ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/342347824

Journées Doctorales en Informatique (JDI'2014)

Book · December 2014

CITATIONS
0

READS
56

2 authors:

Nadjib Kouahla
Université 8 mai 1945 - Guelma
30 PUBLICATIONS **28** CITATIONS
SEE PROFILE

Brahim Farou
Université 8 mai 1945 - Guelma
40 PUBLICATIONS **38** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Internet of Things: Motion Tracking in Collaborative Smart Cameras Network View project

Wireless Multimedia Sensor Network (WMSN) View project

All content following this page was uploaded by Brahim Farou on 21 June 2020.

The user has requested enhancement of the downloaded file.

Preface

Ces actes regroupent les articles présentés lors des quatrièmesJournées Doctorales en Informatique (JDI'2014) qui se sont déroulées les 03-04 Décembre 2014 à l'Université 8 Mai 1945 de Guelma, Algérie.

JDI est une manifestation scientifique qui a pour objectifs de regrouper les jeunes thésards en informatique avec leurs différentes problématiques de recherche et dans divers domaines d'application en informatique. Cette activité offre un cadre convivial aux participants pour présenter leurs travaux, confronter leurs idées et débattre leurs contributions dans les différents domaines de recherche. La présence de chercheurs séniors permet d'orienter et d'enrichir les idées de participants.

La manifestation couvre un spectre diversifié de thématiques de recherche, telles que:

- **Intelligent Systems**
- **Machine Learning**
- **Data, Image and Video mining**
- **Bio-informatics**
- **Multimedia Databases**
- **Mobile and Web Technology**
- **Networks and Distributed systems**

Le programme de cette édition 2014 comporte à la fois des articles contenant des travaux théoriques ainsi que des applications pratiques qui font souvent appel à diverses techniques issues des différents thèmes couverts par cette journée. Le comité de programme a procédé à l'évaluation des articles soumis et il a sélectionné dix-sept pour une présentation orale et six pour des sessions posters; soit un taux de sélection de 50%.

Nous tenons à remercier l'ensemble des auteurs qui ont manifesté un intérêt pour cette manifestation, les conférenciers invités qui n'ont pas hésité à apporter leurs contributions, les membres du comité de programme pour leur effort d'évaluation et les membres du comité d'organisation pour leur disponibilité.

Le Président du Comité d'Organisation

Dr. Mohamed Nadjib KOUAHLA

Table of Contents

Compression d'images dans les réseaux de capteurs sans fil 1
Chaouki Araar and Mouhammed Benmouhammed

Un modèle pour lapprentissage collaboratif du diagnostic médical..... 7
Mohamed Abderraouf Ferradji and Abdelmadjid Zidani

PBST*-based Hierarchical Protocol for Wireless Sensors Networks..... 13
Nassima Merabtine, Djamel Eddine Zegour and Djamel Djenouri

A Fuzzy Cognitive Maps-based Collective Decision in Cooperative Information Systems. 19
Maziz Asma and Zarour Nacer Eddine

Modélisation des Situations de Besoins dAssistance dans un EIAH : Le Modèle IMAC 25
Nadia Beggari and Taher Bouhadada

DEACP: An Energy Efficient Hierarchical Clustering Algorithm for Wireless 30
Gherbi Chirihane, Aliouat Zibouda and Mohammed Benmohammed

Amélioration du modèle de filtrage collaboratif hybride relatif 36
Abderrahmane Kouadria and Omar Nouali

An MDA Based Derivation process for e-Learning Software Product Lines..... 42
Nesrine Lahiani and Djamal Bennouar

An approach for coalition formation among autonomous robots
-Application to a rescue scenario- 48
Farouq Zitouni, Ramdane Maamri and Imane Boussebough

Les techniques de data mining dans les systèmes de recommandation..... 56
Abida Benahcene and Yacine Lafifi

Fuzzy Rules based classification: Analysis of the effect of Membership functions and
Rule Weights 63
Halima Salah, Mohamed Nemissi and Hamid Seridi

Efficient compute for complex illumination effects with virtual point light 68
Khemliche Sarra, Babahenini Mohamed Chaouki, Bahi Naima and Zerari Abd El Moumène

Real-Time People Detection..... 71
Houssam Eddine Rouabhia, Brahim Farou and Hamid Seridi

Un modèle pour la génération de situations dapprentissage de pédagogie par projet
collectif pour un environnement dapprentissage collaboratif..... 76
Hana Bendjedid and Tahar Bouhadada

Diff- HEEP: Hybrid Energy Efficiency Protocol with differentiated service 81
Azeddine Bilami and Amal Bourmada

Routing protocol for energy efficiency and fault tolerance in Wireless sensor network 86
Azeddine Bilami and Yasmine Djebaili

An Overview of Question-Answering Systems 90
Benaissa Safa and Omar Nouali

Ombres douces avec source de lumière surfacique 95
Zerari Abd El Moumène, Babahenini Mohamed Chaouki, Bahi Naima and Khemliche Sarra

Un Nouveau Modèle De Négociation Adaptatif Des Systèmes Multi-Agents, Application Dans Le E-Commerce 99
Toumi Chahrazad and Maamri Ramdane

Building Adaptive Mobile Applications Using Composition Mechanism 104
Afrah Djeddar, Hakim Bendjenna and Abdelkrim Amirat

InsMT: Instance Matching System for Ontology Alignment 111
Abderrahmane Khiat

Program Committee

Khababa Abdallah	Computer Science Department of UFAS Sétif
Boukrouche Abdelhani	University of Guelma
Moussaoui Abdelouahab	Professor, University of Sétif 1, Algerie
Abdelmalek Amine	GeCoDe Laboratory, Tahar Moulay University of Saida
Abdelkrim Amirat	LINA Laboratory
Baghdad Atmani	Computer Science Laboratory of Oran (LIO). Oran University
Mohamed Chaouki Babahenini	Computer science Departement University of Biskra
Halima Bahi	University of Badji-Mokhtar Annaba
Mohamed Batouche	Mentouri University, Constantine
Amel Behaz	Université de Batna
Nadjia Benblidia	Université Saad Dahlab de Blida
Chemesse Ennehar Bencheriet	Université 8 Mai 1945 de Guelma
Azeddine Bilami	Université de Batna
Thouraya Bouabana Tebibel	ESI, Algeria
Karim Bouamrane	University of Oran, Computer Science Department, LIO laboratory
Bachir Boucheham	University of Skikda
Houria Boudouda	université de guelma
Mahmoud Boufaida	Lire laboratory, Departement of Computer Science, Mentouri University of Constantine
Abdallah Boukerram	Université de Bejaia
Douadi Bourouaieh	university 8 mai 45 guelma
Beldjilali Bouziane	Department of Computer Science - University of Es-Senia Oran
Allaoua Chaoui	University Mentouri Constantine
Salim Chikhi	MISC laboratory, Mentouri Univerdity
Nadir Farah	Labged Laboratory Annaba University
Hadria Fizazi	university of sciences and technology of Oran
Cherif Foudil	Université de Biskra
Nacira Ghoualmi-Zine	Badji mokhtar university and Networks and systems Laboratory
Salim Haddadi	University of the 8th of May,1945
Reda Hamou	Université de Saida
Okba Kazar	Computer science department
Mohamed Tarek Khadir	Univ Badji Mokhtar
Mohamed-Khireddine Kholladi	University Mentouri of Constantine
Yacine Kissoum	Skikda University

Mohamed Nadjib Kouahla	Université de Guelma
Yacine Lafifi	LabSTIC Laboratory, Guelma University, Algeria
Zidani Madjid	Université de Batna
Smaine Mazouzi	Labo LERI-CReSTIC , IUT de Reims, 51100 France
Kamal E Melkemi	University of Biskra
Ahlem Melouah	University of Annaba
Elkamel Merah	Centre Universitaire Khenchela
Hayet Farida Merouani	University of Badji-Mokhtar Annaba
Djamel Meslati	Université Badji Mokhtar Annaba
Nemissi Mohamed	université de guelma
Laskri Mohamed Tayeb	Université de Annaba
Benmohammed Mohammed	Université de cne
Farid Mokhati	Université d'Oum elbouaghi
Zaidi Sahnoun	University Constantine2
Djamel Eddine Saidouni	Department of theoretical computer science and its applications, University Constantine 2, Algeria
Hassina Seridi	Université Badji Mokhtar Annaba Algérie
Ismahane Souici	Université de Guelma
Labiba Souici-Meslati	Laboratoire LRI, Université Badji Mokhtar, Annaba, Algérie
Bensebaa Tahar	Université Annaba, Algérie
Bouhadada Tahar	University of Annaba
Said Talhi	Université de Batna
Tighiouart	university
Yamina Tlili	university of Annaba
Mohamed Touahria	Ferhat Abas University
Belabbas Yagoubi	University
Nacereddine Zarour	LIRE Laboratory
Samir Zidat	Université de Batna
Boufaida Zizette	University Constantine, laboratory LIRE

Organization Committee

President: Dr. M. Nadjib KOUAHLA

Vice-president: Mr. Brahim FAROU

Membres:

- MeryemCHERAITIA
- Halima SALAH
- Fatima GUESSOUM
- Abida BENAHCENE
- AsmaBOUDRIA
- HoussamEddine ROUABHIA

Author Index

Abd El Moumène, Zerari 1, 104
Amirat, Abdelkrim 63
Araar, Chaouki 7
Asma, Maziz 13

Beggari, Nadia 19
Benahcene, Abida 25
Bendjedid, Hana 30
Bendjenna, Hakim 63
Benmohammed, Mohammed 56
Benmouhammed, Mouhammed 7
Bennouar, Djamal 81
Bilami, Azeddine 36, 42
Bouhadada, Tahar 30
Bouhadada, Taher 19
Bourmada, Amal 36
Boussebough, Imane 111

Chahrazad, Toumi 48
Chirihane, Gherbi 56

Djebaili, Yasmine 42
Djeddar, Afrah 63
Djenouri, Djamel 86

Farou, Brahim 90
Ferradji, Mohamed Abderraouf 68

Khiat, Abderrahmane 71
Kouadria, Abderrahmane 76

Lafifi, Yacine 25
Lahiani, Nesrine 81

Maamri, Ramdane 111
Merabtine, Nassima 86
Mohamed Chaouki, Babahenini 1, 104

Nacer Eddine, Zarour 13
Naima, Bahi 1, 104
Nemissi, Mohamed 99
Nouali, Omar 76, 95

Ramdane, Maamri 48

Rouabhia, Houssam Eddine 90

Safa, Benaissa 95
Salah, Halima 99
Sarra, Khemliche 1, 104
Seridi, Hamid 90, 99

Zegour, Djamel Eddine 86
Zibouda, Aliouat 56
Zidani, Abdelmadjid 68
Zitouni, Farouq 111

Keyword Index

ambient systems 111
answer extraction. 95
apprentissage 30
Apprentissage collaboratif 68
Apprentissage humain 25
assistance active 19

Background subtraction 90

CENTRIST descriptor 90
characteristic function 111
Cluster- based Routing 56
competences 111
Composition cost 63
Composition process 63
compression dimage 7
consommation dénergie 7
Context-aware 63
cooperation 111
Cooperative Information Systems 13

Data mining 25
DCT rapide 7
Decision Support 13
decision trees 111
Diagnostic médical 68
Distributed algorithm 56
dynamic coalition formation of rescue robots 111

E-Commerce 48
EIAH 25, 30
embedded systems 111
Energy saving 56

fault tolerancy 42
Filtrage 1
Filtrage Collaboratif 76
Filtrage Collaboratif hybride 76
Fuzzy Cognitive Maps 13
Fuzzy rule based classification system 99

global illumination (GI) 104
GPU 1

HEEP 36
Heterogeneous software entities 63
Hierarchical protocols 86

indicateurs dapprentissage 19
information research 95
Instance Annotation. 71
Instance Matching. 71

Machine Learning 25
Membership function 99
Metamodeling 63
Mobile apps 63
Model Driven Architecture 81
modèle de négociation. 48
multi-agent systems 111
Multiagent Systems 13

Network lifetime 86
négociation automatique 48

Ombre douce 1
Ontology Alignment. 71

PBST* 86
People detection 90
Personnalisation 25
Product Derivation 81
projet pédagogique 30
Pédagogie médicale 68

Quality of service 36
question analyses 95
question-answering systems 95

Raisonnement clinique 68
rendu 3D réaliste 1
Road Safety 13
Routing 86
routing protocol 42
Routing Protocol 36
Rule Weights 99
Réseaux de capteurs 7

SDDS 86
selection between robots 111
Semantic Interoperability. 71
Semantic Web. 71

Service differentiation 36
situations dassistance 19
socialisation des apprenants 30
Software Product Line 81
Système de recommandation 25
Système de Recommandation 76
système dassistance 19
Systèmes multi-agents 48

traces. 19
tâche collective 30

Urban Search And Rescue (USAR) 111

virtual point light(VPL) 104
visibility. 104

Wireless Sensor Network 56
Wireless sensor network 36, 42
Wireless Sensor Networks 86

Compression d'images dans les réseaux de capteurs sans fil

Mr. ARAAR Chaouki
Département d'informatique
Université Badji Mokhtar – ANNABA
chaouki_wil23@live.fr

Pr. BENMOHAMMED Mohammed
Département d'informatique
Université Constantine 2
ben_moh123@yahoo.com

Résumé – *Avec l'avènement de mini-caméras CMOS, l'engouement pour les réseaux de capteurs d'images est commencé. Comparés aux réseaux de capteurs classiques qui manipulent des données scalaires, les réseaux de capteurs d'images sont particulièrement gourmands en énergie puisque le volume d'information associe à une image est de plusieurs ordres de grandeurs supérieur à une information scalaire simple. Il semble évident de prime abord que le coût énergétique de la transmission de l'image peut être réduit significativement en compressant l'image à la source. Or, la contrainte de la limitation des ressources des nœuds rend impossible ou inefficace en pratique l'exécution des algorithmes de compression standards. Dans ce domaine, on remarque des efforts de plusieurs auteurs pour adapter les algorithmes de compression bien connus comme JPEG, JPEG2000 ou SPIHT aux contraintes particulières des réseaux de capteurs sans fil. Notre proposition, suit aussi cette voie en proposant dans cet article d'intégrer dans une chaîne de compression de type JPEG, une DCT rapide basé sur l'algorithme de Loeffler, de la quantification scalaire, et du codage entropique basé sur le codeur Golomb couplé avec le codeur arithmétique QM. Cela permet de réduire le nombre d'opérations pour calculer les coefficients de la DCT. Ce qui entraîne systématiquement une réduction de la complexité du calcul de la chaîne de compression, et par incidence une réduction de la consommation d'énergie.*

Mots-clés: Réseaux de capteurs, compression d'image, DCT rapide, consommation d'énergie**.**

1 Introduction

Les progrès technologiques incessants ont permis le développement de nouveaux types de capteurs, issue de la convergence de la micro-électronique et des technologies de communication sans fil. Les capteurs, nouveaux venus dans les thèmes de recherche en informatique peuvent être vus comme des systèmes autonomes miniaturisés, équipés d'une unité de traitement et de stockage de données, d'une unité de transmission sans fil et d'une batterie. Ces nouveaux petits composants peu onéreux pouvant s'organiser en réseau autonome pour récolter des données et les remonter à une station de base, avec comme contrainte fondamentale la limitation de leurs ressources de calcul, de stockage et surtout d'énergie.

Parmi toutes les applications potentielles des réseaux de capteurs sans fil, celles utilisant des capteurs d'image sont appréciables pour tout ce qui concerne la reconnaissance, la localisation et le dénombrement d'objets par la vision [1]. Des nœuds capteurs dotés d'une caméra existent déjà, comme Cyclops [2], et le capteur multimédia IMB vendu par Crossbow [3]. Mais les applications basées sur des capteurs d'image sont particulièrement gourmandes en énergie puisque la quantité de données nécessaire pour représenter une image diffère de plusieurs ordres de grandeur en comparaison d'une valeur scalaire classique (une mesure de température par exemple).

Comme l'émetteur-récepteur radio est un des composants les plus gourmands en énergie, il semble évident que le coût d'énergie de la transmission de l'image peut être réduit significativement en compressant l'image à la source. Toutefois, les nœuds capteurs sont par nature très limités en capacité de calcul et de mémoire, c'est une conséquence directe des contraintes de consommation d'énergie, de miniaturisation et de coût de fabrication. Ils ne peuvent donc pas implanter des algorithmes de compression de trop grande complexité. De plus, des travaux comme ceux de Ferrigno et al. [4] ont montré que des algorithmes comme JPEG2000, SPIHT et même JPEG sont inefficaces sur la plupart des plateformes logicielles utilisées dans les réseaux de capteurs, c'est-à-dire que l'opération de compression de l'image coûte plus d'énergie qu'une transmission de l'image non compressée.

Nous proposons dans cet article, d'intégrer dans une chaîne de compression de type JPEG, une DCT rapide basé sur l'algorithme de Loeffler [5], de la quantification scalaire, et du codage entropique basé sur le codeur Golomb couplé avec le codeur arithmétique QM. L'approche DCT rapide utilisée permet de réduire le nombre d'opérations pour calculer les coefficients de la DCT. Elle entraîne systématiquement une réduction de la complexité de calcul de la chaîne de compression, et par incidence une réduction de la consommation d'énergie. Le choix du codeur Golomb couplé avec le codeur

arithmétique présente un bon compromis vitesse d'exécution/taux de compression.

Ce papier est organisé comme suit, dans la section 2 on présente quelques exemples d'applications pour les réseaux de capteurs d'images. Dans la section 3, on discute les principales différences entre les réseaux de capteurs classiques et les réseaux de capteurs d'images. Après cela, on expose les travaux issus de l'état de l'art de la compression d'images dans les réseaux de capteurs d'images sous la contrainte de la limitation des ressources. Dans la cinquième section, on présente la chaîne de compression proposée que nous évaluions par simulation, et on terminera enfin par une conclusion.

2 Applications des réseaux de capteurs d'images

Les réseaux de capteurs d'image concernent un vaste champ d'applications, en fait toutes celles qui touchent à la détection, la reconnaissance, la localisation et le dénombrement d'objets par la vision, en voici quelques exemples :

2.1 Applications militaires

Les réseaux de capteurs d'image peuvent être très utiles pour l'espionnage militaire et la surveillance des champs de bataille. Un réseau de capteurs expérimental a par exemple été déployé en 2003 par une équipe de l'Université de l'état d'Ohio sur la base MacDill de l'US air force à Tampa, Floride [6].

2.2 Vigilance environnementale

Les réseaux de capteurs d'image sont aussi très utiles pour l'observation de la faune et de la flore en milieu naturel. Le projet PODS [7] développé par l'université d'Hawaii avait pour objectif d'identifier les raisons pour lesquelles certaines plantes menacées poussent dans certaines régions mais pas dans d'autres. Le projet BearCam [8] est un autre exemple d'application pour l'observation de la nature. Le but de ce projet était de surveiller les ours bruns à la sortie de leur sommeil hivernal.

2.3 Applications de la surveillance mission-critiques

Des réseaux de capteurs d'image peuvent être utilisées pour la surveillance mission-critiques telles que la détection d'intrusion ou la réaction aux catastrophes [9][10][11]. Les capteurs d'images peuvent être jetés sur la masse en cas de besoin pour la détection d'intrusion ou pour les applications de secours (missions SAR) en cas de catastrophe.

3 Vue d'ensemble sur la compression d'images dans les réseaux de capteurs sans fil

Rappelons que les réseaux de capteurs d'images sont des réseaux distribués qui consistent de petits capteurs équipés de caméras CMOS de basse consommation telle que Cyclops. A l'intérieur d'un tel réseau, chaque nœud-capteur d'image a la capacité d'acquérir, de compresser et de transmettre des images utiles à la station de base, à travers le lien établie entre la source et la station de base (figure 1).

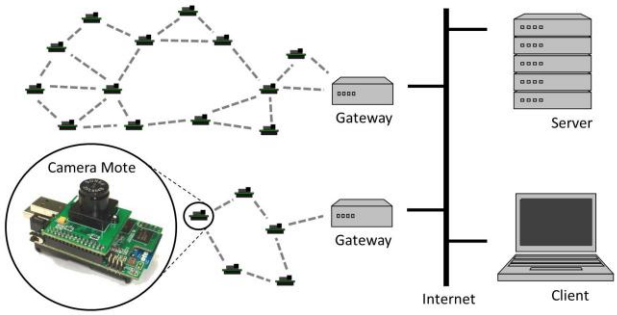


Figure 1. Réseau de capteurs d'image

Comparés aux réseaux de capteurs classiques qui manipulent des données scalaires (mesure de température, d'humidité, etc.), les réseaux de capteurs d'images ont quelques différences très importantes [12] :

- Le volume d'information associée à une image est de plusieurs ordres de grandeurs supérieur à une information scalaire simple.
- Les images naturelles possèdent une certaine tolérance aux pertes de paquet puisqu'il y a des corrélations spatiales entre les pixels voisins. Contrairement aux réseaux de capteurs classiques où la perte de quelques paquets peut affecter la valeur de donnée collectée.
- La plupart des capteurs traditionnels ont un champ de perception omnidirectionnelle, c'est-à-dire que l'orientation du capteur n'a pas d'impact sur la valeur de la grandeur physique mesurée. Le champ de vision des caméras est par contre restreint en direction.
- Les nœuds capteurs d'image voisins surveillant la même petite région locale d'intérêt ont de multiples et de différentes vues de cette scène, comparativement au capteur de données scalaire où une valeur unique (exemple, valeur de température) est collectée par tous les nœuds voisins (situés dans la même région).
- Ressources importantes en mémoire, en traitement et en puissance de calcul sont nécessaires afin de manipuler la quantité importante du flux visuel.
- Des algorithmes de compression efficaces en énergie sont obligatoires pour manipuler le flux d'information.

La plupart des études sur les réseaux manipulant des données scalaires ont typiquement assumées que les coûts computationnels, y compris acquisition et compression, sont insignifiant comparés aux coûts de la communication [13]. Cette supposition peut convenir aux réseaux de capteurs classiques qui manipules des grandeurs physiques légères (tels que la température, l'humidité, la pression) où le coût du calcul de la compression, est négligeable comparé au coût de la communication. Dans le cas de réseaux de capteurs d'images, cette supposition ne peut pas être tenue, vue la quantité énorme du flux visuel.

Habituellement, la transmission d'image précédée par compression est un choix idéal afin de gagner en temps et en énergie. Des algorithmes de compression existants ont fournies d'excellentes performances en termes de rapport débit/distorsion, mais ils ne sont pas éligibles dans le contexte des réseaux de capteurs. Ces algorithmes, en général, sont utilisés à des fins de stockage, ou encore utilisés quand aucune restriction de temps ou d'énergie nécessaires. Par exemple, la compression utilisant Fractals ou JPEG2000 est très consommatrice en temps et en énergie [14]. D'autres travaux aussi [4] ont montré en effet que des algorithmes bien connus comme JPEG ou JPEG2000 ont un coût d'énergie bien supérieur au gain qu'ils amènent sur l'émetteur-récepteur radio. Autrement dit, le capteur d'image épuiserait plus vite son énergie en envoyant des images compressées que des images non compressées.

Les travaux issus de l'état de l'art de la compression d'images dans les réseaux de capteurs sous la contrainte de la limitation des ressources peuvent être classés en deux catégories : ceux traitant de la compression de l'image à la source et ceux portant sur la compression distribuée.

3.1 Compression d'images à la source

Ici la compression de l'image est réalisée par un algorithme local, c'est-à-dire que tout le traitement est supporté par le nœud caméra initiateur de la prise d'image. Cette partie de l'état de l'art est la plus abondante car la compression d'image à la source est ce qu'il y a de plus facile à mettre en œuvre. La plupart des travaux proposent des modifications des algorithmes bien connus comme JPEG [15][16], JPEG2000 [17][18] ou SPIHT [19] pour réduire leur coût de calcul et/ou de mémoire.

3.2 Compression d'images distribuée

Il ne faut pas confondre la compression d'image distribuée, au sens où le traitement d'une image sera réparti sur plusieurs nœuds de manière coopérative, et la compression de source distribuée (DSC), où des images prises par des caméras distinctes mais mutuellement corrélées sont compressées séparément et décodées conjointement. En d'autres mots, il n'y a pas de communication entre les sources et chacune encode ses données indépendamment des autres. Une caractéristique principale de la DSC est que les nœuds sources ont des encodeurs de très faible complexité, la complexité se situant au niveau du décodeur étant donc supportée par le puits. La DSC a été étudiée dans le contexte des réseaux de capteurs et récemment, quelques travaux se sont intéressés au codage d'images et de séquences vidéo [20][21].

La deuxième approche consiste à développer un système de coopération entre les nœuds du réseau pour répartir le coût de la compression d'image. Il s'agit pour le nœud source de découper l'image en plusieurs morceaux et de les envoyer à des nœuds distincts pour qu'ils exécutent, chacun, une partie du traitement [22][23].

4 Chaîne de compression proposée

Les nœuds capteurs sont par nature très limités en capacité de calcul et de mémoire, ce qui rend l'exécution des algorithmes de compression bien connus comme JPEG ou JPEG2000 inefficace en termes d'énergie. Ainsi, Le développement de nouvelles méthodes de compression tenant compte de l'efficacité énergétique est nécessaire pour que le déploiement de réseaux de capteurs d'images puisse être envisagé en pratique.

Plusieurs auteurs ont proposé de nouvelles méthodes de compression satisfaisant aux contraintes posées par réseaux de capteurs sans fil, c'est-à-dire de faible complexité calculatoire et efficace en énergie. La méthode qu'on propose dans cet article repose sur l'adaptation de l'algorithme JPEG aux contraintes particulières des réseaux de capteurs sans fil.

La compression JPEG est une méthode de compression non conservative de l'image. L'algorithme a besoin de réaliser plusieurs étapes afin de pouvoir compresser une image. Premièrement, l'image de l'entrée est divisée en plusieurs blocs de taille fixe 8x8 pixel. Ensuite la transformée 2D-DCT est appliquée sur chaque bloc afin de séparer les basses fréquences des hautes fréquences. Par la suite, chaque bloc DCT est quantifié uniformément. Le résultat de la quantification est réorganisé à nouveau d'une manière zigzag, des fréquences basses aux fréquences hautes. Après cela, le codage RLE (Run Length Encoding) est appliqué pour réduire la longueur des séquences produites. Finalement, le processus de codage entropique réversible (tel que Huffman ou codage arithmétique) est effectué sur les données quantifiées pour produire des mots de code de longueur fixes ou variables [24] (voir figure 2).

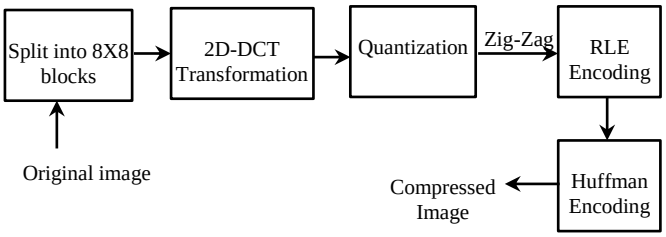


Figure 2. Étapes de la compression JPEG

Dans la section précédente, nous avons présenté les étapes de la compression JPEG, à chaque étape notre choix s'oriente vers la méthode la plus adaptée.

4.1 Transformation

Dans le schéma de compression basé sur la DCT, l'étape de transformée bidimensionnelle représente à elle seule 60% du coût total en énergie de la compression [25].

Beaucoup de différents algorithmes rapides pour le calcul de la DCT ont été présenté dans la littérature [5][26][27][28][29]. La plupart de ces algorithmes nécessitent 12 multiplications et 29 additions pour calculer la DCT d'un vecteur à 8 points (voir tableau). L'algorithme de Loeffler est actuellement le plus efficace ayant été publié (avec 11 multiplications pour la DCT à 8 points).

Auteur Réf.	Chen [26]	Wang [27]	Vetterli [28]	Suehiro [29]	Loeffler [5]
Mult.	16 (13)	13	12	12	11
Add.	26 (29)	29	29	29	29

Tableau 1. Complexité arithmétique des algorithmes rapides de la DCT.

Afin de rendre l'utilisation de la DCT compatible avec les contraintes des réseaux de capteurs, notre choix s'oriente vers l'algorithme de Loeffler pour le calcul de la DCT. La DCT 2-D peut être obtenue en appliquant d'abord la DCT 1-D sur chaque ligne du bloc de pixels puis sur chacune des colonnes du bloc résultant. Avec l'algorithme de Loeffler, la DCT 2-D nécessite 176 multiplications et 464 additions.

4.2 Codage entropique

Puisque la majorité des applications de réseaux de capteurs d'image se focalisent sur la vitesse d'exécution (un temps plus bas correspond à une consommation d'énergie plus basse) ainsi que sur les performances de compression, le codage de Golomb-Rice couplé avec le codeur arithmétique QM présente un bon compromis vitesse d'exécution/taux de compression.

Enfin, la chaîne de compression qu'on propose est illustrée sur la figure 3.

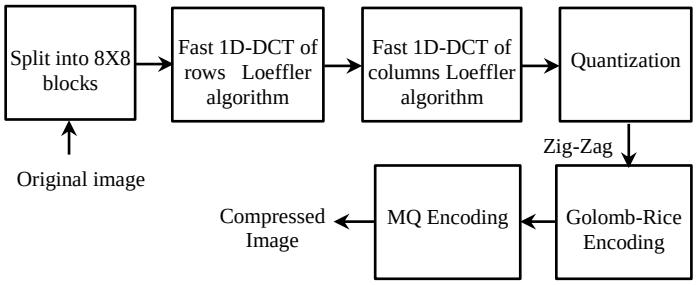


Figure 3. Chaîne de compression proposée.

5 Simulation et performance

Nous avons pris comme référence pour notre simulation les valeurs caractéristiques de nœud-capteur bien connu : Mica2 fabriqué par Crossbow [3]. Il s'appuie sur un microcontrôleur Atmel ATMEGA 128L cadencé à 7, 3728 MHz, avec un émetteur radio ChipCon CC1000 qui opère à la fréquence de 915 MHz et qui fournit un débit de transmission de 38, 4 kbps. La simulation porte sur des images (BMP) en niveaux de gris codées à l'origine sur 8 bits par pixel. Le tableau 2 résume Les paramètres de la simulation pour la plateforme utilisée.

