الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي و البحث العلمي

Université Abou Bekr Belkaid
Tlemcen Algérie

جامعة أبي بكر بلقايد
تلمسان الجزائر

9ème Colloque sur l'Optimisation et les Systèmes d'Information

COSI'2012, 12-15 mai 2012, Tlemcen

Posters

Organisé par les Départements Informatique et Mathématiques
Université Abou Bekr Belkaid de Tlemcen

Actes des posters du Neuvième Colloque sur l'Optimisation et les Systèmes d'Information- COSI'2012

12-15 Mai 2012, Tlemcen, Algérie

Université Abou Bekr Belkaïd-Tlemcen

Faculté des Sciences

Départements de Mathématiques et d'Informatique

Contents

Préface

Ces actes regroupent les posters présentés lors de la 9ème édition du Colloque sur l'Optimisation et les Systèmes d'Information (COSI 2012) qui s'est déroulé à Tlemcen, Algérie, du 12 au 15 mai 2012.

COSI est une manifestation scientifique pluridisciplinaire qui rassemble des chercheurs qui travaillent dans les domaines de : Théorie des Graphes et Combinatoire, Recherche Opérationnelle, Traitement d'Images et Vision Artificielle, Intelligence Artificielle et Systèmes d'Information. Les précédentes éditions de COSI 2012 ont eu lieu à : Guelma (2011), Ouargla (2010), Annaba (2009), Tizi-Ouzou (2008), Oran (2007), Alger (2006), Béjaia (2005) et Tizi-Ouzou (2004).

Le comité de programme a examiné 242 articles soumis à COSI 2012. A la fin du processus d'évaluation, 37 articles longs (soit un taux d'acceptation de 15,29%) et 22 posters étaient acceptés pour publication.

Les articles contenus dans ces actes représentent de façon tout à fait homogène l'ensemble des thématiques couvertes par COSI.

Nous remercions les auteurs pour leurs excellentes contributions, les membres du comité de programme, les relecteurs externes, les membres du comité d'organisation et les sponsors.

Nous sommes particulièrement heureux que quatre chercheurs de très haut niveau aient accepté de nous présenter une conférence invitée :

- Ne votez pas : jugez ! par ***Michel Balinski*** (Ecole Polytechnique et CNRS, France).
- Fouille de données et calcul scientifique dans le cloud : vers un environnement virtuel collaboratif ubiquitaire pour la recherche par ***Karim Chine*** (Cloud Era Ltd, Cambridge, UK).
- On the power of graph searching par ***Michel Habib*** (LIAFA, Université de Paris Diderot - Paris 7, France).
- L'optimisation globale déterministe : méthodes, applications et extensions par ***Frédéric Messine*** (IRIT, ENSEEIHT, Toulouse, France).

Mohand-Saïd Hacid (Président du CP)
Méziane Aïder (Vice-Chair : Théorie des Graphes et Combinatoire)
Mourad Baiou (Vice-Chair : Recherche Opérationnelle)
Nacéra Benamrane (Vice-Chair : Traitement d'Images et Vision Artificielle)
Hélène Fargier (Vice-Chair : Intelligence Artificielle)
Hassina Seridi & Michel Schneider (Vice-Chair : Systèmes d'Information)

Organisation

Université Aboubakr Belkaïd-Tlemcen

Président d'honneur

Professeur Norreddine Ghouali
Université Aboubakr Belkaïd, Tlemcen, Algérie

Comité d'Organisation

Président

Pr. Boufeldja Tabti, Université Aboubakr Belkaïd, Tlemcen, Algérie

Vice-Président

Pr. Benmiloud Mebkhout, Université Aboubakr Belkaïd, Tlemcen, Algérie

Membres

Pr Ghalem Saïd - Vice doyen, Université Aboubakr Belkaïd, Tlemcen, Algérie
Dr Mebkhout Benmiloud -Chef de département de Mathématiques, Université Aboubakr Belkaïd, Tlemcen, Algérie
Mr Bentifour Rachid, Université Aboubakr Belkaïd, Tlemcen, Algérie
Mr Mamchaoui Mohammed, Université Aboubakr Belkaïd, Tlemcen, Algérie
Mr Menouar Mohammed El Amine, Université Aboubakr Belkaïd, Tlemcen, Algérie
Mr Benchaib Abdelatif, Université Aboubakr Belkaïd, Tlemcen, Algérie

Secrétariat

Mme Cherki Samira Ep. Bounacer
Melle Hamouni Kheira

Comité de Pilotage

Mohamed Aidene, Université Mouloud Mammeri de Tizi-Ouzou, Algérie
Nacéra Benamrane, Université des Sciences et Technologie d'Oran, Algérie
Abdelhafidh Berrachedi, Université des Sciences et Technologie Houari Boumédienne, Alger, Algérie
Mohand-Saïd Hacid, Université de Lyon I, France
Lhouari Nourine, Université de Clermont-Ferrand II, France
Brahim Oukacha, Université de Tizi-Ouzou, Algérie
Jean Marc Petit, INSA de Lyon, France
Bachir Sadi, Université de Tizi-Ouzou, Algérie
Lakhdar Saïs, CRIL - CNRS, Université d'Artois, France
Kamel Tari, Université Abderahmane Mira de Bejaia, Algérie

Comité de Programme

Président

Mohand-Said Hacid, LIRIS, Université Lyon I (France)

Vice-Chairs

Méziane Aider, USTHB (Algérie)
Mourad Baiou, LIMOS (France)
Nacéra Benamrane, USTO (Algérie)
Hélène Fargier, IRIT (France)
Michel Schneider, ISIMA (France)
Hassina Seridi, Université Annaba (Algérie)

Membres

Abdelmalek Amine, Université de Saida (Algérie)
Adi Kamel, Université de Québec en Outaouais (Canada)
Ahmed-Nacer Mohaned USTHB (Algérie)
Ahmed-Ouamer Rachid Université de Tizi-Ouzou (Algérie)
Aidene Mohamed Université de Tizi-Ouzou (Algérie)
Aït Haddadene Hacène USTHB Alger (Algérie)
Aït Mohamed Otmane, université Concordia (Canada)
Alimazighi Zaia, USTHB (Algérie)
Arab Ali Cherif, Université Paris 8 (France)
Badache Nadjib CERIST (Algérie)
Barkaoui Kamel, CNAM-Paris (France)
Barki Hichem, INRIA Bordeaux (France)
Barra Vincent, UBP, Clermont-Ferrand (France)
(France)
Belbachir Hafida USTO Oran (Algérie)
Bellatreche Ladjel LISI, Université de Poitier (France)
Benamar Abdelkrim, Université Abou Bekr Belkaid Tlemcen (Algérie)
Benferhat Salem, Université d'Artois, Lens (France)
Benmammar Badr, Université Abou Bekr Belkaid Tlemcen (Algérie)
Ben Yahia Sadok Faculté des Sciences de Tunis (Tunisie)
Benatallah Boualem University of New South Wals (Australie)
Benbernou Salima Université Paris Descartes (France)
Benhamou Belaid Université d'Aix-Marseille I (France)
Berrachedi Abdelhafid USTHB Alger (Algérie)
Bessy Stéphane, Université Montpellier II (France)
Bibi Mohand Ouamer Université de Béjaia (Algérie)
Bouchekif Mohamed, Université Abou Bekr Belkaid,Tlemcen (Algérie)
Bouchemakh Isma USTHB Alger (Algérie)
Boudhar Mourad, USTHB (Algérie)
Boufaida Mahmoud Université Mentouri, Constantine (Algérie)
Boufaïda Zizette, Université Mentouri, Constantine (Algérie)
Boughanem Mohand IRIT, Toulouse (France)
Bouzeghoub Amel Telecom Sud, Paris (France)
Bouzouane Abdenour UQC (Canada)
Champion Thierry Université du Sud Toulon-Var (France)
Chaouki Babahenini Mohamed Univ. Mohamed Khider,

Biskra (Algérie)
Chellali Mustapha, Université Saad Dahlab, Blida (Algérie)
Chikh Mohammed El Amine, Univ. Abou Bekr Belkaid Tlemcen (Algérie)
Crouzeix Jean-Pierre, UBP, Clermont-Ferrand (France)
d'Orazio Laurent Université Blaise Pascal Clermont-Ferrand (France)
Didi Fedoua Université Abou Bekr Belkaid Tlemcen, (Algérie)
Djedi Nouredine Université Biskra (Algérie)
Elghazel Haytham Université Claude Bernard Lyon 1 (France)
Ellaggoune Fateh Université de Guelma (Algérie)
Engelbert Mephu, Université Blaise Pascal, Clermont-Ferrand (France)
Farah Nadir, Université Badji Mokhtar d'Annaba (Algérie)
Gourvès Laurent, Lamsade, (France)
Habib Michel, Université de Paris VII (France)
Hadjali Allel, ENSSAT, Lannion (France)
Hamadi Youssef, Microsoft Research Cambridge (Royaume Uni)
Hantouche Abderrahim CMM, Universidad de Chile (Chili)
Kanté Mamadou, Clermont-Ferrand (France)
Kazar Okba Université de Biskra (Algérie)
Kechadi Tahar UCD (Irlande)
Kedad Zoubida, UVSQ (France)
Khadir Tarek, Université Badji Mokhtar d'Annaba (Algérie)
Kheddouci Hamamache Université de Lyon 1 (France)
Lafifi Yacine Université de Guelma (Algérie)
Lacomme Philippe LIMOS, Université Blaise Pascal (France)
Laskri Mohamed Tayeb, Université Badji Mokhtar, Annaba (Algérie)
Lebbah Yahia Université D?Oran Es-Sénia (Algérie)
Leger Alain, France Télécom (France)
Lehsaini Mohamed, Université Abou Bekr Belkaid Tlemcen (Algérie)
Limouzy Vincent, Université Blaise Pascal Clermont-Ferrand (France)
Mahjoub Ridha, Université Paris Dauphine (France)
Merouani Hayett, Université Badji Mokhtar, Annaba (Algérie)
Missaoui Rokia, Universite de Quebec en Outaouais (Canada)
Nourine Rachid, Université d'Oran Es-Sénia (Algérie)
Ouanes Mohand, Université Mouloud Mammeri de Tizi-Ouzou (Algérie)
Oukacha Brahim Université Mouloud Mammeri de Tizi-Ouzou (Algérie)
Ouksel Aris, UIC (USA)
Petit Jean-Marc, INSA de Lyon (France)
Rao Michael, ENS Lyon (France)
Rabhi Fethi, Université du New South Wales, Sydney (Australie)
Radjef Mohand-Said, Université Abderrahmane Mira de Bejaia (Algérie)
Raspaud André, Université Bordeaux 1 (France)
Rebaïne Djamal, Université du Québec, Chicoutimi (Canada)
Sadi Bachir, Université Mouloud Mammeri de Tizi-Ouzou (Algérie)
Sais Fatiha LRI, Université de Paris Sud (France)
Salleb-Aouissi Ansaf Columbia University (USA)
Sam Yacine, Université de Tours (Algérie)
Seridi Hamid, Université de Guelma (Algérie)
Spiteri Pierre, INP- Toulouse (France)
Taher Yehia, Tilburg University (Netherlands)
Taleb-Ahmed Abdelmalik, Université de Valenciennes (France)
Tchemisova Tatiana, University of Aveiro (Portugal)
Yebdri Mustapha, Université Abou Bekr Belkaid de Tlemcen (Algérie)
Ziou Djemel, Unversité de Sherbrooke (Canada)

Relecteurs additionnels

Mohamed Lehsaini
Katia Abbaci
Hassene Aissi
Samir Aknine
Fatiha Amirouche
Sabeur Aridhi
Mahmoud Barhamgi
Hichem Barki
Sofiane Batata
Yacine BELHOUL
Hattoibe Ben Aboubacar
Abdelkrim Benamar
Fatiha Bendali
Soumia Benkrid
Badr Benmammar
Kheir Eddine Bouazza

Omar Boucelma
Mohammed Bouchekif
Saida Boukhedouma
Abdelmadjid Boukra
Yahia Chabane
Thierry Champion
Jean-Pierre Crouzeix
Ibrahima Diarrassouba
Rahma Djiroune
Philippe Fournier-Viger
Thierry Garcia
Asma Hachemi
Assia Hachichi
Salima Hacini
Mehdi Haddad
Fayçal Hamdi
abderrahim Hantoute
Mounir Hemam
Stéphane Jean
Faiza Khan Khattak
Rania Khefifi
Selma Khouri
Ben Jabeur Lamjad
Ludovic Liétard.
Gaëlle Loosli
Philippe Marthon
Sebastien Martin
Arnaud Mary
Baraa Mohamad
Mohamed Amine Mostefai
Sarah Nait Bahloul
Cheikh Niang
Damien Nouvel
Samir Ouchani
Andre Raspaud
Lynda Said L'Hadj
Rabie Saidi
Idrissa Sarr
Vladimir Shikhman
Grégory Smits
Sahri Soror
Nouredine Tamani
Clovis Tauber
Yamina Tlili
Ronan Tournier
Tarik Zouagui

Application de la Génération de Requêtes de Médiation au Dossier Médical Personnel

Assia Soukane1, Elisabeth Metais2, Frederic Ravaut1

1 Laboratoire d'Analyses et Contrôle des Systèmes Complexes de l'ECE de Paris (LACSC)
37 quai de grenelle, 75015 Paris Cedex 15
{soukane, ravaut}@ece.fr
2 Centre d'Etudes et de Recherche en Informatique du CNAM de Paris (CEDRIC)
292 rue Saint-Martin, F-75141 Paris Cedex 03
metais@cnam.fr

De nos jours, améliorer la qualité des soins dans les milieux hospitaliers est devenu une priorité. Le Dossier Médical Personnel (DMP) est aujourd'hui une solution qui permet aux patients et aux professionnels de santé de partager des informations en vue de faciliter la coordination et d'éviter des examens inutiles [1] et [2].

Par ailleurs, les systèmes multi-sources sont de plus en plus développés. Ils sont définis comme l'intégration de plusieurs sources hétérogènes et distribuées. Parmi ces systèmes d'informations, nous distinguons les entrepôts de données, les systèmes d'informations basés sur le web, les systèmes de bases de données fédérées, ou encore les systèmes de médiation. Notre étude s'inscrit principalement dans le contexte des systèmes de médiation. Ils peuvent servir entre autre d'outils d'acquisition de données pour alimenter un système d'informations médicales.

Un système de médiation est un système qui permet d'intéropérer sur un ensemble de sources hétérogènes et distribuées. Ses composants essentiels sont : le schéma global appelé schéma de médiation, les mappings du schéma global avec les sources, les fonctions de réécriture de requêtes et les fonctions de composition des résultats. Les mappings du schéma global avec les sources sont des requêtes, appelées requêtes de médiation, dont l'expression varie selon l'approche choisie: 1) approche descendante (Global As View ou GAV) où chaque objet du schéma global est défini par une requête sur les sources, 2) approche ascendante (Local As View ou LAV) où chaque objet d'une source de données est défini par une requête sur le schéma global.

Cependant leur mise en œuvre pose un certain nombre de problèmes, en particulier la définition de requêtes de médiation en présence d'un grand nombre de sources de données distribuées et hétérogènes, comme cela est typiquement le cas pour la constitution d'un DMP.

En effet, comme les données sont hétérogènes, il est important de garantir leur conformité mutuelle et leur conformité avec le schéma global. Deux types de conflits liés à l'hétérogénéité des sources sont distingués: 1) les conflits sémantiques liés au schéma dus à l'utilisation d'une terminologie différente pour décrire un même concept (ex. Nom et Nom_Patient ; et 2) les conflits sémantiques liés aux données dus à la provenance de données de diverses origines, leur saisie à des moments distincts par des personnes différentes qui n'ont pas la même perception du réel, et qui utilisent

des conventions différentes. Par exemple, la différence de format de la date (ex. JJ/MM/AA et MM/JJ/AA) entre la valeur de l'attribut date_entrée appartenant à une relation et la valeur de l'attribut date-entrée d'une autre relation. La non prise en compte de ces conflits peut fausser la sémantique des requêtes de médiation et retourner à l'utilisateur des résultats non conformes à ceux définis dans son schéma de médiation. Les opérations qui composent une requête de médiation ne sont valides que si les conflits liés à l'hétérogénéité des données sont détectés et résolus. Par exemple lorsque les valeurs dans une source de données n'ont pas le même format que celui défini dans le schéma de médiation, ou encore une jointure de deux relations source sur l'attribut date peut retourner un résultat incorrect quand les dates sont représentées dans des formats différents.

