences,).

For the second solution concept, Palfrey and Srivastava (1991) provided a sufficient condition for implementing a SCC in undominant strategies. However, their condition is based on a certain property, called Holaucaust alternative (each agent has a worst alternative), which is very restrictive and is not verified in many applications. Hence, we would like to address this shortcoming by proposing new sufficient condition that can be applied to more general domains.

To apply, the candidate must choose one of the two next topics:

- 1) Mathematical games theory: implementation in dominated strategies.
- 2) Mathematical games theory: implementation in **undominated** strategies.

References

- [1] E. Maskin. Nash equilibrium and welfare optimality. *Review of Economic Studies*, 66:23–38, 1999.
- [2] H. Palfrey and J. Srivastava. Nash implementation using undominated strategies. *Econometrica*, 59:479–5001, 1991.

- [3] H. Yao and J. Yi. Social choice rules implemented in dominant strategies. *Economics Letters*, 97:197–200, 2007.

Développement de nouvelles techniques d'optimisation et de modélisation pour un mix énergétique efficace.

Sujet proposé par :

Pr. Mohamed OUZINEB, INSEA.

Ce sujet est en collaboration avec le professeur Mustapha Nour El Fath de l'université Laval. Possibilité de stage de 3 à 6 mois par an à l'université de Laval.

L'objectif de la thèse est de développer des algorithmes efficaces de gestion de l'énergie. Ces algorithmes permettent l'adéquation optimale entre la disponibilité des sources et la consommation énergétique, tout en minimisant les coûts totaux sujets à des contraintes économiques et environnementales. Ces algorithmes seront basés sur les prévisions de la demande et la disponibilité des technologies. Afin d'assurer la pertinence des propositions, l'étude sera basée sur des exemples concrets s'appuyant sur des firmes situées dans des régions marocaines pour lequel des données de conception et d'exploitation sont disponibles.

Dans une première phase, il sera effectué une analyse de la consommation énergétique. Le candidat recruté doit recueillir d'abord toute l'information disponible relative à la consommation énergétique auprès des différentes entités concernées au Maroc : Ministère de l'Énergie, des mines et du développement durable et l'Office National de l'Électricité et de l'Eau Potable.

Dans une deuxième phase, il sera effectué une caractérisation des disponibilités des sources énergétiques ainsi que des éléments de stockage. Ces éléments seront alors modélisés avec un niveau d'abstraction cohérent avec celui des fonctions consommatrices. Ces différents modèles seront intégrés avec des algorithmes d'optimisation et de contrôle qui seront proposés afin de maximiser l'efficacité énergétique tout en satisfaisant la demande.