Nous avons utilisé comme simulateur, Castalia [30], qui se base sur OMNET++ et qui offre un modèle radio intéressant basé sur les travaux de Zuniga et al [31]. Castalia gère la consommation d'énergie des capteurs et de processeur, l'utilisation de la mémoire et du temps CPU.

	Mica2
Interfacer Radio	• Chipcon CC1000 • Puissance d'émission : Max (10dbm) Min (-20dbm)
Mémoire	• 4k de RAM • 512K de mémoire flash
Format de l'image	BMP en niveaux de gris (8bpp)
Taille de l'image	
Facteur de qualité	
Taille des paquets	256 octets
Batterie	18720 joules (2 Piles alcalines)

Tableau 2. Paramètres de simulation pour Mica2

Les figures 4 et 5 montrent la consommation en énergie de la compression et transmission d'images pour le mote Mica2 dans le cas où son émetteur radio est mis à sa puissance minimale (-20dbm) et maximale (10dbm).

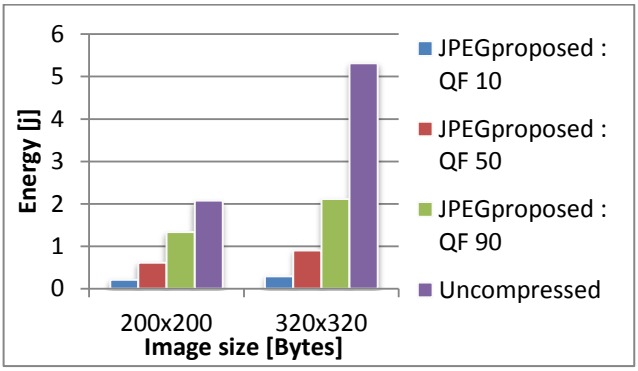


Figure 4. Consommation en énergie de la compression et transmission d'images à (-20dbm) avec différents facteurs de qualité : 10, 50, 90

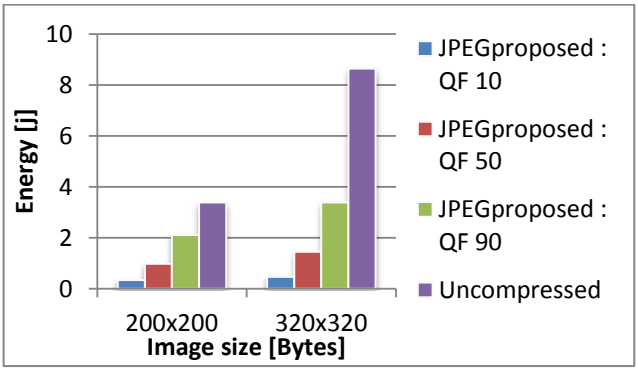


Figure 5. Consommation en énergie de la compression et transmission d'images à (10dbm) avec différents facteurs de qualité : 10, 50, 90

6 Conclusion et perspectives

Nous avons présenté dans cet article un cas spécifique de réseau de capteurs sans fil : les réseaux de capteurs d'images. Comparés aux réseaux de capteurs classiques qui manipulent des données scalaires, les réseaux de capteurs d'images traitent un volume d'informations beaucoup plus grand. Nous nous sommes intéressés à la compression d'images pour diminuer le volume des données à transmettre afin réduire la consommation d'énergie. Ainsi, Nous proposons une nouvelle méthode de compression d'image satisfaisant aux contraintes posées par les réseaux de capteurs sans fil, c'est-à-dire de faible complexité calculatoire et efficace en énergie.

Nous avons évalué les performances de la chaîne de compression proposée par simulation et nous avons pu obtenir un gain important en énergie. Les résultats montrent bien que la compression est rentable de point de vue consommation en énergie pour le mote Mica2. Les économies d'énergie deviennent de plus en plus intéressantes en diminuant le facteur de qualité. Les économies sont d'environ 60.87%, 83.16%, 94.54% respectivement qu'on transmet une image (320x320) codée à 2.98 *bpp*, 1.30 *bpp* et 0.42 *bpp* avec une puissance d'émission de l'émetteur radio de 10dbm.

Les résultats sont probants et montrent sans ambiguïtés la faisabilité de notre chaine de compression pour les systèmes limités en ressources tel que le Mica2. Comme perspective, on veut étendre notre simulation pour d'autres plateformes comme le mote MicaZ et imote2.

Références

[1]. S. Soro and W. Heinzelman. A survey of visual sensor networks. *Advances in Multimedia*, 2009: Article ID 640386, 21 pages, 2009.

[2]. M. Rahimi, R. Baer, O. I. Iroezi, J. C. Garcia, J. Warrior, D. Estrin, and M. Srivastava. Cyclops: In situ image sensing and interpretation in wireless sensor networks. *ACM Conference on Embedded Networked Sensor Systems (SenSys)*, pages 192–204, San Diego (CA), USA, November 2005.

[3]. Crossbow Technology Inc. http : //www.xbow.com/.

[4]. L. Ferrigno, S. Marano, V. Paciello, et A. Pietrosanto. Balancing computational and transmission power consumption in wireless image sensor networks. Dans *IEEE Int. Conference on Virtual Environments, Human-Computer Interfaces, and Measures Systems (VECIMS)*, Giardini Naxos, Italy, July 2005. IEEE.

[5]. C.Loeffler, A. Lieenberg, and G.S. Moschytz, "Practical fast 1-D DCT algorithms with 11 miltiplications", IEEE, pp.988-991.1989.

[6]. Arora, A., P. Dutta, S. Bapat, V. Kulathumani, H. Zhang, V. Naik, V. Mittal, H. Cao, M. Demirbas, M. Gouda et al. (2004). A line in the sand: a wireless sensor network for target detection, classification, and tracking. Computer Networks, 46(5), 605–634.

[7]. Biagioni, E.S. et KW Bridges (2002). The application of remote sensor technology to assist the recovery of rare and endangered species. International Journal of High Performance Computing Applications, 16(3), 315–324.

[8]. Wawerla, J., S. Marshall, G. Mori, K. Rothley et P. Sabzmeydani (2009). Bearcam : Automated wildlife monitoring at the arctic circle. Machine Vision and Applications, 20(5), 303–317.

[9]. T. He and al., "Energy-efficient surveillance system using wireless sensor networks," in MobiSys '04: Proceedings of the 2nd international conference on Mobile systems, applications, and services. New York, NY, USA: ACM, 2004, pp. 270–283.

[10]. S. Oh, P. Chen, M. Manzo, and S. Sastry, "Instrumenting wireless sensor networks for real-

time surveillance," in Proc. of the International Conference on Robotics and Automation, May 2006.

[11]. R. Cucchiara and al., "Using a wireless sensor network to enhance video surveillance," J. Ubiquitous Computing and Intelligence, 1(2), 2007.

[12]. A. Mammeri, B. Hadjou, A.Khoumsi. "A survey of image compression algorithms for visual sensor networks".

[13]. B. Krishnamachari, D. Estrin, and S. B. Wicker, "The impact of data aggregation in wireless sensor networks," in *ICDCSW '02: Proceedings of the 22nd International Conference on Distributed Computing Systems*. Washington, DC, USA: IEEE Computer Society, 2002, pp. 575–578.

[14]. A. Mascher-kampfer, H. Stgner, and A. Uhl, "Comparison of compression algorithms impact on fingerprint and face recognition accuracy," in *in Visual Computing and Image Processing VCIP 07, Proceedings of SPIE*, 2007.

[15]. A. Mammeri, A. Khoumsi, D. Ziou et B. Hadjou (2008). Modeling and adapting JPEG to the energy requirements of VSN. In: Computer Communications and Networks, 2008. ICCCN'08. Proceedings of 17th International Conference on. IEEE. pp. 1–6.

[16]. L. Makkaoui, V.Lecuire and J.M. Moureaux, Fast Zonal DCT-based image compression for Wireless Camera Sensor Networks. International Conference on Image Processing Theory, Tools and Applications (IPTA10), page 126-129, France, Juillet 2010.

[17]. Lian, C.J., K.F. Chen, H.H. Chen and L.G. Chen (2003). Analysis and architecture design of block-coding engine for ebcot in jpeg 2000. Circuits and Systems for Video Technology, IEEE Transactions on, 13(3), 219–230.

[18]. Chiang, J.S., C.H. Chang, Y.S. Lin, C.Y. Hsieh and C.H. Hsia (2004). High-speed EBCOT with dual context-modeling coding architecture for jpeg2000. In: Circuits and Systems, 2004. ISCAS'04. Proceedings of the 2004 International Symposium on. Vol. 3. IEEE. pp. III–865.

[19]. Kim, K.L. and S.W. Ra (2004). Performance improvement of the SPIHT coder. Signal Processing: Image Communication, 19(1), 29–36.

[20]. Wagner, R., R. Nowak and R. Baraniuk (2003). Distributed image compression for sensor networks using correspondence analysis and super-resolution. In: Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on. Vol. 1. IEEE. pp. I–597.

[21]. Thirumalai, V., I. Tosic and P. Frossard (2008). Symmetric distributed coding of stereo omnidirectional images. Signal Processing: Image Communication, 23(5), 379–390.

[22]. Wu, H. and A.A. Abouzeid (2004b). Power aware image transmission in energy constrained wireless networks. In: Computers and Communications, 2004. Proceedings. ISCC 2004. Ninth International Symposium on. Vol. 1. IEEE. pp. 202–207.

[23]. Marino, F., V. Piuri et E.E. Swartzlander Jr (1999). A parallel implementation of the 2-D discrete wavelet transform without inter processor communications. Signal Processing, IEEE Transactions on, 47(11), 3179–3184.

[24]. G. K. Wallace, "The JPEG still picture compression standard," *Commun. ACM*, vol. 34, no. 4, pp. 30–44, 1991.

[25]. Marcelloni, F. and M. Vecchio (2008). A simple algorithm for data compression in wireless sensor networks. Communications Letters, IEEE, 12(6), 411–413.

[26]. W. A. Chen, C. Harrison, and S. C. Fralick, "A Fast computational Algorithm for the Discrete Cosine Transform," IEEE Transactions on Communications, Vol. COM-25, No. 9, Sept. 1977, pp. 1004-1011.

[27]. Z. Wang, "Fast Algorithms for the Discrete W-Transform and for the Discrete Fourier Transform," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-32, No. 4, Aug. 1984, pp. 803-816.

[28]. M. Vetterli, H. Nussbaumer, 'Simple FFT and DCT Algorithms with Reduced Number of Operations," Signal Processing (North Holland), Vol. 6, No. 4, Aug. 1984, pp. 267-278.

[29]. N. Suehiro and M. Aatori, "Fast Algorithms for the DFT and other Sinusoidal Transforms," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-34, No. 3, June 1986, pp. 642-644.

[30]. http://castalia.npc.nicta.com.au/index.php.

[31]. Marco Zuniga, Bhaskar Krishnamachari, "Analyzing the Transitional Region in Low Power Wireless Links", First IEEE International Conference on Sensor and Ad hoc Communications and Networks (SECON), Santa Clara, CA, October 2004.

Un modèle pour l'apprentissage collaboratif du diagnostic médical

Mohamed Abderraouf FERRADJI
Département d'informatique Université Hadj Lakhdar
Laboratoire LaSTIC
Batna, Algérie
ferradji.abderraouf@gmail.com

Abdelmadjid ZIDANI
Département d'informatique Université Hadj Lakhdar
Laboratoire LaSTIC
Batna, Algérie
abdelmadjid.zidani@univ-batna.dz

Résumé— **L'établissement d'un diagnostic médical est une tache fondamentalement collaborative qui implique la coordination et la négociation entre plusieurs médecins des différentes spécialités pour répondre aux besoins des malades. Cette collaboration sera d'autant plus efficace lorsque les médecins se familiarisent les uns avec les autres depuis leur formation médicale. L'apprentissage collaboratif du diagnostic médical constitue une alternative pédagogique effective qui peut compléter les insuffisances de l'enseignement classique. Notre objectif dans ce papier est de mettre en valeur l'aspect de la collaboration dans les environnements d'apprentissage collaboratif du diagnostic médical, en proposant le modèle approprié. Notre travail part donc d'une analyse minutieuse centrée autour des études cognitives accomplies dans le contexte de la pédagogie médicale. Notre approche s'appuie sur le fait que l'apprenant est considéré en tant que partenaire actif à part entière du processus d'apprentissage collaboratif du diagnostic médical. Dans ce papier, nous commençons d'abord par introduire la notion de diagnostic médical. Nous revenons ensuite sur les fondements du raisonnement clinique, notamment en discutant la modélisation de ce dernier. Nous explicitons ensuite notre modèle dédié au support de la collaboration entre plusieurs apprenants. Enfin, nous concluons notre papier en spécifiant les perspectives qui s'ouvrent à notre travail de recherche.**

***Mot clés*— Diagnostic médical, Raisonnement clinique, Apprentissage collaboratif, Pédagogie médicale.**

I. INTRODUCTION

Ces dernières années, l'enseignement des sciences médicales a connu des progrès considérables, ce qui a imposé la nécessité de la création des sous spécialités dans la même spécialité médicale. Le souci majeur des établissements devient ainsi la proposition de formations de bonne qualité afin de garantir une prise en charge efficace des malades. Il devient ainsi nécessaire pour les apprenants d'acquérir une masse importante de connaissances durant leur cycle de formation spécialisé. Cependant, la prise en charge des malades requiert une collaboration entre les médecins de différentes spécialités pour aboutir à des diagnostics rigoureux. D'autre part, le processus d'établir un diagnostic qui constitue une tache cruciale reste souvent délicat et difficilement maitrisable [12, 9, 13, 19]. En effet, afin d'élaborer le diagnostic approprié d'un nouveau cas, le médecin doit établir le lien entre les signes et les symptômes déterminés à partir de la plainte énoncée par le patient, les examens cliniques (éventuellement para cliniques) et l'immense quantité de connaissances médicales mémorisées dans sa mémoire à long terme [12, 9, 7].

Les méthodes d'apprentissage collaboratif du diagnostic médical constituent une alternative pédagogique qui peut aider les futurs médecins à surmonter les problèmes relatifs à la complexité du processus du diagnostic, ainsi que ceux de la collaboration imposés par la vie professionnelle médicale.

L'intégration de ces méthodes d'apprentissage dans les facultés de médecine de notre pays nécessite des changements radicaux au niveau des systèmes éducatifs et même sur l'infrastructure des facultés [2]. De plus, cette activité pédagogique se base sur l'apprentissage au sein de petits groupes d'apprenants et son organisation en mode présentiel n'est pas toujours facile [5]. Précisément dans notre pays avec la superficie qui le caractérise (2 381 741 Km^2) et qui ne permet pas au nombre assez limité de médecins expérimentés de contribuer aux activités pédagogiques de toutes nos facultés de médecines. Ce qui limite considérablement les chances des apprenants des zones éloignées du sud de bénéficier de leurs expériences de façon équitable. Ce constat découle directement de notre étude à propos du déséquilibre de disponibilité des médecins spécialistes par rapport au nord du pays. L'orientation vers les outils de support d'apprentissage collaboratif à distance peut offrir plus de flexibilité et de souplesse, en temps et en espace à ce type d'apprentissage [2].

Dans ce papier nous proposons un modèle d'apprentissage supportant la collaboration, la coordination et la communication entre un groupe d'apprenants géographiquement distants, qui partagent une tache commune d'élaboration du diagnostic médical en mode synchrone. Ce modèle vise particulièrement à favoriser le partage de la connaissance et savoir-faire entre les différents participants. Pour assurer la pertinence de notre modèle pertinent, nous avons considéré les études cognitives accomplies dans le domaine de la pédagogie médicale, ainsi que les recherches dédiées à la modalisation des processus du raisonnement clinique. Enfin, il faut également préciser que nous avons tenu compte des travaux entrepris dans le contexte de l'amélioration de la qualité de l'apprentissage dans le cadre académique.

II. LE RAISONNEMENT CLINIQUE

Dans le domaine de la pédagogie médicale, toutes les définitions associées au processus de raisonnement clinique, qui ont été proposées s'accordent sur un point commun. Il s'agit de la perception de ce processus comme un ensemble d'activités mentales qui permettent au clinicien de prendre la bonne décision dans une situation clinique spécifique [12, 1, 17, 18]. En outre, les études qui ont été réalisées dans ce domaine ont révélé que les médecins utilisent différentes méthodes de raisonnement. Ces méthodes peuvent être décomposées en deux grandes classes : La classe des processus analytiques et celle des processus non-analytiques [12].

Suivant une démarche de raisonnement analytique, le médecin peut expliciter son processus de raisonnent [19], à travers lequel il doit établir une relation entre les signes et les symptômes du patient, ainsi que la désignation des catégories des maladies qui leurs sont associés [9]. Ce type de raisonnement est utilisé, en général, par les médecins qui n'ont que très peu d'expérience et qui doivent toujours suivre un ensemble d'étapes qui leur permettent de confirmer le diagnostic mentionné. Ce type de raisonnement est également utilisé dans les cas complexes ou peu rencontrés, où le médecin ne peut pas établir directement un diagnostic et devra fournir des efforts de raisonnement avant de parvenir au diagnostic final [18].

Néanmoins, suivant la démarche non-analytique, qui est aussi dénommée : « *Reconnaissance de Forme* », le médecin ne peut pas expliciter le processus de raisonnement utilisé [19]. Donc, le raisonnement non-analytique peut être utilisé dans les cas typiques, où le médecin peut établir le diagnostic sans fournir d'efforts de raisonnement, lorsque le cas se caractérise par sa grande similarité avec le prototype d'une maladie spécifique. Ce type de raisonnement est utilisé même par les médecins expérimentés qui peuvent ainsi établir leur diagnostic de façon systématique en mobilisant le minimum de leur conscience à travers la projection directe du cas similaire sur celui en cours de traitement. Il faut noter ici que les cas de référence sont mémorisés par les cliniciens dans la mémoire à long terme et structurés sous forme d'un répertoire des cas, au fur et à mesure de l'évolution du processus d'acquisition des connaissances [9,12, 17, 18, 19].

Dans ce papier nous nous contentons d'aborder le processus complexe de raisonnement hypothético-déductif *(Fig. 1)* qui nous intéresse dans notre travail. Ce dernier s'intègre dans le cadre des processus analytiques et reste comme le procédé le plus connu dans le domaine de la pédagogie médicale et le plus utilisé par les médecins expérimentés. Suivant cette démarche analytique, le médecin propose un ensemble d'hypothèses précoces, en se basant sur l'interaction préliminaire avec le patient. Le nombre de ces hypothèses est généralement compris entre 4 et 5 propositions. Après une phase de recherche de données supplémentaires, basée sur les hypothèses proposées dans l'étape précédente, une interprétation de ces données doit être effectuée afin de vérifier la compatibilité entre les données recherchées et l'hypothèse proposée. Cette vérification permet d'établir une évaluation de l'hypothèse qui permet de déterminer si elle doit être retenue, rejetée ou réévaluée. Il faut préciser qu'il est tout à fait possible que des nouvelles hypothèses soient générées et évaluées lors des itérations suivantes. Donc notre démarche itérative peut se dérouler à travers un cycle qui continuera à se dérouler tant que le diagnostic différentiel final n'est pas atteint [12, 4].

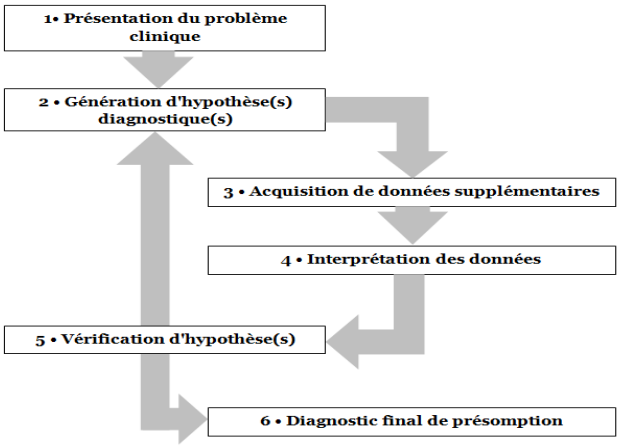


Fig. 1. Processus du raisonnement clinique hypothético-déductif [12]

Il faut prendre en considération que l'efficacité du processus de raisonnement dépend étroitement du type d'organisation des connaissances dans la mémoire à long terme du clinicien [12]. Dans la section suivante nous allons passer en revue les principales méthodes d'organisation de connaissances dédiées au domaine de la pédagogie médicale.

III. MÉCANISMES D'ORGANISATION DE CONNAISSANCES

Les recherches cognitives relatives à ce contexte ont démontré qu'il existe différentes méthodes d'organisation des connaissances au niveau de la mémoire des médecins [12]. Parmi les études les plus marquantes dans ce cadre on distingue celle de Bordage [12, 7], qui a tenté de déterminer les différentes méthodes d'organisation de connaissances médicales, à travers l'analyse des discours des cliniciens en situation de travail lors de la résolution de cas cliniques. C'est ainsi qu'il a proposé 4 types de discours, correspondent à 4 modèles d'organisation de connaissances dans la mémoire du clinicien *(Fig. 2)*:

- **Réduit :** ce discours peut être le résultat du manque d'informations approprié au problème, au niveau de la base de connaissances du clinicien. Il peut même correspondre aux cliniciens qui éprouvent des difficultés d'accès à leurs connaissances médicales.
- **Dispersé :** C'est le cas vraisemblablement de mémorisation au travers d'un apprentissage par cœur. Cependant, bien que la mémoire à long terme de ce type de clinicien peut contenir les informations pertinentes au problème rencontré, mais elles ne sont

pas structurées de façon à lui permettre de les réactiver et les mobiliser dans la résolution du problème.

- **Élaboré :** dans ce type de discours, on constate qu'il y a beaucoup de comparaisons et de contraction, ce qui reflète la bonne organisation de connaissances, qui se caractérisent par des liens riches et fonctionnels établis entre les données de la mémoire.
- **Compilé ou Encapsulé :** il représente le discours des cliniciens expérimentés, qui utilisent beaucoup plus le processus de raisonnement non-analytique. Ce type de médecins utilise quelques mots clés qui peuvent résumer tout le détail autour de la situation rencontrée.

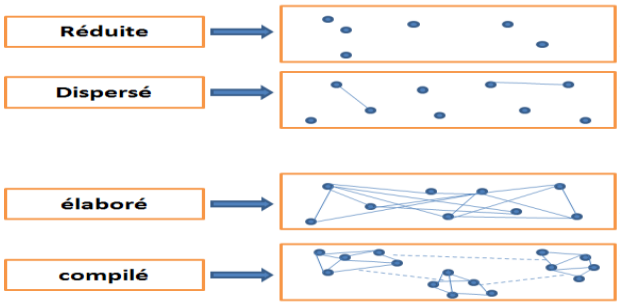


Fig. 2. Les quatre modèles d'organisation de connaissances dans la mémoire des cliniciens [15]

Ces différents mécanismes d'organisation des connaissances, nous permet de déduire que les cas cliniques rencontrés durant l'expérience médicale offrent aux médecins, ainsi que les futurs médecins, la possibilité de structurer leurs connaissances pour construire des réseaux de connaissances pertinents, qui peuvent être activés de façon automatique et inconsciente à chaque fois qu'ils sont confronté à une nouvelle situation clinique [15].

A partir de cette étude cognitive, on distingue qu'une méthode d'apprentissage du diagnostic médical devient pertinente dès lors qu'elle permet aux apprenants de surmonter facilement les défis relatifs au processus du raisonnement, à structurer leurs connaissances de manière efficace [14] et de collaborer avec les autres de façon similaire au contexte réel de leur profession.

En ce qui concerne notre modèle qui concerne la conception des environnements d'apprentissage collaboratif du diagnostic médical, il semble clair que nous devons adopter une démarche analytique afin de stimuler l'interaction et la négociation entre les apprenants, ainsi que pour favoriser et expliciter la collaboration et la co-construction des connaissances. Par conséquent, notre choix s'est porté sur la méthode analytique hypothético-déductif qui représente à notre sens, le procédé pédagogique et le plus convenable. En effet, cette méthode offre aux étudiants la possibilité de se familiariser naturellement avec le processus du raisonnement utilisé par les médecins expérimentés, dont la résolution collaborative du problème clinique permet à la fois, la construction et la réorganisation de connaissances. Elle permet même de générer des situations de conflits sociocognitifs qui étendent le processus d'apprentissage vers des activités sociales telles qu'elles se déroulent dans les milieux cliniques.

IV. LE MODÈLE D'APPRENTISSAGE COLLABORATIF DU DIAGNOSTIC MÉDICAL

L'objectif principal de ce papier est de décrire notre modèle d'apprentissage qui traite le processus d'acquisition de connaissances médicales en tant qu'activité collaborative et qui vise à mettre en valeur l'aspect de collaboration dans les environnements d'apprentissage du diagnostic médical. Cette proposition s'inspire de l'étude réalisée par Stahl [8], dont l'objectif est d'expliciter la démarche d'acquisition de connaissances qui est adoptée par les apprenants. De même que notre travail tient compte également des résultats des recherches relatives aux processus du raisonnement clinique, qui ont été menées dans le domaine de la pédagogie médicale.

L'étude effectuée par Stahl prend en considération le processus d'acquisition de connaissance dans le cas général. Cependant, dans notre travail nous avons mis l'accent sur le domaine médical, pour construire un modèle qui peut qui peut servir en tant que plateforme pour la conception des environnements d'apprentissage collaboratif du diagnostic médical. Concernent le domaine de la pédagogie médicale, les recherches associées aux processus du raisonnement clinique focalisent souvent leur attention autour de l'apprenant de manière individuelle. Dans ce contexte, nous ne distinguons pas pour l'instant de travaux qui ont considéré la modélisation des processus du raisonnement dans un contexte de groupe. En conséquence, notre tâche sera sans aucun doute difficile et longue en l'absence totale de cadre de référence par rapport au développement des stratégies d'apprentissage collaboratif, à la fois, en mode présentiel et à distance.

Notre proposition reflète le point de vue que l'on partage avec Stahl, qui considère l'apprentissage comme une activité qui dépasse l'échelle individuelle et la projette dans un cadre plus large à savoir, le procédé social de construction de connaissances, qui favorise l'interaction et la communication entre les apprenants. Selon Stahl, la notion de 'construction de connaissance' est plus significative que 'l'acquisition de connaissance' quand on parle de l'apprentissage collaboratif.

Du fait que l'acquisition et l'organisation de connaissances représentent deux processus conjoints dans le domaine médical, il semble indispensable d'intégrer ce dernier processus, à la fois, dans l'apprentissage individuel et social. Ceci implique que l'impact des activités sociales dans la formation médicale peut être étendu pour inclure la structuration de connaissance au niveau de la mémoire de l'apprenant. De plus, ces activités d'interaction sociales telles que la négociation, la confrontation et la contraction ne génèrent pas forcement de nouvelles connaissances, lors de la démarche d'apprentissage, alors qu'elles peuvent fournir

justement une opportunité idéale de renforcement et de structuration des anciennes connaissances.

Le modèle que nous proposons démontre que l'apprentissage du diagnostic médical repose en premier lieu sur un cycle d'activités individuelles d'acquisition et d'organisation de connaissances. Le résultat de ce cycle représente la matière première en termes de connaissances du processus d'apprentissage collaboratif. Ces connaissances acquises devront être exprimées et confrontées avec les connaissances des autres, dans le but de les raffiner et structurer d'avantage afin de parvenir à une solution collaborative d'un problème clinique spécifique.

et qui se passent de façon consciente dans les environnements de formation, avec des objectifs d'apprentissage prédéfinis. Une autre activité d'apprentissage informelle, dépend souvent de la pratique dans le milieu clinique. Cette activité d'apprentissage inconsciente repose toujours sur ce qui a été acquis durant l'activité formelle et représente la source primordiale d'acquisition de l'expertise médicale et d'affinement des certitudes personnelles. L'efficacité de la stratégie formelle est optimale au début du processus d'apprentissage où l'apprenant a besoin d'être motivé et orienté pour atteindre des objectifs prédéfinis. L'influence de ce type d'apprentissage va être atténuée, au fur et à mesure de l'évolution de l'expérience de l'apprenant. Ensuite l'apprentissage informel prend place naturellement jusqu'à ce qu'il devienne la méthode la plus utilisée comme c'est le cas par exemple, des médecins expérimentés [11, 3].