Nous proposons un système de génération automatique de requêtes de médiation. A partir de la description relationnelle d'un ensemble de sources de données médicales distribuées et hétérogènes, notre algorithme produit un ensemble de requêtes de médiation possibles qui calculent une instance du DMP. Notre système compte essentiellement six modules : (1) interface graphique ; (2) recherche des relations de mapping ; (3) recherche des relations de transitions ; (4) recherche des opérations de jointures ; (5) recherche des chemins de calcul ; (6) génération de requêtes de médiation. Les méta-données nécessaires à la génération sont stockées dans une méta-base [3] et [4].

Notre objectif principal est de fournir aux patients et aux professionnels de santé un outil adapté aux petits et aux grands systèmes. Notre outil dispose aussi de deux modules supplémentaires : Générateur de jeux d'essais et Evaluateur. *Générateur de jeux* d'essais a pour objectif de générer automatiquement des jeux d'essais ; *Evaluateur* a pour but de tester la conformité et la cohérence des requêtes produites par rapport aux attentes des utilisateurs, et d'évaluer les performances de notre outil pour mesurer son passage à l'échelle. Compte tenu d'une évaluation théorique et pratique, les résultats des tests montrent qu'il est possible de générer des requêtes de médiation dans des temps raisonnables (en moyenne 30 secondes) pour des graphes connexes acycliques de grandes tailles, et pour des graphes connexes cycliques de tailles moyennes.

1. Alliance. (2008, April 28). The National Alliance for Health Information Technology, on Defining Key Health Information Technology Term. Report to the Office of the National Coordinator for Health Information Technology. Retrieved February 17, 2009, from http://www.hhs.gov/healthit/documents/m20080603/10_2_hit_terms.pdf.
2. ASIP Santé. (Février,2011). Publication des cahiers des charges de la DMP, http://esante.gouv.fr/dmp
3. Bouzeghoub M., Kedad Z., Soukane A., « génération de requêtes de médiation intégrant le nettoyage de données », RSTI série ISI-NIS Ingénierie des Systèmes d'Information, volume 7, N°3, p.39-66, 2002.
4. Kostadinov D., Peralta V., Soukane A., Xue X., « Intégration de données hétérogènes basée sur la qualité», Actes du XXIIIème Congrès INFORSID, 2005.

The Use of Biology to Describe Bio-Inspired System: Case of Multi Agent Systems

[1]Seif Eddine Mili, [2]Vincent Rodin, [3]Djamel Meslati

[1,3]Laboratoire de Recherche en Informatique(LRI) Badji Mokhtar University Annaba, Algeria
[2]Laboratoire d'Informatique des Systèmes Complexes(LISyC)European University of Brittany,France
seifmili@gmail.com, rodin@univ-brest.fr, meslati_djamel@yahoo.com

Keywords: Model Driven Architecture, Bio Inspired System, Multi Agent System, Agent Group Role, Architectural Units.

Abstract. Biologically inspired complex systems are increasing. As the abstractions presented by biologically inspired systems, systems architects will be required to include the abstractions in their architecture in order to communicate the design to system implementers. The paper argues that in order to correctly present the architectures of bio inspired system a need of bio inspired views will be required. The paper describes a new formalism based on biology and Model Driven Architecture (MDA) in order to find a new and easy way to design and understand (reverse engineering) a complex bio inspired system. The paper also describes then a set of bio inspired views which are used when describing bio inspired system. Finally we use the set of bio inspired views to describe Multi Agent System (MAS).

1 Introduction

Multi-agent systems, artificial neural networks, artificial immune system, swarm intelligence, genetic algorithms, cellular system, artificial ontogeny, and cognitive intelligent have one thing in common. Each of them presents a biologically inspired computing paradigm[1].

The purpose of the paper is to present a new additional manner for describing bio inspired system based on MDA and biology.

The novel contribution of the paper consists of describing the multi agent system with the biological inspired view points.

2 The Proposed Approach

Our approach is inspired from both: POE model (for more details see [2] and MDA approach, the Architectural Units (AU) are the central points of the method. Our purpose is to define several Architectural Units to describe the ontogenetic, epigenetic and phylogenetic views.

We can formulate the Architectural Unit like a function as show below:

Name_of_AU(IM1,IM2,...,IMn)→ OM1,OM2,...,OMk.

This form of description will be used below to describe the ontogenetic view and the epigenetic view of multi agent systems.

There are three parts involved in all bio-inspired systems: the processes, the structures and the environment where the system is designed to operate [2]. Therefore, characterizing a system comes to characterize each part.

3 Description of MAS

To describe the organization of MAS we need to use the Class Diagram from UML 2.0, and the figure 1 show the structural form of MAS, inspired by the works of Ferber [3] upon AGR.

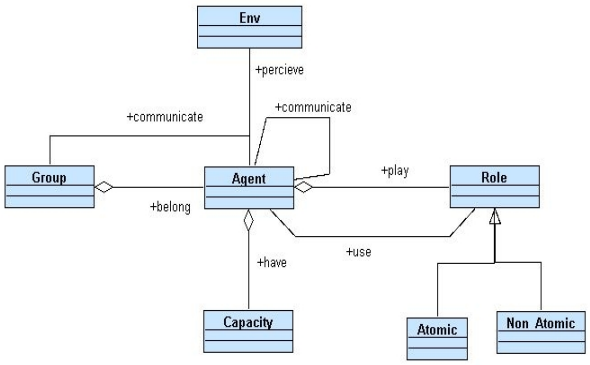


Fig. 1. A view of generic meta-model of MAS

In the organization of the restaurant there is several roles that means several agents like customers, servers, cooks and other one, this agents can form a groups in purpose to achieve a task. These agents interact among other to achieve a service; it is possible to describe what we are waiting from an agent who play roles. The server agent for example has a capacity to bring menu, take order… every one of this service can be executed by any server agents.

Below we give the bio inspired views of restaurant's example.

3.1 The Ontogenetic View

The MAS, first decompose the global goal into a set of sub-goals. For this, we use the transformation Develop with global goal like input model and we obtain sub-goals of output model, and we wrote Develop(D,M)→ M' where: D,M,M' are models

For example in the restaurant the role server can be split in sub-role: take menu, take order, bring dishes, bring the invoice.

The ontogenetic process is iterative and therefore we need the iterative AU, and we wrote Iterate(C, Develop(D,M))→M' where: C show the condition of end of iteration (until all sub-goals were atomic).

The Ontogenetic view describes the initialisation part of MAS and how the structure of the system can change.

3.2 The Epigenetic View

The second part concern the execution of the system and its view like epigenetic process, two transformations are involved to describe this, Interpret and Ajust.

An agent interprets first what he receives from role and according to his capacity he accept or deny.

The epigenetic view is wrote Ajust(A,AU)→ M' where:

A: describe the adjustment model and the update of agent's capacity, the adjustment can be added, deleted of a capacity.

AU: is a composed architectural unit who characterize the transformation of interpretation, and it can be written: AU= Interpret(I,M); I is the interpretation model, it specify how agents interpret the received data, in other word a set of condition that the agent must do for playing role. For example if customer claim invoice, an agent before accepts must interpret this request according to his capacity of making calculate and give back currency.

The epigenetic process is iterative and therefore we need the iterative AU, and we wrote Iterate(C, Ajust(A,Interpret(I,M)))→M' where: C show the condition of end of iteration.

The epigenetic view describes how environmental feedback changes the architecture.

3.3 Multi-Agent Systems Criteria Set

After describing the ontogenetic view and the epigenetic view, we show below some criteria set concerning the Multi-Agent Systems.

Role: the agent plays the role of individual and all system is considered like a population

Description type: the agent can have several description type, he can be genome, neural, antibody

Element/Set: the MAS can be conceded like a set of package and that are agent, role and environment

Granularity: the granularity of agent is variable because he can represent one or several behaviours

Alterability: the agents are alterable

Composition: the MAS are a juxtaposition of two transformations that are ontogenetic and epigenetic.

4 Conclusion

The paper highlights the need for a new formalism to describe bio inspired systems. The paper argues that the need of biological inspired point view is the best opportunity to describe bio inspired systems.The paper shows

the contribution of Model Driven Architecture to the MAS engineering. For example the models provide an advantage for agents to have a flexible (adaptable) behaviour.

References

1 T.L. van Zyl and E.M. Ehlers. A Need for Biologically Inspired Architectural Description: The Agent Ontogenesis Case. 10th Pacific Rim International Conference on Multi-Agents (PRIMA 2007), Bangkok (Thailan), 21-23 November, 2007. Lecture Notes in Computer Science, Springer, Vol. 5044, Agent Computing and Multi-Agent Systems, pages 146-157, April 2009.

2 D. Meslati, L. Souici-Meslati and S.E Mili. Software Evolution and Natural Processes: A Taxonomy of Approaches. 3rd International Symposium on Innovation and Information and Communication Technology (ISIICT 2009), pages 48-57, Philadelphia University, Amman (Jordan), 15-17 December 2009.

3 J. Ferber, O. Gutknecht and F. Michel. From Agents to Organizations: an Organizational View of Multi-Agent Systems. In Proceedings of the 4th International Workshop on Agent-Oriented Software Engineering (AOSE 2003), Melbourne (Australia), 15 July 2003. Lecture Notes in Computer Science, Springer, Vol. 2935, pages 214-230, December 2003

L'INTEGRATION DE L'INDICE NDVI AVEC LES INDICES TASSELED CAP POUR LA DETECTION DE CHANGEMENT DANS LES IMAGES SATELLITAIRES

RIFFI Mohamed Amine, FIZAZI Izabatene Hadria

Laboratoire SIMPA, Département d'Informatique, Faculté des sciences, USTO

r.amine78@gmail.com, hadriafizazi@yahoo.fr

De nombreuses problématiques abordées aujourd'hui en télédétection sont liées à la détection de changement, c'est-à-dire à la caractérisation et à la localisation des zones qui ont évolué entre deux instants ou deux périodes données et cela à partir de deux observations ou séquences d'observation, sur une même scène.

La classification des images satellitaires est une étape centrale dans le processus de la télédétection. Plus précisément, la classification supervisée qui consiste à chercher une partition de l'ensemble des pixels, qui constituent l'image à classer, en K classes, de nombreuses méthodes ont été apparues pour résoudre ce problème, qui peut à la fois faire appel à des principes heuristiques et mathématiques.

Parmi ces méthodes on trouve Les machines à vecteurs support (SVMs), qui on été introduites en 1995 par Cortes et Vapnik. Elles sont utilisées dans de nombreux problèmes d'apprentissage : reconnaissance de forme, catégorisation de texte ou encore diagnostic médical. Les SVM reposent sur deux notions : celle de marge maximale et celle de fonction noyau. Elles permettent de résoudre des problèmes de discrimination non linéaire. Nous allons voir comment un problème de classification se ramène à résoudre un problème d'optimisation quadratique.

Les changements induits sont donc de différents types, d'origines et de durées variées. De nombreuses méthodes de détection de changement à partir des séries temporelles d'images satellite ont été proposées dans la littérature. Ces méthodes peuvent être regroupées en deux grandes familles **:** les méthodes bi-temporelles et les méthodes d'analyse des profils temporels ou encore dites méthodes multi-temporelles.

On s'est intéressé plus précisément à la méthode bi-temporelle. Cette dernière est appliquée pour la détection des changements d'une zone géographique à partir d'images acquises à deux dates différentes. Ces méthodes reposent sur des algorithmes variés de traitements d'images. On a opté pour l'utilisation de deux méthodes distinctes et qui sont: la transformation Tasseled Cap et l'intégration de l'indice NDVI auprès des indices Tasseled Cap.

La transformation du Tasseled Cap est un outil utile pour comprimer des données spectrales en quelques bandes liées aux caractéristiques physiques de la scène. Cette transformation permet de ressortir trois informations majeures notamment l'indice de brillance, l'indice d'humidité et l'indice de verdure. Les indices de la Tasseled Cap ont été calculés à partir des données des six bandes TM du satellite Landsat.

Cette approche a été choisie afin de réduire les erreurs d'omission. Elle consiste d'abord à effectuer un seuillage du changement sur l'image obtenue après la transformation. Ensuite, elle crée un masque de changements pour extraire les zones ayant changé et celles qui n'ont pas changé. Ce masque de changements permet de se concentrer sur la classification des zones qui ont été changées. De ce fait. L'exactitude de la classification est augmentée.

L'intégration de l'indice de végétation NDVI avec les indices Tasseled Cap à conduit a l'amélioration du taux de changement dans les classes végétales par contre il n'a pas influencé sue

le taux de changement des autres classes: l'indice NDVI n'est intervenu rien que pour les classes de type végétation, les autre classes dépendent des autre indices (brillance et humidité).

La fiabilité des résultats est étroitement liée à la méthode adoptée. La transformation Tasseled Cap donne des bons résultats pour les classes qui sont représenté par les deux indices brillance et humidité et des résultats moins bons pour les classes de végétation. L'intégration de l'indice NDVI avec les indices Tasseled Cap a amélioré le taux de changement des classes de végétation sans influencer sur le taux de changement des autres classes de brillance et humidité.

Pour les études futures, nous proposons d'utiliser des images à plus fine résolution spatiale pour une zone extensive, comme celle d'Oran. Par exemple, des images ASTER de TERRA (résolution spatiale de 15 m), qui sont faciles à obtenir pour un pays en développement comme l'Algérie, pourraient être utilisées. Aussi, l'utilisation des cartes topographique des zones étudié afin de validé nos résultats de classification et de changement. Et surtout l'intégration d'autres indices de végétations pour améliorer de plus le taux de changement dans les classes végétales.

Complexity analysis predictor-corrector interior-point algorithm for convex nonlinear programming

El Amir Djeffal[a,1] **and Lakhdar Djeffal**[a]

[a] *Departement of Mathematics, Faculty of Sciences University Hadj Lakhdar, Batna, Algeria*
E-mail Address: djeffal_elamir@yahoo.fr (Corresponding Author. El Amir Djeffal)

Abstract. In this paper, we present a complexity analysis predictor-corrector interior-point algorithm for convex nonlinear programming based on a modified predictor-corrector interior-point algorithm. In this algorithm, there is only one corrector step after each predictor step, where Step 2 is a predictor step and Step 4 is a corrector step in the algorithm. In the algorithm, the predictor step decreases the dual gap as much as possible in a wider neighborhood of the central path and the corrector step draws iteration points back to a narrower neighborhood and make a reduction for the dual gap. Moreover, we derive the currently best known iteration bound for the algorithm, namely $O(\sqrt{n}L)$.

Keyword(s). Quadratic programming, Convex nonlinear programming, Interior point methods
AMS subjet classification**.** 90C30, 90C51

1. Introduction

The convex nonlinear programming (CNP) problem is an important issue in mathematical programming. After the appearance of the Karmarkar's algorithm [1], the researchers introduced extensions for the convex quadratic programming [4]. The most famous of these extensions is that of [2][3][5]. We propose in this work an interior point method modified version, its corrector steps not only draw iteration points back to a narrower neighborhood but also make a reduction for the dual gap. Based on this modified predictor-corrector interior-point algorithm we present a polynomial predictor-corrector interior-point algorithm in this paper. In our algorithm, step 2 is a corrector step and step 4 is a predictor step. At the corrector step we try to decrease the value of μ^k as much as possible in a wider neighborhood of the central path. And at the predictor step, we put iteration points to a narrower neighborhood of the central path $\rho(\frac{1}{4})$ and make a reduction for the dual gap. Our algorithm then generates a sequence (x^k, y^k, s^k) in the neighborhood $\rho(\frac{1}{4})$ starting from an arbitrary initial point $(x^0, y^0, s^0) \in \rho(\frac{1}{4})$ and stopping at an approximate optimal solution (x^k, y^k, s^k) satisfying $\frac{(x^k)^t s^k}{n} < \varepsilon$. It is shown that the method is a polynomial-time algorithm and has iteration complexity of order $O(\sqrt{n}L)$.