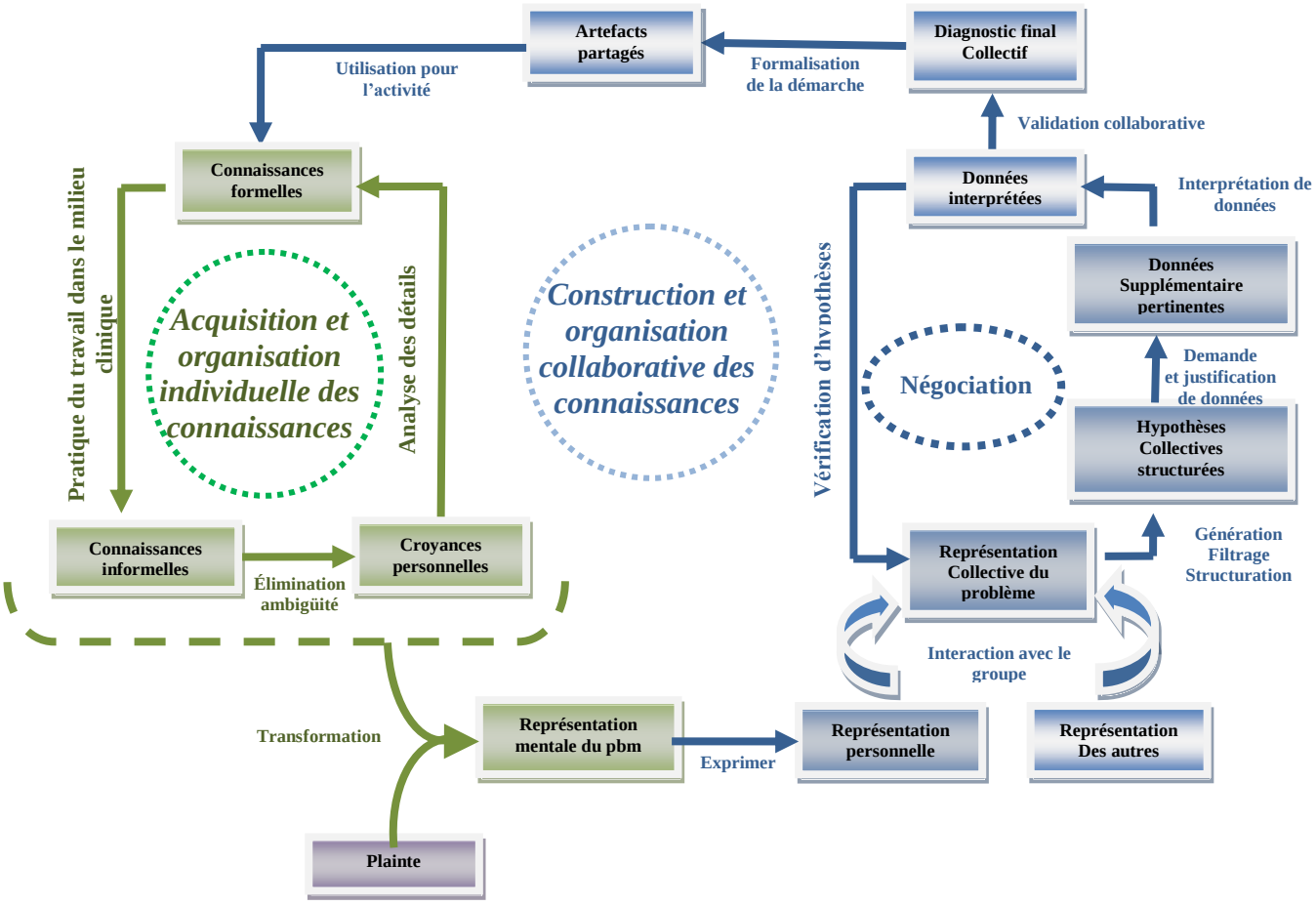


Fig. 3. Le modèle de l'apprentissage collaboratif du diagnostic médical

A. Processus d'apprentissage individuel

Le processus d'apprentissage individuel constitue une activité formelle basée sur l'acquisition et l'organisation des connaissances. Cette démarche est généralement issue des situations d'apprentissage classiques qui nous sont familières

B. Processus d'apprentissage collaboratif

Concernant l'aspect d'apprentissage collaboratif du diagnostic médical, notre modèle se base sur le processus hypothético-déductif qui est considéré comme la méthode pédagogique la plus efficace dans le domaine de l'éducation

médicale. En effet, ce processus permet de modéliser la démarché du raisonnement clinique utilisée par les médecins expérimentés au niveau individuel, alors que dans notre travail nous avons essayé de réfléchir afin de l'adapter aux situations d'apprentissage collaboratif à distance qui se base sur l'interaction synchrone. Cette adaptation vise à favoriser l'aspect social, dans les environnements d'apprentissage, et qui réside dans les activités d'interaction, de négociation et de confrontation des différents points de vue des apprenants. Ainsi, elle permet d'expliciter le raisonnement collaboratif et la co-construction de connaissances. Sans oublier qu'elle offre aux apprenants la possibilité de maitriser le processus de raisonnement le plus utilisé dans le milieu clinique.

Dans ce scenario d'apprentissage collaboratif, un problème clinique doit être présenté aux membres du groupe, sous forme d'une plainte annoncée spontanément par un patient. Cette plainte est préparée par le tuteur qui va superviser la démarche de résolution du problème, en tenant compte des objectifs d'apprentissage qu'il projette d'atteindre, ainsi que le degré de complexité du problème qui doit être adapté au niveau des apprenants et la durée de la séance.

1) Représentation du problème clinique

Il est largement reconnu dans le domaine médical, que l'utilisation directe de la plainte du patient dans la résolution du problème clinique a une influence négative sur la démarche du raisonnement ainsi sur la qualité du diagnostic proposé. Ce qui requiert de donner un sens médical à cette plainte et de générer ce qu'on appelle axes sémantiques (homme/femme, unilatérale/bilatérale, aigue/chronique) dont les données médicales sont comparées et contractées. Cette transformation sémantique (représentation mentale du problème) permet d'optimiser la recherche autour des hypothèses pertinentes au niveau de la mémoire du clinicien. En effet, des études dans ce contexte ont démontré qu'il y a une relation entre le taux de transformation des données de la plainte et la justesse du diagnostic. Tel que les médecins qui ont élaboré le bon diagnostic sont caractérisés par un taux élevé de transformation sémantique des données du patient [12]. Par rapport aux médecins expérimentés, les novices éprouvent souvent des difficultés pour construire une représentation mentale pertinente du cas rencontré. Ce qui conduit à une génération aléatoire d'hypothèses et une recherche incertaine sur les données clinique. Par conséquent, il semble très importent de fixer la représentation du problème comme l'un des objectifs clés de l'apprentissage dans le milieu clinique [12, 10]. Dans notre stratégie d'apprentissage collaboratif, la discussion des différentes représentations mentales générées par les apprenants peut améliorer cette compétence chez l'apprenant. De plus, elle permet d'aboutir à un consensus sur la représentation la plus pertinente qui constituera ainsi la base de départ à partir duquel le groupe entame la résolution collaborative du problème clinique.

2) Génération, filtrage et structuration d'hypothèses

La génération précoce d'hypothèses, dans notre modèle d'apprentissage, permet de structurer le problème clinique et d'en limiter le nombre de propositions associées. Cette sélection d'hypothèses offre aux apprenants la possibilité d'optimiser l'utilisation de leur mémoire de travail[1] qui ne risque pas d'être surchargée avec beaucoup de détails inutiles, et qui peuvent influer sur le degré de concentration des apprenants [12]. En outre, dans la démarche d'apprentissage collaboratif, il est fortement susceptible d'avoir une génération d'hypothèses qui ne sont pas convenables avec le problème en cours, d'où le besoin de les filtrer dès le début avant de passer aux étapes suivantes. Ce qui permet de centrer l'attention des apprenants exclusivement sur les hypothèses les plus probables, qui peuvent améliorer la qualité de la démarche de résolution du problème.

Dans un processus d'apprentissage du diagnostic médical, il est toujours recommandé de structurer les hypothèses en partant de la plus générale jusqu'à la plus spécifique. De même qu'il convient de ne traiter qu'une hypothèse à la fois en commençant par la plus générale [6]. Tandis que cette structuration en mode collaboratif favorise l'activation des anciennes connaissances, ainsi que la réorganisation des pathologies au niveau de la mémoire de l'apprenant afin que son activation face à une nouvelle situation soit simple et rapide [12]. Cette structuration peut être effectuée selon une approche probabiliste, à travers laquelle les apprenants doivent affecter à chaque hypothèse une valeur qui représente la probabilité qui la caractérise par rapport aux signes et symptôme déterminés. Dans notre modèle d'apprentissage collaboratif, le principe de probabilité permet aux apprenants de discuter les relations existantes entre les signes et les symptômes ainsi les maladies qui leurs sont associées, dans la mesure où nous favorisons la communication entre eux.

3) Demande et justification de données supplémentaires

Après avoir bien structuré les hypothèses générées, la recherche collaborative des données complémentaires représente une étape primordiale dans notre modèle. Ces données regroupent les examens cliniques, paracliniques et les questions qu'il faudrait poser au patient afin de confirmer ou infirmer les hypothèses retenues. En raison des coûts élevés de certains tests, parfois on est contraint d'ignorer l'ordre d'hypothèses lors du traitement en attribuant la priorité à celles dont leurs coûts sont raisonnables [6]. Dans le cas de la collaboration de notre modèle, les apprenants doivent justifier le besoin des données supplémentaires demandées, ce qui permet d'aboutir à un consensus sur les tests qui sont effectivement nécessaires pour la vérification des hypothèses. Ainsi on peut éviter une recherche exhaustive des données, qui peut influer sur l'efficacité de la démarche du raisonnement collaboratif [12].

Contrairement aux séances d'apprentissage du raisonnement clinique où le rôle du patient est toujours affecté à un apprenant du groupe, dans notre scenario d'apprentissage collaboratif on préfère attribuer ce rôle au tuteur qui est considéré comme un médecin expérimenté. Son expérience dans le milieu clinique lui permet de jouer parfaitement ce rôle, et répondre aux interrogations des apprenants comme un

[1] La mémoire de travail est la capacité de retenir des informations à court terme, quelques secondes ou quelques minutes, pour exécuter mentalement les opérations associées.

véritable patient. Ceci leur permet d'améliorer leurs compétences d'interaction avec les patients et améliore de façon générale le contact social. De plus le tuteur veillera sans aucun doute à retenir l'attention de tous les membres du groupe équitablement durant la résolution du cas posé.

4) Interprétation de données

Lors de l'interprétation de données associées à une hypothèse spécifique comme l'anémie par exemple, les apprenants analysent la compatibilité entre cette hypothèse et le taux d'hémoglobine dans le sang qui doit être inferieur à 130 grammes par litre chez l'homme, à 120 g/l chez la femme en dehors de la grossesse et 115 g/l chez la femme enceinte. L'interprétation, qui se base sur la communication entre les apprenants du groupe, offre la possibilité d'activer ce genre de connaissances, et de les réorganiser en établissant le lien entre ces connaissances théoriques et celles qui sont issues de cas réels cliniques.

5) Validation du diagnostic final

Les hypothèses vérifiées en mode collaboratif ont plus d'opportunités d'être évaluées convenablement. Cette vérification se base sur les différents points de vue des apprenants. Elle exploite les résultats de l'interprétation précédente afin de déterminer si l'hypothèse en cours de traitement doit être retenue ou rejetée. De nouvelles hypothèses peuvent être progressivement générées et avoir un impact sur la structuration des hypothèses et la représentation collaborative du problème. En se focalisant sur une négociation coordonnée par l'environnement d'apprentissage collaboratif, notre démarche itérative peut se répéter plusieurs fois jusqu'à se que les apprenants parviennent à un consensus sur le diagnostic différentiel le plus approprié. La formalisation de ce processus itératif (sous forme d'un texte, tableau, carte conceptuelle, etc.) permet d'avoir une structure simple et lisible, qui résume l'essentiel de notre démarche collaborative. De plus, elle constitue un terrain idéal pour la réorganisation et l'activation de connaissances.

V. CONCLUSION

Dans ce papier, nous avons longuement discuté notre modèle d'apprentissage qui dédié au support du processus d'acquisition de connaissances et favorisant l'amélioration des compétences dans le milieu clinique. Le modèle est dédie en premier lieu à la conception des environnements d'apprentissage collaboratif du diagnostic médical, en mode synchrone. Cette proposition se base sur l'étude accomplie par Stahl, qui interprète la démarche de l'apprentissage comme un procédé qui dépasse l'échelle individuelle et évolue vers un processus collaboratif de construction sociale des connaissances. Afin d'avoir un modèle d'apprentissage pertinent, nous avons essayé d'exploiter les études cognitives réalisées dans le domaine de la pédagogie médicale qui ont tenté de cerner les méthodes du raisonnement clinique, ainsi les facteurs qui peuvent les influencer. Cependant, au stade actuel de l'évolution des recherches, il subsiste encore beaucoup de zones d'ombres qu'il convient de clarifier.

Notre modeste démarche a quand même le mérite de poser la question pertinente de recourir aux outils collaboratifs qui ont fait leur preuve afin de construire des environnements dédiés au support de l'apprentissage collaboratif du diagnostic médical. Notre modèle qui est loin d'être complet requiert également beaucoup d'efforts pour l'affiner d'avantage et l'expérimenter en milieu réel. Cette démarche nous permettra certainement de mesurer objectivement la portée des choix qui ont été adoptés.

Concernant, l'étape suivante de notre travail de recherche, nous comptons expérimenter les prototypes que nous avons développés avec un groupe d'étudiants en faculté de médecine. Ce travail, nous permettra de collecter les enseignements nécessaires pour travailler en profondeur sur la question pertinente de modélisation de l'interaction entre les apprenants.

RÉFÉRENCES

[1] Bernard Charlin, Georges Bordage, et Cees Van Der Vmeuten '' L'évaluation du raisonnement clinique'' 2003.

[2] Céline Joiron '' Une contribution aux systèmes supports de Formation Médicale Continue à distance et d'apprentissage entre pairs: conception et expérimentation du forum DIACOM (Discussions Interactives à base de Cas pour la formation Médicale)'' 2002.

[3] Clark Quinn ''Social Networking: bridging formal and informal learning'' 2009.

[4] D. vanpee, J.B. gillet, et V. Godin '' Séance d'apprentissage au raisonnement clinique: une méthode potentiellement intéressant pour l'enseignement de la médecine aiguë'' 2002.

[5] E. Medelez Ortega , Y Lessard, A Burgun, P Le Beux '' Virtu@l Consult@tion: an Interactive and Multimedia Environment for Remote Clinical Reasoning Learning in Cardiology'' 2005.

[6] Froduald Kabanza, Guy Bissonb, Annabelle Charneaua, et Taek-Sueng Jang '' Implementing tutoring strategies into a patient simulator for clinical reasoning learning'' 2006.

[7] G. Bordage '' La prise de décision en médecine : quelques mécanismes mentaux et des conseils pratiques'' 2005.

[8] Gerry Stahl '' A Model of Collaborative Knowledge-Building'' 2000.

[9] Kevin W. EVA '' Ce que tout enseignant devrait savoir concernant le raisonnement clinique'' 2005.

[10] Malcolm Cox, M.D., and David M. Irby, Ph.D., Editors '' Educational Strategies to Promote Clinical Diagnostic Reasoning'' 2006.

[11] Mangenot François ''Apprentissages formels et informels, autonomie et guidage'' 2011.

[12] M. Nendaz, Bernard Charlin, Vicki Leblanc et Georges Bordage '' Le raisonnement clinique: données issues de la recherche et implications pour l'enseignement'' 2005.

[13] M. Nendaz '' Éducation et qualité de la décision médicale : un lien fondé sur des preuves existe-t-il ?'' 2010.

[14] Martine Chamberland '' Les séances d'apprentissage du raisonnement clinique (ARC): description de la méthode pédagogique '' 2007.

[15] Peter H. Harasym, Tsuen-Chiuan Tsai, et Payman Hemmati '' current trends in developing medical students' critical thinking abilities'' 2006.

[16] Pierre Pottier, Bernard Planchon, Jacques Henri Barrier, Christelle Volteau et Alain Mouzard '' Étude d'impact de séances d'entraînement à des exercices de dossiers cliniques simulés sur les performances d'étudiants à un examen national'' 2006.

[17] P. Pottier, B. Planchon '' Les activités mentales au cours du raisonnement médical diagnostique'' 2010.

[18] T. Pelaccia · J. Tardif · E. Triby · C. Ammirati · C. Betrand · B. Charlin '' Comment les médecins raisonnent-ils pour poser des diagnostics et prendre des décisions thérapeutiques ? Les enjeux en médecine d'urgence'' 2010.

[19] Vincent Paul '' étude du raisonnement du clinicien expérimenté et de l'étudiant : apports de la psychologie cognitive'' 2009.

PBST*-based Hierarchical Protocol for Wireless Sensors Networks

Nassima Merabtine
Computer Science Higher School
ESI, ex. INI
Algiers, Algeria
n_merabtine@esi.dz

Djamel Eddine Zegour
Computer Science Higher School
ESI, ex. INI
Algiers, Algeria
d_zegour@esi.dz

Djamel Djenouri
CERIST Research Center
Algiers, Algeria
ddjenouri@acm.org

***Abstract*—Managing a large and dense deployment is a challenging issue in energy-constrained Wireless Sensor Networks (WSN). Hierarchical Organization of WSNs has proven its efficiency in ensuring scalability and enhancing network lifespan. The work described in this paper presents an original hierarchical protocol based on the Partitioned Binary Search Tree (PBST*). PBST* is a new tree-based scalable distributed data structure. The principal feature of our protocol is the organization of WSN as a balanced tree. This means that all nodes can achieve base station on log(N) hops, where N is the number of nodes in the network. By this way, we aim achieving longer network lifetime and enhancing transmission delay.**

Keywords— Hierarchical protocols, Routing, SDDS, PBST*, Wireless Sensor Networks, Network lifetime.

I. INTRODUCTION

Currently, a large variety of applications based on environmental parameters have resort to WSNs including medical [1], industrial [2] and military applications [3]. WSN can be defined as a collection of small and inexpensive devices called micro-sensors, having the ability of sensing physical parameters. Micro-sensors are also capable of processing collected data and communicate it using wireless links. Constrained resources, especially in energy supply, is the main feature of WSN.

To achieve their ultimate goals, several WSN's applications require deployment of large number of micro-sensors. Managing such large deployment is a scalability issue, which can be defined as the network capability to cope to large scale deployment. That is, ensuring good functioning while maintaining network resources. Indeed, a large number of nodes means a significant amount of messages circulating in the network, which rapidly hampers network performance and depletes the node-batteries. Further, deploying dense and large networks complicates some native tasks such as routing.

As response to these issues, clustering protocols have been proposed. Clustering consists on dividing the network into a number of groups called clusters. In each one, a particular node is designated as cluster-head (CH). CHs collect data from their members and forward it to base station (BS). Further, CHs may perform other tasks such as aggregating collected data or scheduling media access for their cluster members. The key advantage of hierarchical protocols is localizing messages transmission within clusters, and between CHs and BS. This has a good repercussion on optimizing energy dissipation and hence extending network lifespan. Facilitating routing process is another benefit of network clustering. In fact, in a hierarchical organization of the network, each node transmits collected data to its CH. Only CHs relay packets to the BS using either direct or multi-hops communication.

Clustering protocols which consider multi-hops communication between CHs and BS are called tree-based protocols. This category has the advantage of achieving longer lifetime regarding to direct-communication based clustering algorithms [4]. In fact, enhancing energy dissipation is warranted by the fact that CHs will communicate data with some other selected nearest CHs, instead of spending a big amount of energy when using direct communication. Moreover, intermediate CHs may do some data aggregation before transmitting packet for longer lifespan.

Nevertheless, organizing nodes as a random tree will have negative impact on routing process. In fact, some tree branches will be very long. That is, long routing paths. This means more energy consumed and bad end-to-end delay, which implies the importance of organizing the network as a balanced tree. Some existing hierarchical protocols deal with this issue such as [5].

This work proposes a new protocol for organizing network as a balanced tree. Our proposition is based on PBST*, which is a novel scalable distributed data structure designed specifically to store and access large amount of data in distributed systems.

The remainder of this paper is organized as follows. Second section resumes some background and state-of-the-art protocols of both SDDS and clustering, with focus on PBST*. Third section presents our proposed solution. Fourth section discusses the advantages of our protocol. Fifth section concludes the paper and outlines future research directions.

II. BACKGROUNDS AND RELATED WORKS

*1. SDDS and PBST**

Scalable Distributed Data Structure (SDDS) is a new generation of data structures proposed specifically for distributed systems. The motivations behind its proposition are:

- overcoming some weaknesses in classical data structure, such as centralization of address calculation;
- Taking advantage of benefits offered by distributed systems, e.g. accumulated storage capability, possibility of parallel and distributed process, etc.

The principal asset of this data structures is scalability. In fact, SDDS are capable to manage large amounts of data while maintaining almost steady performances. We refer here by performances to data access time and number of control messages exchanged to maintain the SDDS file.

In SDDS, data are distributed over special nodes called servers, and accessing by autonomous clients. Initially, the SDDS file begins on one server, where all operations are carried out. Other servers will be designed when the file size exceeds the initial server capacity. In general, the number of servers dynamically grows and retracts with the file size.

SDDS model is characterized by the absence of any main directory. Therefore, each client should keep an image of how data is distributed. To preserve network resources, this image is not updated synchronously. Thus, addressing errors may occur, i.e. clients may contact a server which does not contain the required data. Nevertheless, the principles of SDDS guarantee the forwarding of the incorrect request to the appropriate server. This is by the way of some information called server indexes. Hence, an IAM "Image Adjustment Message" is sent to the client who makes the error, to correct its image.

SDDS can be classified into two main categories, hashing-based, LH* [6] and trees-based, RP* [7]. SDDS based on trees have the advantage of maintaining the file order.

Several SDDS schemes have been proposed, with different purposes. Some deal with the high-availability for supporting server failures [8], other schemes favorite the k-d partitioning [9], etc.

Partitioned Binary Search Tree (PBST*) is a tree-based SDDS [10]. This method respects the basics axioms of the SDDS along with some specifications. In fact, file evolution is the same as described in the class RP*. That is, clients realize their operations on the initial server, qualified as root. Every server contains a set of records organized in a balanced binary search tree and an interval] α, β [which defines the range of keys stored in the tree. Images of clients are also organized as binary search trees, where each node is a triple (a, α, β), such as, 'a' is the server address already referenced by this client,]α, β[is the range of keys which belong to the server S. by this manner, clients can autonomously access servers.

PBST* inherits SDDS's benefits such as data distribution over server and parallel treatment possibility. Moreover, since PBST* belongs to RP* category, it permits order-preserving which promotes range queries and ordered traversals.

The novelties of this method regarding to the RP* category are summarized below:

A) Internal organization

PBST* proposes binary search trees for storing data in servers as well as clients images and server indexes. This further improves data access time.

B) File evolution

Like any SDDS, data is stored in servers. Initially on root server with an interval]-inf, +inf[.This interval is the range of keys of data stored in this server. In each server, a threshold is predefined called partitioning parameter. As soon as a server overloads (reaches this threshold) due to data inserts, a splitting will take place. If the overloaded server is the root, then two new servers are allocated. One of them receives the right sub-tree with the interval]-inf, median-Key[and the other receives the left one, with]median-Key, +inf [. The median record is kept in root node. In case of non-root server saturation, only one server is created which receives the right sub-tree of saturated server and the range value]median-Key, b[. The median record rises to the splitting server father's. This can cause cascading splitting.

C) Maintenance operations

To maintain load balancing between servers, i.e. load factor, PBST* provides two operations: fusion and co-balancing. These named operations occur between two brother servers (having the same parent in the server tree). To do this, two new parameters are defined: minimum threshold "MIN" and intermediate threshold "INTER". These parameters are checked whenever a modification (insert or delete) is occurred on a server. If the load factor of that server reaches the MIN, it consults one of its brothers. Two cases can happen. If the brother load factor is greater than INTER, some data of this server will be transferred to the first one. This operation is called co-balancing. The second case occurs when the brother load factor is lower than INTER. Here, a fusion operation takes place. That is, the server reaching the MIN is released after sending its tree to his brother server.

D) Search protocol

Clients can execute different kind of requests on PBST* file, namely, insertion, search, deletion, etc. All these requests begin by a search in client image for the server address which might execute this request. As seen before, client image may be incorrect or incomplete. Therefore, each server receives a request verifies if it is the appropriate server to execute the request. If not, the request is redirected to the parent server. This latter repeats the same verification. If it is the appropriate server, it first searches the desired data in its local data tree. If the data don't belong to its tree, then the request is redirected to one of its son servers according to the interval. This procedure is repeated until reaching the appropriate or a leaf server. The querying process ends either by executing the request on that server and returning the result, if any, to the client, or by concluding that desired data don't exist (in case of search request).

It is to note that IAM messages are sent throughout all redirections path, from the first server receiving the request to the correct server.

2. *Clustering protocols*

Due to their efficiency in ensuring scalability and prolonging WSN lifetime, clustering protocols have been a hot research area for years. This leads to the proposition of numerous solutions [5; 13–23]. Main differences between those solutions can be resumed on:

A. How they deal with some design decisions?

In general, clustering process consists of two separated phases: head election and clusters formation. The first design decision is how to implement each phase. For the head election process, some protocols designate the nodes with the biggest ID as CHs, or node with greatest connectivity degree like in [11]. These approaches have the facilitating advantage but they lack efficiency. Other protocols use more adaptive criteria such as energy, or even a combination of criteria like [12].

Another design decision is the type of communication used for transmitting data between nodes and their CH (intra-cluster) and from CHs to BS (inter-cluster). Direct communication is simple but can cause the worst of big amount of energy, especially for remote node, since energy consumed in communication is proportional to the distance. Multi-hop communication is more practical to preserve energy and permit some aggregations, but add the overhead of paths creation and maintenance. Example of clustering protocols using intra-cluster and inter-cluster direct communication and multi-hop one are: [13], [14], [15] and [16], respectively.

Whether base station manages the clustering algorithm or not is another choice. In some centralized protocols [17], BS is the only responsible of CH election and clusters formation when other algorithms are fully distributed [18]. In the second case, a node decides to join a cluster or becomes CH based on its local view (adjacency information.). Centralized algorithms have the privilege of electing optimal and well distributed CH since they have a complete view of topology. However, this is at the cost of a high complexity and convergence time contrary to distributed ones which are fast but may have a non-optimal CH choice.

B. Which clustering challenge they treat?

Clustering algorithms are faced to many challenging issues which should be carefully treated to maintain clustering advantages. Load balancing is one of the important challenges. In fact, because CHs have extra charge comparing to the other simple nodes in the network, they will exhaust their energy more quickly. Furthermore, if the number of nodes assigned to CHs is not balanced, the high rate CHs are likely to die faster and prematurely making end to the network lifetime. Another issue is clusters stability, which reflects the ability of clustering protocol to tolerate network dynamicity caused by node failure, mobility, etc. Quality of service is also an important issue to deal with it. Indeed, many clustering protocols consider duty-cycle to improve energy consumption. The major drawback of duty-cycle based approaches is the delay introduced by the extra phase required to synchronize senders and receivers active periods. This latency is not tolerated in some real time applications.

Because all challenges cannot be achieved together, each proposed solution treats a subset of challenges. For instance, [19] and [20] are clustering protocols which consider load balancing issue. LEACH [13] has proposed the mechanism of rotating the role of CH between cluster members. Cluster stability is treated in [21]. BOD-LEACH [22] proposes a solution to broadcast delay-sensitive information in LEACH.

In the following, we resume some works from the state of the art. We focus on tree-based protocols since our proposition belongs to this category.

The work in [23] proposes a new clustering protocol for hierarchical control in multi-hop ad hoc and sensor networks. Two versions of the protocol are presented. The first one is a centralized version which considers unit or bounded disk graphs. This version is used to formally validate some protocol properties, such as correctness. The distributed version has two phases: cluster creation and cluster maintenance. The first phase begins by creating a spanning tree using a distributed BFS algorithm. Clusters with size comprises between K and 2K are then created over that tree. Maintenance phase has the aim of guaranteeing a good quality of clustering in the presence of network dynamicity. The main contribution of this work is the exploitation of some geometric properties of wireless networks to perform clustering with a set of desired properties such as: designing connected clusters, imposing clusters size to be between K and 2K, etc.

Another tree-based clustering protocol (TBC) is presented in [24]. Enhancing network lifetime comparing to some state-of-the-art clustering protocols, which are: LEACH, PEGASIS and TREEPSI, is the contribution of this work. The authors argue that this advantage is obtained by replacing intra-cluster direct communication, used in LEACH and PEGASIS, by the multi-hop one. Regarding to TREEPSI, TBC reduce the routing path computation load. The operation of TPC can be divided into two phases, CHs election and intra-cluster tree construction. CHs are elected by the same manner as LEACH [13]. The second phase aims to construct a tree inside each cluster, with CH as root. For this, CHs broadcast an invitation message and then receive a Join_REQ from each node that decides to join this CH' cluster. This message includes the node's ID and location information, which is the basis for determining the tree level of member node. Finally, CHs define a TDMA schedule which should be respected by each member.

SOLC (Sink-Oriented Layered Clustering) [5] focalizes on load balancing challenge in tree-based clustering to maximize the network lifetime. In fact, in tree-based network organization, CHs closer to the BS will have the responsibility of relaying data for those farther away so they tend to die faster. To face this problem, SOLC divides the sensor field

into concentric rings around the BS. Initially, Optimal Rings' width and the number of clusters by ring are estimated. Then, intra-ring clustering and inter-ring routing are performed. Clustering consists of two steps, CHs election which is based on an energy criterion and clusters formation. In this second step, each CH broadcasts an invitation request. Simple nodes may receive many invitation requests; they choose the closest CH according to invitation message received signal strength. After clustering, each CH broadcasts a message to advertise itself as a candidate relay for its downstream CHs. This is to ensure a multi-hop routing to BS. SOLC guarantees that a data packet is transmitted to the sink in at most R/r + 1 hops, where R is the radius of the sensor field and 'r' is rings width.

III. PROPOSED SOLUTION

A. Assumptions

We assume a large network of battery-powered stationary wireless sensor nodes uniformly distributed on a grid zone. This topology is supposed in many clustering approaches [3; 15]. All nodes are homogeneous and have the same capabilities. Each node is assigned a unique identifier (ID).

Initiator of the algorithm is supposed to be located in a network central region; it can be either sink or a simple node. Initiator will be the root of the tree and the first CH in the network.

B. Description

The proposed algorithm organizes the network's nodes as a balanced tree by following the general idea of PBST*. Initially, at network deployment, all nodes exchange HELLO messages between their one-hop neighbors. That is, each node will have a complete neighboring list. This information will be used later for selecting optimal and well distributed CHs. After this step, the initiator node advertises itself as the first CH in the network. This is by broadcasting an invitation message (INVIT) to its 1-hop neighbors. Each node receiving the INVIT message, for the first time, replies to the initiator by a JOIN message. The JOIN contains a sender ID and its neighboring list. At the receipt of a JOIN by the initiator, it first registers the sender as a member of its cluster and retrieves the neighboring list. After receiving JOIN messages from all neighbors, initiator compares the number of its attached nodes with the threshold. This latter is a predefined integer defining the number of attached nodes allowed by every CH. If the threshold is not reached yet, the initiator rebroadcasts the INVIT message, but this time, to its two hops neighbors. In general, initiator will rebroadcast the INVIT upon former number of hops, over which the INVIT has been propagated in the last round, plus one. This process should be repeated until attaining the threshold.