In this paper we consider the following convex nonlinear program.

$$\min \{f(x) : \ Ax = b, \ x \geq 0\} \tag{P}$$

where $f : \Re^n \to \Re$ is a convex nonlinear function, $x \in \Re^n$, $A \in \Re^{m \times n}$, $Rank(A) = m < n$. Its dual problem is

$$\max \ \left\{L(x, y, s) : \ -\nabla f(x) + A^t y + s = 0, s \geq 0\right\} \tag{D}$$

where $L(x, y, s)$ is a lagrangian function, $y \in \Re^m$ are the dual variables and $s \in \Re^n$ are the slack variables.

In this paper we denote
$\mathcal{F} = \{(x, y, s) : \ Ax = b, \ -\nabla f(x) + A^t y + s = 0, \ x \geq 0, \ s \geq 0\}$,and define its strict feasible set:
$\mathcal{F}^+ = \{(x, y, s) : \ (x, y, s) \in \mathcal{F}, x > 0, \ s > 0\}$

If $(x, y, s) \in \mathcal{F}^+$, we call x and (y, s) as interior points of (P) and (D), respectively, and denote $X = diag(x)$, $S = diag(s)$, and $e = (1, 1, ..., 1)$.

For every $(x, y, s) \in \mathcal{F}^+$, let $\mu = \frac{x^t s}{n}$ which can form a path of centers. Then we can define a neighborhood of the path of centers

$\rho(\beta) = \{(x, y, s) \in \mathcal{F}^+ : \|Xs - \mu e\|_2 \leq \beta\mu,\ \beta > 0\}$.

Algorithm

Step 0. $(x^0, y^0, s^0) \in \rho(1/4)$ is the initial solution, $\varepsilon > 0,\ k = 0$.

Step 1. If$(\frac{(x^k)^t s^k}{n} < \varepsilon)$ then

Stop. (x^k, y^k, s^k) is $\varepsilon-$approximative solution of (P) and (D).

Step 2. Compute the solution $(\triangle x^p, \triangle y^p, \triangle s^p)$ of the system (1) of equations to obtain predictor direction

$$\begin{cases} A\triangle x^p = 0 \\ -\nabla^2 f(x)\triangle x^p + A^t\triangle y^p + \triangle s^p = 0 \\ S^k\triangle x^p + X^k\triangle s^p = \frac{2}{3}\mu^k e - X^k s^k \end{cases} \tag{1}$$

Step 3. Choose $\alpha_k = \min\left\{\frac{1}{8\sqrt{n}}, \sqrt{\frac{\mu^k}{8\|\triangle x^p \triangle s^p\|}}\right\}$, and compute:

$(\widehat{x}^k, \widehat{y}^k, \widehat{s}^k) = (x^k, y^k, s^k) + \alpha_k(\triangle x^p, \triangle y^p, \triangle s^p)$

compute: $\widehat{\mu}^k = (1 - \alpha_k)\mu^k$

Step 4. If ($\left\|\widehat{X}^k\widehat{s}^k - \widehat{\mu}^k e\right\|_2 \leq \frac{1}{2}\widehat{\mu}^k$) then

Choose $r = 1 - \frac{1}{4\sqrt{2n}}$

Compute the solution $(\triangle x^c, \triangle y^c, \triangle s^c)$ of the system (2) of equations to obtain corrector direction

$$\begin{cases} A\triangle x^c = 0 \\ -\nabla^2 f(x)\triangle x^c + A^t\triangle y^c + \triangle s^c = 0 \\ \widehat{S}^k\triangle x^c + \widehat{X}^k\triangle s^c = r\widehat{\mu}^k e - \widehat{X}^k\widehat{s}^k \end{cases} \tag{2}$$

Step 5. Compute the new iteration

$(x^{k+1}, y^{k+1}, s^{k+1}) = (\widehat{x}^k, \widehat{y}^k, \widehat{s}^k) + (\triangle x^c, \triangle y^c, \triangle s^c)$

let $k = k + 1,$ and go back to **Step 1.**

References

[1] Karmarkar N. A new polynomial time algorithm for linear programming. Combinatorica, 1984,4: 373-393

[2] Ye Y, Tse E. A polynomial algorithm for convex programming. Working Paper. Department of Engineering Economic Systems, Stanford University, Stanford, CA, 1986

[3] Mizuno S, Todd M J, Ye Y. On adaptive-step primal-dual interior point algorithms for linear Programming. Math Oper Res, 1993, 18: 964-981

[4] Zhang Juliang, Zhang Xiangsun. A predictor-corrector interior-point algorithm for covex quadratic programming. J Sys Sci Math Sci, 2003, 23(3): 353-366

[5] Gao Binsong. A modified predictor-corrector interior-point algorithm for linear programming. Mathematics in Economics, 1998, 15: 61-64

[6] Ximing L, Jixin Q. A predictor-corrector interior-point algorithm for convex quadratic programming. Numerical mathematics. A Journal of Chinese Universities(Eng1ish Series), 2002, ll: 52-62

Applications des Forêts Aléatoires pour la Sélection de Descripteurs de Données Médicales

Nesma SETTOUTI, Meryem SAIDI, Mohamed Amine CHIKH

University Abou Bekr Belkaid Tlemcen, Algeria
Laboratory of GBM "Génie Biomédical"
nesma.settouti@gmail.com, miryem.saidi@gmail.com,
mea_chikh@mail.univ-tlemcen.dz

L'algorithme Random Forests introduit par Breiman 2001 est l'un des derniers aboutissements de la recherche consacrée à l'agrégation d'arbres randomisés. Il génère un jeu d'arbres doublement perturbés au moyen d'une randomisation opérée à la fois au niveau de l'échantillon d'apprentissage et des partitions internes. Chaque arbre du jeu est ainsi généré au départ d'un sous-échantillon bootstrap du jeu d'apprentissage complet, de manière similaire aux techniques de bagging. Ensuite, l'arbre est construit en utilisant la méthodologie CART, à la différence près qu'à chaque nœud la sélection de la meilleure partition, basée sur l'indice de Gini, s'effectue non pas sur le set complet de M attributs mais sur un sous-ensemble sélectionné aléatoirement au sein de celui-ci. L'arbre est ainsi développé jusqu'à sa taille maximale, sans élagage. Lors de la phase de prédiction, l'individu à classer est propagé dans chaque arbre de la forêt et étiqueté en fonction des règles CART.

Dans ce travail nous nous intéressons plus particulièrement à la capacité de sélection de variables dans les forêts aléatoires. Notre procédure de sélection de variables se base sur l'indice d'importance calculé par la forêt aléatoire, afin de traiter deux problèmes distincts : trouver toutes les variables reliées à la variable réponse (interprétation) ; et trouver un ensemble de variables suffisants pour prédire la variable réponse (prédiction). Nous avons testé cette approche sur des bases données UCI (machine learning), les résultats obtenus sont très prometteurs.

Mots clés Sélection de variables, forêts aléatoires, indice d'importance, bases données UCI.

Un groupware basé web pour améliorer la coordination lors de la planification des tâches dans un environnement médical

Fouzi Lezzar[1], Abdelmadjid Zidani[2], Atef Chorfi[3]

[1 2 3] Laboratoire de sciences et technologies de l'information et de la communication, Université de Batna, Algérie
[1]Lezzar.fouzi@gmail.com
[2]abdelmadjid.zidani@univ-batna.dz
[3]Atefchorfi@yahoo.fr

Résumé. Les services d'urgences des hôpitaux sont extrêmement complexes à gérer, et posent des risques sérieux pour la santé des patients. Les tâches connexes qui sont principalement centrés autour de la gestion des patients sont essentiellement réalisées grâce à un esprit de coopération qui implique plusieurs professionnels des soins médicaux. Dans ce papier, nous allons d'abord discuter de notre étude, qui a été réalisée dans un service de maternité pour mieux comprendre la façon sous laquelle les tâches sont effectivement réalisées. Ensuite nous allons présenter l'architecture de notre système collaboratif CPlan.

M**ots-clés :** Modélisation des tâches médicaux, travail coopératif, artefacts partagés, interaction synchrones/asynchrones, planification et coordination.

1 Introduction

En Algérie, les établissements de santé à travers le pays souffrent de dysfonctionnements multiples, et ça malgré les réformes engagées par les autorités. L'étude que nous avons menée sur cette question vitale a révélé divers problèmes qui sont principalement liés à la mauvaise gestion et au manque de la coordination entre le personnel médical [2] [3]. Ces problèmes de coordination devraient être considérablement réduits par l'exploration des technologies de l'information [1] [3]. L'objectif principal de ce document est de présenter les concepts conceptuels de base de la planification de notre système coopératif CPlan.

Notre étude nous a conduites à choisir une maternité Algérienne pour mieux comprendre la manière habituelle avec laquelle le personnel médical accomplissent leurs tâches. Nous avons concentré notre attention sur les processus de collaboration entre le personnel médical, parce que nous avons remarqué qu'une grande partie des tâches sont réalisées de manière collaborative, et un manque de coordination peut conduire à une dégradation dans la chaine de soins des malades. Ensuite nous avons identifié les artefacts utilisés qui sont souvent des tableaux de plan classiques ou des feuilles de papier qui sont parfois perdues, et non mises à jour.

2 Architecture logicielle

Le système développé est un groupware synchrone qui est un logiciel (dans notre cas a lancé sur un navigateur) qui permet la collaboration en temps réel entre les membres d'un groupe géographiquement distribués. L'architecture proposée est basée sur la technologie AJAX Push également connue sous le Server Push ou Comet qui permet la création des Applications Internet Riches.

Fig. 1. Architecture logicielle de CPlan

3 Conclusion et perspectives

Pour mesurer l'efficacité de CPlan, nous avons implémenté un premier prototype de l'outil collaboratif, et l'évaluation de la version actuelle sur un réseau local, nous a apportée des idées riches sur les opportunités de la collaboration et de la coordination fournies. Être conscient du grand intérêt de l'expérimentation de CPlan dans des situations plus réelles, nous la prévoyons à l'étape suivante de notre travail de recherche pour recueillir des informations sur les activités du personnel médical.

References

1. Ren, Y., S. Kiesler, S. Fussell, P. Scupelli. Multiple Group Coordination in Complex and Dynamic Task Environments: Interruptions, Coping Mechanisms, and Technology Recommendations. Journal of Management Information Systems / Summer 2008, Vol. 25, No. 1, pp. 105–130.
2. Kuziemsky, C.E., Varpio, L. 2011. A Model of Awareness to Enhance Our Understanding of Interprofessional Collaborative Care Delivery and Health Information System Design to Support it, International Journal of Medical Informatics, doi:10.1016/j.ijmedinf.2011.01.009, forthcoming 2011
3. Bardram JE, Hansen TR. Context-based workplace awareness concepts and technologies for supporting distributed awareness in a hospital environment**.** *Computer Supported Cooperative Work.* 2010;*19:105–138*

Intégration d'ordonnancement et de réplication dans les grilles de données

Salima Behdenna, Hafida Belbachir,

Lab LSSD, Université des Sciences et de la Technologie Mohamed Boudiaf Oran, Algérie
{selma_salima, h_belbach}@yahoo.fr

Résumé

Le concept des grilles de calcul et de stockage a commencé vers les années quatre vingt avec l'idée de partager des ressources de calcul et de stockage à l'échelle planétaire. Les grilles consistent en un réseau d'ordinateurs faiblement couplés et ont pour but d'offrir une très grande puissance de calcul à leurs utilisateurs de la façon la plus transparente que possible. Ces ordinateurs peuvent être des supercalculateurs, des clusters ou des stations de travail ordinaires reliés par un réseau à très grande échelle, le plus souvent Internet.

Une exploitation rationnelle et efficace des grilles de calcul passe nécessairement par une gestion efficace de leurs ressources. Parmi les aspects liés à la gestion des ressources, l'utilisation optimale des ressources de calcul et de communication d'une grille est un aspect très important. Cette optimalité nécessite une répartition de la charge de travail sur les différents nœuds d'une grille. Il faut en effet éviter, dans la mesure du possible, les situations où certains nœuds sont surchargés alors que d'autres sont sous-chargés ou complètement libres.

En général, dans les environnements de grilles de calcul, la gestion de la réplication des données et l'ordonnancement des tâches sont deux tâches indépendantes. Dans quelques cas, les gestionnaires de réplication sont interrogés pour déterminer l'emplacement des meilleurs réplicas en termes de coût d'accès. Mais le choix du meilleur réplica devrait être fait en même temps que l'ordonnancement des requêtes de calcul.

La méthode que nous proposons est basée sur la stratégie **IRS** (Integrated Replication and Scheduling) qui consiste à étudier le couplage entre l'ordonnancement de requêtes et la réplication des données. Dans ce cas, l'ordonnancement est calculé en fonction des informations fournies par le système de réplication et les réplications sont faites en fonction d'informations fournies par l'ordonnanceur, afin d'optimiser l'exécution des requêtes de calcul des utilisateurs. Pour cela, il est indispensable de connaître les caractéristiques, en termes de volume de stockage et de puissance de calcul, des différents composants de la grille ainsi que des informations sur les requêtes d'exécution qui seront soumises à cette grille.

Dans notre méthode, l'ordonnancement est fait à la soumission d'une requête et il est basé sur le coût. Ce coût est calculé pour chaque site en fonction du temps requis pour le transfert des données nécessaires T_d, le temps d'attente dans la file du site T_f et le temps d'exécution de la requête elle-même sur le site T_e .

$$\text{Coût} = T_d + T_f + T_e \quad (1)$$

La requête est alors envoyée au site qui obtient le coût le plus faible. À chaque soumission de requête, la demande des données est sauvegardée dans une table d'historique. Celle-ci servira de base pour la phase de réplication.

Mais tous les calculs de placement des données et des informations d'ordonnancement sont faits avant la soumission réelle des requêtes dans la plate-forme. De ce fait, la stratégie est incapable de s'adapter à un changement des schémas d'utilisation, puisqu'une fois que le placement est fait, il n'est Jamais remis en cause. Ce qui nécessite une opération de rééquilibrage de charge.

L'opération de rééquilibrage de charge entre les nœuds se déroule donc en plusieurs phases. Tout d'abord, il faut déterminer les nœuds qui sont en surcharge de travail par rapport aux autres. Ensuite, il faut trouver quelles sont les causes de cette surcharge et enfin trouver des solutions à apporter pour corriger ce point, et améliorer les performances globales en fonction des conditions actuelles. Pour tous ces points nous utilisons différentes heuristiques inspiré de la stratégie **SRA** (Scheduling and Replication Algorithm).

L'algorithme suivant résume de façon générale les différentes étapes à effectuer pour que notre système puisse s'adapter de façon dynamique à des variations de charge.

Algorithme général d'adaptation dynamique de la charge

```
1:Tan qu'il y a des requêtes à traiter
2: Regarder la charge des serveurs qui détiennent les
    données sollicitées par la requête;
3: S'il y a des serveurs en surcharge Alors
4:  Choisir le serveur P_surcharge le plus surchargé;
5:  Déterminer la donnée D_charge cause de la surcharge;
6:  Choisir un serveur P_sous- charge en sous-charge;
7:  S'il y a assez de place pour répliquer D_charge sur
       P_sous- charge Alors
8:    Répliquer
9:  Sinon
10:   Tant qu'il n'y a pas assez de place sur P_sous-charge
        pour D_charge
11:     Choisir une donnée à supprimer;
12:     S'il n'y a toujours pas assez de place Alors
13:       Restaurer les données supprimées;
14 :      Choisir un autre serveur pour P_sous-charge  et
           recommencer à partir de 7;
15:Si des données ont été répliquées ou supprimées
     Alors
16:Mettre à jour l'ordonnancement;
17:Resoumettre les requêtes concernées par ces
modification ;
```

Conditions de Swensen (1985) adaptées au processus Autorégressif à coefficients aléatoires périodiques d'ordre un Propriété de Normalité Locale Asymptotique

Nabila OUARTIOU

Faculté de Mathématiques, Département de Recherche Opérationnelle, Université USTHB, BP 32, El-Alia 16111 Alger
E-mail: nouartiou@yahoo.fr

Résumé

Dans ce travail, nous présentons les conditions suffisantes de Swensen (1985) que nous avons adaptées au contexte d'un modèle autorégressif périodique d'ordre un à coefficient aléatoire. La propriété de normalité locale asymptotique est établie. C'est un préalable majeur pour l'estimation adaptative et les tests adaptatifs.

Modèle périodique, modèle autorégressif à coefficients aléatoires, propriété LAN.

1 Introduction

De nombreuses séries chronologiques, révèlent un comportement saisonnier. Cette constatation a été le point de départ qui a donné naissance aux modèles $SARIMA$ proposés par Box et Jenkins (1970). Les limites de ces modèles sont qu'ils ne prennent pas en charge une éventuelle périodicité dans les autocorrélations. Grâce aux travaux de Cramer (1961) qui a généralisé la décomposition de Wold (1938) aux processus non stationnaires, sont nés les modèles linéaires à coefficient évolutifs dans le temps. $ARMA_t(p_t, q_t)$ le lecteur intéressé trouvera plus de détail dans Bentarzi (1995). Le souci du statisticien de trouver un cadre mathématique qui reflète de plus en plus fidèlement l'évolution du processus qui a donné naissance aux données d'observations a conduit à des modèles périodiques à structures de plus en plus complexes.

Dans un contexte non linéaire, l'hypothèse de normalité est irréaliste et la densité des innovations est traitée comme un paramètre de nuisance. L'approche semi paramétrique a été adoptée par Koul et Schik (1996) qui ont prouvé dans ce contexte, la propriété de normalité locale asymptotique pour un modèle RCA(1) stationnaire et ergodique.

Akharif et Hallin (2003) ont établi la normalité locale asymptotique non standard pour le modèle plus général RCA(p), en se basant sur les conditions de Swensen (1985).

Dans la section suivante, nous nous intéressons à l'établissement de la propriété LAN pour l'un de ces modèles : le modèle autorégressif d'ordre un à coefficient aléatoire périodique (PRCA(1)). La propriété LAN est un préalable majeur pour l'estimation adaptative et les tests adaptatifs, le qualificatif, adaptatif, étant utilisé pour dire que l'optimalité des estimateurs des paramètres et/ou des tests d'hypothèses est garantie pour toute classe de densités bien définie.

2 Préliminaire

- $\{X^{(ms)}(t),\ t \in \mathbb{Z}\}$ une suite de réalisation d'un processus.
- $(\mathbf{H}^{(ms)}_{f_{\sigma_i}}(\underline{\theta}^{(ms)}))_{ms}$ suite d'hypothèses sous laquelle le processus $\{X^{(ms)}(t),\ t \in \mathbb{Z}\}$ est engendré par un modèle autorégressif d'ordre un à coefficient aléatoire périodique (PRCA(1)) suivant :

$$X^{(ms)}(t) = (\beta_t + B_t(t))\ X^{(ms)}(t-1) + \varepsilon^{(ms)}(t) \tag{2.1}$$

Il est périodique de période s.

où

- f_{σ_t} est la densité des erreurs $\{\varepsilon^{(ms)}(t), t \in \mathbb{Z}\}$ qui supposée suite de variables aléatoires i.p.d centrées, et de variance finie périodique $\sigma_t^2 = \sigma_{t+s\tau}^2, \forall t, \tau \in \mathbb{Z}$ On suppose que f_{σ_t} est non spécifiée.
- $\beta_t = \beta_{t+s\tau},\ \forall t,\ \tau \in \mathbb{Z}$, est dépendant du temps et périodique de période s.
- $\{B_t(t), \forall t \in \mathbb{Z}\}$ est un processus i.p.d centré et de variance $\sigma_t = \sigma_{t+s}, \forall t \in \mathbb{Z}$.

Posons $t = i + s\tau; i = \overline{1,s}, s \in \mathbb{N}^* - \{1\}, \tau \in \mathbb{Z}$.

- Sous l'hypothèse nulle $\mathbf{H}^{(ms)}_{f_{\sigma_i}}(\underline{\theta})$ ou $\mathbf{H}^{(ms)}_{f_{\sigma_i}}(\underline{\beta}', \underline{0}')'$ $\underline{\beta} = (\beta_1, \ ..., \ \beta_s)', \ \underline{\theta} \in \mathbb{R}^s$ $X^{(ms)}$ est une réalisation finie de l'équation aux différences stochastique linéaire suivante :

$$X^{(ms)}(i+s\tau) = \beta_i\ X^{(ms)}(i-1+s\tau) + \varepsilon^{(ms)}(i+s\tau), \quad i = \overline{1,s},\ \tau \in \mathbb{Z}. \tag{2.2}$$

(Modèle PAR(1), s-périodique)

Posons $t = i + s\tau; i = \overline{1,s}, s \in \mathbb{N}^* - \{1\}, \tau \in \mathbb{Z}$.

- L'hypothèse alternative locale est

$$\mathbf{H}^{(ms)}_{f_{\sigma_i}}\left(\underline{\theta} + \frac{\underline{\tau}^{(ms)}}{\sqrt{ms}}\right) = \mathbf{H}^{(ms)}_{f_{\sigma_i}}\left(\underline{\beta} + \frac{\underline{\alpha}^{(ms)}}{\sqrt{ms}}, \frac{\underline{\delta}^{(ms)}}{\sqrt{ms}}\right) \tag{2.3}$$

où

- le coefficient $\underline{\beta}$ est localement perturbé de façon périodique.
- $\frac{\underline{\delta}^{(ms)}}{\sqrt{ms}}$ désignant la variance de la perturbation aléatoire du modèle, elle est périodique et appartient au voisinage de zéro.
- Pour des raisons techniques, nous supposons que la suite de vecteurs $\underline{\tau}_i^{(ms)} = (\alpha_i^{(ms)}, \ \delta_i^{(ms)})'; \quad i = 1, \ ..., \ s$ est telle que

$$\sup_m \ \| \underline{\tau}_i^{(ms)} \|^2 < \infty; \quad i = 1, \ ..., \ s \tag{2.4}$$

- Sous alternative locale $\mathbf{H}^{(ms)}_{f_{\sigma_i}}\left(\underline{\beta} + \frac{\underline{\alpha}^{(ms)}}{\sqrt{ms}}, \frac{\underline{\delta}^{(ms)}}{\sqrt{ms}}\right)$ $X^{(ms)}$ est une réalisation finie de l'équation aux différences stochastique non linéaire suivante :

$$\begin{aligned} X^{(ms)}(i+s\tau) &= (\beta_i + (ms)^{-1/2}\alpha_i^{(ms)})X^{(ms)}(i-1+s\tau) \\ &+ (ms)^{-1/4}\delta_i^{(ms)1/2}V(i+s\tau)X(i-1+s\tau) + \varepsilon^{(ms)}(i+s\tau) \end{aligned} \tag{2.5}$$

- $\{V(i+s\tau);\ i = \overline{1,s};\ \tau \in \mathbb{Z}\}$ est i.i.d centré réduit, de densité g, et indépendant du processus des innovations.

$$\begin{aligned} &E_V\left[f_{\sigma_i}\left(X^{(ms)}(i+s\tau) - \beta_i X^{(ms)}(i-1+s\tau) - (ms)^{-1/2}\alpha_i^{(ms)}X^{(ms)}(i-1+s\tau)\right.\right. \\ &\qquad \left.\left. - (ms)^{-1/4}\delta_i^{(ms)\frac{1}{2}}V(i+s\tau)X^{(ms)}(i-1+s\tau)\right)\right] \\ &= \int_{\mathbb{R}} f_{\sigma_i}\left(X^{(ms)}(i+s\tau) - \beta_i X^{(ms)}(i-1+s\tau) - (ms)^{-1/2}\alpha_i^{(ms)}X^{(ms)}(i-1+s\tau)\right. \\ &\qquad \left. - (ms)^{-1/4}\delta_i^{(ms)\frac{1}{2}}X^{(ms)}(i-1+s\tau)\ v\right)g(v)dv \end{aligned} \tag{2.6}$$

Conditions de régularité :

Afin de prouver la propriété de normalité locale asymptotique (LAN), on impose à la densité f_{σ_i} avec $t = i + s\tau$ de vérifier les conditions de régularité suivante.

(CR1) Le processus PAR(1) est causal, c'est à dire que $\prod_{j=1}^{s} \beta_j < 1$.

(CR2) Les variables aléatoires $\varepsilon^{(ms)}(t), t = i + s\tau; i = \overline{1,s}; \tau \in \mathbb{Z}$ sont i.p.d, de densité f_{σ_t} qui qui est périodique et n'est pas nécessairement gaussienne centrée et de variance finie σ_t^2.

(CR3) $f_{\sigma_t}(.)$ est trois fois différentiable.

$$\phi_{f_1}(.) = -\frac{\dot{f}_1(.)}{f_1(.)}; \quad \psi_{f_1}(.) = -\frac{\ddot{f}_1}{f_1(.)};$$

(CR4) On suppose que $E\ (V^8(t)) < \infty; \quad E\left(\phi_{f_1}^8\left(\frac{x}{\sigma_t}\right)\right) < \infty;$

$$E\left(\psi_{f_1}^4\left(\frac{x}{\sigma_t}\right)\right) < \infty; \quad E\left\{\frac{\ddot{f}_1\left(\frac{x}{\sigma_t}\right)}{f_1\left(\frac{x}{\sigma_t}\right)}\right\}^2 < \infty.$$

Normalité locale asymptotique:

- La fonction de vraisemblance exacte sous $\mathbf{H}^{(ms)}_{f_{\sigma_i}}(\underline{\beta}', \underline{0}')'$

$$\begin{aligned} &\ell^{(ms)}_{\underline{\theta}, f_{\sigma_i}}(X(0), \ X^{(ms)}) \\ &= h_{\underline{\theta}}(X(0)) \prod_{i=1}^{s}\prod_{\tau=0}^{m-1} f_{\sigma_i}(X(i+s\tau) - \beta_i\ X(i-1+s\tau)) \\ &= h_{\underline{\theta}}(X(0)) \prod_{i=1}^{s}\prod_{\tau=0}^{m-1} f_{\sigma_i}(u(i+s\tau)) \end{aligned} \tag{2.7}$$

$u(t)$ coïncide avec $\varepsilon(t)$ sous l'hypothèse nulle $\mathbf{H}^{(ms)}_{f_{\sigma_i}}(\underline{\beta}', \underline{0}')'$.

- La fonction de vraisemblance exacte sous $\mathbf{H}^{(ms)}_{f_{\sigma_i}}\left(\underline{\theta} + \frac{\underline{\tau}^{(ms)}}{\sqrt{ms}}\right)$

$$\begin{aligned} \ell^{(ms)}_{\underline{\theta} + \frac{\underline{\tau}^{(ms)}}{\sqrt{ms}};\ f_{\sigma_i}}(X(0), \ X^{(ms)}) &= h_{\underline{\theta} + \frac{\underline{\tau}^{(ms)}}{\sqrt{ms}}}(X(0)) \ \times \\ \prod_{i=1}^{s}\prod_{\tau=0}^{m-1}\int_{\mathbb{R}} f_{\sigma_i}\left[u(i+s\tau) - (ms)^{-\frac{1}{2}}\alpha_i^{(ms)}\ X(i-1+s\tau)\right. \\ \left. - (ms)^{-\frac{1}{4}}\delta_i^{(ms)\frac{1}{2}}\ v\, X(i-1+s\tau)\right]g(v)dv \end{aligned} \tag{2.8}$$

- Le logarithme du rapport de vraisemblances respectivement sous $\mathbf{H}^{(ms)}_{f_{\sigma_i}}\left(\underline{\theta} + \frac{\underline{\tau}^{(ms)}}{\sqrt{ms}}\right)$ et $\mathbf{H}^{(ms)}_{f_{\sigma_i}}(\underline{\theta})$

$$L^{(ms)}_{f_{\sigma_i};\ \underline{\theta} + \frac{\underline{\tau}^{(ms)}}{\sqrt{ms}}\ |\underline{\theta}}(X(0), \ X^{(ms)}) = \frac{\ell^{(ms)}_{\underline{\theta} + \frac{\underline{\tau}^{(ms)}}{\sqrt{ms}};\ f_{\sigma_i}}(X(0), \ X^{(ms)})}{\ell^{(ms)}_{\underline{\theta},\ f_{\sigma_i}}(X(0), \ X^{(ms)})}, \tag{2.9}$$

$$\Lambda^{(ms)}_{f_{\sigma_i};\ \underline{\theta} + \frac{\underline{\tau}^{(ms)}}{\sqrt{ms}}\ |\underline{\theta}}(X(0), \ X^{(ms)}) = \log\ L^{(ms)}_{f_{\sigma_i};\ \underline{\theta} + \frac{\underline{\tau}^{(ms)}}{\sqrt{ms}}\ |\underline{\theta}}(X(0), \ X^{(ms)}) \tag{2.10}$$

3 Résultats

Lemme (Swensen (1985) adapté au cas d'un PRCA(1))

Si les conditions suivantes sont satisfaites,

(c1) $\sum_{i=1}^{s}\sum_{\tau=0}^{m-1} E\left[X_{ms}(i+s\tau) - Z_{ms}(i+s\tau)\right]^2 \longrightarrow 0$ quand $m \to \infty$

(c2) $\sup_m \sum_{i=1}^{s}\sum_{\tau=0}^{m-1} E\ (Z_{ms}^2(i+s\tau)) < +\infty$

(c3) $\max_{1\le i\le s}\max_{0\le\tau\le m-1} |Z_{ms}(i+s\tau)| \xrightarrow{p} 0$ quand $m \to \infty$ sous $\mathbf{H}^{(\mathbf{ms})}_{f_{\sigma_1}}(\underline{\beta}, \underline{0})$

(c4) $\sum_{i=1}^{s}\sum_{\tau=0}^{m-1} Z_{ms}^2(i+s\tau) - \frac{V^{(ms)2}}{4s} \xrightarrow{p} 0$ quand $m \to \infty$ sous $\mathbf{H}^{(\mathbf{ms})}_{f_{\sigma_1}}(\underline{\beta}, \underline{0})$

(c5) $\sum_{i=1}^{s}\sum_{\tau=0}^{m-1} E\left(Z_{ms}^2(i+s\tau)\mathbf{1}_{[|Z_{ms}(i+s\tau)|>\frac{1}{2}}\ |\ \mathcal{F}_{ms}(i-1+s\tau)\right) \xrightarrow{p} 0$ $\mathbf{H}^{(\mathbf{ms})}_{f_{\sigma_1}}(\underline{\beta}, \underline{0})$

(c6) $E\left(Z_{ms}(i+s\tau)\ |\ \mathcal{F}_{ms}(i-1+s\tau)\right) = 0.$

$$X_{ms}(i+s\tau) = \frac{\left[\int_{\mathbb{R}} f_{\sigma_i}(u(i+s\tau) - (ms)^{-\frac{1}{2}}\alpha_i^{(ms)}X(i-1+s\tau) - (ms)^{-\frac{1}{4}}\delta_i^{(ms)\frac{1}{2}}X(i-1+s\tau)v)\ g(v)dv\right]^{1/2}}{(f_{\sigma_i}(u(i+s\tau)))^{1/2}} - 1$$

$$Z_{ms}(i+s\tau) = \frac{1}{2\sqrt{ms}}\left[\alpha_i^{(ms)}\psi_{f_{\sigma_i}}(u(i+s\tau))\ X(i-1+s\tau) - \frac{\delta_i^{(ms)}}{2}\psi_{f_{\sigma_i}}(u(i+s\tau))\ X^2(i-1+s\tau)\right]$$

- La vérification de ces conditions, nous a permis d'établir la propriété de normalité locale asymptotique suivante :