In this case, the initiator CH is declared as saturated. Consequently, its cluster is splitted into two clusters and two new CHs would be elected. Both CHs election and member's assignment are done by saturated-CH. For this, saturated-CH uses its local topological base constructed by the neighboring lists sent by cluster members.

The two nodes designated as CHs should satisfy two criteria:

1. Newer CHs are those having the max hop count from the saturated CH: This is because CHs will have the role of continuing tree construction. So, it must be guaranteed that these CHs have neighbors that are not member of a any cluster yet.
2. To ensure a well distribution of CHs over the network, the hop count between these two nodes should be the maximum possible.

In order to designate these two CHs, saturated CH calculates the hop count of shortest paths to each of its attached node, using the well known Djikstra algorithm. The nodes having the biggest values of hop count are then selected. Once again, hop count between each pair of previously selected nodes are determined by the saturated-CH. Finally, the pair (CHf, CHs) of nodes with the max value of hop count are elected as CHs. Due to grid topology and algorithm logic, several pairs of nodes will meet this last criterion. Saturated CH will choose one of them arbitrary.

Afterwards, each attached node to the saturated CH must be assigned to one of the two new CHs. Following the principle of PBST*, each new CH should have threshold /2 attached nodes. We add a second criterion wich is: nodes must be assigned to the nearest CH in term of hop count. These two criteria can be achieved together since we assumed a grid topology with uniform distribution of nodes. At worst, we can tolerate a shifting of one or two nodes instead of the perfect threshold /2 value. To realize this, we refer to Djikstra to calculate the shortest path (SP) between each node and the elected CHs. For K nodes, the results will be two sets of distances: {(di,CHf), …, (dk,CHf)} and {(di,CHs), …, (dk,CHs)}. (di, CHf) represents hop count of the SP from node i to the first elected CHf, while (di,CHs) is the SP from the same node to the second CH. These distances are used to assign nodes to CHs. The algorithm searches the pair with minimal distance from the two lists alternatively, starting with the first list. The pair choosed will be joined, i.e. when selecting the pair (di, CHf), i will be a member of CHf'f cluster.

At last, an information message "INFO" is broadcasted by the saturated CH over h hops, where h is the number of hops over which the INVIT message has been propagated. INFO message contains a set of couples (i, j) representing the node ID and its CH, respectively. By this way, all nodes participating in the previous phase will be aware whether they are elected as CH or to which CH are assigned.

At this level, the network will contain a tree with three CHs: the initiator as the tree root and two children, namely CHf and CHs.

The algorithm continues its execution by a bottom-up fashion, i.e. nodes closer to the initiator join clusters more early. To do this, both CHf and CHs repeat the same process described above, except for what follow: When a non-root CH is splitted, it loses its role as CH and becames a simple node. Further, one of the saturated-CH's members, let *M*, will join its parent. This is to guarantee the construction of balanced

tree. *M* must be the nearest node to the parent. Obviousely, the saturated CH will join one of the newer elected CHs.

As our solution is a tree-based clustering, the CHs should be logically interconnected; i.e. each CH knows its parent in the hierarchy. In our protocol this is done in spliting phases. We separate two cases: root-CH (initiator) splitting and non-root-CH spliting. In the first case, the new elected CHs will be the children of the root-CH. In case of non root-CH splitting, newer selected CHs will be the childern of the saturated-CH's parent. This information (the parent of each CH) is carried in the INFO message. Following the principals of PBST*, the number of childern of each parent in the tree is limited. That is, at the reception of INFO message, CHs designated as parents compare the number of their children with the threshold. If a parent-CH attains the threshold, a cascading splitting will occur. Cascading splitting consists of the election of two new index-CHs. CHs that are assigned to the saturated-parent-CH will be affected to these new index-CHs, i.e. they will be the children of these index-CHs in the hierarchy. CHs election and member affectation process are the same as described previously. CHs categorized as Index are different from the ordinary ones by the following: index-CHs don't broadcast the INVIT message; they only receive members at each splitting of their children CHs. An information message must be sent after each cascading splitting to the concerned CHs.

Further, next-hop of the SP between each node and its CH can be calculated locally, at each splitting, and piggybacked in the INFO message. This is to allow intra-cluster multihop transmission, without the need of an extra step for searching paths.

The algorithm execution ends when no splitting is triggered.

IV. ANALYSIS

The proposed protocol organizes nodes as PBST* tree, i.e. partionned and balanced tree. Data is transferred by each node to its CH. Every CH send its data to its CH parent until reaching BS. CHs may perform some aggegation before sending data. Some other algorithm properties are listed below:

Lemma 1: All nodes can achieve base station on log(N) hops.

Since the tree created in the network is balanced, all nodes can achieve base station on log(N) hops, where N is the number of nodes in the network. This leads to a more balancing energy consumption and hence longer network lifetime. This also permits the enhancement of transmission delay since all nodes will have the opportunity of sending their data on log(N*T), with T is the approximated channel transmission delay.

Lemma 2: For a network with N nodes, the proposed protocol consumes O(n) messages to perform clustering.

The proposed protocol uses four kind of messages: HELLO, INVIT, JOIN and INFO. The Hello message is brodcasted by each node one time at network deployement, that is, O(N) messages. The INFO message is broadcasted by only CHs over 'h' hops. 'h' is a constant propotional to the threshold. So, the comlexity of this message is O (number of CH * h). The complexity of the join message is O(N*h) since it is broadcasted by each node in the network once over 'h' hops. The complexity of INFO message is the same as INVIT message. The protocol complexity is the sum of complexities explicated above, thus, O(n).

Lemma 3: The proposed protocol creates a disjoint clusters.

Moreover, the logic of the proposed solution guarantees the construction of disjoint clusters, that is, each node is assigned to exactly one CH. This is because each node responses to the first invitation message only and discards the other invitation messages, if any. Only CHs will be interconnected in order to permit inter-cluster routing.

V. CONCLUSION AND FUTURE WORKS

A new tree-based clustering protocol for wireless sensor networks has been presented in this paper. The protocol is based on PBST* which is a scalable distributed data structure. The proposed protocol constructs multi-hops clusters with one node elected as CH. CHs are well distributed on the network. Each node is assigned to its closest CH in term of hop count. The hierarchy created in the network ensures log(N) hops transmission between each node and BS. This protocol is hybridization between centralized and distributed clustering protocols; therefore, it permits the election of optimal and well distributed CHs with a relatively low complexity.

This work is in the beginning stage and many perspectives are to be investigated. Simulation of the proposed solution through a network simulator and its comparison with some tree-based existing solutions are for capital importance. The second perspective is load balancing. Like clear, the clusters formed by the proposed solution have all approximately the same number of nodes. This is due to the predefined threshold. In multi-hop many-to-one traffic, forming clusters of equal size causes hot spot problem in the CHs nearest to the base station. To tackle this problem, the threshold value should be decreased after each splitting. Consequently, the number of CH's members will increase relatively the distance to the BS.

Additionally, how to fix the threshold is a crucial question to avoid the creation of large or very small clusters. Also, fixing well value of threshold will avoid the rebroadcasting of invitation message, which enhances protocol complexity.

For the best of our knowledge, this is the first work inspired from distributed scalable data structure to resolve clustering/routing problems in WSN. Also, this is the first hybridization between centralized and distributed clustering protocols.

REFERENCES

[1] Yan, H., Xu, Y., and Gidlund, M: " Experimental e-Health Applications in Wireless Sensor Networks". In Proceedings of the International Conference on Communications and Mobile Computing, 563-567. 2009.

[2] V. Gungor, G. Hancke, “Industrial wireless sensor networks: challenges, design principles and technical approaches”. IEEE Trans. on Industrial Elect., Vol.56, No. 10, pp 4258-4265, 2009.

[3] Milica Pejanović ðurišić, Zhilbert Tafa. Goran Dimić, Veljko Milutinović: “A Survey of Military Applications of Wireless Sensor Networks”. Mediterranean Conference on Embedded Computing (MECO). 2012.

[4] Farooq, M.O. ; Dogar, A.B. ; Shah, G.A. “MR-LEACH: Multi-hop Routing with Low Energy Adaptive Clustering Hierarchy”. Fourth International Conference on Sensor Technologies and Applications, Italy. 2010.

[5] Yijun Mo · BangWang · Wenyu Liu · Laurence T. Yang: "A Sink-Oriented Layered Clustering Protocol forWireless Sensor Networks". Mobile Networks and Applications, Volume 18, Issue 5, pp 639-650, Springer publisher. October 2013.

[6] W. Litwin, M.A. Neimat, d. Schneider: “LH*: A Scalable Distributed Data Structure”. Submitted for journal publ.1993.

[7] Witold Litwin, Marie-Anne Neimat, Donovan Schneider, “RP* : A Family of Order-Preserving Scalable Distributed Data Structures”. Proceedings of the 20th VLDB Conference Santiago, Chile, 1994.

[8] W. Litwin, M.A. Neimat: "High-availability LH* schemes with mirroring". IFCIS International First Conference on Cooperative Information Systems, pp. 196 – 205. 1996.

[9] W. Litwin, M.A. Neimat: "k-RP*S : A Scalable Distributed Data Structure for High-Performance Multi-Attribute Access". IEEE Intl. Conf. on Parallel & Distr. Inf. Systems. 1996.

[10] A. Chikhaoui, D.E Zegour, W.K Hidouci: “PBST*: une nouvelle variante pour les SDDS”. Rencontre sur la recherche (R2I), Tiziouzou, Algérie. Juin 2011.

[11] C. Wei, J. Yang, Y. Gao, and Z. Zhangn “Cluster-based Routing Protocols in Wireless Sensor Networks: A Survey. International Conference on Computer Science and Network Technology, China. 2011.

[12] Ameer Ahmed Abbasi, Mohamed Younis, “A survey on clustering algorithms for wireless sensor networks”. Computer Communications vol. 30, no. 14-15, pp. 2826–284. 2007.

[13] Wendi Rabiner Heinzelman, Anantha Chandrakasan, and Hari Balakrishnan: " Energy-Efficient Communication Protocol for Wireless Microsensor Networks". Proceedings of the 33rd Hawaii International Conference on System Sciences. 2000.

[14] A. Youssef, M. Younis, M. Youssef, A. Agrawala: "Distributed formation of overlapping multi-hop clusters in wireless sensor networks". In: Proceedings of the 49th Annual IEEE Global Communication Conference (Globecom’06), San Francisco, CA. November 2006.

[15] Mao Ye, Chengfa Li, Guihai Chen, Jie Wu: " EECS: An Energy Efficient Clustering Scheme in Wireless Sensor Networks". 24th IEEE International Conference on Performance, Computing, and Communications IPCCC. 2005.

[16] Vivek Katiyar, Narottam Chand, and Surender Soni: "Efficient Multilevel Clustering for Large-scale Heterogeneous Wireless Sensor Networks". International Conference, ICAC3, Mumbai, India, pp. 293-299. January 2011.

[17] E. Ilker Oyman, Cem Ersoy: "Multiple sink network design problem in large scale wireless sensor networks". In: Proceedings of the IEEE International Conference on Communications (ICC’04), Paris. June 2004.

[18] O. Younis, S. Fahmy, HEED: "A Hybrid, Energy-Efficient, Distributed clustering approach for Ad Hoc sensor networks”. IEEE Transactions on Mobile Computing, pp. 366–379. 2004.

[19] Chengfa Li, Mao Ye, Guihai Chen and Jie Wu: "An Energy-Efficient Unequal Clustering Mechanism for Wireless Sensor Networks ". IEEE International Conference on Mobile Adhoc and Sensor Systems. 2005.

[20] Longpeng Zheng, Zhiping Jia*, Ruihua Zhang, Hui Xia, Lei Ju, Chuanhao Qu: " Context-Aware Routing Algorithm for WSNs Based on Unequal Clustering”. 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications. July 2013.

[21] K. Xu, M. Gerla: "A heterogeneous routing protocol based on a new stable clustering scheme". In: Proceeding of IEEE Military Communications Conference (MILCOM), Anaheim, CA. October 2002.

[22] M. Khiati, D. Djenouri. BOD-LEACH: broadcasting over duty-cycled radio using LEACH clustering for delay/power efficient dissimilation in wireless sensor networks. International Journal of Communication Systems, Early view (Corrected Proof), John Wiley \& Sons Ltd, DOI: 10.1002/dac.2669. 2013.

[23] Suman Banerjee, Samir Khuller: "A Clustering Scheme for Hierarchical Control in Multi-hop Wireless Networks". Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Anchorage, AK.2001.

[24] Kyung Tae Kim, Chang Hoon Lyu, Sung Soo Moon and Hee Yong Youn: " Tree-Based Clustering (TBC) for Energy Efficient Wireless Sensor Networks". IEEE 24th International Conference on Advanced Information Networking and Applications Workshops. 2010.

A Fuzzy Cognitive Maps-based Collective Decision in Cooperative Information Systems

Maziz Asma
Mathematics & Computer Science Department
University Larbi Ben Mâhidi
Oun El Bouaghi, Algeria
maziz.asma@yahoo.com

Zarour Nacer eddine
Software Technologie & Information Systems Department
University Constantine 2
Constantine, Algeria
nasro.zarour@univ-constantine2.dz

***Abstract*— The problem of decision making is not evident in the single-firm supported by a traditional information system. It is becoming increasingly complex in cooperative information systems (CIS), all which are by their nature, distributed, autonomous and heterogeneous systems (inter-company). Few works approached this issue in spite of its importance. Our proposed approach is based on the concept of fuzzy cognitive maps (FCM) and the multi-agent paradigm that will express the strength of the causal relationship in a CIS. Each group of human experts is modeled by a group of agents where each one of them uses fuzzy cognitive map. This group will try to make a collective decision to solve a given distributed problem. To put this approach into practice, we have chosen the field of road safety to see how to reduce this crisis which grows more and more in the south of the Mediterranean.**

Keywords— Decision Support, Fuzzy Cognitive Maps, Multi-agent Systems, Cooperative Information Systems, Road Safety.

I. INTRODUCTION

Nowadays, the modern enterprise is characterized by the globalization of trade and the diversification of actors that made it. To persist in such an environment, the company must support a set of solid business partners and thus integrate an external structure that allows it not only to survive in its environment but also ensure sustainable development [1].

All companies generally take the form of a business network. Each actor who makes up this network supports the implementation of actions that benefit one party or the entire network. Interest increasingly growing for managing corporate networks is due to the constant development of information technology and computer networks such as the Internet. They allow, in fact, accelerate informational and decision-making processes and support decisions on more extensive data. These networks allow a more diversified choice of partners (suppliers and / or customers) of a normal business. Also, the need to rely on stable and competitive partners for a company today is a major strategic issue [1].

In support of a conventional information system, we should solve the problem by ensuring the aspects of reliability, coordination, control and decision-making. These aspects become more difficult to hang in charge when it comes to inter-company cooperation. In our context, we are interested in network of companies where the problems of decision making increase. The main difficulty of the decision lies in the fact that the company does not control all the specificities of different business partners. This implies the need to find ways of coordination that could limit the "harmful" effects, i.e. inconsistencies arising from the decision.

To improve and eliminate inconsistencies, the collective decision-making ensures consistency in cooperative information systems (CIS) with specific business partners. Our approach is based on the cognitive maps system. The effectiveness of such systems depends on the universe in which the decision maker or virtual human moves. In general, any decision maker uses a decision support system to facilitate the work of analyzing the problem. The main role of these systems is to complement, enhance and help the decision maker to reduce the time execution and error rate.

In the proposed approach we propose to associate each information system (IS) (whether individual, company, etc.) a cognitive map fuzzy type (FCM) which will be considered as a communication aid system, help in the analysis of ideas and also as a good support system for decision-making to solve a given problem. To model the CIS, the nearest area that contains most of the characteristics of a CIS is the Multi Agent Systems (MAS). So, the combination of CIS, MAS and cognitive mapping tool provides us the necessary elements for the relevant modeling of behaviors and interactions of human agents. In other words, we will model a CIS by MAS, where each agent has a fuzzy cognitive map in order to make the right decision and therefore solve the global problem.

This paper is organized as follows. Section II presents essential related works using CIS, MAS, and FCM. In Section III, we provide motivation and describe the proposed approach. Section IV shows how the proposed approach is implemented and illustrated by the road safety example. In section V we present some experimentation. Section VI concludes this work and gives some perspectives.

II. STATE OF THE ART

According to [2] "the cooperative information systems (CIS) is a complex system, instead of multiple interactions, open to many interconnected systems, distributed, heterogeneous, autonomous, and permanently interacting with its environment. The emergence of new knowledge is the result of the cooperation of its components". The proposed approach

is located in the intersection of three major areas that are CIS, MAS and FCM. In what follows, we will try to mention some work in each combination of these three domains.

A. CIS and MAS

Many works will combine these two paradigms (CIS and MAS). We cite for example Magali proposes the use of MAS in the CIS to resolve semantic conflicts from heterogeneous databases [3]. Almost in the same context Zarour et al propose to combine the CIS and MAS in a particular field of application, i.e, the medical information systems [4]. Another work [2] proposes architecture for CIS based on MAS to support home care. Cardon [5] suggests a model CIS applied to the management of emergency situations and civil crises. Kamoun proposes an agent-based architecture of a CIS in search multimodal routes [6]. In [7], authors describe a Multi-Agent Architecture for the implementation of the Semantic Interoperability in the CIS to meet the need of uniform and cooperative access to multiple sources of information.

B. FCM and MAS

The cognitive maps have proven their performance in many areas such as control systems, policy, supervision and management of companies. We use cognitive maps as a mechanism for gathering information on the evolution of emotions following an event, as well as a tool for decision making. A cognitive map is a model of the world (or system) which is in the form of a graph whose nodes represent concepts in the world, and arcs represent the causal relationships between these concepts. However, it is not a static representation of the world; a cognitive map is an "inferential" system in which calculations will be performed to assess the consequences of a state of the system (making predictions) [2]. An FCM is a dynamic modeling tool in which the resolution of the system representation can be increased by applying further mapping [8,9].

The FCM were used for different modeling. For example in [10], authors used to represent the FCMs supervision activity power plant. The contribution in [11] was used the FCM in the context of trade. On the other hand, studies the origin of the use of FCMs, [12] [13] [14] [15] have demonstrated a new approach: the use of FCMs for modeling actors in a virtual world. In [16] author proposes the use of causal maps in a multi-agent environment as an alternative to represent knowledge in a qualitative manner and to manage decision making of an agent. Acampora and al introduced a novel methodology of AmI systems design that exploits multi-agent paradigm and a novel extension of FCMs theory benefiting on the theory of Timed Automata. This methodology allows creating a collection of dynamical intelligent agents that use cognitive computing to define actionsô patterns able to maximize environmental parameters as, for instance, userôs comfort or energy saving [17]. Rodin and al developed a fuzzy influence graph to model cell behavior in systems biology through the intracellular biochemical pathway [18]. Stula et al developed an agent-based fuzzy cognitive map (ABFCM) for injecting the concept of multi-agent system (MAS) into the FCM and the different inference algorithms in each node enabled the simulation of systems with diverse behavior concepts [19].

Throughout the period of the state of the art study, we do not find works that combine the three areas CIS, MAS and FCM.

III. An Approch To Integrate FCM into CIS

The cooperative information systems (CIS) support collaborative work between different types of entities that are individuals, services or businesses. This work will enable cooperation based on collective problem solving (joint response to a tender, innovation project, etc.).

The problem of decision making is not apparent in the single-firm. It is becoming increasingly complex in cooperative information systems (CIS). Among the problems in need of rapid decision-making, is the case of military, emergency medical, diagnostic of industrial installations, meteorology, etc.

To improve the decision making process and ensure consistency of decisions taken in the CIS, a lot of systems for decision support are in charge. There are two types, quantitative or qualitative. Since the reasoning of the human being is more qualitative than quantitative, we proposed to use the system of qualitative type that called cognitive map. This support system has many characteristics, for example it is easy to use and to build, flexible in performance, low time to perform, transparent to non-specialists, deals with complex issues related to trigger and knowledge management, etc.

[20] proposes a multi-agent model for conception of collective decision support system. In other words, he is interested to decisions taken in a collective framework. We are inspired from this work to apply it in the context of CIS.

In CIS, the information system (IS) can act alone to make decisions, communicate, be autonomous, be proactive, negotiate, etc. To represent an IS under the machine, it must be transformed into its language. These characteristics can be found in the concept of agent. The field of MAS is a very good domain to modeling, it offer many advantages for their design and development. So we model each IS by an agent.

In our work, we associate each agent representing an IS a fuzzy cognitive map that will describe his view and his decision to solve a global problem. The manipulation of information between the FCMs of agents in MAS requires the use of union and / or intersection operations. According to our approach we will only use the union operation. This operation used when it is necessary to gather in single card the information from several cognitive maps drawn by different experts on the same domain.

On the one hand, Zhang [21] specifies that to achieve composition (union) of cognitive maps, we connect pairs of identical concepts of each FCM with bidirectional links whose weight is (+1). It also establishes that the decomposition of these cards can be achieved by changing the link weights of 1 to 0. Kosko [22] proposes the union between fuzzy cognitive maps by combining the adjacency matrices which denote E1... En. This is achieved by adding the matrices increased of cards A1... An with the traditional addition of matrix. The condition to use the union between two cards is that they contain at least

one concept in common, which allows us to make the connection between the two cards.

We choose the methodology of Kosko to make the union between the fuzzy cognitive maps unless we make the union in an incremental manner, in other words, with only two cognitive maps, more specifically, with the adjacency matrix of two maps E1 and E2. We limit ourselves to two cards at a time to use the rules of union **[24].** So from two adjacency matrices E1 and E2 we get the augmented matrices A1 and A2. Each Ai increased matrix with **n** rows and **n** columns, **n** being the total number of concepts used by the various agents (IS). Then, we apply the union operation of A1 and A2. The latter is performed at each pair of common concepts between the two cards, taking the maximum value between them **[23].**
After we construct the MAS, where each agent has a FCM and a degree of expertise that reflects his experience on the field. The resulting cognitive map is obtained by gradually incrementing the following way (see Figure.1):"Partial integrations (using union operation) are made between each pair of agents selected randomly then the result of the partial integration is assigned to the agent with the highest level of expertise. "

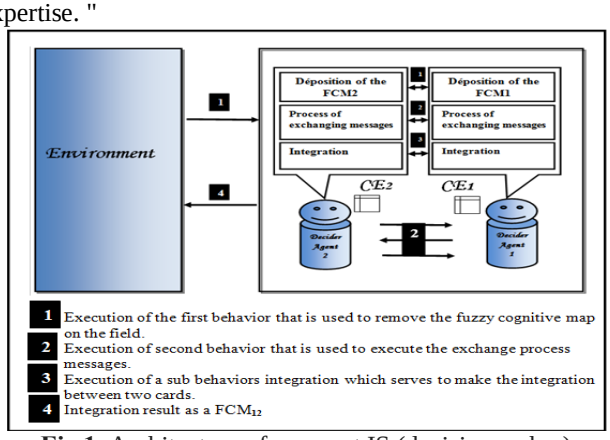


Fig.1. Architecture of an agent IS (decision maker)

After discussing of the various agents (IS) on the resulting FCM. The latter is automatically associated to the supervisor agent represents each sub CIS. So we have a group of agents, where each has its own FCM. The problem encountered is the assortment of agents which have different specificities. So, how to master these specificities?

We proposed to use another parameter named coefficient of emergence. The difficulty lies in deciding the value of this coefficient, since several factors are considered. These factors may include, for example, if each CIS has: -a different perspective -a different gain percentage, a different percentage of participation - a chance to make a return for participation in the global cooperative information system. The difficulty therefore is how to enjoy the use of FCMs since we have agents with different sub-goals because they are intermediaries and competitors together to achieve the overall goal.

In our approach, we consider a new supervisor agent to decide the value of this coefficient. After this agent (which represents the global CIS) distributes the coefficient to each agent (which represents each sub CIS). These agents cooperate together to decide how to resolve the given problem. This cooperation is performed using the union operation.

The result of union operation is a new global fuzzy cognitive map. Once the resulting cognitive map is obtained, analysis is performed. This is made possible through other types of agents called analyzers. The work associated with each analyzer agent is to analysis and inference on the resulting collective FCM. In other words, we must ask each analyzer agent to draw some typical behaviors. The inference process is the same as that described in [24], except at the threshold function (normalized) [23]. We have summarized this process by the following algorithm:

Begin Algo
- Suppose we have a cognitive map of n concepts
1. Extract the adjacency matrix M_{ij}.
2. Preparing exciter vector a_1.
/ * Initialize the vector exciter to zero (0). * /
For i = 1 to n do
a_1 [i]←0 ;
endFor
/ * Each value i represents the state of concept i. * /
-Set to 1 the desired concepts. - Add a_1 to working memory WM.
i ← 1;
3. While i <= n do
$a_{i+1} = a_i \times M_{ij}$;
If a_{i+1} ú WM then a_{i+1} is typical behavior.
else For i = 1 to n do
If $a_{i+1}[i]>0.5$ then $a_{i+1}[i]$← 1 ;
else $a_{i+1}[i]$←0 ;
endif
endFor
add a_{i+1} in WM
endif
i ←i +1;
endWhile
end Algo

We summarized our approach by figure (2).

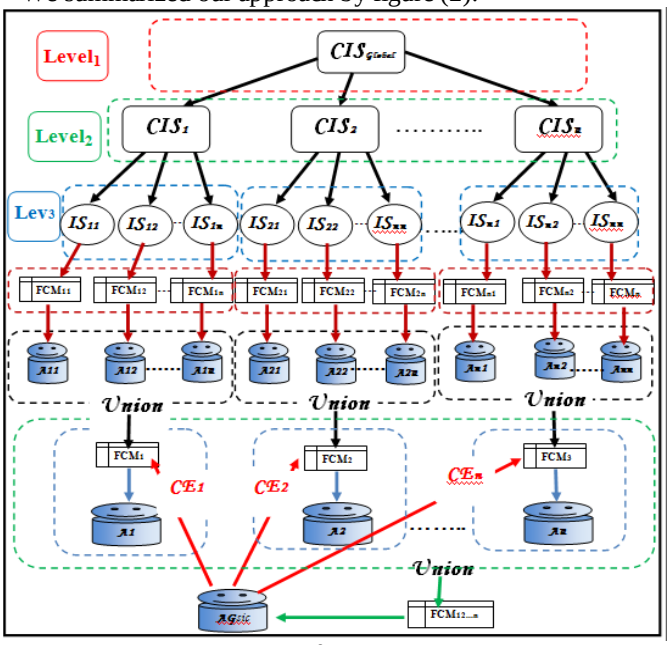


Fig.2. Architecture of the proposed approach

IV. Example and Implementation

A. Description of the example

To illustrate our approach, we considered the field of road safety as an example of application. We are among the top three countries in the world where the percentage of fatal accidents increases in an infinite way. Very few people recognize the predominance of human factor either directly in the element that caused the accident (drunkenness, lack of vigilance, etc.) or indirectly (technical inspection not performed, profile inappropriate conduct, etc.).The balance of road accidents recorded since the beginning of the year 2014 revealed 961 people dead, 13,815 injured in 9,047 road accidents, said the Director General of the National Centre for Prevention and Road Safety El Hashemi Boutalbi [25]. The awareness of people to crisis of roads and looking for a methodology to minimize the number of fatal accidents is the project of the century. This is the main goal of proposed example. The figure.3 summarizes this example with the proposed approach.

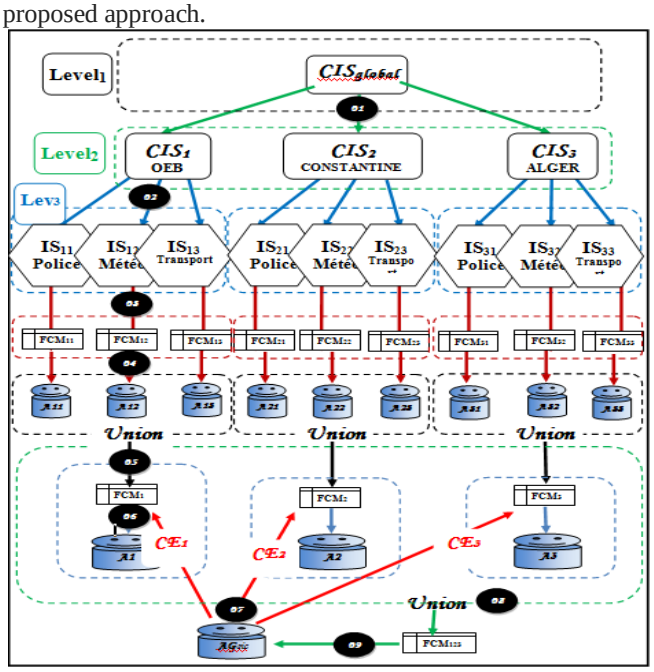


Fig.3. Instantiation of the proposed architecture by the example

01 The CIS is comprised of a plurality of different Centers for Prevention and Road Safety in Algeria (center 1, center 2 ...).

02 Each sub CIS is also formed of a set of ISs.

03 Each IS will file its own FCM on the field Road Safety.

04 Associations of each card each agent.

05 Result of union [a new fuzzy cognitive map].

06 Associating each resulting map to each agent model in a sub CIS.

07 Distribute the emergence coefficient to each agent.

08 Result of union

09 Associate the global collective FCM to the global agent AG_{SIC}.

The first level is the global CIS (represents the National Center for the Prevention and Road Safety of Algeria NCPRS). The problem is how to educate people to reduce the number of fatal accidents. The answer to this question is the goal of each center of every wilaya. The second level will therefore expose these centers. In the third level each center will depose the problem to each sub IS. The decision of each information system (police, meteorology, transportation ...) is interpreted by a FCM. To put the example in application, we modeled each IS by an agent. So, each center is a MAS of three agents (we associate the corresponding FCM to each agent). These agents will discuss to obtain a new FCM. The resulting map will describe the decision of the corresponding center.