Proposition

Supposons les conditions **(CR1)-(CR4)** vérifiées. Alors sous les conditions de Swensen (1985) et sous $\mathbf{H}^{(\mathbf{ms})}_{f_{\sigma_1}}(\underline{\beta}', \underline{0}')$ quand $m \to \infty$,

$$\Lambda^{(ms)}_{f_{\sigma_i};\ \underline{\theta} + \frac{\underline{\tau}^{(ms)}}{\sqrt{ms}}\ |\underline{\theta}}(X(0), \ X^{(ms)}) = \underline{\tau}^{(ms)\prime}\ \Delta(\underline{\theta}) - \frac{1}{2}\ \underline{\tau}^{(ms)\prime}\ \Gamma(\underline{\theta})\ \underline{\tau}^{(ms)}$$

$$\Delta(\underline{\theta}) = \begin{pmatrix} \frac{1}{\sqrt{ms}}\frac{1}{\sigma_1}\sum_{\tau=0}^{m-1}\phi_{f_1}\left(\frac{u(1+s\tau)}{\sigma_1}\right)X(s\tau) \\ -\frac{1}{2\sqrt{ms}}\frac{1}{\sigma_1^2}\sum_{\tau=0}^{m-1}\psi_{f_1}\left(\frac{u(1+s\tau)}{\sigma_1}\right)X^2(s\tau) \\ \frac{1}{\sqrt{ms}}\frac{1}{\sigma_2}\sum_{\tau=0}^{m-1}\psi_{f_1}\left(\frac{u(2+s\tau)}{\sigma_2}\right)X(1+s\tau) \\ -\frac{1}{2\sqrt{ms}}\frac{1}{\sigma_2^2}\sum_{\tau=0}^{m-1}\psi_{f_1}\left(\frac{u(2+s\tau)}{\sigma_2}\right)X^2(1+s\tau) \\ \vdots \\ \frac{1}{\sqrt{ms}}\frac{1}{\sigma_s}\sum_{\tau=0}^{m-1}\psi_{f_1}\left(\frac{u(s+s\tau)}{\sigma_s}\right)X(s-1+s\tau) \\ -\frac{1}{2\sqrt{ms}}\frac{1}{\sigma_s^2}\sum_{\tau=0}^{m-1}\psi_{f_1}\left(\frac{u(s+s\tau)}{\sigma_s}\right)X^2(s-1+s\tau) \end{pmatrix}$$

et $\Gamma(\underline{\theta})$ une matrice de dimension $(2s \times 2s)$ donnée par

$$\Gamma(\underline{\theta}) = \begin{pmatrix} B_1 & & & 0 \\ & B_2 & & \\ & & \ddots & \\ & & & \ddots \\ 0 & & & B_s \end{pmatrix}$$

où $\forall\ i = \overline{1,s}$

$$B_i = \begin{pmatrix} \frac{1}{\sigma_i^2}I_\phi(f_1)E(X^2(i-1+s\tau)) & -\frac{1}{2\sigma_i^3}I_{\phi\psi}(f_1)E(X^3(i-1+s\tau)) \\ -\frac{1}{2\sigma_i^3}I_{\phi\psi}(f_1)E(X^3(i-1+s\tau)) & \frac{1}{4\sigma_i^4}L_\psi(f_1)E(X^4(i-1+s\tau)) \end{pmatrix}$$

Bibliographie

[1] Akharif, A and Hallin, M. (2003), "Efficient detection of random coefficients in autoregressive models". *Ann. Statist.* **31** *675-704.*

[2] Bentarzi, M. (1995), "Modèles de séries chronologiques à coefficients périodiques ($ARMA_t(p_t, q_t)$)". *Thèse de Doctorat d'État.*

[3] Koul, H.L. et Schick, A. (1996), "Adaptative estimation in a random coefficient model", *The Annals of statistics.* **24** *1025-1052.*

[4] Nicholls, D. F et Quinn, B. G. (1982), "Random Coefficient Autoregressive Models : An Introduction". *Springer-Verlag.*

[5] Swensen, A. R. (1985), "The Asymptotic Distribution of the likelihood ratio for autoregressive time series with a regression trend". *Journal of Multivariate. Analysis* **16**, *54-70.*

A two-stage flow shop with dedicated machines and batch transfer

Nacira CHIKHI and Moncef ABBAS

Laid3,Faculty of Mathematics, USTHB University,
BP 32 Bab-Ezzouar, El-Alia 16111, Algiers , Algeria
nacira.chikhi@gmail.com
mabbas@usthb.dz

Abstract. This paper deals with a two-stage hybrid flow shop scheduling problem. The objective is the minimizing of maximum completion time of all the jobs. Each job is defined by two operations processed on the two stages in series. Depending on the job type, the job is processed on either of the two machines at stage1 and must be processed on the single machine at stage2. The jobs are transported between the stages by a conveyor in batches. After problem formulation, we present lower bounds for the objective function. We then discuss a few polynomially solvable cases of the problem. Since the problem is strongly NP-hard, heuristic algorithms are developed to find approximate solutions to the problem. Computational experiments are done on a number of randomly generated tests, and the test results are reported.

Keywords: Two-stage flow shop, Makespan, Batch transfer, Heuristics

1 Introduction

Our problem may defined as follows. We are given a set N of n independent jobs distributed in two disjoint subsets have to be scheduled on a two-stage flow shop in series, the first stage contain two dedicated machines and the second stage contain one common single machine with unlimited buffer spaces. Additionally, transportation times are considered. We assume that all transport operations have to be done by a single transport robot which can carry up to c jobs in one shipment. The transportation time from one machine to the other machine is denoted by t (the round-trip requires 2t).

Having dedicated machines at stage1 is common in real-world situations. For example, this problem may arise in a manufacturing environment in which two different products are initially processed on two independent machines but each product must go through a final common machine after their transportation, such as an inspection and testing station. In this paper, we study two-stage flow shop production problem with dedicated machines at the first stage and one single common machine at the second stage which is different from those machine environment studied in a simple two stage productions problem. Furthermore, transportation time and conveyor capacity are considered.

2 Main results

We proposed a mixed integer linear programming model for the general problem in order to determine the sequence of jobs that minimizes the total makespan criterion and we give three lower bounds for the objective function.

Recall that the general problem is NP-hard, we developed two heuristics for its solution and we analyzed some polynomially solvable cases. We also proved that some particular cases of the general problem are NP-Hard.

To study the practical value of the proposed methods, the two heuristics are empirically evaluated. For each combination of number of jobs, we randomly generate 100 problem instances for the performance test of the heuristic algorithms. As there are 12 combinations for each heuristic, the total 2400 instances are generated.

For each problem instance, the heuristic makespan, denoted by C(i), and lower bound of makespan, denoted by LB are computed, as seen in figure Fig.1.

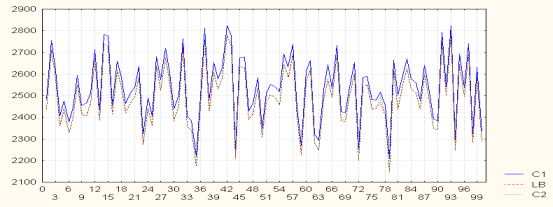


Fig. 1. Comparaison between Cmax(H) and the lower bound LB.

3 Conclusion

We studied a two-stage hybrid flow shop scheduling problem with dedicated machines. The stages are connected by a conveyor. We presented and modeled our problem as a linear program in integer and binary variables. We also proposed lower bounds for the objective function. Some polynomially solvable cases of the problem are analyzed. We developed two heuristics to find approximate solutions to the general problem with computational experiments.

References

1. Besbes, W., Loukil, T. , Teghem, J.: A Two-stage flow shop problem with parallel dedicated machines. MOSIM10, Tunisia,(2010).
2. Ceyda, O., Bertrand, M. T .:Two-stage Flowshop Scheduling with a Common Second-stage Machine, (1998).
3. Chen, ZL., Lee, CY.: Machine scheduling with transportation considerations. Journal of scheduling, 4 :3-24, (2001).

eXtreme Designing : Vers une méthode efficace de conception rapide de système d'information à base de briques logicielles

T. Milon, M. Kazmierski, M. Deshayes, M. Teisseire

UMR TETIS IRSTEA 500 rue Jean-François Breton 34093 Montpellier Cedex 5

1 Introduction

Les systèmes d'informations sont identifiés par leur rôle de manipulation de l'information, passant selon les cas par la récupération, la classification, le traitement et la diffusion de l'information [3]. La conception de système d'information est régie par les besoins et contraintes définis par les acteurs du projet ciblant des traitements et des données à traiter. Ceci détermine directement la méthode adoptée [1]. De nombreuses méthodes de conception [4,6] ont permis une meilleure maîtrise de la complexité, des coûts et des délais dans la réalisation des projets informatiques. Les lacunes des méthodes de conception traditionnelles ont été soulignées par leurs utilisations dans un contexte industriel : monolithique et séquentielle, peu réutilisable, sans aide dans la mise en œuvre des activités de développement.

2 Réponse aux enjeux de conception

Dans le cadre de projets multi-acteurs qui demandent beaucoup de flexibilité et de réactivité, avec des délais courts et des restrictions budgétaires, les méthodes classiques de conceptions de systèmes d'informations posent certains problèmes. Le principal vient du fait qu'elles sont basées sur l'hypothèse que l'exhaustivité des besoins et problèmes sont identifiables en début de projet. Dans notre cas, les données sont récupérées tout au long de l'action sans connaissance a priori de leur structure, encodage, format et contenu, sans compter le manque de métadonnées et la variabilité de casse et d'orthographe. L'absence d'un cahier des charges précis et complet sous-tendait l'évolution des fonctionnalités selon les interventions du commanditaire. L'équipe était restreinte (2 personnes) avec des délais très courts pour en assurer la conception et la réalisation d'un SI basé sur des logiciels gratuits & libres de droits. Les enjeux ainsi décrits nous imposaient une grande flexibilité dans la conception du SI.
Les travaux qui ont inspiré la méthode que nous avons adaptée et adoptée étaient ceux des méthodes agiles [2]. En effet, il était nécessaire de mettre en œuvre une gestion de projet itérative du type eXtreme Programming (XP) mais l'adapter à notre contexte, en le rebaptisant eXtrem Designing (XD). Pour cela, nous avons choisi de développer une architecture du SI orientée composant, à partir de briques logicielles complexes afin de respecter l'ensemble des enjeux précédemment décrits. l'XP propose 4 piliers qui, s'ils sont suivis, faciliteront la gestion de projet informatique : la communication, les réactions, la simplicité et le courage. Dans le contexte de l'action "Portrait", ces piliers sont tout à fait adaptés mais deux autres points essentiels sont nécessaires : la curiosité (veille continue sur les briques logiciels) et les standards (interopérabilité intergiciel et des données).
D'autre part, l'intérêt de travailler au niveau de la brique logiciel est d'avoir un grand nombre de fonctionnalités complexes à disposition pour composer rapidement les services souhaités sans la nécessité de produire de code à l'intérieur de

la brique. Cet intérêt est réel si et seulement si ces briques sont i) facilement manipulables, ii) paramétrables pour répondre aux besoins métiers spécifiques et iii)sont interopérables entres elles.

3 Évaluation de la méthode

Misra [5] propose 5 critères de succès dans le cadre du développement logiciel. Quatre sont recherchés avec la méthode de l'XD : réduire le temps de livraison, optimiser la capacité à répondre aux besoins, optimiser la réactivité face au changement, améliorer la réutilisation. L'XD va dans ce sens : l'apprentissage et l'ouverture sont deux points clés de l'XD, tout autant que la satisfaction du client, la collaboration et l'implication de celui-ci. Ceci a permis de réagir de nombreuses fois aux changements de manière efficace. L'utilisation de briques logicielles complexes permet de minimiser le besoin en compétences techniques poussées et le temps de développement tout en assurant la capacité à répondre aux besoins spécifiques.

4 Discussion

Dans un contexte de contraintes temporelles et financières fortes sur un projet informatique comme celui des "Portraits de la Biodiversité Communale", l'XD répond aux besoins. Il facilite le développement de système d'information modulable au cours du temps de par l'utilisation de briques logicielles complexes interopérables. L'XD est tout a fait transposable dans le cadre d'autres projets similaires. Elle présente cependant plusieurs limites : historiquement, l'XP est loin d'être une méthode nouvelle et elle ne s'est pas imposée de manière massive dans le monde industriel. Les méthodes de développement agile sont adaptées pour les petites équipes, mais pour les grands projets, d'autres processus sont plus appropriés. Concernant la spécificité d'utilisation des briques logiciels, si certaines fonctionnalités ne sont pas existantes dans l'environnement des briques proposées, un développement est nécessaire. Dans ce cas, les principes de l'XD sont moins applicables. De plus, l'intégration d'une nouvelle brique et l'évolution des briques même peuvent perturber parfois l'architecture du SI. La limite de complexité d'une brique logicielle est contenue lorsque les outils les plus « universels » possible sont favorisés. Cette idée est dans la droite lignée de celle d'interopérabilité entre les SI environnementaux européens notamment (INSPIRE) ou du critère de «modularité d'un SI» qui apparaît comme en étant une exigence fondamentale aujourd'hui.

Références

1. Conboy, K. : Agility from first principles : Reconstructing the concept of agility in information systems development. Information Systems Research 20(3), 329–354 (September 2009)
2. Conboy, K., Fitzgerald, B. : Method and developer characteristics for effective agile method tailoring : A study of xp expert opinion. ACM Trans. Softw. Eng. Methodol. 20, 2 :1–2 :30 (July 2010)
3. Courcy, R.D. : Les systèmes d'information en réadaptation. Réseau international CIDIH et facteurs environnementaux 1-2(5), 7–10 (1992)
4. LeMoigne, J. : Les Systèmes d'information dans les organisations. PUF (1973)
5. Misra, S.C., Kumar, V., Kumar, U. : Identifying some important success factors in adopting agile software development practices. The Journal of Systems and Software 82, 1869–ï¿½ï¿½1890 (2009)
6. Reix, R. : Systèmes d'information et Management des organisations. Vuibert (2004)

Approche basée sur les mixtures de gaussiennes et les agents guetteurs pour l'extraction des objets en mouvement

Farou Brahim[1], Seridi Hamid[1]

[1] Département d'informatique, Université 8 mai 1945 Guelma, B. P. 401, Algérie
LabSTIC, Université 8 mai 1945 Guelma, B. P. 401, Guelma 24000, Algérie

farou@ymail.com, seridihamid@yahoo.fr

Résumé. La détection automatique des objets en mouvement est une tâche très importante dans les systèmes de vidéo de surveillance. Les anciens systèmes basés sur les compétences des agents de sécurité ont montré leurs limites. Effectivement, beaucoup de séquences visuelles importantes peuvent leur échapper. Pour cela un système de vidéo surveillance s'impose permettant de seconder les agents dans leurs travaux. Les infrastructures existantes des systèmes de surveillance actuels offrent seulement la possibilité de capturer, de sauvegarder et de distribuer les vidéos afin de les traiter manuellement par des policiers ou par des agents spécialistes dans la détection des anomalies et des comportements suspects des humains. Ce domaine a connu une expansion en termes de méthodes utilisées pour résoudre le problème de la vidéo surveillance dans ses différents niveaux, à savoir la segmentation, la détection, la reconnaissance et le suivi des objets en mouvement. Ce processus commence d'abord par l'extraction des objets qui n'appartiennent pas à l'arrière-plan. La deuxième étape consiste à classifier les objets extraits de la première étape. Cette phase utilise des techniques telles que la détection des contours et les descripteurs spécifiques d'un objet. La troisième étape est le suivi des objets à travers la scène. Les mixtures de gaussiennes sont parmi les méthodes les plus utilisées pour la modélisation de l'environnement. Elles permettent de donner des résultats extraordinaires dans le cas où la vidéo ne contient pas de variation locale où un changement de luminosité. L'apprentissage d'un arrière-plan encombré et contenant plusieurs objets en mouvement reste le problème majeur de cette méthode. Dans ce papier, nous allons nous intéresser à la détection et l'extraction des objets en mouvement dans une vidéo de surveillance à travers une caméra fixe et en utilisant les mixtures de gaussiennes. Pour pallier le problème de la variation locale, nous avons proposé un système qui permet d'abord la segmentation de l'image initiale en plusieurs zones de taille égale ou chaque zone sera affectée à un agent qui permet de détecter les changements au niveau de la zone. Seuls les segments qui ont subi un changement seront mis à jour par les GMM

Mots clés: vidéo surveillance, mixture de gaussienne, modélisation de l'arrière-plan, traitement d'images.