Each resulting map is associated with a new supervisor agent (which delegates each center). New officers will be discussed by the methodology described in Part III. The result of the discussion is a new global FCM. Finally, the decision is achieved by dynamics of this card. In other words, the analysis and inference established by another group of agents analyzers. The following titles will explain the scenario described.

B. Implementation

We used the platform "Jade" to develop MAS and to implement the distributed systems. Together with "JBuilder 2008" which makes our application more intuitive and user-friendly. In the proposed application, the global CIS (represents the National Center for the Prevention and Road Safety of Algeria NCPRS) comprises of three sub CIS (represent the centers in different wilayas of Algeria). Each sub CIS comprises also three IS (possibly more, just to see the communication between agents in Jade and JBuilder 2008).

To obtain a collective fuzzy cognitive map it must first build an initial cognitive map (Figure 4), where we will offer to different agents ISs the basic concepts relating to the field of road safety, with unknown causality relations. The value of these relationships is generated by different agents ISs.

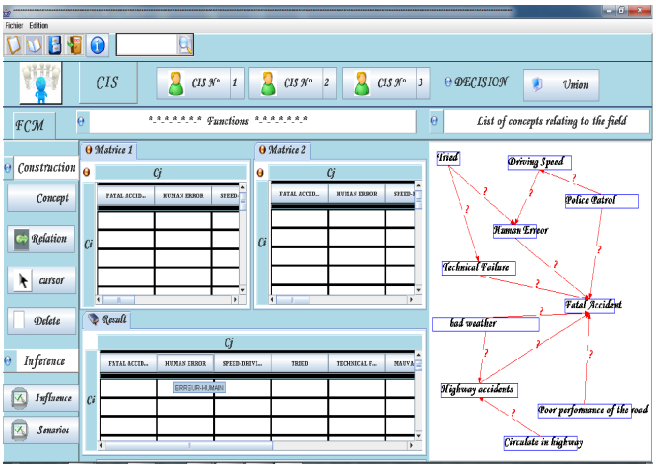


Fig. 4. Cognitive map on the field of road safety

Each button (named CIS) will launch three agents makers (IS) who will discuss and will make the union described in the third part. To launch the union operation we click the Union button. The cognitive maps of different agents (ISs) are integrated into a card that is subsequently associated to sub CIS. Now we have three cards of the three CIS which will also discuss to find the global FCM. The result of union in JBuilder 2008 is as follows (Figure 5):

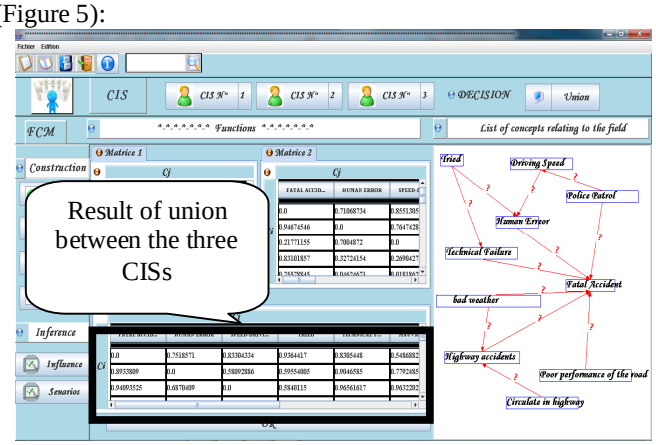


Fig.5. Result union in JBuilder 2008

And the result in jade is as follows (Figure 6):

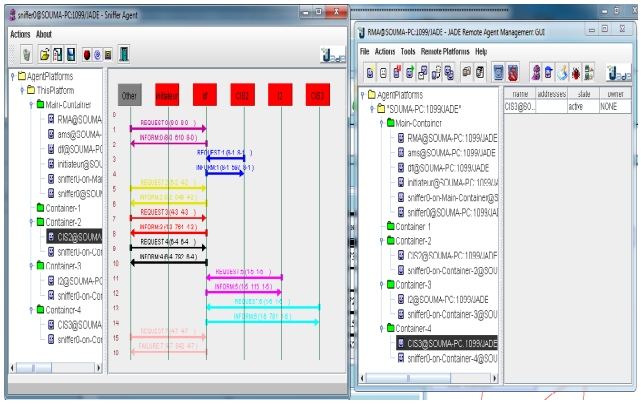


Fig.6. Process exchange messages between the groups of agents

The collective fuzzy cognitive map obtained is associated to the second group of analyzers agents, for the analysis and inference, whose goal is to reach a collective decision to solve the crisis problem. In our application, the "Scenarios" button will launch a global analyzer agent (will execute all the typical behaviors). The figure (7) shows the effect of "Scenarios" button:

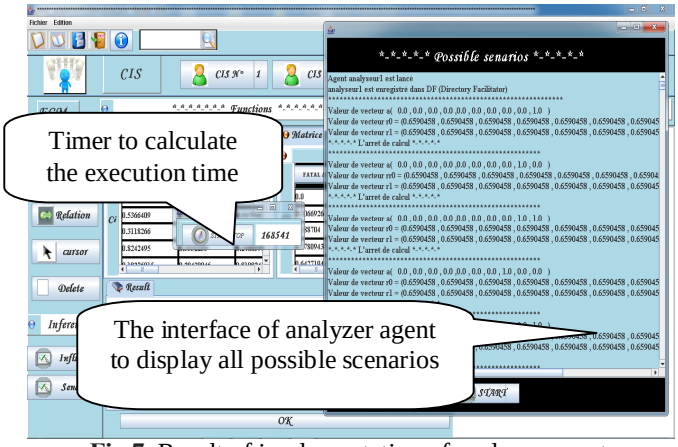


Fig.7. Result of implementation of analyzer agent

V. EXPERIMENTATION

In our example, we used ten concepts, so the second group analyzers agents will pull 2^{10}-2 typical behaviors equal to 1022 scenarios in few minutes, giving the advantage of really optimize in terms of time and term memory, since at the agent DF (Directory Facilitator) of the JADE platform, launched each agent is registered in the DF, it performs these specific behaviors and after its implementation, the DF automatically unregisters.

Collective decision of CIS is obtained by the dynamics of the collective cognitive map, i.e. the analysis and inference on this map. For example, if we increase the number of fatal accidents then activates the Fatal Accident concept and observed the inference result by the first analyzer (Figure 8).

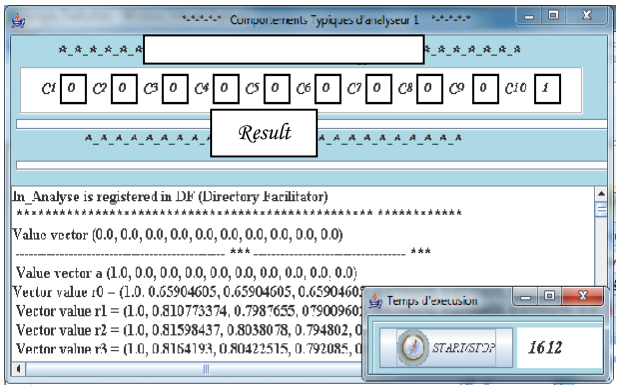


Fig.8. Execution result of analyzer agent

Regarding the result, we can deduce that if we activate the concept "fatal accident" then we obtained an increase of almost all other concepts because it is an engine concept. I.e., the number of fatal accidents increases if human error (0.8164193), fatigue (0.79520863), technical failure (0.76096), driving speed (0.8124419) ... increase and decrease if for example the number of police patrol (0.68057156) increases.

VI. Conclusion

In this paper, we proposed the use of fuzzy cognitive map tool in cooperative information systems (CIS) to improve and eliminate inconsistencies in the collective decision-making in inter-CIS. CIS are modeled by the MAS. So the proposed approach is a mix of the three paradigms that are CIS, MAS and FCM. Each agent according to its role, we associate it a fuzzy cognitive map that will describe its decision. After the union between the FCMs of ISs and union between the FCMs of different CISs agents, we obtain the collective fuzzy cognitive map. This card is automatically associated to the second group of analyzer agents for parallel analysis and inference in order to reach a collective decision to solve a given problem. We applied our approach in the field of road safety where we saw the collective decision that minimizes the crisis of roads. The development of MAS with the JADE platform offers the advantage through the *Sniffer* agent which is running the process of exchanging messages, or the inference process at each analyzer agent, where each analyzer agent will pull a some typical behaviors which save the cost both in terms of time and in terms of memory space. There are areas where the concept of time is present, for example, "Meteorology". So the proposed CIS will decide for example what the weather is tomorrow or next week. Here we suggest adding time concept in FCM and apply it, this makes more accurate and more realistic result.

References

[1] S. Khouider, ñ Outils Dôaide A La Décision Pour La Prise De Commandes Imprévues,ò Ph.D. Thesis. Engineering Sciences. Metz University. France, 2008.

[2] K. Zarour, ñVers Un Système Dôinformation Coopératif Pour La Prise En Charge Des Soins à Domicile : Une Architecture Basée Agent,ò M.S. Thesis, Mentouri University, Constantine, Algeria.

[3] M. Séguran, ñ Résolution Des Conflits Sémantiques Dans Les Systèmes D'information Coopératifs : Proposition D'un Modèle Dôinteraction Entre Agents,ò. Jean Moulin University, Lyon3, 2003.

[4] K. Zarour, N. Zarour and S. Khalfi, ñ Inter-Agent Interaction In Medical Information System: Case Homecare,ò Journal of Theoretical & Applied Information Technology. Vol. 11 Issue 1/2, 13p. 1 Black and White Photograph, 11 Diagrams, 1 Chart. ,2010, 130-142.

[5] A. Cardon, ñ Modélisation De La Couche Communicationnelle Dôun SIC Civil Coopératif,ò LirINSA Rouen, LIP6 Paris VI. Faculty of science, France.

[6] M. Kamoun, G. Uster and S. Hammadi, ñ Système DôInformation Coopératif A Base dôAgents Pour La Recherche dôItinéraires Multimodaux,ò Ecole Centrale de Lille / Institut National de Recherche sur les Transports et Leur Sécurité Villeneuve dôAscq, France.

[7] M. Zekiouk and M. Boufaida, ñ Une Architecture Multi-Agents Pour La Mise En Oeuvre De lôInteropérabilité Sémantique Dans Les Systèmes dôInformation Coopératifs,ò unpublished.

[8] Y. Miao, Z.Q. Liu, C.K. Siew and C.Y. Miao, ñ Dynamical Cognitive Network - An Extension Of Fuzzy Cognitive Map. Fuzzy Systems,ò IEEE Transactions on, vol. 9, 2001, 760-770.

[9] G. Glykas, ñ Fuzzy Cognitive Maps: Theory, Methodologies, Tools and Applications,ò Springer Verlag , July 2010.

[10] C. Stylios and P. Groumpos, ñThe Challenge Of Modeling Supervisory Systems Using Fuzzy Cognitive Maps,ò Journal of Intelligent Manufacturing , 1998, 339-345

[11] J. Jung and J. Geun-Sik, ñ CRMaps: Managing Customer Relationships Based On Ontological Fuzzy Cognitive Maps,ò Intelligent Ecommerce systems lab, School of Computer Science and Engineering Inha University, Incheon Korea , 2000,

[12] B. Kosko, ñ Fuzzy Cognitive Maps,ò International Journal of Man Machine Studies, 1986, pp. 65-75.

[13] B. Kosko, ñ Neural Networks And Fuzzy Systems: A Dynamical Systems Approach To Machine Intelligence,ò Engelwood Cli_s, Prenticehall Edition ,1992.

[14] J. Dickerson and B. Kosko, ñVirtual Worlds As Fuzzy Cognitive Maps,ò Presence, 1994,3, 173,189.

[15] J. Dickerson and B. Kosko, ñVirtual Worlds As Fuzzy Dynamic Systems,ò In Technology for Multimedia, B. Sheu editor IEEE presses. ,1996.

[16] I.F.M. Avila,. ñ Qualitative Reasoning In Multi-agent Systems Based On Causal Maps,ò Laval ,2002.

[17] G. Acampora and V. Loia, ñA Dynamical Cognitive Multi-Agent System for Enhancing Ambient Intelligence Scenarios,ò IEEE Inter Confer on Fuzzy Systems, art. no. 5277303, 2009, pp. 770-777.

[18] V.Rodin, G. Querrec, P.Ballet, F. Bataille, G. Desmeulles and Jeanï Francois Abgrall. ñ Multi-Agents System To Model Cell Signaling By Using Fuzzy Cognitive Maps. Application to computer simulation of Multiple Myeloma,ò Proc. Ninth IEEE International Conference on Bioinformatics and Bioengineering, 2009, pp. 236-241

[19] M.Stula, D. Stipanicev and L.Bodrozic, ñ Intelligent Modeling With Agent-Based Fuzzy Cognitive Map,ò Inter. Jour of Intelligent Systems. Vol 25, Issue 10.,October 2010, 981ï 1004

[20] B. Nachet,ñ Modéle multi agent pour la conception de systems dôaide à la decision collective,ò Thesis submitted to obtain a doctoral degree in Computer Science. Oran University of Oran, Algeria. ,2014.

[21] W.-R. Zhang, S.-S.Chen, W. Wang and R. King, ñ A Cognitive Map-Based Approach To The Coordination Of Distributed Cooperative Agents,ò IEEE Transactions on systems, man, and cybernetics, 22(1) ,1992, pp.103ï 113.

[22] B. Kosko, ñNeural Networks and Fuzzy Systems, a dynamical Systems Approach to Machine Intelligence,ò Prentice-Hall, USA, 1992.

[23] L.Chauvin, ñ Modèles De Cartes Cognitives Etendues Aux Notions De Contexte Et Dôéchelle,ò Faculty science. Angers University ,2010. 116-117

[24] S. Merazga, ñ Un outil de manipulation et d'analyse des cartes Cognitives,ò Oum El Bouaghi University. Algeria ,2010.

[25] http://www.lematindz.net/news/14349-accidents-routiers-en-algerie-961-tues-en-quatre-mois.ht, 05/05/2014, 09:39:00.

Modélisation des Situations de Besoins d'Assistance

dans un EIAH : Le Modèle IMAC

Nadia Beggari
Groupe de Recherche en e-Learning (GReLearn)
Laboratoire de Recherche en Informatique *(LRI)*
Université Badji-Mokhtar, Annaba 23000, Algérie.
nadiabeggari@hotmail.fr

Tahar Bouhadada
Groupe de Recherche en e-Learning (GReLearn)
Laboratoire de Recherche en Informatique *(LRI)*
Université Badji-Mokhtar, Annaba 23000, Algérie.
bouhadada.tahar@univ-annaba.org

***Résumé*— Dans cet article nous nous intéressons à la conception d'un système d'assistance active aux apprenants dans un EIAH. Notre problématique porte sur l'identification des situations de difficultés rencontrées par les apprenants, pour la prise en main et l'utilisation de l'environnement d'apprentissage. Nous proposons le modèle IMAC qui décrit et modélise les situations problématiques (d'assistance) par un indicateur de traces avec sa règle de calcule , une Modalité d'intervention (proactive permanente, proactive occasionnelle, réactive), un Aspect(cognitif, métacognitive, méthodologique , organisationnel, instrumentale technique et/ou pédagogique et /ou socio-affective ,disciplinaire) , et une Catégorie d'interventions (Situation de blocage, Situation de non accomplissement de tache, le cas du passage d'une tâche à une autre, les demandes d'aide, les situations des difficultés...)**

***Mots clés*— assistance active, indicateurs d'apprentissage, système d'assistance, situations d'assistance, traces.**

I. INTRODUCTION

LES environnements informatique d'apprentissage humain « EIAH » sont des dispositifs de formation conçus pour fournir à l'apprenant les bonnes activités avec les bons outils. Ils se caractérisent par les très riches possibilités qu'ils offrent aux apprenants. Néanmoins, devant les nombreuses fonctionnalités offertes par ces dispositifs, et leur interface souvent chargée, l'apprenant peut renoncer à les utiliser ou les sous-exploiter [4], limitant ainsi leur intérêt et leur richesse. L'adjonction d'un système d'assistance est l'une des solutions pour pallier ces difficultés de prise en main et d'utilisation. Cependant le développement de tel système est une tâche complexe et coûteuse, posent plusieurs problèmes, sur le contenu de l'assistance fournie à l'apprenant, et sur la méthode et la manière d'offrir cette assistance.

Notre travail consiste à concevoir et à mettre en place un système d'aide à l'apprentissage individuel capable de fournir aux apprenants une assistance active et personnalisée, l'aide dont ils ont réellement besoin au moment où ils en ont besoin. Dans ce papier, nous nous intéressons particulièrement aux deux principales questions suivantes : *Comment identifier les situations de besoin d'assistance ? Et comment apporter une assistance appropriée à ces besoins?*

Dans cet article, nous présentons une approche qui vise à guider et à faciliter l'identification des situations d'assistance (les moments de difficultés). L'approche proposée est fondée sur un modèle conceptuel destiné à caractériser les situations de difficultés dans le cadre d'une formation à distance. Ce modèle est basé sur l'analyse de traces d'interactions des apprenants avec le système.

II. ETAT DE L'ART

Dans un système d'apprentissage, le rôle du système d'assistance est d'analyser les interactions entre l'apprenant et le système et d'apporter une aide active et intelligente à l'apprenant. Le terme « aide active » désigne que non seulement l'apprenant obtient sur sa demande une aide ciblée sur ses activités, mais le système peut également décider d'intervenir pour proposer une assistance, lorsqu'il lui semble que l'apprenant a de trop grandes difficultés avec la tâche en cours [17].

La conception d'un système d'aide active fait l'objet de plusieurs travaux. En particulier, Dans [17], les auteurs proposent un système conseiller actif fondé sur l'analyse du modèle de la tâche et une typologie de conseils. Cette approche vise à guider et à faciliter l'identification des conseils sur la tâche. Plus récemment, des chercheurs proposent un modèle générique d'assistance aux acteurs du téléapprentissage [7]. Il est constitué d'un modèle de la tâche, d'un modèle de l'usager et du groupe, ainsi que d'un ensemble d'interventions d'assistance. Chaque intervention est associé a un ensemble de métadonnées (identifiant, but…) et à un ensemble de règles du type « Si Condition d'assistance, Alors Action d'assistance ».

D'autres travaux proposent des consignes ouvertes qui peuvent être accompagnées d'artefacts cognitifs destinés à faciliter la compréhension de la consigne et l'appropriation des ressources [15]. Ces assistants méthodologiques sont intégrés dès la conception aux scénarios mis en œuvre dans les parcours (cas dans Pairform@nce). Aussi dans [21], les auteurs proposent un système d'aide adaptatif à base de trace d'interaction. Il utilise les traces des usagers comme des sources de connaissances que le système peut exploiter pour s'adapter au profil utilisateur. Il se sert des expériences des utilisateurs pour aider un autre utilisateur se trouvant dans un contexte similaire.

Enfin dans [11], les auteurs proposent un modèle générique d'assistance aux utilisateurs d'applications informatiques permettant la spécification et l'exécution de systèmes d'assistance personnalisée dans une application-cible existante, sans avoir à la modifier. Puis l'exécution de l'assistance spécifiée afin de répondre aux besoins des utilisateurs. Ce modèle s'appuie sur un formalisme des règles d'assistance inspiré de [7] pour identifier les besoins d'assistance.

La plupart de ces systèmes proposent des aides proactives qui sont mises en place à partir des prédictions et des conceptions de concepteurs d'assistance. Cependant plusieurs études ont démontré que les systèmes d'aide sont généralement ignorés ou rarement consultés pour diverses raisons [5], parmi lesquelles, l'écart entre le point de vue de l'utilisateur et celui du concepteur de l'aide. Ainsi selon les auteurs de [14], l'origine de cet écart vient du fait qu'il est difficile d'anticiper de manière très précise les compétences des utilisateurs cibles, en particulier si l'aide est envisagée dans un système grand public, et par conséquent, ces compétences supposées peuvent être superflues ou sous-estimées.

Nous proposons une approche d'assistance active basée sur une fonction d'accompagnement à travers l'utilisation des expériences de concepteurs d'assistance et les techniques d'analyse de traces d'interactions des apprenants avec l'environnement d'apprentissage. Dans cet article nous présenterons le modèle IMAC, un modèle conceptuel destiné à caractériser et modéliser les situations de difficultés rencontrées par les apprenants au cours de sessions d'apprentissage.

III. PROPOSITION

Pour être utile et efficace, un système d'assistance doit être capable de fournir à l'utilisateur l'aide ou l'assistance dont il a réellement besoin au moment où il en a besoin. Pour déterminer ces moments, il est nécessaire d'identifier les difficultés que l'apprenant va rencontrer au cours de son apprentissage et comme il n'est pas possible de prédire toutes les difficultés au moment de la conception de système [12], Il est envisageable de calculer des indicateurs révélateurs de ces difficultés au fil du déroulement des sessions d'apprentissage à partir des traces d'interactions des apprenants.

Le modèle IMAC proposé à pour objectif d'identifier les situations et les moments où l'apprenant à besoin d'une assistance. Le modèle décrit et modélise les situations problématiques (d'assistance) par un ensemble de caractéristiques : un Indicateur de trace ; une Modalité d'intervention; un Aspect ; et une Catégorie d'interventions.

A. Indicateur de trace

Afin d'aider l'assistant à suivre l'activité de l'apprenant et déterminer les conditions d'interventions proactives pour pouvoir l'assister, des indicateurs d'apprentissage sont proposés. Chaque indicateur est défini par une règle de calcul et un domaine de valeurs permettant d'identifier une situation critique qui peut être une situation d'assistance. Dans ce cas, l'assistant intervient si la valeur de l'indicateur appartient à leur domaine de valeurs. Par exemple : *un indicateur de temps d'inactivité est supérieur à un seuil (défini par un expert) peut être un indicateur d'une situation d'assistance de type blocage. Si un apprenant clique sur de nombreuses zones de l'interface pendant un temps relativement long, cela pourrait signifier qu'il est à la recherche de quelque chose , qu'il se soit perdu ou qu'il soit désorienté.*

B. Modalité d'intervention

Dans un dispositif de formation en ligne, on trouve deux modalités d'intervention, proactive et réactive. L'intervention est proactive quand l'assistant ou le tuteur intervient spontanément auprès de l'apprenant, cependant l'intervention réactive est une réaction à une demande de l'apprenant. Ainsi, la modalité proactive se décompose en deux sous-modalités : une systématique dont les interventions sont planifiées lors de certains moments bien définis considérés comme des moments critiques d'apprentissage comme le début et la fin de la formation; et une opportune consistant à intervenir lorsque les apprenants semblent être en difficultés comme le changement d'activité, la situation du passage d'une phase de travail à une autre.

Dans le présent travail, nous nous basons particulièrement sur la modalité proactive afin d'anticiper les difficultés rencontrées par les apprenants, et ainsi réduire les perturbations qui pourraient y être associées. Ainsi, nous avons identifié trois types d'assistance, selon les trois modalités d'interventions :

- Une assistance réactive qui représente les sollicitations de l'apprenant envers son assistant ou ses pairs et inversement, permettant à ces derniers d'avancer dans leur tâche.

- Une assistance proactive qui représente les moyens et les stratégies prévus par l'assistant en amont, et avant toute utilisation des acteurs, afin de médiatiser ce processus [13].

- L'assistance proactive peut être systématique ou opportune.

  - Dans une assistance proactive **systématique**, l'assistant décide d'intervenir de manière proactive organisée lorsque, par exemple, à l'issue d'une étape, l'assistant apporte des informations signifiantes pour la suite de l'activité.

  - Dans une assistance proactive **opportune**, un indice déclencheur (une erreur particulière commise, un choix non pertinent, etc.) est l'occasion pour l'assistant de réorienter l'utilisateur vers des stratégies d'apprentissage plus efficientes. Dans notre système l'indice

déclencheur est un indicateur de traces proposé dans la première phase.

C. Un Aspect

La situation d'assistance peut être caractérisée par un Aspect ou une dimension, qui détermine le contenu qu'elle apporte. Beaucoup de recherches se sont penchées sur cette caractérisation, tel que dans [3], où les auteurs s'intéressent à la dimension didactique de l'aide, et donc à tout ce qui porte sur les contenus. Dans [15], les auteurs s'intéressent pour leur part à la dimension cognitive alors que d'autres, à la dimension technologique [21].

De façon plus spécifique, une intervention d'assistance est toujours faite en fonction d'un objet ou d'un ensemble d'objets [7]. Elle fait partie du système d'apprentissage (traitement du contenu, traitement d'information dans le but de satisfaire une fonction représentationnelle), du processus de socialisation, d'une tâche technique ou d'organisation. Ainsi, Sur la base de la topologie de la dimension d'aide proposée dans [13], nous avons associé à la situation d'assistance des aspects :

1) Aspect cognitif

Nous rattachons à l'aspect cognitif, toute situation critique dont le contenu est porté sur l'activité de compréhension et d'appropriation. Par exemple: la mise en relation des concepts abordés, l'analyse du problème envisagé, la confrontation des points de vue et l'incitation à appliquer concrètement les notions découvertes, blocages dans les apprentissages théoriques, d'une difficulté de mémorisation ou de raisonnement.

2) Aspect métacognitif

Dans l'aspect *métacognitif,* nous regroupons les situations critiques comportant des problèmes sur des activités de construction de nouvelles connaissances, et qui nécessitent une réflexion métacognitive. A cet égard, l'assistant offre un soutien par étayage le conduisant à réaliser avec l'apprenant les activités que celui-ci ne peut effectuer seul [*19*].

3) Aspect organisationnel

Nous ajustons toute situation critique qui porte sur l'organisation du travail des apprenants dans une catégorie distincte, qualifiée, selon des auteurs, d'« organisationnelle », de « régulative » ou de « managériale »[19]. Par exemple : les situations de problèmes de la gestion du temps qui facilite la planification des activités afin de respecter les échéances fixées [16, 2, 22,6].

4) Aspect technique

Nous rattachons à l'aspect technique, toute situation relevant de la technique comme par exemple : signaler un problème technique ou un dysfonctionnement sur sa machine, un problème de connexion, des problèmes liés à l'utilisation des logiciels liés à la formation : demander comment faire fonctionner un logiciel donné, …etc.

5) Aspect socio-affectif

Il est d'usage de regrouper des situations d'assistance destinées à soutenir l'assistance socio-affective d'ordre social. Comme par exemple : les situations de difficultés pour établir des relations positives entre les utilisateurs (apprenant, assistant et enseignant), les problèmes du contact, et les problèmes d'utilisation d'espace en commun et de partage des outils de communication.

6) Aspect méthodologique

Pour cet aspect, nous nous intéressons particulièrement aux situations destinées aux principaux problèmes méthodologiques que rencontrent les apprenants. Par exemple : des problèmes de la gestion de temps ou de la gestion des ressources d'apprentissage, la recherche des méthodes et des approches afin de réaliser certaine activité.

7) Aspect instrumental

Une situation est dite instrumentale lorsque l'apprenant recherche l'information minimale qui lui manque pour réussir la tâche par ses propres moyens [9]. On peut déterminer ce type de situation grâce à un indicateur du but de maîtrise où l'apprenant recherche à développer ses compétences. Plusieurs chercheurs ont montré que plus l'apprenant poursuit un but de maîtrise, plus il est porté par la recherche de l'aide instrumentale, donc, il est dans une situation d'assistance [20].

D. Catégorie d'interventions

Le but des systèmes d'assistance, en général, est de répondre à un besoin particulier, spécifique à chaque utilisateur du système d'apprentissage qui se trouve en situation d'échange, de blocage ou encore d'interrogation face à plusieurs choix qui s'offrent à lui [18].

Ces situations sont liées à un domaine précis, nous avons essayé dans le cadre de ce travail, de généraliser afin d'établir cette topologie pour les situations d'assistance :

1) *Situation de blocage*

Nous parlerons de blocage lorsqu'après une période normale d'essais plus ou moins fructueux, un apprenant ne réussit toujours pas à progresser. De ce fait, son comportement est caractérisé par un refus, une incapacité apparente et provisoire de poursuivre un apprentissage, de réagir à une situation (GDP Larousse).

2) Situation de non-accomplissement de tache

Dans le cadre d'une session d'apprentissage, l'apprenant doit apprendre, améliorer et développer ses connaissances et ses compétences à partir des tâches professionnelles qui sont des unités de travail bien définies. Chaque unité est divisée en séquence ordonnée d'étapes pour les tâches de type procédure (Exemple : créer un tableau dans Microsoft Word) ; ou en lignes directrices qui doivent être appliquées pour les tâches fondées sur des principes (Exemple : organiser une conférence). De ce fait, l'apprenant peut être bloqué dans une étape ou dans une ligne directrice qui l'empêche d'accomplir sa tâche correctement et atteindre l'objectif du cours.

3) Le cas lors du passage d'une phase de travail à une autre

Le contenu de la formation en ligne peut inclure des ressources d'apprentissage, des leçons en ligne interactives; des simulations électroniques ; des outils de travail. Ces différents composants mettent l'apprenant dans une situation de confusion où il ne peut pas choisir la méthodologie la plus optimale de navigation et ne pas pouvoir répondre aux questions du genre: « quel chemin vais-je choisir ? », « quel composant, vais-je consulter ? », « que dois-je faire maintenant ? », « quelle est la prochaine étape». Ainsi,

l'assistant doit pouvoir intervenir pour fournir à l'apprenant des directives afin de lui permettre d'atteindre les objectifs d'apprentissage.

4) Les demandes d'aide

Au sein de sessions d'apprentissage, l'apprenant réagit en fonction du contexte et de la tâche. Poser des questions sur la consigne, les étapes et le temps qui lui est alloué relève de l'organisationnel, puisqu'à travers de telles questions, l'apprenant cherche les balises à suivre pour réaliser le travail qu'il a à faire. Et poser des questions sur la méthode, les approches possibles, relève pour sa part du méthodologique [13].