FLERE: Un Système de Relaxation Avancé de Requêtes Flexibles

Amine BRIKCI-NIGASSA[1], Allel HADJALI[2]

[1] Département d'Informatique, Faculté des Sciences de l'Ingénieur
Université Abou Bakr-Belkaid , Tlemcen, Algérie
nh2@LibreTlemcen.org

[2] IRISA/Enssat, Université Rennes 1
22305 Lannion Cedex, France
hadjali@enssat.fr

1 Introduction

Dans le contexte des requêtes flexibles à prédicats graduels (avec des conditions floues), le *problème des réponses vides (PRV)* concerne les cas où aucune des données présentes ne satisfait *un tant soit peu* les conditions de la requête soumise. La *relaxation* de ces requêtes peut être faite en appliquant sur les prédicats flous une transformation basée sur une relation de tolérance, afin d'obtenir des requêtes retournant des réponses qui satisfont les conditions avec un degré non nul.

Ce prototype a été réalisé au cours d'un projet de Magister à l'École Doctorale STIC [1] qui entre dans le cadre de l'étude de l'interrogation flexible des bases de données par des requêtes à prédicats graduels. Il se base sur les travaux de l'équipe PILGRIM de l'IRISA (France). Les fonctions de calcul des MFS proviennent d'un prototype réalisé au sein de cette équipe [2].

2 Approche utilisée

La plupart des approches proposées pour traiter le PRV sont fondées sur un mécanisme de relaxation. Pour une requête flexible conjonctive Q, une transformation élémentaire $T^{\uparrow}$ est appliquée à un ou plusieurs des prédicats flous de Q.

Nous avons retenu l'approche basée sur une *relation de tolérance* (ou *relation de proximité*) *relative* [9]. L'application de $T^{\uparrow}$ sur un prédicat P donnera le prédicat relaxé P', de f.a.t. $(A, B, a + \Delta_l(\varepsilon), b + \Delta_r(\varepsilon))$ où $\Delta_l(\varepsilon) = A \cdot \varepsilon$ *et* $\Delta_r(\varepsilon) = B \cdot \varepsilon/(1 - \varepsilon))$ sachant que ε est la *valeur de tolérance.* Afin de ne pas dépasser une limite sémantique de validité, une *relaxation maximale*, notée $T^{\uparrow max}(P)$, d'un prédicat P a été déterminée. Elle correspond à la valeur de tolérance $\varepsilon_{max} = (3 - \sqrt{5})/2$ [3].

L'ensemble des requêtes modifiées correspondant à $Q=P_1 \wedge \ldots \wedge P_N$ forme une structure de treillis. Si chaque prédicat dans Q peut être relaxé au plus ω fois, alors le nombre d'étapes de relaxation autorisé est $\omega \times N$

Afin de rester cohérente, la relaxation doit agir sur les différents prédicats dans des proportions équivalentes ; c'est la propriété de l'*Égalité de l'Effet de Relaxation* (ÉER), qui considère les rapports entre les longueurs des supports des prédicats. Les valeurs de ε_i pour tous les prédicats P_i doivent vérifier la propriété de l'ÉER sans qu'aucune relaxation maximale ne dépasse la limite sémantique précédemment fixée.

On parcourt le treillis de manière efficace en exploitant les sous-requêtes à réponse vide minimales (*Minimal Failing Subqueries, MFS*) de la requête originale *Q* (les nœuds dont au moins une MFS de leur parent est préservée ne sont pas évalués).

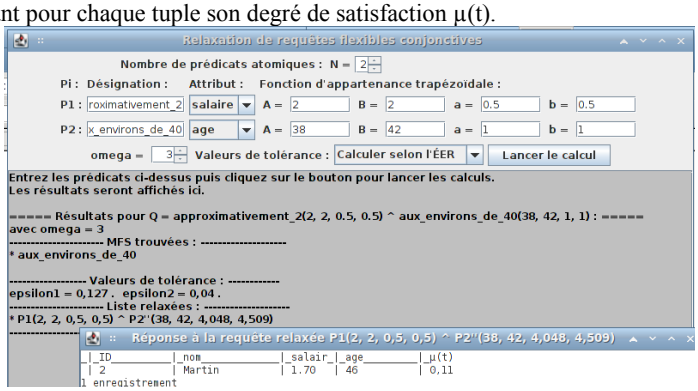
La meilleure relaxation, c'est-à-dire la requête la plus proche de *Q*, sémantiquement parlant, qui produit des réponses non-vides peut être retournée en utilisant la mesure de proximité sémantique entre les requêtes déduite de la distance de Hausdorff entre les prédicats flous et les prédicats relaxés [4].

3 Fonctionnement et implémentation

FLERE est implémenté avec Java SE, Swing et JDBC pour permettre sa connexion à la plupart des SGBD actuels. Pour chaque prédicat, l'utilisateur remplit la désignation du critère flou, sélectionne l'attribut concerné, remplit les valeurs (A, B, a, b) définissant la fonction d'appartenance trapézoïdale.

Le paramétrage de la relaxation se fait en choisissant ω (omega) ; les ε_i pourront au choix être calculées selon l'ÉER, fixées à $\varepsilon_{max} / \omega$ ou à des valeurs personnalisées.

Les résultats s'affichent en 4 sections : liste des MFS utilisés pour la relaxation, valeurs de tolérance, liste des requêtes relaxées à réponse non vide (en indiquant le nombre de relaxations de chaque prédicat et leurs nouvelles f.a.t.), meilleure relaxation (obtenue par le calcul de la distance de Hausdorff). S'ils aboutissent à une requête relaxée satisfaisante, le résultat de cette requête est ensuite affiché, en indiquant pour chaque tuple son degré de satisfaction μ(t).



Références

1. A. Brikci-Nigassa, Étude de la relaxation de requêtes dans un contexte flexible, *mémoire de Magister école doctorale STIC option SIC, Tlemcen,* 2011.
2. B. Joseph, Réponses Coopératives et Interrogation Flexible de Bases de Données, *Rapport de Projet d'Ingénieur Logiciels et Systèmes,* 2008.
3. A. Hadjali, D. Dubois and H. Prade, Qualitative reasoning based on fuzzy relative orders of magnitude. *IEEE Transactions on Fuzzy Systems*, Vol. 11, No 1, pp. 9-23, 2003.
4. A. Brikci-Nigassa, A. Hadjali, Finding the Best Relaxations of Flexible Database Queries, *COSI'2009* Annaba Algeria, 2009.

SVM Regularization of satellite images K-means Clustering

Akkacha Bekaddour[1,1], Abdelhafid Bessaid [2], Fethi Tarek Bendimerad [1],

[1] Faculty of technology, Department of electronic, telecommunication laboratory
University of Abou Bekr Belkaid, Tlemcen,13000, Algeria

[2] Faculty of technology, Department of electronic, biomedical laboratory
University of Abou Bekr Belkaid, Tlemcen,13000, Algeria
{okacha.bekkadour, a_bessaid, ftbendimerad, LNCS}@mail.univ-tlemcen.dz

Abstract. A lot of supervised and unsupervised classification techniques have been used in the remote sensing image processing field to meet human needs. Image regularization is another way of classification which is based on a two layer system, one for a first image segmentation and the other for correcting classification results. In this paper, we used a well known clustering method k-means as first segmentation, and we followed it by SVM regularization. The main contribution of this work is to show how we can adopt a supervised classification technique (SVM) to make a regularization of an unsupervised (K-means) segmentation results by choosing good training area when using SVM classification. Finally, we used two satellite images representing the same area but in a two different time, the area is the OUARGLA oasis located in the south of Algeria, and it was acquired in 1972 in first time, and in 2000 in second.

Keywords: Classification, Segmentation, Clustering, Kmeans, Support vector machine.

1 Summary

Today, an estimated of 15 to 20 satellites will be launched by 2015 with the possibility of managing, measuring and observing Earth [1]. All that interest on remote sensing has helped to develop more and more applications. In remote sensing, most applications aim is to collect earth observation images in order to use them in several professional and scientific fields. More than that, new approaches and visions of remote sensing appeared this last years, like strategic, juridical, and social approaches [2]. This expands the use of remote sensing not only in a lot of fields but in a lot of ways also. All this need of remote sensing leads the researchers to perform more and more efficient tools to meet the new remote sensing challenges. Image classification is one of the important tools used to identify and detect most pertinent information in satellite images. One of the first classification techniques used in remote sensing fields is kmeans method. It gives the user a first figure about the different segments of the image. However, this results need to be refined by another classification methods in order to get more and more accurate vision of the different classes composing the image. SVM method is one of the important techniques used to perform a distinction between image pixels basing on support vectors separating the non similar pixels according to a certain parameters.

The main goal of this paper is to show how to combine two image classification methods in order to get optimized classification results .The two methods are used in a two stage approach, the first one is a first segmentation using a well known clustering methods: k-means, and the second one is a learning machine technique: SVM (Support Vector Machine) which is used to correct the K-means segmentation results. The idea of combining the two methods to optimize classification results is not new .For example; we find in literature the use of HMM (Hidden Markov Model) technique to regularize maximum likelihood classification results using pixels neighborhood homogeneity. Also , the two classification methods are even used in many remote sensing classification projects and article , but the originality of this paper is the special use of the SVM method to correct (regularize) the K-means clustering results using training area selected and defined by the user as a regularization criterion. The **figure 1** shows the k-means segmentation and SVM regularization results of the OUARGLA region which is a well know city with a nice oasis and a very important irrigation projects

[1] Please note that the LNCS Editorial assumes that all authors have used the western naming convention, with given names preceding surnames. This determines the structure of the names in the running heads and the author index.

of the vegetation in the area. This region is acquired by two satellite sensors at two different times, the first one in September 1976 and the second one in December 2000. The changing affecting vegetation and urban area is well visualized in the figure.

2 Results show

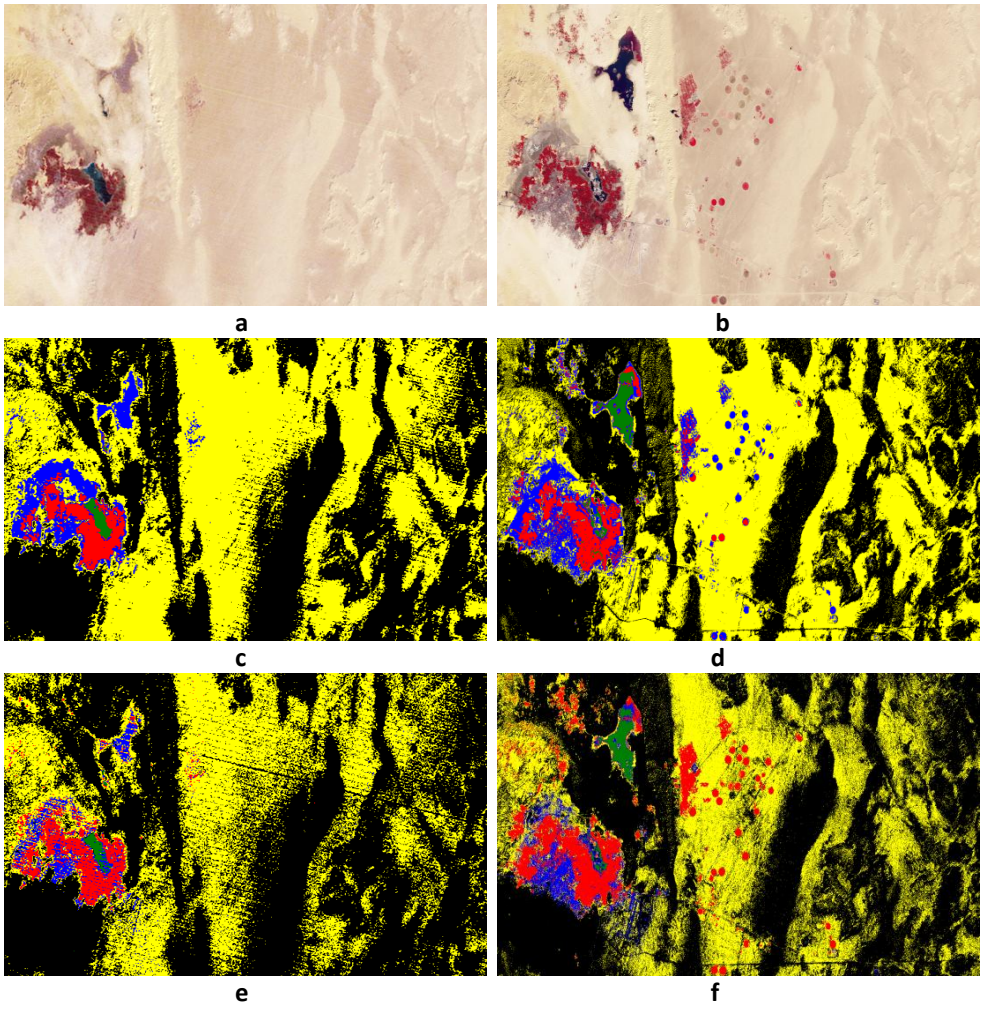


Fig. 1. a. OUARGLA image acquired by the LANDSAT 2 on January, 16 1976. **b.** OUARGLA image acquired by the LANDSAT 7 on December, 20 2000. **c.** Kmeans segmentation result of OUARGLA OASIS (January 1976), k=5. **d.** Kmeans segmentation result of OUARGLA OASIS (December 2000), k=5. **e.** SVM regularization result of OUARGLA OASIS (January 1976), C=500. **f.** SVM regularization result of OUARGLA OASIS (December 2000), C=500.

In this paper, we used a specific image classification approach that is necessary for detecting the different pertinent areas in remote sensing images. In this work, we consider LANDSAT 2 and 7 remote sensing images of the same area. The classification techniques used are a combination of two classification methods which are kmeans and SVM techniques. Even if kmeans segmentation results give us a first understanding of the different component of our images, we need to optimize these results using SVM techniques by introducing training areas in order to separate conflict regions from each other. The results show that we can separate vegetation from urban landscape. Finally, we can say that the choice of training areas is one of the most important points to enhance and improve the initial classification results.

Découverte de sous-graphes fréquents

Nour el islem Karabadji[1],Hassina Seridi[2]

[1,2]Electronic Document Management Laboratory (LabGED), Badji Mokhtar University, BP 12 Annaba, Algeria.
[1]Karabadji@labged.net,[2]Seridi@labged.net

Résumé La fouille des graphes est devenue un domaine de recherche très important. Beaucoup de méthodes de fouille (données transactionnels ou complexes) comme le clustering, la classification et la découverte des sous motifs fréquents... ont été étendues aux données structurées comme les graphes. Dans ce travail, nous nous sommes intéressés à la découverte des sous graphes fréquents, où un nombre important de méthodes ont été développées. Nous donnerons un apercu des différentes méthodes et localiserons les difficultés dans le processus de fouille.

Mots clés: Fouille de données, Fouille de graphes.

1 Introduction

Ces dernières années l'information est devenue de plus en plus volumineuse et relationnelle. L'un des problèmes les plus connus de la fouille de données est la découverte des motifs fréquents et nous citons les itemsets, les arbres et les graphes.... Ces problèmes s'inscrivant dans un problème plus général qui est la découverte des motifs intéressants, formulé par Mannila et Toivonen [1]. Cette derniere discipline consiste en la récupération d'un ensemble de motifs. De ce fait quelque soit la spécialisation de cette théorie, les principales questions aux quelles nous devrons trouver des réponses, définir un ordre entre les motifs, générer les motifs l'un à partir de l'autre, tester la satisfaction du prédicat P.

2 Problèmes et Motivations

La recherche(découverte) des motifs intéressants peut être déterminé principalement par le processus suivant (1)Génération des candidats et (2)Énumération des candidats intéressants par validation du prédicat P.Le processus de découverte est lié au type des motifs à gérer (items, arbres,graphes...). La canonisation étant une principale solution liée a l'étape de génération des candidats. La canonisation des items étant selon l'ordre lexicographique et l'interdiction de la jointure que dans le cas où l'item $X_k < Y_k$. Cela est dû à la notion d'identité constatée de la théorie des ensembles. De même le calcul de la fréquence est très coûteux dû au nombre énorme d'accès en mémoire cela a été contourné par l'élagage basé sur la propriété anti-monotonie. Pour le cas des arbres le problème est plus difficile : la canonisation est plus difficile dans le cas des arbres non ordonnées

étant donnée celles qui sont ordonnées nous pouvons garder l'ordre orignal de l'arbre. L'étape la plus coûteuse est le calcul du support, elle varie selon le type des sous arbres à confrondre, qui peut même être de complexité NP-complet.

Pour les graphes le problème est largement plus difficile que les deux premiers cas. Une solution pour la redondance consiste à explorer un sous graphe seulement si sa représentation est canonique (code-DFS [5] CAM[6], BBP[2], Ordre tringler lexicographique[4],...). En dèpit de tous les efforts le problème reste encore complexe, parce que le fait de trouver la représentation canonique d'une paire de graphes est équivalent à résoudre le problème d'isomorphisme de graphe entre ses deux derniers graphes. Le problème de sous graphes isomorphes est l'un des pires cas du problème d'évaluation du prédicat d'un processus de découverte des motifs intéressants. Pour le calcul de support beaucoup de techniques ont été créés afin de réduire l'espaces de recherche en utilisant principalement l'élagage[4][5][6] ou de partir des structures faciles vers les graphes exemple [3].

3 Conclusion

Nous avons décrit la difficulté du problème de découverte des motifs intéressants qui est liée principalement au type des données dont les motifs sont représentés (item, séquence, arbre, graphe,...). Ce qui nous oriente automatiquement vers l'étude de la difficulté du problème de fouille des motifs fréquents selon le domaine d'application qui a une influence direct sur la représentation des données. Les méthodes de découverte des sous graphes fréquents souffrent de l'énumération combinatoire des motifs et de la redondance.

Références

[1] Mannila et Toivonen. (1997). Levelwise search and borders of theories in knowledge discovery. Data Mining and Knowledge Discovery, 1(3), 241-258.

[2] Tamas Horvathyz, Jan Ramon and Stefan Wrobel. Frequent Subgraph Mining in Outerplanar Graphs, Proceedings of the ACM SIGKDD International Conference, pages 197-206, ACM Press, New York, NY, 2006.

[3] S. Nijssen and J. Kok. A quickstart in frequent structure mining can make a difference. In Proc. 2004 ACMSIGKDD Int. Conf. Knowledge Discovery in Databases (KDD04), pages 647-652, 2004.

[4] A. Inokuchi, T. Washio, H. Motoda. An Apriori-based Algorithm forMining Frequent Substructures from Graph Data.PKDD Conference, pages 13-23, 2000.

[5] X. Yan, J. Han.Gspan : Graph-based Substructure Pattern Mining. ICDM Conference, 2002.

[6] Luke Huan, Wei Wang, Jan Prins. Efficient Mining of Frequent Subgraphs in the Presence of Isomorphism. In : Proceedings of the 2003 International Conference on Data Mining (ICDM2003)

Note on Roman k-domination number in graphs

Ahmed Bouchou* and Mostafa Blidia†

Abstract

For an integer $k \geq 1$ and a graph $G = (V, E)$, a set S of V is k-dominating if every vertex in $V \setminus S$ has at least k neighbors in S. A Roman k-dominating function on G is a function $f : V(G) \longrightarrow \{0, 1, 2\}$ such that every vertex u for which $f(u) = 0$ is adjacent to at least k vertices $v_1, v_2, ..., v_k$ with $f(v_i) = 2$ for $i = 1, 2, ..., k$. The weight of a Roman k-dominating function (Rk-DF) is the value $f(V(G)) = \sum_{u \in V(G)} f(u)$. The k-domination number $\gamma_k(G)$ is the minimum cardinality of a k-dominating set of G and the minimum weight of a Roman k-dominating function on a graph G is called the Roman k-domination number of G, denoted by $\gamma_{kR}(G)$. Since every $(k+1)$-dominating set is k-dominating set and every Roman $(k+1)$-dominating function is roman k-dominating function, the sequences (γ_k) and (γ_{kR}) are weakly increasing.

In this paper we give relations between γ_{kR} and γ_R and some characterization of extremal trees. Also we give the structure of $\{K_{1,3}, K_{1,3}+e\}$-free graphs and we characterize the special $\{K_{1,3}, K_{1,3}+e\}$-free graphs in which the sequence (γ_{kR}) is strictly increasing.

Keywords: k-domination, Roman k-dominating function, Roman k-domination number.

References

[1] M. Blidia, A. Bouchou, *On sequences* (β_k) *and* (γ_k) *in regular graphs*, Studia Informatica Universalis, **9** (2011) 7-18.

*University Yahia Fares of Médéa, Algeria. E-mail: bouchou.ahmed@yahoo.fr.

†Dept. Mathematics, University of Blida, B.P. 270, Blida, Algeria. E-mail: m_blidia@yahoo.fr.

[2] M. Chellali, O. Favaron, A. Hansberg and L. Volkmann, *k-Domination and k-Independence in Graphs: A Survey, Graphs and Combinatorics* (2012) **28**: 1-55.

[3] E. W. Chambers, B. Kinnersley, N. Prince, and D. B. West, *Extremal problems for Roman domination*, SIAM J. Discr. Math. **23** (2009) 1575-1586.

[4] T. W. Haynes, S. T. Hedetniemi, and P. J. Slater, *Fundamentals of Domination in Graphs*, Marcel Dekker, New York, 1998.

[5] K. Kämmerling and L. Volkmann, *Roman k-domination in graphs*, J. Korean Math. Soc. **46** (2009) 1309-1318.

[6] E. J. Cockaynea, P. A. Dreyer Jr., S. M. Hedetniemi, and S. T. Hedetniemi, *Roman domination in graphs,* Discrete Mathematics **278** (2004) 11-22.

Algorithme génétique pour l'ordonnancement sous contraintes de préparation

Wafaa LABBI[1], Mourad BOUDHAR[1] et Ammar OULAMARA[2]

[1] Laboratoire LAID3, Université USTHB,
B.P. 32, Bab-Ezzouar, El-Alia 16111, Alger, Algérie
fawalab@yahoo.fr, mboudhar@yahoo.fr
[2] Laboratoire LORIA, Ecole des Mines de Nancy,
Parc de saurupt, 54042, Nancy Cedex, France
ammar.oulamara@loria.fr

Résumé

Nous considérons le problème d'ordonnancement qui consiste à traiter un ensemble de n tâches indépendantes $T = \{T_1, T_2, ..., T_n\}$ sur deux machines identiques M_1 et M_2 en présence de k ouvriers spécialisés $O = \{O_1, O_2, ..., O_k\}$. Chaque tâche $T_i (i = 1, ..., n)$ est caractérisée par un temps d'exécution p_i et un temps de préparation s_i, durant le temps de préparation, la tâche T_i a besoin de b_i ouvriers spécialisés avec $b_i \leq k$. Nous supposons que toutes les tâches sont disponibles à l'instant $t = 0$. L'interruption des tâches n'est pas autorisée. L'objectif est de trouver un ordonnancement minimisant la date de fin de traitement de l'ensemble des tâches notée C_{max}. Le problème ainsi défini est noté $P2, Sk|s_i, b_i|C_{max}$ et il est NP-difficile car le problème $P2//C_{max}$ en est un cas particulier.

Considérons le graphe $G = (V, E)$, où V est l'ensemble des tâches et où une paire de tâche (T_i, T_j) est dans E si et seulement si $b_i + b_j \leq k$. On peut décomposer cet ensemble V de tâches en deux sous-ensembles $S \cup C$ tels que : S (stable) est l'ensemble des tâches qui vérifient la relation $b_i + b_j > k \ \forall \ T_i, T_j \in S$, C (clique) est l'ensemble des tâches qui vérifient la relation $b_i + b_j \leq k \ \forall \ T_i, T_j \in C$. Nous avons $BI = BI_S + \max\{0, \lceil \frac{2BI1_C - IT}{2} \rceil\}$ est une borne inférieure pour le C_{max}, où BI_S=$\max\{BI1_S, BI2_S\}$, $BI1_S$=$\sum_{T_i \in S} s_i + \min_{T_i \in S}\{p_i\}$ et $BI2_S$=$\lceil (\sum_{T_i \in S}(s_i + p_i) + \min_{T_i \in S}\{s_i\})/2 \rceil$ représentent des bornes inférieures dans le cas où l'ensemble des tâches forment un stable. $BI1_C$ = $\lceil (\sum_{T_i \in C}(s_i + p_i))/2 \rceil$ est une borne inférieure dans le cas où les tâches forment une clique et $IT = \max\{0, 2BI_S - \sum_{T_i \in S}(s_i + p_i)\}$ le temps mort laissé par les tâches du stable exécutées sur les deux machines.

Pour la résolution du problème, nous proposons un algorithme génétique dont les différentes phases sont : (i) Le codage choisit est de longueur égale au nombre total de tâches n. (ii) La population initiale est composée des séquences de tâches utilisées dans les heuristiques définies dans [5] et de séquences générées aléatoirement dont la taille $T_{int} = 50$. (iii) La procédure de sélection que nous avons utilisé est la sélection aléatoire, donc chaque individu a une probabilité uniforme $\frac{1}{T_{int}}$ d'être sélectionné. (iv) Les individus survivants à la phase de sélection

vont subir le croisement avec une probabilité $P_{crois} = 0.8$, nous utilisons le croisement en 2-points. (v) Les individus issus du croisement vont ensuite subir un processus de mutation avec une probabilité $P_{mut} = 0.1$, l'opérateur de mutation utilisé est l'opérateur d'échange réciproque. Si β, généré aléatoirement, appartient à $[0, P_{mut}]$ nous mutons l'enfant i pour avoir un enfant $imut$, et si $\beta > P_{mut}$ nous mutons l'enfant j pour avoir un enfant $jmut$. (vi) Pour chaque chromosome, nous appliquons les méthodes sérielle et parallèle définies dans [5], ensuite nous choisissons la meilleure des deux solutions trouvées. (vii) L'algorithme génétique s'arrête après un certain nombre fixé de génération noté $Itermax = 100$.

Les résultats de l'algorithme génétique ont été comparés avec la borne inférieure BI. Pour évaluer notre méthode, nous avons généré aléatoirement suivant la loi uniforme plusieurs instances. Pour chaque instance générée, nous calculons le pourcentage de la déviation moyenne et celui de la déviation maximale. Nous avons considéré plusieurs cas possibles : temps de traitement et de préparation générés dans un même intervalle, temps de traitement inférieurs aux temps de préparation et temps de traitement supérieurs aux temps de préparation.

D'aprés les tests réalisés, nous avons remarqué que le temps d'exécution est raisonnable (ne dépassant pas 20 secondes pour 1000 tâches). Dans le cas o*ù* les temps de traitement sont, soit inférieurs aux temps de préparation, soit générés aléatoirement dans un même intervalle, nous avons constaté que les solutions obtenues par l'algorithme génétique sont très proches de la borne inférieure si le nombre de tâches est inférieur à 100. Par contre, si le nombre de tâches dépasse 100, les déviations moyenne et maximale sont inférieures à 1.5% si les temps de traitement et de préparation sont générés dans un même intervalle, et inférieures à 0.3% si les temps de préparation sont inférieurs aux temps de traitement. Dans le cas o*ù* les temps de traitement sont supérieurs aux temps de préparation, la déviation moyenne ne dépasse pas les 5% et la déviation maximale est inférieure à 5% si on a plus de 100 tâches.

Références

1. Abdelkhodaee, AH., Wirth, A. : Scheduling parallel machines with a single server : some solvable cases and heuristics, Computers and Operations Research 29(3) : 295-315, 2002.
2. Aronson, JE. : Two heuristics for the deterministic, single operator, multiple machine, multiple run cyclic scheduling problem. Journal of Operations Management, 4(2) : 159-73, 1984.
3. Golumbic, M.C. : Algorithmic Graph theory and Perfect graph. Academic Press, 1980.
4. Koulamas, CP., Smith, ML. : Look-ahead Scheduling for minimising machine interference. International Journal of Production Research 26 : 1523-33, 1988.
5. Labbi, W., Boudhar, M., Oulamar, A. : Scheduling with preparation constraints. The Second International Symposium on Operational Research (ISOR'11), p. 159-160, Mai 30-02 June, 2011, Algeries, Algeria.

Automatic rules extraction using Artificial Bee Colony algorithm for breast cancer disease

Fayssal Beloufa, M. A. CHIKH
Biomedical Engineering Laboratory Tlemcen University Algeria
E-mail: beloufa.fayssal@gmail.com, mea_chikh@mail.univ-tlemcen.dz

Résumé

Un classifieur flou peut être un outil pratique dans le processus de diagnostic. Il se compose de règles linguistiques qui sont faciles à interpréter par l'expert humain. Ce classifieur n'est pas une boîte noire au sens où il peut contrôler la vraisemblance. Ceci est très important pour les systèmes de prise de décision, car les experts n'acceptent pas une évaluation sur ordinateur, à moins qu'ils comprennent pourquoi et comment une recommandation a été donnée. Une règle floue (ou relation floue) peut être interprétée comme un prototype vague des données crées par le classifieur flou. Un résultat de classification est donné par les degrés de vérité (activations) de plusieurs règles. Les règles floues pour des cas similaires peuvent se chevaucher, delà un exemple peut appartenir à plusieurs classes à différents degrés d'appartenance. Cette information peut être exploitée pour évaluer la qualité du résultat de la classification. Les avantages des classifieurs flous peuvent être principalement vus dans leur interprétabilité linguistique, la manipulation intuitive et la simplicité, qui sont des facteurs importants pour l'acceptation et l'utilisation d'une solution. Ce travail aborde le problème de la création d'un classifieur flou par apprentissage à partir des données, et notre objectif principal est d'obtenir une lisibilité du classifieur pour le diagnostic du cancer du sein. L'application de la colonie d'abeille a montré son intérêt dans la répartition des données et permet ainsi la régularisation des contraintes qui s'appliquent sur les paramètres des fonctions d'appartenance floues. Cet algorithme élimine la similarité par la fusion des partitions floues afin d'améliorer leurs distinctions et de minimiser le nombre de règles pour une connaissance ciblée. Delà donc automatiser et optimiser la structure et les paramètres des fonctions d'appartenance. Les résultats obtenus dans ce travail de recherche nous permettent d'apporter une valeur ajoutée dans les systèmes d'Aide au Diagnostic Médical avec comme objectif principal la transparence du modèle. Chaque cas classifié est justifié par une ou plusieurs règles floues distinctes. Cette qualité absente chez la majorité des classifieurs de données, peut convaincre les médecins à utiliser largement les outils d'aide au diagnostic Médical Intelligents.

Analyse des sessions de navigations du site web de l'Université Ferhat Abbes de Sétif par les algorithmes de réseaux sociaux

Yacine SLIMANI et Abdelouahab MOUSSAOUI

Laboratoire de Recherche en Informatique Appliquée,
Université Ferhat Abbas Sétif, Algérie.

Résumé Dans les dernières années il y a eu une croissance exponentielle du nombre de sites web et de leurs usagers. Cette croissance phénoménale a produit une quantité de données énorme liées aux interactions d'utilisateurs avec les sites web, stockés par les serveurs web dans des fichiers d'accès (Logs). Ces fichiers peuvent être utilisés par les administrateurs de sites web pour découvrir les intérêts de leurs visiteurs afin d'améliorer le service par l'adaptation du contenu et de la structure des sites à leurs préférences [1]. Dans ce papier, on propose une méthode d'analyse des fichiers logs en modélisant les sessions de navigation des internautes sous formes de graphes et ensuite on analyse ces graphes pour détecter les communautés partageant des caractèristiques communes. Cette approche est constituée de trois phases.

Dans la première phase du prétraiement on transforme le fichier log de sa forme textuelle brute, en une forme structurée. Ensuite un nettoyage de données est prévu pour supprimer les enregistrements inutiles afin de maintenir seulement les actions liées aux comportements de navigation des utilisateurs [2],[3]. Enfin, des sessions de navigation sont identifiés, sachant qu'une sessions est composée de l'ensemble de pages visitées par le même utilisateur durant la période d'analyse [4].

Dans la deuxième phase de découverte des communautés, on transforme les sessions de navigation en un graphe. Les nœuds du graphe sont les pages du site et les arêtes sont les navigations des internautes [5]. Ensuite, deux approches pour la détection des groupes de nœuds en fonction de la modularité des graphes sont utilisées . La première approche permet la découverte des communautés existantes dans un graphe non pondéré tandis que la deuxième utilise les graphes pondérés pour mieux discriminer ces communautés [6].