5) Les situations de difficulté

Être en difficulté d'apprentissage, c'est avoir des problèmes au niveau de la perception, de la compréhension et (ou) de l'utilisation des concepts. Ces problèmes causent un retard de développement et (ou) des difficultés avec un ou tous les aspects suivants : organisation dans le raisonnement, l'attention, la mémoire, le raisonnement, la coordination, la communication, la lecture, la rédaction, l'orthographe, le calcul, l'habileté sociale et la maturité affective.

6) Désorientation

Etre désorienté, en termes de progression dans le système d'apprentissage, c'est avoir des difficultés pour identifier sa position dans la structure d'ensemble d'information. Dans de telle situation, l'apprenant doit reconstruire le parcours qui l'a mené à ce nœud, discerner les choix possibles qui s'offrent à lui, choisir une destination et générer un chemin vers un nœud dont il sait qu'il existe [10]. Aussi, dans [8] les auteurs définissent la désorientation par la difficulté à extraire des informations en vue de réaliser une tâche ou à effectuer des traitements sur le contenu comme la compréhension, la sélection ou l'apprentissage.

IV. CONTEXTE DE LA MISE EN ŒUVRE

Pour appliquer et valider notre modèle, nous avons développé un système d'assistance personnalisable comporte deux parties, une partie concerne la construction, dans laquelle le concepteur d'assistance doit spécifier ses préventions sous forme des situations de difficultés associées à des propositions d'assistance. Et une partie concerne la mise en place de ce système d'assistance dans l'environnement d'apprentissage.

La partie de construction permet au concepteur de présenter ses préventions sous forme d'une liste de situations de difficultés anticipées. Cette présentation s'appuie sur le modèle IMAC qui modélise la situation de difficulté par un ensemble des caractéristiques. Initialement, le concepteur doit définie les indicateurs révélateurs des difficultés des apprenants proposés par le système. Ces indicateurs sont associés à des situations problématiques nécessitant une intervention d'assistance et à des composants de l'environnement d'apprentissage. Les situations identifiées sont classées selon leur aspect et leur catégorie du problème.

Ainsi le concepteur doit associe à chaque situation de difficulté une proposition d'assistance. Pendant l'exécution de système d'assistance dans l'environnement d'apprentissage. Une fois le critère d'activation satisfait, le système va lancer le processus de calcule de l'indicateur. Selon la valeur de ce dernier. Le système va décider, si l'apprenant dans un cas de besoin d'assistance ou non. Si, c'est le cas, le système va proposer à l'apprenant l'assistance associée à la situation de difficulté.

Dans notre approche d'assistance, les indicateurs proposés sont classifié selon trois dimensions : indicateurs de temps « La dimension temporelle », indicateurs sur les interactions des apprenants « L'aspect d'interaction », et indicateurs de production des apprenants « L'aspect de production et de progression).

V. CONCLUSION

Dans cet article, nous avons décrit notre proposition d'assistance active à l'apprentissage individuel dans un EIAH. Le modèle IMAC présenté apporte une réflexion sur la modélisation des situations de besoins d'assistance. Chaque situation est caractérisé par : (i) un Indicateur de trace avec sa règle de calcul et une ou plusieurs m-traces, (ii) une Modalité d'intervention (proactive permanente, proactive occasionnelle, réactive), (iii) un Aspect (cognitif, métacognitive, méthodologique, organisationnel, instrumentale technique et/ou pédagogique et /ou socio-affective ,disciplinaire), (iiii) et une Catégorie d'interventions (Situation de blocage, Situation de non-accomplissement de tache, le cas de passage d'une tâche à une autre, les demandes d'aide, les situations de difficultés...).

Une tentative d'intégration du modèle IMAC dans notre système d'assistance AssistAgent[1] est en cours pour valider la pertinence des paramètres définis.

REFERENCES

[1] N .Beggari, T. Bouhadada"Assistance pédagogique dans un environnement d'apprentissage à distance basé web services". in proceeding of 5th Guide International Conference, Università degli Studi Guglielmo Marconi, Italy, 18-19 November 2011.

[2] Z.L.Berge "Facilitating Computer Conferencing: Recommendations From the Field", in: *Educational Technology*, 35 (1), 1995, p. 22-30.

[3] R.Bruderman, « *L'aide à l'apprentissage du vocabulaire à distance : effets des outils de la CMO ». Dans : Revue internationale des technologies en pédagogie universitaire, Vol.9, Issue 3 ; p.25-42 . 2012*

[4] A.Cordier, M.Lefevre, S.Jean-Daubias, N.Guin « Concevoir des assistants intelligents pour des applications fortement orientées connaissances : problématiques, enjeux et étude de cas ». Dans 2010

[5] A.Capobianco, N .Carbonell « Aides en ligne à l'utilisation de logiciels grand public : problèmes spécifiques de conception et solutions potentielles ». Dans *Intellectica 2*(44). 2006

[6] A. Daele,. , F .Docq " Le tuteur en ligne, quelles conditions d'efficacité dans un dispositif d'apprentissage collaboratif à distance «». in : *Regards pluriels et critiques sur les pratiques*, 19e colloque de l'Association internationale de pédagogie

universitaire sur les méthodes actives dans l'enseignement supérieur, Louvain-la-Neuve, 29 au 31 mai 2002.

[7] A.Dufrene, J.Basque, G.Paquette, M.Léonard, K.Lundgren-Cayrol, S.Promtep, « Vers un modèle conceptuel générique de système d'assistance pour le téléapprentissage». Sticef, p 57-88. 2003

[8] W.Elm, D.Woods, " Getting lost: A case study in interface design". In proceedings of the human factors society 29th Annual Meeting, p. 927-931. 1985.

[9] N.Fabrice, N.Huet, Ch.Escribe, J-Ch.Sakdavong, O.Catteau « Buts d'accomplissement de soi et jugement métacognitif des aides en EIAH ». Environnements Informatiques pour l'Apprentissage Humain, Lausanne 2007.

[10] P. Fastrez, "*Navigation entailments as design principles for structuring hypermedia"*. In: *Education, Communication and Information, Special Issue on Hypertext and Hypermedia*, Vol.2, no.1, p.7-22 (2002). http://hdl.handle.net/2078.1/69317.

[11] B.Ginon, S.Jean-Daubias, P-A.Champin « Mise en place d'un système d'assistance personnalisée dans une application existante » 24ièmes Journées francophones d'Ingénierie des Connaissances, Lille, 3-5 juillet, Lille, France, 2013.

[12] D.Gaonac'h, « *Les différentes fonctions de la mémoire dans l'apprentissage des langues étrangères* ». Conférence plénière au XXVIIe Congrès de l'APLIUT, IUT de Toulon Sud (Var), France. 2005.

[13] S-H. Kasdali, « Le soutien : entre situation d'apprentissage synchrone et situation asynchrone, quelle différence chez l'apprenant en ligne ». Congrès AREF 2013, Actualité de la recherche en éducation et en formation, 2013.

[14] J.Leplat, « A propos des procédures ». Performances humaines & techniques, n°94, Mai-Juin, p. 6-15. 1998

[15] C.Loisy, Ch. Péllissier « *Des aides pour une consigne ouverte : assistants cognitifs dans Pairform@ance* ». Revue internationale des technologies en pédagogie universitaire, volume : 9, Issue : 3 ; pp.43-54.

[16] R.Mason, "*Moderating Educational Computer Conferencing* Deonews, 1(19),1991.http://www.ed.psu.edu/acsde/deos/deosnews/deosnews1_19.asp.

[17] G.Paquette, " TeleLearning Systems Engineering. Towards a new ISD model". *Journal of Structural Learning*, 14 (4), p. 319-354, 2001.

[18] Ch. Pélissier « *Notion d'aide dans un dispositif de formation en ligne : approches théoriques et méthodologiques* ». Revue internationale des technologies en pédagogie universitaire, Volume 9 ; Issue 3; p.71-83, 2012.

[19] J.-J.Quintin « *Accompagnement tutoral d'une formation collective via Internet. Analyse des effets de cinq modalités d'intervention tutorale sur l'apprentissage en groupe restreints* ». Thèse de doctorat, Université de Mons-Hainault, 2008.

[20] Ryan, Pintrich" Should I Ask for Help? The Role of Motivation and Attitudes in Adolescents Help Seeking in Math Class". *Journal of Educational Psychology*, Vol. 89, n°2, p.1-13, 1997.

[21] K.Sahaba « Système d'aide adaptatif à base de traces. International Journal of Technologies in Higher Education, IJTHE 9(3), 2012.

[22] J.Vermont, N .Verloop" Congruence and friction between learning and teaching". *Learning and Instruction* (9), p.257-280. 1999.

DEACP: An Energy Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks

Chirihane Gherbi
Dept of Mathematics and Computer Science, University of Oum El Bouaghi OEB, Algeria
Gherbi88@gmail.com

Zibouda Aliouat
Department of science
Ferhat Abbes University
Setif, Algeria
Aliouat_zi@yahoo.fr

Mohammed Benmohammed
Dept of computer science
Mentouri University
Constantine, Algeria
Ben_moh123@yahoo.com

***Abstract*— Hierarchical routing is an efficient way to lower energy consumption within a cluster, performing data aggregation and fusion in order decrease the number of transmitted messages to the BS. In this paper, a novel hierarchical approach called distributed energy efficient adaptive clustering protocol with gathering data (DEACP) is proposed for Wireless sensor network. Since nodes in a sensor network have limited energy, prolonging the network lifetime and improving scalability become important. we have proposed (DEACP) approach to reach the following objectives: reduce the overall network energy consumption, balance the energy consumption among the sensors and extend the lifetime of the network, the clustering must be completely distributed, the clustering should be efficient in complexity of message and time, the cluster-heads should be well-distributed across the network, the load balancing should be done well, the clustered WSN should be fully-connected.**
As a result transmission power of the node is reduce which subsequently reduces the energy consumption of the node. Our proposed work is simulated through Network Simulator (NS-2). Simulations show that (DEACP) clusters have good performance characteristics.

Keywords- Energy saving; Distributed algorithm; Cluster-based Routing; Wireless Sensor Network.

I. INTRODUCTION

Networks of Wireless Sensor devices are being deployed to collectivity monitor and disseminate information about a variety of phenomena of interest. A wireless sensor device is a battery- operated device, capable of sensing physical quantities. In addition to sensing, it is capable of wireless communication, data storage, and a limited amount of computation and signal processing. Advances in integrated circuit design are continually shrinking the size, weight and cost of sensor devices, while simultaneously improving their resolution and accuracy. At the same time, modern wireless networking technologies enable the coordination and networking of a large number of such devices.

A wireless sensor network (WSN) consists of a large number of wireless capable sensor devices working collaboratively to achieve a common objective. One or more sinks (or base-stations) which collect data from all sensor devices. These sinks are the interface through which the WSN interacts with the outside world. Challenges in wireless sensor network arise in implementation of several services, there are so many controllable and uncontrollable parameters [4] by which the implementation of wireless sensor network affected such as: *Energy conservation*.

As we know that the sensor node has small size due this small size the battery has low capacity and the available energy is very less. When the wireless sensor network replaced the single macro sensors, it gained advantage in extended range of sensing, fault tolerance, improved accuracy and lower cost than its predecessors. But as the number of nodes increases in the WSN to increase the coverage range and accuracy, energy management becomes a major constraint since all these nodes are battery powered. And in that situation the refilling or replacing of battery is impossible. It is a costly attempt. In order to avoid this problem some more energy efficient protocol are design so that the sensor node communicate efficiently by increasing both throughput and network capacity.

In this paper, we propose a Distributed Energy Efficient Adaptive Clustering Protocol with gathering data. Our goal is to achieve better cluster size balance and obtain clusters such that each has the minimum energy topology.

Clustering is an efficient way to reduce energy consumption and extend the life time of the network, doing data aggregation and fusion in order to reduce the number of transmitted messages to the BS [2]. Nodes of the network are organized into the clusters to process and forwarding the information, while lower energy nodes can be used to do the sensing the target, DEACP makes no assumptions on the size and the density of the network. The number of levels depends on the cluster range and the minimum energy path to the head.

The proposed protocol reduces the number of dead nodes and the energy consumption, to extend the network lifetime. The rest of the paper is organized as follows:

An overview of related work is given by section 2. In section 3, propose an energy efficient level based clustering routing protocol DEACP. Simulations and results of experiments are discussed in the section 4. In

section 5, concludes the work presented in this paper and the scope of further extension of this work.

II. Related work:

The growing interest in wireless sensor networks and the continual advancements in micro electronics and wireless communication technologies have inspired some previous efforts. Various issues in the design of wireless sensor network design of low-power signal processing architectures, low-power sensing interfaces, energy efficient wireless media access control and routing protocols. In 2004, G. Smaragdakis, I. Matta and A. Bestavros [9] proposed Stable Election Protocol (SEP) protocol. This protocol is an extension to the LEACH protocol. It is a heterogeneous aware protocol, based on weighted election probabilities of each node to become cluster head according to their respective energy. This approach ensures that the cluster head election is randomly selected and distributed based on the fraction of energy of each node assuring a uniform use of the nodes energy. In this protocol, two types of nodes (two tier in-clustering) and two level hierarchies were considered. In the Linked Cluster Algorithm [6], a node becomes the clusterhead if it has the highest id among all nodes within two hops. In PEGASIS(Power-Efficient Gathering in Sensor Information Systems), a near optimal chain-based protocol that is an improvement over LEACH. In PEGASIS, each node communicates only with a close neighbor and takes turns transmitting to the base station, thus reducing the amount of energy spent per round. In updated LCA [7], those nodes with the smallest id become cluster head. All the other nodes which are 1-hop to the heads become children of the heads. In [8], those nodes with highest degree among their 1-hop neighbors become cluster heads. In [10], the authors propose two load balancing heuristics for mobile ad hoc networks, where one is similar to LCA and the other is degree-based algorithm. The Weighted Clustering Algorithm (WCA) [11] elects cluster-heads based on the number of surrounding nodes, transmission power, battery-life and mobility rate of the node. WCA also restricts the number of nodes in a cluster so that the performance of the MAC protocol is not degraded. The Distributed Clustering Algorithm (DCA) uses weights to elect clusterheads [12]. These weights are based on the application and the highest weight node among its one hop neighbors is elected as the clusterhead. All of the above algorithms generate 1-hop clusters, require synchronized clocks and have a complexity of *O(n)*, where n is the number of sensor nodes. This makes them suitable only for networks with a small number of nodes. All the previous protocols require either knowledge of the network density or homogeneity of node dispersion in the field. Hybrid Energy Efficient Distributed clustering (HEED). HEED does not make any assumptions about the network, such as, density and size. Every node runs HEED individually. At the end of the process, each node either becomes a clusterhead or a child of a clusterhead. Residual energy of a node is the first parameter in the election of a clusterhead, and the proximity to its neighbors or node degree is the second. HEED generates a 1-level hierarchical clustering structure for intra-cluster communication. Ming Liu et. al. [ming] has presented An Energy-Aware Routing Protocol in Wireless Sensor Networks, The authors present EAP, a novel energy efficient data gathering protocol with intra-cluster coverage. EAP clusters sensor nodes into groups and builds routing tree among cluster heads for energy saving communication. In addition, EAP (Energy Aware Routing Protocol) introduces the idea of area coverage to reduce the number of working nodes within cluster in order to prolong network lifetime. Also, DEEAC protocol Distributive Energy Efficient Adaptive Clustering (DEEAC) protocol. This protocol is adaptive in terms of data reporting rates and residual energy of each node within the network.

III. Distributed Energy Efficient Adaptive Clustering Protocol with Gathering Data for Wireless Sensor Networks. (DEACP):

A. Radio model

We used the following equations for calculating the Communication energy dissipation. The free space (d^2 power loss) channel model is used, depending on the distance between the transmitter and the receiver. The energy spent for the transmission of the k-bit packet over distance d is given by (1) and (2)

$$ETx(k,d) = kEelec + kd^{\alpha} = KE_{elec} + KE_{fs}\, d^2 \;\; d < d_0 \quad (1)$$

The electronics energy, *Eelec*, depends on factors such as digital coding and modulation, whereas the amplifier energy $KE_{fs}\, d^2$ depends on the transmission distance and the acceptable bit-error rate. To receive this message, the radio consumes energy as follows:

$$ER_x = KEelec \quad (2)$$

Mean energy consumed by node 'i' is:

$$MEi = e_s + e_g * r_i + e_r * r_j + r_i * e_{ij} \quad (3)$$

where, e_s is the energy for sensing, e_g is the energy for generating the data, r_i and r_j are the rates of interference and e_{ij} is the energy needed for transmitting the single bit data from the ith node to the jth node. Residual Energy (RE) is calculated as follows:

$$RE = E_{initial} - ME_i \quad (4)$$

Where, *Einitia*l is the initial energy and *MEi* is the mean energy consumed by node 'i'.

B. DEACP protocol:

One of the important factors to improve lifetime of wireless sensor network is the design of network.
In this section we describe the proposed *DEACP* approach. The *DEACP* approach utilizes adaptive clustering scheme. A clustering scheme is called an adaptive scheme, if over time, the number of clusters varies and the nodes membership evolves. In *DEACP*, the BS is assumed to have unlimited energy residues and communication power. It is also assumed that the BS is located at a fixed position, either inside or away from the sensor field.
Nodes with special high "Cch" condition can act as CH to burden the pressure of data transmission. In order to prevent early death due to excessive energy expenditure,

all nodes should be alternately take turns to become CH, CH election need to consider many factors , In *DEACP* the following factors is considered: node-weight, residual energy, condition distance between nodes, condition distance with SB.
The *DEACP* protocol achieves a good distribution of clusters (unresolved problem with many protocols). The *DEACP* protocol takes place in "rounds" that represent time intervals determined in advance. Each round is consisting of four phases, the initialization phase, phase decision, phase of group formation and phase transmission.
Proposed clustering algorithm is divided into three phases; initialization, setup phase and steady state. Our proposed approach is explained using these three phases:

1) INITIALIZATION PHASE:

Each node in order to calculate the node-weight sends a message "Discov-neigh-msg" which contains its identifier. Any node receiving the message sends immediately a "Discov-neigh-msg" message the same type, then each member has its neighbors table, allowing him to know its cost is that the size of the latter.

2) DECISION PHASE:

The most important part of each clustering scheme is the cluster-head election. For the cluster-head election, the proposed *DEACP* uses a hybrid scheme of residual energy and distance among the cluster-heads, distance between node and Bs, Weight- node. The cluster-head election phase is done in two steps: the local competition and the distance condition.

- *Local competition*: The most important part of sensor network protocol is the selection of the head; therefore the design should be such that the node selected as head must dissipate the minimum power for transmission. In our proposed method the nodes compete in a competition scheme to be elected as the cluster head candidate. At first the condition of each node being selected as the cluster-head candidate is found. To do so, this condition "*Cch*", is determined proportional to the remainder energy of node "Ni" as:

$$\beta 1(i,j) = 1 - \alpha 1\left(1 - \frac{DBs,i}{DBs,j}\right)$$

$$\beta 2(i,j) = 1 - \alpha 2\left(1 - \frac{Er,i}{Er,j}\right)$$

$$\beta 3(i,j) = 1 - \alpha 3\left(1 - \frac{Nw,i}{Nwj}\right)$$

$$Cch(i,j) = Max[\,1 - \sum_{i,j=0}^{n} \beta 1(i,j), \beta 2(i,j), \beta 3(i,j)]$$

Cch(i,j): condition to be cluster head for node «CH ».
DBs,i: the distance between the node "i" and the base station Bs.
DBs,j : the distance between the node neighbor "j" and the base station Bs.
Er,i: Residual energy of node « i ».
Er,j: Residual energy of node neighbor « j ».
Nwi: Node-weight of node « i ».
Nwj: Node-weight of node neighbor « j ».
α1, α2, α3: constant coefficient « 0 » or « 1 ».

Each sensor node (Ni) in the network calculates its condition Cch, and then broadcasts a message to the other nodes; called CH-ADV. this message includes the node ID and the value of condition Cch. In the proposed competition scheme, we define a competition range called Rcomp; *Cluster range (cluster radius*), RcompThis parameter specifies the radius of a cluster, i.e., the farthest a node inside a cluster can be from the clusterhead. The cluster radius is a system parameter and is fixed for the entire network, this range should be reasonable, that it is should not be too long to overload the network and should not be too short to increase the number of cluster-head candidate advertisements.
The node i wait for t_{wait} seconds and receives the message Cch-msg from all its neighbors. Then it compares its condition value with that of its neighbors.
If it found its condition Cch greater than Cch value of all its neighbors, then it elects itself as cluster-head candidate. Else, it sends a join message to the neighbor that has the highest condition Cch to become a member of the cluster.
The number of selected cluster heads varies according to the specified cluster radius. The smaller the radius, the larger the required number of cluster heads to fully cover the entire network
The pseudo code protocol:

The pseudo algorithm of DEACP is given as follow:

1. *Phase 1: "Initialization phase"*
2. *Each sensor node Ni;*
3. *Broadcast "Disc-neigh-msg";*
4. *wait for twait seconds to receive message;*
5. *IF the "Disc-neigh-msg" message is received THEN*
6. *Add identifier "id" and send msg "Disc-neigh-msg";*
7. *End IF*
8. *Create table of neighbors;*
9. *Weight-node=size of neighbors table;*
10. *wait for twait seconds;*
11. *Phase 2: "Decision phase"*
12. *Each sensor node sends msg {WN, Er, Dbs};*
13. *Each sensor calculates the Cch(i);*
14. *Evaluate Cch (i);*
15. *For each node neighbor*
16. *If* $\forall i; Cch(i) > Cch(j)$ *Then // Cch(i) is great than all condition Cch of nodes*
17. *Ni ← CH(i): // Ni is the cluster head*
18. *Else*
19. *Nj ← CH(j); // Nj is the cluster head*
20. *EndIF*
21. *Cluster head "Ch" is defined*
22. *Phase 3: "Steady phase"*
23. *Cluster head node fixes and sends a TDMA to all its cluster members.*

24. *Each cluster member sends its data packets in allocated TDMA s*
25. *CH collects the data from all the nodes in its cluster multihope communication inter-cluster r.*
26. *CH transmits the data with multihope communication inter-cluster to the Sink node.*

3) STEADY PHASE:

After the cluster formation phase, the cluster-heads establish a TDMA protocol, and send a time schedule to each of their members. In this protocol, a transmission time slot is assigned to each node, during which the nodes can send their messages.

The TDMA (Time Division Multiple Access) based scheduling protocols make the nodes to be in inactive mode, until their allocated time slots. The TDMA based protocols [1] are designed such that the shortest path for communication will be found out and only a particular link will be in wake up mode for a transmission. "Fig1, Fig.2"

Any scheduling protocol for the WSN for Medium Access Control (MAC) should have the following: Narrowband Modulation techniques, Good throughput efficiency, and moderately low Data transfer rate. Less Hardware complexity. Low access delay, low transmission delay and low overhead. The TDMA based scheduling allocates separate time slot for each node to access the medium to send the sensed data or to forward the aggregated data.

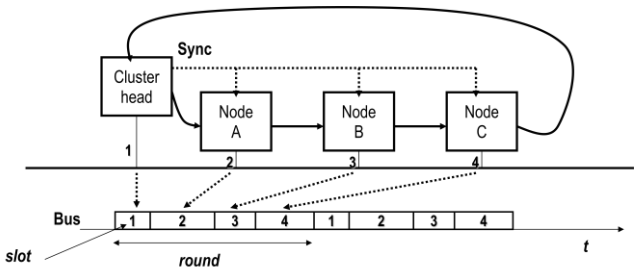


Fig. 1. Slots et round

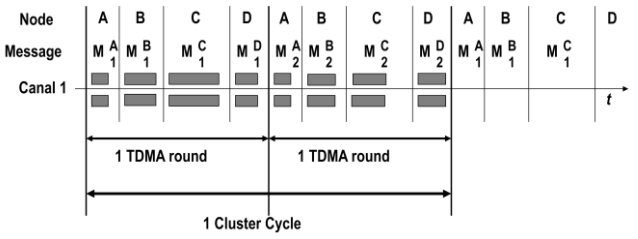


Fig.2 Cluster Cycle

- DEACP NETWORK TRANSMISSION TIME (NTT):

One the clusters are formed and TDMA schedule is fixed in all clusters of the network. In NTT all nodes send their data to their CCHs, in assigned time slots. Cluster-heads receive the data from its cluster and aggregate the data. Data aggregation is key technique in order to compare data amount. Cluster-heads only send meaningful information to BS in order to prolong the battery lifetime.

From a routing perspective, clustering allows to split data transmission into intra-cluster (within a cluster) and inter-cluster (between cluster-heads and every cluster-head and the sink) communication. This separation leads to significant energy saving since the radio unit is the major energy consumer in a sensor node. In fact, member nodes are only allowed to communicate with their respective clusterhead, which is responsible for relaying the data to the sink with possible aggregation and fusion operations. Moreover, this separation allows to reduce routing tables at both member nodes and cluster-heads in addition to possible spatial reuse of communication bandwidth.

The sink is usually located far away from the sensing area and is often not directly reachable to all nodes due to signal propagation problems. A more realistic approach is multi-hop inter-cluster routing that had shown to be more energy efficient. "Fig.3" In DEACP protocol, sensed data are relayed from one clusterhead to another until reaching the sink. Inter-cluster communication in several proposals is achieved through organizing the cluster-heads in a hierarchy Fig. allows better energy distribution and overall energy consumption. However, maintaining the hierarchy could be costly in large and dynamic networks where nodes die as soon as their y supply is completely discharged.

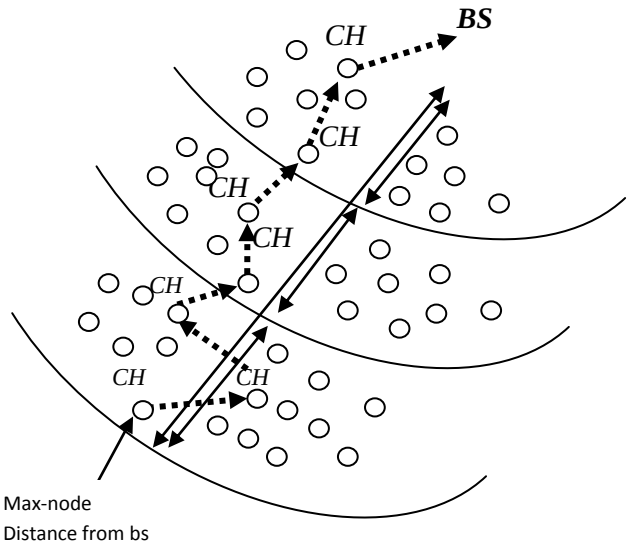


Fig.3 Multihop communication in WSN

The pseudo algorithm of Inter cluster routing is given as follow:

1. For each (level i)
2. for each CH
3. CH receives the data from the cluster member
4. Aggregate the data.
5. If (i ==1)
6. CH transmits data to the BS.
7. Else
8. CH broadcasts data in the next level.
9. End if
10.End for
11.End for

DEACP uses Intra-cluster communication is contention less using TDMA. Each parent node polls its direct children and forwards the data to its parent node until the data reaches the clusterhead. The parent node may combine several data packets from its children together with its own data into one packet.

IV. PERFORMANCE EVALUATION

In this section we evaluate the performance of the proposed DEACP via several simulation experiments. At first the simulation setup is explained and then the results are presented.

A comparison between the simulation results in DEACP, DEEAC, and PEGASIS algorithms is performed via NS2 simulator. We use two scenarios for simulations. In the first scenario, 100 nodes are uniformly and randomly dispersed in a field of size 100m_100m. To study the effect of scale on the performance of DEACP, in the second scenario, 200 nodes are uniformly and randomly dispersed in a field of size 200m_200m.We assume that the BS is located at the center of the field. The other simulation parameters are summarized in TABLE1.

TABLE1- Parameters of simulation

Parameter	Value
Area	100m X 100m
Data packet size	4000 bits
Control packet size	512 bits
Number of sensor nodes	100/200
Initial energy	2 Joule
Base station location	(50,50)
Distance d_0	87m
E_{elec}	50nj/bit

- ENERGY CONSUMPTION OF SENSOR NODES:

Fig.4 shows the results for the energy consumed by sensor nodes in DEACP, DEEAC and PEGASIS protocols The energy consumed by sensor nodes for each round in DEACP is much lower than that in DEEAC and PEGASIS. According to the data presented in this figure, DEACP has less energy consumption than the other two protocols, because this protocol periodically selects cluster heads according to a hybrid of their residual energy distance between node and BS, weight-node: number of neighbor nodes , such as node proximity to its neighbors or node degree. The main reason for this result is the suitable number and distribution of the clusters in the network. As expected, PEGASIS has variant energy consumption, relevant to the pendulous number of its clusters in consecutive rounds. Although, DEEAC has distributed clusters across the network properly, as the number of clusters in DEEAC is large, energy consumption in the whole network increases. Therefore, DEACP has the lowest energy consumption among the two protocols and has more energy consumption in contrast with the other protocols.

Other main reason, DEACP uses a multihop communication inter-cluster and intra cluster. Each parent node polls its direct children and forwards the data to its parent node until the data reaches the clusterhead, and a multihop communication between cluster-head and base station.

- LIFE TIME FOR WIRELESS SENSOR NETWORK

The network lifetime for three protocols is depicted in figures 1.The result between the number of nodes alive and the number of rounds is shown by "Fig.5, Fig.6". The result obtained by measuring of time until the first node dies to time until the last node dies appear the DEACP has a better lifetime than the other protocols DEEAC and PEGASIS because the DEACP method elects the nodes with the highest condition cluster head 'Cch' Also, in this approach, the load balancing in the network is performed properly, which provides a longer time between the beginning of the operations until the time the first node dies.

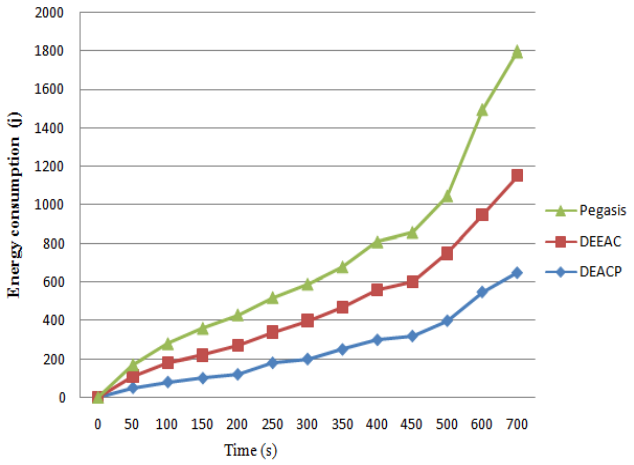


Fig.4 Comparison of Energy consumption

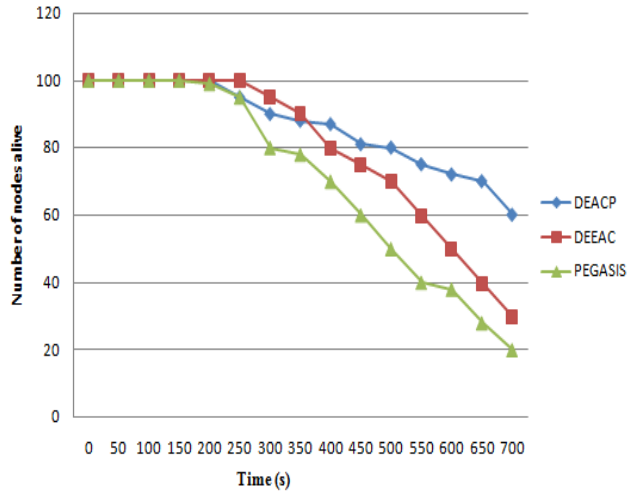


Fig.5 Comparison of network life time with 100 nodes

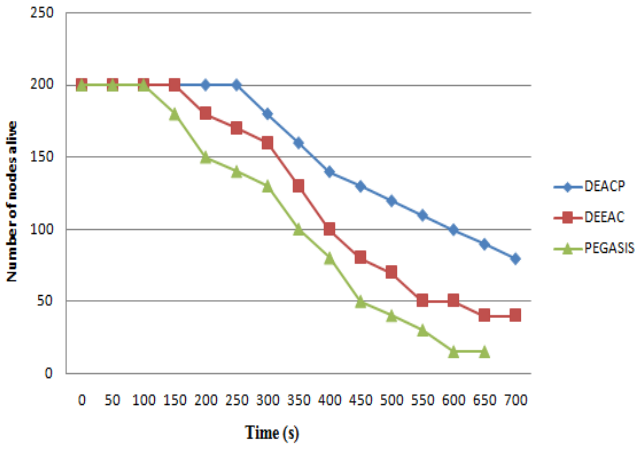


Fig.6 Comparison of network life time with 200 nodes

V. CONCLUSION AND FUTURE WORK

In this paper, we proposed a Distributed Energy Efficient Adaptive Clustering Protocol with Gathering Data for Wireless Sensor Networks. DEACP operates with only realistic assumptions. The algorithm constructs multilevel clusters and the nodes in each cluster reach the clusterhead by relaying through other nodes. DEACP is well distributed, which is a major advantage in a power constrained sensor network. In this method the distance among the cluster-heads has been utilized to reach a well-distributed clustered WSN with suitable size clusters.

Our simulations demonstrated that DEACP generates well balanced clusters. Both intra-cluster and inter-cluster energy consumption is greatly improved over clusters generated by DEACP algorithm. In the future, in addition to the energy-efficiency, we will try to design the DEACP approach in a way that it meets the other WSN requirements, like the full coverage of the monitored area? And involves real-time implementation using wireless sensor motes and secure routing protocol in terms of memory.

REFERENCES:

[1]. Ataul Bari, Shamsul Wazed, Arunita Jaekel, Subir Bandyopadhyay.: A genetic algorithm based approach for energy efficient routing in two-tiered sensor networks, Ad-Hoc Networks 7, pp 665-676 (2009).

[2]Younis O, Fahmy S. HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Trans on Mobile Computing*, 2004, 3(4), p. 366–379.

[3]. Bara'a A.Attea, Enan A.Khalil.: A new evolutionary based routing protocol for clustered heterogeneous wireless sensor networks, Applied Soft Computing.J.(2011).

[4] An Energy-Aware Routing Protocol in Wireless Sensor Networks Ming Liu 1, Jiannong Cao 2, Guihai Chen 3 and Xiaomin Wang Sensors 2009.

[5]Al-Karaki JN, Ul-Mustafa R, Kamal AE. Data aggregation in wireless sensor networks-exact and approximate algorithms. *In Proc of Workshop on High Performance Switching and Routing*, 2004, p. 241–245.

[6] D. J. Baker and A. Ephremides, "The Architectural Organization of a Mobile Radio Netwrok via a Distributed Algorithm", IEEE Transactions on Communications, Vol. 29,

[7] A. Ephremides, J. E. Wieselthier and D. J. Baker, "A Design concept for Reliable Mobile Radio Networks with Frequency Hopping Signaling", Proceedings of IEEE, vol. 75, No.

[8] A. K. Parekh, "Selecting Routers in Ad-Hoc Wireless Networks", in Proceedings of ITS, 1994.

[9] G. Smaragdakis, I. Matta, A. Bestavros. 2004. SEP: A Stable Election Protocol for clustered heterogeneous wireless sensor networks. In Second International Workshop on Sensor and Actor Network Protocols and Applications (SANPA).

[10] A. D. Amis, and R. Prakash, "Load-Balancing Clusters in Wireless Ad Hoc Networks", in Proceedings of ASSET 2000, Richardson, Texas, Mar. 2000.

[11] M. Chatterjee, S. K. Das, and D. Turgut, "WCA: A Weighted Clustering Algorithm for Mobile Ad Hoc Networks", Cluster Computering, pp. 193-204, 2002

[12] S. Basagni, "Distrbuted Clustering for Ad Hoc Networks", in Proceedings of International Symposium on Parallel Architectures, Algorithms and Networks, pp. 310-315 2002

[13]. Bara'a A.Attea, Enan A.Khalil.: A new evolutionary based routing protocol for clustered heterogeneous wireless sensor networks, Applied Soft Computing.J.(2011), doi: 10.1016/j.asoc.2011.04.007.

[14] Fuad Bajaber, Irfan Awan.: Adaptive decentralized re-clustering protocol for wireless sensor networks, Journal of computer and Systems sciences, doi:10.1016/j.jcss.2010.01.007.

[15]Heinzelman WB, Chandrakasan AP, Balakrishnan H. Energy-efficient communication protocol for wireless microsensor networks. *In Proc of IEEE Conf on System Sciences*, 2000, p. 3005–3014.

[16] Younis O, Fahmy S. HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Trans on Mobile Computing*, 2004, 3(4), p. 366–379.

[17] J. Sinha and S. Barman. Energy e_cient routing mechanism in wireless sensor network. In Recent Advances in Information Technology (RAIT), 2012 1st International Conference on, pages 300 –305, march 2012.

[18] R. Soua and P. Minet. A survey on energy e_cient techniques in wireless sensor networks. In Wireless and Mobile Networking Conference (WMNC), 2011 4th Joint IFIP, pages 1 –9, oct. 2011.

[19] J. Stankovic and T. He. Energy management in sensor networks. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 370(1958):52–67, 2012.

[20] J. Suhonen, M. Kohvakka, V. Kaseva, T. D. H¨am¨al¨ainen, and M. H¨annik¨ainen. Low-Power Wireless Sensor Networks - Protocols, Services and Applications. Springer, Berlin, Heidelberg, 2012. aufl. edition, 2012

Amélioration du modèle de filtrage collaboratif hybride relatif au système de recommandation

Abderrahmane Kouadria,
Département d'Informatique
Université Ibn Khaldoun Tiaret – Algérie
ab_kouadria@esi.dz

Omar Nouali,
Division théorie & ingénierie des systèmes informatiques,
CE.R.I.S.T.
Rue des 3 frères Aïssiou, Ben Aknoun, Alger, Algérie
onouali@cerist.dz

***Résumé*— Les systèmes de recommandation sont des techniques utiles pour résoudre le problème de la surcharge d'information. Le Filtrage Collaboratif (FC) représente l'une des techniques de recommandation les plus populaires, dont la plupart de leurs méthodes fondent leur approche sur la prédiction de notes pour générer les recommandations. Beaucoup d'algorithmes récents ont été développés pour améliorer les divers aspects du filtrage collaboratif, cependant, la plupart d'entre eux ne prennent pas en compte les données communes des utilisateurs et les informations des items ou leurs caractéristiques. Dans cet article, nous proposons une nouvelle amélioration du système de recommandation basé sur le modèle de filtrage collaboratif hybride qui combine entre l'algorithme de FC basé-item et l'algorithme de FC basé-utilisateur, dont les similarités entre les items et les utilisateurs ont été modifiés et les évaluations communes prennent totalement l'avantage de la matrice utilisateur-item, pour modifier le calcul de la similarité et la prédiction des notes. Les résultats expérimentaux montrent la supériorité de notre proposition par rapport aux autres approches de l'état de l'art.**

***Mots-Clés*— Système de Recommandation, Filtrage Collaboratif.**

I. INTRODUCTION

Face à la très grande masse d'information disponible sur le Web, il reste très difficile pour un utilisateur de trouver l'information qui l'intéresse. Au lieu de laisser ce dernier dépenser son temps à chercher l'information pertinente, les recherches actuelles visent à concevoir des outils qui permettent de lui faciliter cette tâche, en lui faisant parvenir continuellement l'information souhaitée. Les systèmes de recommandation (SR) ont été largement adoptées et sont devenus progressivement important pour soulager le problème de surcharge de l'information et de fournir aux utilisateurs des informations et des services qui conviennent. Les systèmes de recommandation sont des applications logicielles utilisés pour suggérer les bons items (produits ou services) à des utilisateurs particuliers en fonction de leurs préférences explicites et implicites [4, 9]. Ils sont couramment utilisés par des sites Web de commerce électronique pour aider les consommateurs à prendre leur décision d'achat. Leur utilisation peut augmenter les ventes de e-commerce par: (a) la conversion de visiteurs en acheteurs, (b) l'augmentation des ventes croisées en proposant des produits supplémentaires et fidélisation de la clientèle, (c) la création d'une relation de valeur ajoutée entre le site et le client [13].

Actuellement, la recommandation personnalisée représente le domaine le plus attractif pour l'utilisateur et le plus intéressant économiquement pour le fournisseur de recommandations.

Plus concrètement, la recommandation personnalisée a pour objectif de filtrer des articles afin de ne conserver que les plus pertinents pour un utilisateur donné. L'idée principale est de prédire l'opinion qu'un utilisateur portera sur les articles qu'il ignore, afin de ne lui proposer que ceux qu'il est susceptible d'apprécier, ou tout du moins, qui auront une grande chance de l'intéresser.

Les systèmes de recommandation sont généralement classés en trois catégories en fonction de leur approche de recommandation: approche à base de contenu [11], collaboratif [14], et hybride [7].

Dans les prochaines sections, nous résumons l'état de l'art et les méthodes clés liés à notre proposition, à savoir : algorithme de FC basé-utilisateur et basé-item. Ensuite, nous présentons notre proposition ainsi que la validation expérimentale, et enfin, nous terminons par une conclusion et une présentation des perspectives de ce travail.

II. TRAVAUX CONNEXES

Le filtrage collaboratif est l'une des approches les plus populaires des systèmes de recommandation. Leur principe est de suggérer de nouveaux articles ou de prédire l'utilité des articles inconnus pour un utilisateur donné, en se basant sur les opinions et les évaluations d'un groupe d'utilisateurs afin de proposer un ou plusieurs articles. Jusqu'à présent, la majorité des méthodes de filtrage collaboratif fondent leur approche sur la prédiction de notes. Elles prennent toutes les notes disponibles en entrée et leur but étant de prédire les notes

inconnues en sortie. Le système propose à l'utilisateur actif les articles dont les notes estimées sont les plus élevées.

Selon Breese, Heckerman et Kadie [17], les algorithmes de filtrage collaboratif peuvent être regroupés en deux catégories : Les algorithmes basées mémoire, et les algorithmes basées modèle.

Les algorithmes basées mémoire consistent à utiliser la totalité des informations disponibles pour générer les prédictions, les exemples les plus analysés dans ses algorithmes comprennent des approches basés-utilisateurs [6, 14] et basées-items.

Les algorithmes basés modèle utilisent la base de données des évaluations des utilisateurs pour créer ou apprendre un modèle de prédiction via un processus d'apprentissage, telles que : le clustering, les réseaux bayésiens, les arbres de décision, le modèle Latent semantic et factorisation matricielle [5].

A. Algorithme de FC basé-utilisateur :

Dans cet algorithme la prédiction d'un utilisateur évaluant un item est fondée sur les estimations, sur cet item, des plus proches voisins. Donc une mesure de similarité entre les utilisateurs doit être définie avant qu'un ensemble de voisins les plus proches ne soit choisi. De plus, une méthode pour combiner les évaluations de ces voisins sur l'item prévu doit être choisie [16].

Etape 1: Calcul de similarité :

Il existe plusieurs façons pour calculer la similarité entre les items et les utilisateurs, parmi lesquels : la corrélation de Pearson et la similarité vectorielle. Dans cet article, nous choisissons la corrélation de Pearson.

La corrélation de Pearson est une méthode très utilisée dans le domaine des systèmes de recommandation pour mesurer la similarité entre deux utilisateurs. Car elle peut être facilement mis en œuvre et peut atteindre une grande précision lors de la comparaison avec d'autres méthodes de calcul de similarité [8,18]. La similarité est calculée par la formule suivante :

$$Sim(a,u) = \frac{\sum_{i \in I}(R_{a,i} - \bar{R}_a)(R_{u,i} - \bar{R}_u)}{\sqrt{\sum_{i \in I}(R_{a,i} - \bar{R}_a)^2}\sqrt{\sum_{i \in I}(R_{u,i} - \bar{R}_u)^2}} \quad (1)$$

Où *Sim(a, u)* désigne la similarité entre l'utilisateur actif a et l'utilisateur *u*. *I* l'ensemble d'items ayant été noté à la fois par l'utilisateur *a* et *u* , $R_{a,i}$, $R_{u,i}$ est l'estimation de l'item *i* par l'utilisateur *a* et l'utilisateur *u* respectivement ; $\bar{R}_a$, $\bar{R}_u$: moyenne respective des notes des utilisateurs *a* et *u*.

Etape 2: Sélection des voisins similaires :

La sélection des voisins similaires est une étape très importante dans la prédiction des notes manquantes. Il existe deux alternatives: seuil de corrélation [3] qui s'agit de sélectionner seulement les utilisateurs dont la similarité avec l'utilisateur actif dépasse un seuil donné et le nombre maximum de voisins [18] qui s'agit de sélectionner *k* utilisateurs qui sont les plus semblables à l'utilisateur actif, où *k* est un paramètre de l'algorithme. Dans ce document, nous utilisons le nombre maximum de voisins.

Étape 3: La prédiction de note d'un item *j* évaluée par l'utilisateur *u* est calculée à l'aide de la somme pondérée des estimations des voisins les plus proches qui ont déjà estimé l'item *j* :

$$P_{u,i}(user) = \bar{R}_u + \frac{\sum_{k=1}^{K} sim(u,k)X\left(R_{k,i} - \bar{R}_k\right)}{\sum_{k=1}^{K}(|sim(u,k)|)} \quad (2)$$

Où *K* : le nombre d'utilisateurs présents dans le voisinage de *u*, ayant déjà noté sur l'item *i* ; $R_{k,i}$: la note de l'utilisateur *k* pour l'item *i*; $\bar{R}_u, \bar{R}_k$: moyenne respective des notes des utilisateurs *u et* k.

B. Algorithme de FC basé-item :

Les Méthodes basées-item, telles que [10, 12] sont similaires aux approches basées-user, la différence est que les méthodes basées-item emploient la similarité entre les articles à la place des utilisateurs. La similarité entre deux items est calculée par la formule suivante :

$$Sim(i,j) = \frac{\sum_{u \in U_i}(R_{u,i} - \bar{R}_i)(R_{u,j} - \bar{R}_j)}{\sqrt{\sum_{u \in U_i}(R_{u,i} - \bar{R}_i)^2}\sqrt{\sum_{u \in U_j}(R_{u,j} - \bar{R}_j)^2}} \quad (3)$$

Où *Sim (i, j)* est la similitude entre l'item *i* et *j*, U_i l'ensemble d'utilisateurs qui ont notés pour les 2 items *i* et *j*, $R_{u,i}$, $R_{u,j}$: l'évaluation de l'utilisateur *u* sur l'itemt *i* et l'itemt *j* respectivement. $\bar{R}_i$, $\bar{R}_j$: la moyenne respective des notes des items *i* et *j*.

Une fois que la similarité parmi les items a été calculée, la prochaine étape est de prédire pour l'utilisateur cible *u*, une valeur pour l'article actif *i*. Une manière commune est de capturer comment l'utilisateur *u* évalue les items similaires. La valeur prévue est basée sur la somme pondérée des estimations de l'utilisateur ainsi que les déviations des estimations moyennes et peut être calculée à l'aide de la formule suivante :

$$P_{u,i}(item) = \bar{R}_i + \frac{\sum_{k=1}^{K} sim(i,k)X\left(R_{u,k} - \bar{R}_k\right)}{\sum_{k=1}^{K}(|sim(i,k)|)} \quad (4)$$

Où *K* : le nombre d'items présents dans le voisinage de item *i*, ayant déjà été noté par l'utilisateur *u* ; $R_{u,k}$: la note de

l'utilisateur *u* pour l'item *i*; $\bar{R}_i$, $\bar{R}_k$: moyenne *respective* des notes pour les items *i et k*.

C. Algorithme de FC hybride :

Modèle FC hybride combine l'algorithme de FC basé-item avec l'algorithme de FC basé-utilisateur, et deux matrices de similarités (matrice de similarité item et matrice de similarité utilisateur) contribuer à la prédiction de l'évaluation finale. Deux relations de voisinage, la relation utilisateur-utilisateur et de la relation item-item, sont conservées dans le même temps. Cette combinaison de ces deux algorithmes permet non seulement de récompenser une précision de prédiction améliorée, mais aussi pour apporter une plus grande robustesse au système de recommandation pour le problème de rareté [1].

La première étape de cet algorithme consiste à calculer la similarité entre les utilisateurs selon la formule (1) et construire une matrice de similarité utilisateur, de la même façon on doit calculer la similarité entre les items (formule 3) et de construire matrice de similarité item.

Étape 2: Pour l'item actif, on doit sélectionner les k items plus proches voisins respectivement pour l'utilisateur actif.

Étape 3: La prédiction de note manquante est calculée par une somme pondérée de la combinaison des évaluations des k proches voisins de l'item et celles de l'utilisateur. Cette prédiction est définie comme suit:

$$p_{u,i}(hybrid) = \alpha\, X\, p_{u,i}(user) + (1-\alpha)\, X\, p_{u,i}(item) \quad (5)$$

Où α est un paramètre de réglage, $p_{u,i}(user)$ et $p_{u,i}(item)$ calculés respectivement par les équations (2) et (4).

III. L'approche proposée

Dans cet article, nous proposons un filtrage collaboratif amélioré basé sur FC hybride qui intègre l'algorithme basé-item et l'algorithme basé-utilisateur dans un cadre unifié, où les similarités entre les items et les utilisateurs ont été modifiés et les évaluations communes prennent totalement l'avantage de la matrice utilisateur-item, pour modifier le calcul de la similarité et la prédiction de notes.

A. Calcul modifié de la prédiction :

Afin d'améliorer les divers aspects du filtrage collaboratif, plusieurs algorithmes récents ont été développés, mais beaucoup d'entre eux ne prennent pas en compte les données communes des utilisateurs et les informations des items ou leurs caractéristiques. Wang et al [2] ont proposés d'ajouter une pondération qui modifie le calcul de la similarité pour l'approche basée-item et la prédiction des notes. De la même façon, nous appliquons cette pondération pour l'approche basée-utilisateur. Cette pondération permet d'améliorer la précision et la qualité de la prédiction.

En prenant comme exemple la matrice utilisateur-item (tableau 1), nous aurons maintenant besoin de prévoir la valeur de user3 sur item2. En conséquence, nous calculons la similarité entre les utilisateurs avec l'équation (1).

	Item1	**Item2**	**Item3**	**Item4**	**Item5**	$\bar{R}_u$
User1	3	5				4
User2	5	4		3	3	3.75
User3	4	**?**	5	2	4	3.75
User4		4			3	3.5
User5	3	2	4		5	3.5
$\bar{R}_i$	3.75	3.75	4.5	2.5	3.75	

Tableau1 : *Matrice utilisateur-item*

$$Sim(3,1) = \frac{(4-3.75)(3-4)}{\sqrt{(4-3.75)^2}\sqrt{(3-4)^2}} = 1$$

$$Sim(3,2) = \frac{(4-3.75)(5-3.75)+(2-3.75)(3-3.75)+(4-3.75)(3-3.75)}{\sqrt{(4-3.75)^2+(2-3.75)^2+(4-3.75)^2}\sqrt{(5-3.75)^2+(3-3.75)^2+(3-3.75)^2}} = 0.491$$

$$Sim(3,4) = \frac{(4-3.75)(3-3.5)}{\sqrt{(4-3.75)^2}\sqrt{(3-3.5)^2}} = 1$$

$$Sim(3,5) = \frac{(4-3.75)(3-3.5)+(5-3.75)(4-3.5)+(4-3.75)(5-3.5)}{\sqrt{(4-3.75)^2+(5-3.75)^2+(4-3.75)^2}\sqrt{(3-3.5)^2+(4-3.5)^2+(5-3.5)^2}} = 0.406$$

Enfin vous découvrirez que la similarité entre user3 et user1 ou user4 est plus élevée que la similarité entre user3 et user2 ou user5, ce qui n'est pas logique, car user3 et le user2 ou user5 devraient avoir plus de similarité pour prédire note de user3 sur item2, qui ont le nombre de notes plus fréquent que user1 ou user4. Dans ce cas, nous croyons que lorsque nous prévoyons le taux de similarité de l'utilisateur actif, nous devrions prendre en compte le nombre de notes communes entre chaque utilisateur comme un poids.

	User 1	**User2**	**User3**	**User4**	**User5**
User1	0	2	1	1	2
User2	2	0	3	2	3
User3	1	**3**	0	1	3
User4	1	2	1	0	2
User5	2	3	3	2	0

Tableau2 : *Matrice utilisateur-utilisateur du nombre de notes communes*

Nous calculons tout d'abord le nombre de notes communes de chaque utilisateur avec les autres et formons une matrice utilisateur-utilisateur de nombres d'évaluations communes entre les utilisateurs (tableau 2). Nous avons proposé une équation qui améliore le calcul de similarité entre les utilisateurs de la manière suivante:

$$W_{user}(a,u) = \begin{cases} 0 & Si\ a = u \\ \frac{|U_a \cap U_u|}{\max(|U_a \cap U_{all}|, |U_u \cap U_{all}|)} & Si\ a \neq u \end{cases} \quad (6)$$

Où $w_{user}(a,u)$ est un poids de l'utilisateur *a* et l'utilisateur *u* pour modifier le calcul de similarité basé-utilisateur; $|U_a \cap U_u|$ est le nombre de notes communes entre l'utilisateur *a* et l'utilisateur *u*; $|U_a \cap U_{all}|$ est le nombre de notes communes entre l'utilisateur *a* et les autres utilisateurs, $|U_u \cap U_{all}|$est le nombre de notes communes entre l'utilisateur *u* et les autres utilisateurs.

Avec la formule (6) nous calculons les poids qui modifient la similarité entre les utilisateurs.

$$W_{user}(3,1) = \frac{1}{\max((1,3,0,1,3),(0,2,1,1,2))} = \frac{1}{3}$$

$$W_{user}(3,2) = \frac{3}{\max((1,3,0,1,3),(2,0,3,2,3))} = \frac{3}{3}$$

$$W_{user}(3,4) = \frac{1}{\max((1,3,0,1,3),(1,2,1,0,2))} = \frac{1}{3}$$

$$W_{user}(3,5) = \frac{3}{\max((1,3,0,1,3),(2,3,3,2,0))} = \frac{3}{3}$$

Et après avoir obtenu la valeur de $W_{user}(3,1)$, $W_{user}(3,2)$, $W_{user}(3,4)$, $W_{user}(3,5)$, nous pouvons modifier la similarité entre les utilisateurs, et en fait, user2 et user5 sont plus proche de user3 que user1 et user4.

$$W_{user}(3,1)\ X\ Sim(3,1) = 1\ X\ \frac{1}{3} = 0.333$$

$$W_{user}(3,2)\ X\ Sim(3,2) = 0.491\ X\ \frac{3}{3} = 0.491$$

$$W_{user}(3,4)\ X\ Sim(3,4) = 1\ X\ \frac{1}{3} = 0.333$$

$$W_{user}(3,5)\ X\ Sim(3,5) = 0.406\ X\ \frac{1}{3} = 0.406$$

Nous procédons de la même manière :

- Formation d'une matrice item-item qui contient les nombres de notes communes de chaque item avec les autres items. (tableau 3).

	Item1	Item2	Item3	Item4	Item5
Item1	0	3	2	2	3
Item2	3	0	1	1	3
Item3	2	**1**	0	1	2
Item4	2	1	1	0	2
Item5	3	3	2	2	0

Tableau 3 : *Matrice item-item du nombre de notes communes*

- Calcul des poids qui modifient les similarités entre les items par la formule suivante:

$$W_{item}(i,j) = \begin{cases} 0 & Si \quad i = j \\ \frac{|I_i \cap I_j|}{max(|I_i \cap I_{all}|, |I_j \cap I_{all}|)} & Si \quad i \neq j \end{cases} \quad (7)$$

Où $w_{item}(i,j)$ est un poids de l'item *i* et l'item *j* pour modifier le calcul de similarité basé-item; $|I_i \cap I_j|$ est le nombre de notes communes entre l'item *i* et l'item *j*; $|I_i \cap I_{all}|$ est le nombre de notes communes entre l'item *i* et les autres items, $|I_j \cap I_{all}|$est le nombre de notes communes entre l'item *j* et les autres items.

Nous prévoyons un utilisateur à moindre évaluations ou un nouvel utilisateur, le nombre maximum de notes communes de l'utilisateur actif avec d'autres utilisateurs n'est que très peu. Dans ce cas, la similarité sera une grande valeur, mais il fera l'inexactitude de prédiction si nous prenons cette similarité pour prédire la note. Donc, nous utilisons les fonctions de poids (6 et 7). Et le calcul de la prédiction de note manquante sera une somme pondérée de la combinaison des évaluations des k proches voisins de l'item et celles de l'utilisateur. La fonction de prédiction est définie comme suit:

$$p_{u,i}(hybrid) = \alpha\ X\ \left(\bar{R}_u + \frac{\sum_{k=1}^{K} w_{user}(u,k) X sim(u,k) X \left(R_{k,i} - \bar{R}_k\right)}{\sum_{k=1}^{K} (|w_{user}(u,k) X sim(u,k)|)}\right) +$$
$$(1-\alpha)\ X\ (\bar{R}_i + \frac{\sum_{k=1}^{K'} w_{item}(i,k) X sim(i,k) X \left(R_{u,k} - \bar{R}_k\right)}{\sum_{k=1}^{K'} (|w_{item}(i,k) X sim(i,k)|)})$$

IV. Expérimentation et discussion des résultats

Nous utilisons pour l'expérimentation, le jeu de données réelles du système de recommandation de films MovieLens. Ce jeu contient 100 000 évaluations (la note donnée à un film est sur une échelle de 1 à 5), fournies par 943 utilisateurs sur 1682 films. Ainsi, il existe 93,7% de données manquantes (matrice creuse). Nous avons divisé les données en un ensemble d'apprentissage et un ensemble de test. Nous choisissons au hasard 80% des données pour l'apprentissage et les 20% restants pour le test qui est généralement adoptée dans la plupart des littératures.

A. Mesures d'évaluation

Nous utilisons MAE (Mean Absolute Error) dans notre évaluation, qui est également une mesure d'erreur permettant de calculer la moyenne des erreurs absolues entre les prédictions et les valeurs réelles fournies par les utilisateurs. En d'autres termes, elle permet de déterminer la qualité de ces prédictions. Rappelons que les prédictions sont générées à partir d'un ensemble d'apprentissage.

La mesure MAE se calcule de la manière suivante :

$$MAE = \frac{\sum_{i=1}^{N} |p_{ui} - R_{ui}|}{N}$$

Où R_{uj} : Représente la mesure réelle de l'utilisateur u sur l'item i,

p_{ui} : La mesure prédite par le système,

N : Le nombre total de mesures prédites

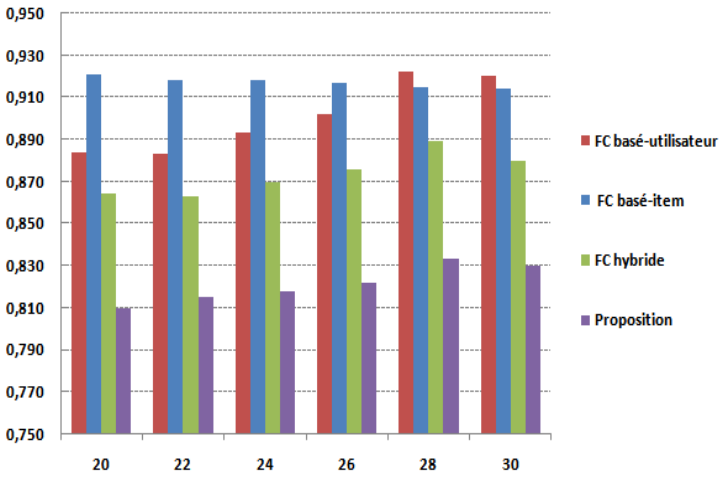


Figure 01: *L'erreur absolue moyenne (MAE) de prédiction de notes de quatre modèles différents de FC avec le paramètre de voisinage k (20 à 30).*

K	FC basé-utilisateur	FC basé-item	FC hybride	Proposition
20	0,884	0,921	0,864	0,810
22	0,883	0,918	0,863	0,815
24	0,893	0,918	0,870	0,818
26	0,902	0,917	0,876	0,822
28	0,922	0,915	0,889	0,833
30	0,920	0,914	0,880	0,830

Tableau 04 : *Comparaison des quatre modèles de FC*

Afin de montrer les performances de notre proposition et sont efficacité, nous l'avons comparé avec les algorithmes: algorithme basé-utilisateur, basé-item et hybride. Nous avons fixé le paramètre α de la fonction de notre proposition à 0,4 avec différents paramètres de voisinage k, tel que comme illustré sur la figure 01. Les valeurs MAE détaillant les quatre modèles de FC sont représentées dans tableau ci-dessus. On constate que notre nouvelle approche améliore de façon significative la qualité de filtrage collaboratif de recommandation, et surpasse les trois autres approches (basé-utilisateur, basé-item et hybride).