L'algorithme dans sa version non pondérée a identifié trois (3) communautés : les visites effectuées aux pages Web de formation et de pédagogie, pages des manifestations scientifiques et pages de valorisation et recherche. , les facultés, les liens utiles, la recherche bibliographique. Alors que l'algorithme dans sa version pondérée a identifié six(6) communautés dont leurs contenues est une subdivisions des premières communautés. Ce qui signifie que l'utilisation de l'information du poids sur le graphe a été très utile pour mieux extraire les intérêts des internautes.

Enfin cela a permis au web master du site d'avoir une vision plus globale des ces internautes pour mieux adapter notre site à leurs besoins.

Keywords: Fouille de données, fouille de données d'usage du web, Session de navigation, Graphe, Réseaux sociaux, Modularité

Références

1. D. Pierrakos, G. Paliouras, C. Papatheodorou, and C.D. Spyropoulos. Web usage mining as a tool for personalization : A survey. *User Modeling and User-Adapted Interaction*, 13(4) :311–372, 2003.
2. D. Tanasa and B. Trousse. Advanced data preprocessing for intersites web usage mining. *Intelligent Systems, IEEE*, 19(2) :59–65, 2004.
3. P.N. Tan and V. Kumar. Discovery of web robot sessions based on their navigational patterns. *Data Mining and Knowledge Discovery*, 6(1) :9–35, 2002.
4. R.W. Cooley. *Web Usage Mining : Discovery and Application of Interestin Patterns from Web Data*. PhD thesis, University of Minnesota, 2000.
5. G.W. Flake, S. Lawrence, C.L. Giles, and F.M. Coetzee. Self-organization and identification of web communities. *Computer*, 35(3) :66–70, 2002.
6. M.E.J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physical review E*, 69(2) :026113, 2004.

The Social Semantic web between the meaning and the mining

Samia Beldjoudi[1], Hassina Seridi[1], Catherine Faron-Zucker[2],

[1] Laboratory of Electronic Document Management LabGED, Badji Mokhtar University Annaba, Algeria
{beldjoudi, seridi}@labged.net
[2] I3S, Université Nice - Sophia Antipolis, CNRS 930 route des Colles, BP 145, 06930 Sophia Antipolis cedex, France
catherine.faron-zucker@unice.fr

Abstract. The social semantic web becomes in these recent years an attractive research field where the contributions attached to this domain know a real interest either by the researchers' community or even by the net users. We will present in this paper an original approach concerning a powerful technology which has recognized a great success in the social web area, we talk about folksonomies. The aim of our contribution is to provide another view about the semantics missed in the social web technologies and showing how we can use Data mining methods to extract the tags meaning in folksonomies. Especially we will present how we can analyze user profiles according to their tags in order to personalize recommendation of pertinent resources.

Keywords: Folksonomies, Web 2.0, Semantics, Association Rules, Resource Recommendation, Tags Ambiguity, Spelling Variations.

1 Introduction

Among the powerful technologies of Web 2.0 we find folksonomies, this term has recently appeared on the net to describe a system of classification to collaboratively creating and managing tags to annotate a set of web resources. This technology continues every day to gain popularity; however, the use of folksonomies for finding information on the web poses some problems [4], [3], [1]. Recently data mining has proved its efficiency in the Web field; in which data quantities become very large [5], [2]. We aim in this contribution to recommend resources which are annotated with tags suggested by association rules, in order to enrich user profiles with these resources. Also we will try to give a significant analysis to the relation between tags which are extracted by the association rules in order to exploit these ones to overcome the problem of spelling variations and let users benefit from all the resources annotated by a set of equivalent tags. In the next section we will give an overview about the design of our approach. Conclusion and future works are discussed in the last section.

2 Approach description

The objectives of our approach are twofold: first, we aim to treat the problem of tags ambiguity and realize a technique of resource recommendation. Second, we want to show how we can tackle the problem of semantics lacks in folksonomies in order to allow any user in the system benefit from the most possible number of relevant resources tagged by tags equivalent to the one written by this user. We have chosen to realize this idea a powerful method in the Data Mining field; which is the Association Rules.

So, in this approach we have used the association rules to achieving two different aims: The first one is to recommend a set of pertinent resources to each user in the community according to his profile basing on the community effect and the social interactions between society's members. The second is to use the association rules to overcome the problem of semantics lack in folksonomies by giving a logical classification for the linked tags in order to allow any user benefiting from the most possible number of relevant resources in these systems.

3 Conclusion

We have proposed a new technique based on the association rules of data mining and the social interactions between the folksonomies users, in which our objective is creating a consensus among the actors of a same system in order to teach them how they can organize their web resources with a correct and optimal manner. We have tested this approach on two datasets where we have obtained good results. Now, in order to expand and improve this work we propose to enrich the generated association rules by other measures which will help to specify the relevance of each rule to each user and we wish also to validate our approach on other databases.

References

1. S. Beldjoudi, H. Seridi and C. Faron-Zucker. Ambiguity in Tagging and the Community Effect in Researching Relevant Resources in Folksonomies. In Proc. of ESWC workshop User Profile Data on the Social Semantic Web, 2011.
2. S. Beldjoudi, H. Seridi and C. Faron-Zucker. Improving Tag-based Resource Recommendation with Association Rules on Folksonomies. In Proc. of ISWC workshop on Semantic Personalized Information Management: Retrieval and Recommendation, 2011.
3. P. De Meo, G. Quattrone, and D. Ursino. A query expansion and user profile enrichment approach to improve the performance of recommender systems operating on a folksonomy. User Modeling and User-Adapted Interaction, 20(1), 2010.
4. P. Mika. Ontologies are us: A unified model of social networks and semantics. In Proc. of 4th Int. Semantic Web Conference (ISWC 2005), Galway, Ireland, volume 3729 of LNCS. Springer, 2005.
5. C. Schmitz, A. Hotho, R. J□aschkee, and G. Stumme. Mining association rules in folksonomies. In Proc. of IFCS 2006 Conference: Data Science and Classification, Ljubljana, Slovenia. Springer, 2006.

Approche multi-agent pour l'ordonnancement job shop à deux machines sans attente

Ahmed Kouider, Samia Ourari, Brahim Bouzouia
Centre de Développement des Technologies Avancées, BP 17, Baba Hassen, 16303, Alger
{akouider, sourari, bbouzouia}@cdta.dz

Mots clés: Ordonnancement distribué; Job shop à deux machines sans attente; système multi-agent.

1 Introduction

Le champ d'application de l'ordonnancement considéré dans cette étude est celui des systèmes de production automatisés. Nous considérons le problème d'ordonnancement job shop classique avec une contrainte additionnelle sur le délai d'attente imposé nul entre les opérations d'un même job, ou le No-Wait Job Shop Scheduling (*NWJSS*). Nous nous intéressons au cas de deux ressources avec comme critère la minimisation du makespan, problème noté $J_2|nwt|C_{max}$ et qui est défini par un ensemble de n travaux (jobs) indépendants $\{J_1, J_2, \ldots, J_n\}$ à ordonnancer sur un ensemble de deux machines $\{M_1, M_2\}$. Chaque travail J_i, $i=1, \ldots, n$ est disponible à la date $r_i=0$, et possède sa propre gamme de production composée de deux opérations devant être exécutées selon un ordre bien défini et sans attente sur les deux machines de l'atelier. Nous supposons que les opérations sont non-préemptives, que les machines sont disjonctives, et que l'exécution d'une opération nécessite une durée connue à priori. Une solution du problème d'ordonnancement consiste alors à définir pour chaque opération de chaque travail une date de début d'exécution qui satisfait l'ensemble des contraintes citées plus haut et qui optimise le critère. Le problème *NWJSS* est connu pour être NP-difficile au sens fort pour le cas général [1], et demeure NP-difficile s'il est réduit à deux machines [2]. Schuster dans [3] estime il y a peu de contributions dans le domaine traitant de ce problème. Une approche métaheuristique est proposé par Liaw [4] pour résoudre le problème $J_2|nwt|C_{max}$.

De manière classique, la résolution du problème d'ordonnancement est assimilée à un problème de décision global et centralisé car la fonction ordonnancement gère l'organisation de la totalité des ressources. De point de vue pratique, et en raison du fait qu'un ordonnancement n'est jamais réalisé comme prévu, il plus avantageux d'adopter une démarche de résolution coopérative [5] afin de prendre compte l'autonomie de certaines ressources. Dans cet article, l'ordonnancement est considéré comme une fonction distribuée où la solution globale résulte de la négociation entre les différentes ressources.

2 Approche de résolution multi-agent

Comme évoqué en introduction et contrairement à une approche centralisée de résolution où les décisions d'ordonnancement sont calculées de manière globale, nous proposons de résoudre ce problème par une approche coopérative en considérant que les entités décisionnelles sont décentralisées au sein d'une organisation distribuée. Le problème d'ordonnancement No-Wait Job Shop Scheduling (NWJSS) est alors décomposé en deux sous problèmes d'ordonnancement locaux. Un agent est alors associé à chaque machine qui contrôle sa propre ressource et dispose de sa propre autonomie décisionnelle tout en

conservant le caractère confidentiel de ses données.

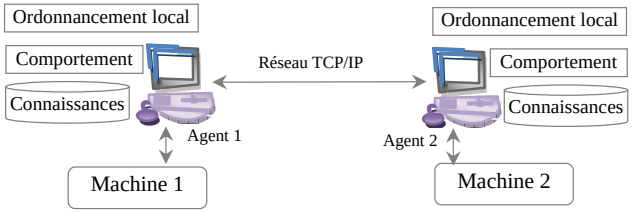


Fig 1. Architecture multi-agent pour l'ordonnancement distribué

Le système d'ordonnancement s'appui donc sur une architecture distribuée composée de deux agents ordonnanceurs qui coopèrent et interagissent, par échange de messages, suivant une approche décentralisée pour construire un ordonnancement local pour chaque machine, et de manière à converger au mieux vers une solution globale qui satisfait des exigences correspondant à des règles de coopération adoptées collectivement, cf. Figure 1. Le principe général du mode d'interaction entre les agents est le suivant : à chaque réception d'une nouvelle information par un agent, concernant la date de fin d'exécution d'une opération prédécesseur appartenant à l'agent amont, l'agent en question lance un processus de réordonnancement local afin d'établir son nouveau programme. Dans ce processus, l'agent veille à satisfaire toutes les contraintes du problème tout en minimisant le temps d'inoccupation de la ressource qu'il contrôle. L'interaction dans le système proposé consiste en la mise en relation dynamique entre les deux agents *Ordonnanceurs* par le biais d'un ensemble d'actions inter-changées (algorithmes locaux, échange de messages). Cette mise en relation dynamique résume le comportement de l'agent dans le système, et conduit à la fin de la coopération à une convergence vers une solution globale satisfaisante qui minimise le makespan. Le mécanisme de coopération adopté est donné par un algorithme distribué dont l'idée de base se résume comme suit : chaque agent transmet à son voisin aval chargé de réaliser une opération successeur une proposition sur la date prévue de livraison de l'opération, et en reçoit deux messages ; les informations véhiculées par les deux messages concernent respectivement une proposition sur la date de disponibilité d'une opération choisie par son voisin émetteur, et une information sur l'espace d'inoccupation résultant à son niveau s'il accepte la proposition de son voisin.

Afin de valider l'approche, des expérimentations sont réalisées sur des problèmes tests dont les résultats sont comparés à ceux trouvés par l'approche distribuée multi-agent pour la résolution du problème job shop proposée dans [6] dont les résultats constituent une borne inférieure pour notre problème.

[1] Lenstra, J. K., Rinnooy Kan, A. H. G., & Brucker, P. *Complexity of machine scheduling problems*. Annals of Discrete Mathematics, 1, 343–362 (1977)

[2] Sahni, S., & Cho, Y. *Complexity of scheduling shops with no wait in process*. Mathematics of Operations Research, 448, 448–457 (1979)

[3] Schuster CJ. *No-wait job shop scheduling:Tabu search and complexity of subproblems*. Mathematical Methods of Operations Research. 63 (3). 473-491 (2006)

[4] Liaw C.F. *An efficient simple metaheuristic for minimizing the makespanin two-machine no-wait job shop*, Computer and operations research, vol 35 (10). 3276-3283 (2008)

[5] Briand C., Ourari S. Bouzouia B. *Une approche coopérative pour l'ordonnancement sous incertitudes*, MOSIM'10, Paris-France, 2010.

[6] Kouider A., Bouzouia, B. *Multi-agent Job Shop Scheduling System based on Cooperative Approach of Idle Time Minimization*. International Journal of Production Research DOI:10.1080/00207543.2010.539276 (2011).

The Use of WordNets for Multilingual Text Categorization: A Comparative Study

Mohamed Amine Bentaallah and Mimoun Malki

EEDIS Laboratory, Department of computer sciences
Djillali Liabes University
Sidi Bel Abbes, 22000. ALGERIA
mabentaallah@univ-sba.dz
malki-m@yahoo.com
http://www.univ-sba.dz

Abstract. The successful use of the Princeton WordNet for Text Categorization has prompted the creation of similar WordNets in other languages as well. This paper focuses on a comparative study between two WordNet based approaches for Multilingual Text Categorization. The first relates on using machine translation to access directly the princeton WordNet while the second avoids machine translation by using the WordNet associated for each language.

Key words: Multilingual, Text Categorization, WordNet, Ontology, Mapping

1 Introduction

The growing popularity of the Princeton WordNet as a useful resource for English and its incorporation in natural language tasks has prompted the creation of similar WordNets in other languages as well. Indeed, WordNets for more than 50 languages are currently registered with the Global WordNet Association [1]. In this paper we try to answer the question: " ***Will the use of these WordNets in Text Categorization guarantee good results better than those obtained by the Princeton WordNet ?***".

2 Description of the two proposed approaches & Results

The two proposed approaches are shown in Figure Fig1. The results of the experimentations are presented in Table2, Concerning the profiles size, the best performances are obtained with size profile k = 900 for the two approaches. Indeed, the performances improve more and more by increasing the size of profiles. Comparing the results of the two approaches, the first approach largely outperform the second approach.

[1] http://www.globalwordnet.org

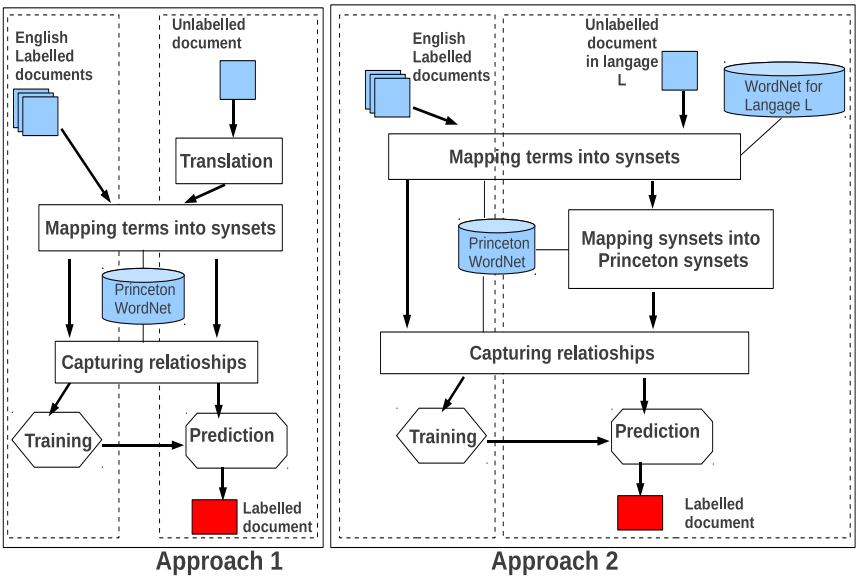


Fig. 1. The two compared approaches

3 Conclusion

In this paper, we have compared two approches for using WordNets for MTC, The first approach is based on using machine translation to use the Princeton WordNet while the second approach is based on replacing the use of machine translation by incorporating a WordNet for each language.

Table 1. Comparison of F-score results on the two approaches

Size of profiles	Approach1	Approach2
k=100	0.586	0.201
k=400	0.621	0.219
k=700	0.634	0.221
k=900	0.639	0.268

Avec le soutien de

Université Abou Bekr Belkaid (Tlemcen)
&
Agence Nationale pour le Développement de la Recherche Universitaire

---

Autres Sponsors