V. Conclusion

Dans cet article, nous avons proposé une nouvelle amélioration du filtrage collaboratif hybride qui combine l'algorithme de FC basé-item avec l'algorithme de FC basé-utilisateur, où les deux similarités (similarité entre les items et les utilisateurs) ont été modifiés et les évaluations communes prennent totalement l'avantage de la matrice utilisateur-item, pour modifier le calcul de la similarité et la prédiction des notes. Les résultats expérimentaux ont prouvé la faisabilité de cette proposition afin d'atténuer les problèmes de rareté, tout en réalisant une meilleure précision par rapport aux autres approches de filtrage collaboratif de l'état de l'art.

Pour les travaux à venir, nous prévoyons d'étendre notre approche et l'appliquer sur les systèmes de recommandation multicritère et les comparer avec d'autres approches.

REFERENCES

[1] H. Ji, L. Jinfeng, C. Ren, M. He "Hybrid Collaborative Filtering Model for improved Recommendation" 978-1-4799-0530-0/13/$31.00 ©2013 IEEE.

[2] W. Weijie. Y. Jing, H. Liang " An Improved Collaborative Filtering Based on Item Similarity Modified and Common Ratings," 978-0-7695-4814-2/12 $26.00 © 2012 IEEE.

[3] Buhwan Jeong, Jaewook Lee, Hyunbo Cho. "Improving memory-based collaborative filtering via the updating and prediction modulation". Information Sciences 180, 602-612, 2010.

[4] Q. Shambour and J. Lu, "A framework of hybrid recommendation system for government-to-business personalized e-services," in Proceedings of the 7th International Conference on Information Technology: New Generations, USA, pp. 592-597, 2010

[5] Koren, Y., Bell, R., and Volinsky, C., "Matrix factorization techniques for recommender systems". IEEE Computer, 42, 8, 30-37. 2009.

[6] Ding, S., Zhao, S., Yuan, Q., Zhang, X., Fu, R. and Bergman, L., "Boosting collaborative filtering based on statistical prediction errors". In RecSys '08, 3-10. 2008.

[7] .Burke R., "Hybrid web recommender systems," In The Adaptive Web, Springer Berlin Heidelberg, pp. 377–408, 2007.

[8] Miha Grcar, Dunja Mladenic, Blaz Fortuna and Marko Grobelnik, Data Sparsity Issues in the Collaborative Filtering Framework ,Lecture Notes in Computer Science,Volume 4198,pp58-76, 2006.

[9] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," IEEE Transactions on Knowledge and Data Engineering, vol. 17, pp. 734-749, 2005.

[10] Deshpande, M., and Karypis, G., 2004. Item-based top-N recommendation algorithms. ACM TOIS, 22, 1, 143-177.

[11] Matthew R. McLaughlin, Jonathan L. Herlocker, "Content-based filtering & collaborative filtering: A collaborative filtering algorithm and valuation metric that accurately model the user experience", In Proc. Of the 27th ACM SIGIR Conf, pp. 329-336, 2004.

[12] Sarwar, B., Karypis, G., Konstan, J., and Reidl, J., "Item-based collaborative filtering recommendation algorithms". In WWW ′01, 285-295, 2001.

[13] Schafer J.B., Konstan J. A., and Riedl J., "E-commerce recommendation applications," Data Mining and Knowledge Discovery, 5(1/2): pp.115–153,2001.

[14] Herlocker, J., Konstan, J., Borchers, A., and Riedl, J.," An algorithmic framework for performing collaborative filtering" . In SIGIR ′99, 230-237 . 1999.

[15] B.Sarwar, G.Karypis, J.Konstan, J.Riedl, "Analysis of recommendation algorithms for E-commerce", In: ACM Conference on Electronic Commerce, pp.158-167, 2000.

[16] J. S. Breese, D. Heckerman, and C. Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In Proc. Of UAI, 1998.

[17] J. S. Breese, D. Heckerman, and C. Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence (UAI-98), pages 43–52, July 1998.

[18] U. Shardanand and P. Maes. Social information filtering: Algorithms for automating word of mouth. In Proc. of SIGCHI Conference on Human Factors in Computing Systems, 1995.

An MDA Based Derivation process for e-Learning Software Product Lines

Nesrine LAHIANI
Departement of Computer Science
Saad Dahlab University
Blida, Algeria
lahiani.nesrine@gmail.com

Djamal BENNOUAR
Departement of Computer Science
Akli Mohand OulHadj University
Bouira, Algeria
dbennouar@gmail.com

***Abstract*— Platforms such as Moodle aims to ease and improve the teaching-learning process by means of taking advantage of internet technologies. All existing e-learning platforms are pretty similar the concepts of activity, assignment, deliverable or grade. But also a wide range of differences among them exists. Software Product Line (SPL) has as goal the effective production of similar software systems. .Product derivation represents a fundamental aspect in SPL. It is also the main challenge that SPL faces. Despite its importance, there is only a little research on product derivation compared to the large work on developing product lines. In addition, the few available research reports guidance about how to derive a product from a product line. In this paper we describe a combination of SPL and MDA which both fit perfectly together in order to build applications in cost effective way. We proposed an approach for product derivation that adopts MDA with its organized layers of models to achieve SPL goals.**

***Index Terms* - Software Product Line, Product Derivation, Model Driven Architecture.**

I. INTRODUCTION

A software product line (SPL) is as a set of software-intensive systems that share a common, managed set of features satisfying the specific needs of a particular market segment or mission and that are developed from a common set of core assets in a prescribed way [1]. A feature [2] is a system property or functionality that is relevant to some stakeholder and is used to capture commonalities or discriminate among systems in SPLs.

The SPL approach makes a distinction between domain engineering where a product is derived based on the platform components and application engineering where individual products using the platform artefacts are constructed [6]. The process of creating these individual products from a product line of software assets is known as product derivation [4].

Derivation of a product from an SPL seems to be an easy step since it's relied on reuse. Actually the derivation activity represents one of the main challenges that SPL faces. A number of publications reported clearly the difficulties associated with this activity.

As an example, Deelstra and al. reported in [4] the following assertion: "Contrary to popular belief, deriving individual products from shared software assets is a time-consuming and expensive activity".

Product derivation has been defined in many different ways. In [4] Deelstra, Sinnema, and Bosch define product derivation by, "A product is said to be derived from a product family if it is developed using shared product family artifacts. The term product derivation therefore refers to the complete process of constructing a product from product family software assets" and also by "Product derivation is a key activity in application engineering. It addresses the construction of a concrete product from the product line core assets".

Model Driven Architecture (MDA) [7] defines three layer of software model specification: the CIM layer, the PIM layer and the PSM layer. In the context of MDA, a software system is produced after series of model transformation which starts from the CIM layer model. The CIM model is transformed to a PIM model. This later is finally transformed in a PIM model.

The work presented in this paper deal with an SPL derivation process based on MDA concepts. The main idea is to represent each steps of Application engineering in SPL with a model of MDA starting from requirement engineering until the product implementation. Indeed, there are needs for decisions model and transformations rules to get at the end a running application that satisfy customer's wishes.

The remainder of the paper is structured as follows: we first outline related work (Section 2). We then describe how we combine application engineering and MDA (Section 3), the process of our product derivation approach (Section 4). We finish with case study (Section 5) and finally, conclusions and futures directions are presented (Section 6).

II. RELATED WORK

In [12] Kim and al. proposed an overview of a complete method called Dream stands for DRamatically Effective Application development Methodology, which integrates both SPL engineering and model-driven architecture. DREAM, that adopts the key activities of SPL and model transformation feature of MDA. The process consists of 9

phases, and each phase was specified with work instructions utilizing UML and representation scheme utilizing PIM and PSM of MDA. However, there is little support for the derivation process other than a high level description of the activities required. A similar approach has been proposed by Haugen et al. [10] who present a conceptual model for SPL engineering aligned with MDA and serves as the basis for both modeling and product derivation. To derive a product the process is as follows: first, the "product model" is expressed using Computation Independent Model (CIM), which is the same formalism as the product-line model is defined (modeled in terms of UML 2.0 use cases). Then, a model transformation taking both product and product line models as parameters transform the core assets so that the resulting model, "Product/System Model", correspond to the PIM model of the product (modeled in terms of UML 2.0 composite structures). Finally product implementation is obtained after several refinements at the Product Specific Model or PSM level.

FIDJI is a flexible product derivation process [5], part of an overall model-driven SPL based development methodology [11]. The FIDJI process consists in writing a model transformation, using a set of predefined transformation operations that will reuse core assets' models to build the product. This transformation is written by the product engineer and checked against instantiation constraints. Hence, the FIDJI PD process offers the flexibility required to support product-specific requirements by supporting them via transformation operations while controlling their realization through instantiation constraints.

Ziadi et al. [13] modeled Core assets in terms of UML. Class diagrams are used to represent the static part and sequence diagrams to represent the behavioral part. The decision model which is a set of requirements and engineering decisions that an application engineer must resolve in order to describe and construct a product and determine the extent of variation that is possible among the systems of the domain was defined by Ziadi et al. as a class diagram (as shown in figure 1) which exposes variants as stereotyped elements. Product derivation is formalized by using a UML model transformation. An algorithm is given to derive a static model for a product and an algebraic approach is proposed to derive product-specific statecharts from the sequence diagrams of the product line. Based on product engineer's choices, relevant classes are selected and a model transformation removes unused variants as well as optimizes the model.

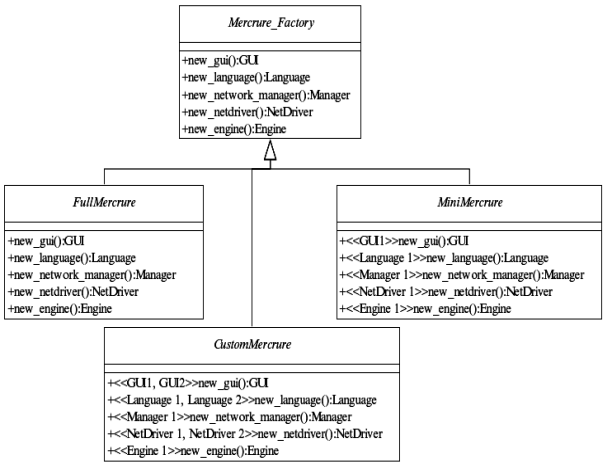


Figure 1. The Abstract Factory as a decision model for the Mercure SPL

III. INTEGRATE MDA AND SPL

Application engineering is the second process of SPL which comprises: (1) *Application requirement engineering* identifies the specific requirements for an individual product. Then, (2) *Application Design* derives an instance of the feature model, which conforms to the requirements identified in the previous step. In parallel product-specific requirement are captured and also modeled. (3) *Application designed detailed* focus on refining the design model, by considering platform specific characteristics such as programming language, middleware and component platform. Finally, (4) *Application Realisation* develops the final product by using the design detailed model.

The main idea is to represent all the 4 phases of application engineering by MDA model as shown in Figure 2. The requirements for the system are modeled in a computation independent model, CIM describing the situation in which the system will be used [7]. After that, application design is transformed in a platform independent model, a PIM, is built. It describes the system, but does not show details of its use of its platform [7]. Then integrate both models into one PIM model. Application detailed design is modeled in PSM the platform specific model produced by the transformation is a model of the same system specified by the PIM; it also specifies how that system makes use of the chosen platform [7]. Finally, the PSM obtained contained all the information necessary to produce computer program code.

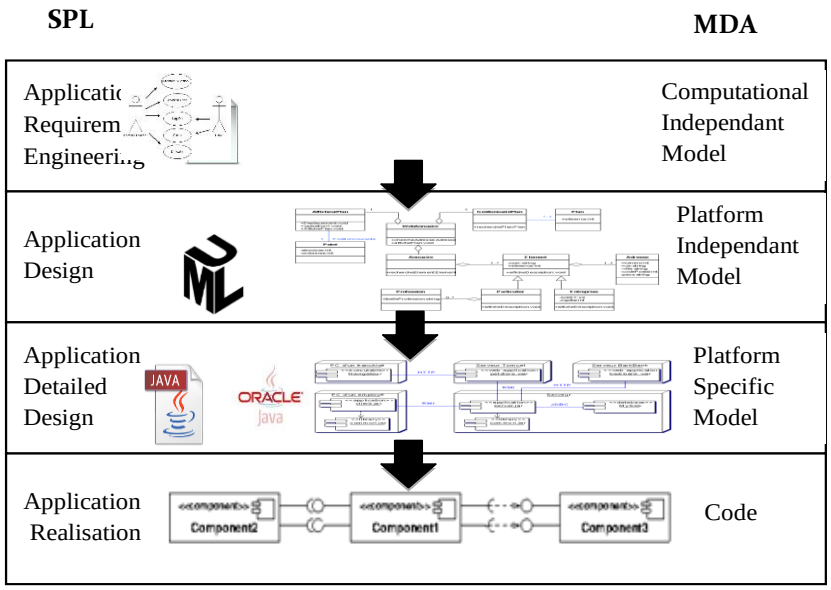


Figure2. Combination of MDA and SPL

IV. DERIVATION PROCESS

The product derivation process consists of six activities: (1) engineers starts with feature model for *pre-configuration* to select feature relevant to customer's wishes. (2) Users defined in the feature instance model perform the actual *configuration* by taking decisions visible to them. (3) In parallel, they capture product-specific requirements in aim to do a *specific-asset implementation*. (4) *Product integration* uses generic and specific model to obtain at the end one single model. (5) *Product development* uses the final model to develop an executable application. (6) Finally, *Product testing* passes a test to the final application. Eventually, newly captured product-specific requirements are added to the product line. Figure 3 depicts the activities of our product derivation approach.

A. *Pre-Configuration*

Initially step1 uses feature models [8] as input to select the feature relevant for customer's requirements to build the product and identify the specific-assets of the product. Once the selection is checked and validated by the engineer the output at this stage is a specialized version of feature model (instance) based on [3]. In parallel we represent this instance in MDA with CIM with UML to represent the model but it can be represented in any form as long as the semantics of this model is well preserved.

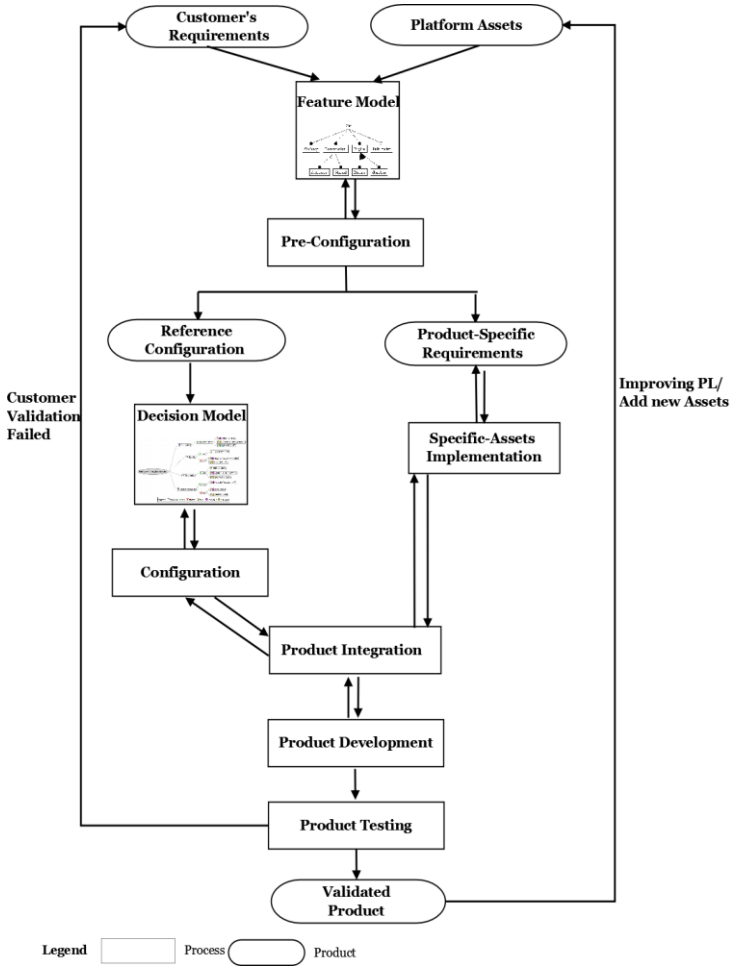


Figure3. Product derivation process

B. Configuration

In this step we start with a reference configuration as input which is a partial configuration designed as basis for the development of the new product that includes all parameters setting. Reselecting or mapping of customer features according to the final instance of feature model. After that, we use the decision model as input to take decision and customize assets throw answering questions, our decisions model is based on [9]. Based on this taken decisions a configuration is generated which is also the output of this phase. The representation in MDA at this stage could be represented as PIM (Platform Independent Model).

C. Specific-Asset Implementation

After capturing and identifying customer's specific requirements others decision must be taken for these specific-assets. As input we used the output of the previous phase that contains taken and open decisions in order to complete those opens decisions. Based on decisions and information about their relationship with the available assets, if it's possible we just modify an already existing asset to adapt the new customer's wishes, else we develop completely a new asset from the scratch. This new developed assets must be tested individually to make sure that they work before integration. These newly developed/modified assets are the output of this phase. The representation in MDA will be also a PIM specific model to facilitate integration in a later phase.

D. Product Integration

In this step, we should integrate the two separated PIM specific-asset model and PIM generic model into one single PIM model in order to facilitate implementation and derivate one single coherent product that satisfy customer's requirements. After that, we use this final PIM to transform it into a PSM (Platform Specific Model) that contain all details about the final application such as component platform, a specific programming language…etc. The process of mapping PIM to PSM is automated by using a set of predefined transformation rules. The output at this stage is PSM model that represent all detailed design and includes platform decisions.

E. Product Development

Based on what we obtained from the previous phase (PSM model), we use it as an input to produce an executable application code.

F. Product Testing

After the product development, which is finished when we obtained the application code, this final product must pass a test before delivering to the customer. Testing means that if the integration works properly and the final product satisfies all customers' requirements, then the product is validated by the engineer. If the customer validation is failed we must repeat all the previous activities until the customer is satisfied.

V. CASE STUDY

In this section we exemplify our approach on a simplified "e- Learning application" used in educational institutions. The application aim to ease and improve the teaching-learning process by means of taking advantage of internet technologies.

Figure 4 shows a part of the feature model we constructed for e-Learning. This feature model specifies that the e-Learning application has four main features: (1) the Human Machine Interface that can supply (*HMI*); (2) the kind of courses that can the platform provided (*Courses*); (3) the collaboration tool it uses (*Collaboration tool*); and (4) *Connexion BD*

The *HMI* could contain or not *Theme* which is an optional feature but must contain only one of three different *languages* (*FR*, *EN*, and *AR*) since these features are mutually exclusive alternatives. Two mandatory features must be used (1) *lesson* which users can read the content of the lesson online (Online reading) or simply *download* it. (2) *Evaluation* the second mandatory feature of *courses* could be done by two different kind of *Exercise* (*Online Test* or *Work*).So, if we select work which is an optional feature only one of the two features must be selected (*Individual* or and *AR*) since these features are mutually exclusive alternatives. Two mandatory features must be used (1) *lesson* which users can read the content of the lesson online (Online reading) or simply *download* it. (2) *Evaluation* the second mandatory feature of *courses* could be done by two different kind of *Exercise* (*Online Test* or *Work*).So, if we select work which is an optional feature only one of the two features must be selected (*Individual* or *Group*).

The last mandatory feature is about connexion to the data base so only one feature is allowed to be selected (MYSQL or SQL Server).

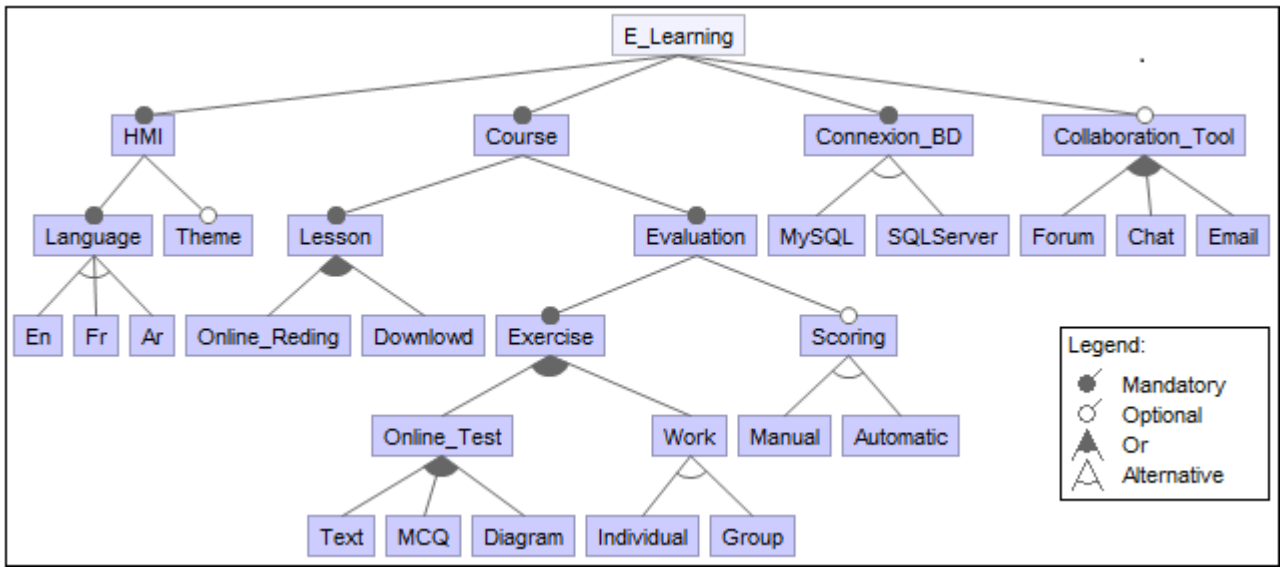


Figure4. Feature Model for e-Learning application

In order to evaluate the described approach, we performed the case study of "e-Learning application" following the organization proposed in section 4 in aim to create MDA models shown previously in Figure 2 (CIM, PIM, PSM and Code)

Pre-Configuration require a feature model which exposes in a concise way features and their variants supported by SPL's core assets. We use the feature model in figure 4 in aim to select the variants related to the customer's requirements. Then, once the selection is checked we create an instance of the feature model based on [3] which can be represented in MDA as CIM. The pre-configuration step ends when the combination of the selected features is validated. After that, the configuration phase used the output of the pre-configuration phase, during this steps decisions must be taken by the engineer. Based on this taken decisions a configuration is generated and could be represented in MDA as PIM.

Decision model rely on describing the decisions that need to be made to derive a specific product from the product line. Decisions are typically represented in form of questions with a defined set of possible answers. Decision-based variability models are often represented in tables containing decisions, their attributes, and dependencies. Products are derived from a decision model by setting values to the decisions through answering questions and following the sequence defined by the decisions' dependencies. Our product derivation process is also based on decisions Figure 5 shows a decision model for the e-Learning application in a tabulation notation.

Name	Relevance	Question	Range	Cardinality	Constraints
Online test		Which type of exercise do you want to be able to use in your platform?	Text, MCQ, Diagram	1-3	
Collaboration tool		Which collaboration tool you want your platform has?	Forum, Chat, Email	1-3	
Scoring	Exercise=>yes	Do you want to be able to score student's evaluation?	Yes,no	1	Scoring=>Exercice

Figure5. Example decision model based on [9]

In parallel, specific-requirements are captured to be designed and integrated later. The model design of specific-requirements is represented as a PIM model. Application specific PIM is identical to the generic PIM except the content of PIM is only relevant to a specific application.

After implementing two separate models (generic and specific) product integration phase integrates those two models into one single model which will still be a conventional PIM. Next, the obtained model after integration is detailed in order to map PIM to PSM that

contain specific characteristics such as programming language, middleware and component platform. Finally, product realisation takes PSM to produce an executable application. Our derivation process is implemented based on EMF Ecore plugin of the Eclipse open development platform available at (http://eclipse.org/modeling/emf/)

VI. CONCLUSION

In this paper we have proposed a product derivation process using MDA approach. MDA with its organized layers of models achieving the SPL goals with more benefits and the generative natures of MDA makes it a useful approach to derive product for SPL .MDA consists in the separation of platform dependent and platform independent models, which distinct between business (CIM), applications (PIM), and technology (PSM). The main idea is to represent each phase of the application engineering (application requirement engineering, application design, application detailed design, and application realisation) by MDA models. The derivation process proposed consists of six activities, each activities output was represented by MDA. We also illustrated each step of the process with an e-Learning application.

As future work, we will add more features and also intend to build a set of components to this e-learning application.

At the tool level, improvements may concern the visual representation of feature models (via the Ecore reflexive editor provided by Eclipse). A possibility is to develop our metamodel and generate this later using tool such as GMF (http://www.eclipse.org/gmf/)

References

[1] Clements,P., Northrop,L. Software Product Lines: Practices and Patterns. The SEI series in software engineering. Addison-Wesley, Boston, 2002.

[2] Czarnecki, K., Helsen, S.: "Feature-Based Survey of Model Transformation Approaches", IBM Systems Journal, 45, 3, 621-64, 2006.

[3] Czarnecki, K., Helsen, S., Eisenecker, U.: "Staged Configuration Using Feature Models". In Proceedings of the 3rd Software Product-Line Conference (SPLC'04), September 2004.

[4] Deelstra, S., M. Sinnema, and J. Bosch, Product Derivation in Software Product Families: A Case Study. Journal of Systems and Software, 2005. 74(2): p. 173-194.

[5] Guelfi, N. and Perrouin, G., A Flexible Requirements Analysis Approach for Software Product Lines, I Requirements Engineering: Foundation for Software Quality. 2007, Springer Berlin / Heidelberg. p. 78-92.

[6] Hotz, L., A. Gunter, and T. Krebs, A Knowledge-based Product Derivation Process and some Ideas how to Integrate Product Development, in Proc. of Software Variability Management Workshop. 2003: Groningen, The Netherlands.

[7] J. Mukerji, and J. Miller, "MDA Guide," 2003.

[8] Kang, K., Cohen, S., Hess, J., Novak, W., Peterson, A.: "Feature-oriented domain analysis (FODA) feasibility study", Technical Report CMU/SEI-90-TR-021, SEI, Pittsburgh, PA, November 1990.

[9] K. Schmid and I. John. A Customizable Approach to Full-Life Cycle Variability Management. Journal of the Science of Computer Programming, Special Issue on Variability Management, 53(3), pp. 259-284. 2004.

[10] Ø.Haugen, B. Møller-Pedersen, J. Oldevik, and A. Solberg. An MDA-based Framework for Model-Driven Product Derivation. In SEA, pages 709–714. ACTA Press, 2004.

[11] Perrouin, G., Klein, J., Guelfi, N., and Jezequel, J.M. Reconciling Automation and Flexibility in Product Derivation. in 12th International Software Product Line Conference (SPLC). 2008.

[12] S. D. Kim, H. G. Min, J. S. Her, and S. H. Chang. DREAM: A Practical Product Line Engineering Using Model Driven Architecture. In Information Technology and Applications. (ICITA), pages 70–75, Washington, DC, USA, 2005.

[13] T. Ziadi and J.-M. J´ez´equel. Product Line Engineering with the UML: Deriving Products. In Families Research Book. Springer, 2006.

Une approche pour la formation de coalitions de robots autonomes
-Application à un scenario de sauvetage-

ZITOUNI Farouq[1], MAAMRI Ramdane[2], BOUSSEBOUGH Imane[3]

LIRE Laboratory, University of Abdelhamid Mehri - Constantine

[1]Farouk_Zitouni@yahoo.fr

[2]rmaamri@yahoo.fr

[3]iboussebough@yahoo.fr

Abstract— **Dans ce papier, nous proposons une approche pour la formation de coalitions parmi un ensemble de robots autonomes. Dans ce fait, nous supposons que chaque robot possède un ensemble de compétences où chacune requiert un ensemble de ressources. Aussi, nous supposons qu'aucune information sur les compétences et les ressources que possèdent les robots n'est partagée, c'est-à-dire que chaque robot ne connaît que ses propres compétences et ressources. Comme cadre d'application pour notre approche, nous avons adopté le cas où nous avons un ensemble de victimes, mises en groupes, qui sont dans un environnement altéré par une catastrophe naturelle. En réalité, chaque victime a besoin d'un ensemble de compétences et de ressources pour son évacuation. Les victimes sont organisées en groupes en fonction de certains critères, et nous devons pour chaque groupe de victimes former une coalition de robots sauveteurs afin d'assurer les différentes opérations de sauvetage. Dans le présent papier, nous allons présenter une approche pour la formation de coalitions de robots de sauvetage en utilisant la négociation basée sur des valeurs d'utilité qui sont calculées dynamiquement.**

Keywords— **formation dynamique de coalitions de robots sauveteurs, systèmes multi-agents, coopération, négociation, fonctions d'utilité.**

I. INTRODUCTION

Dans les systèmes multi-agents, le problème d'allocation de tâches et de compétences a été toujours considéré comme un goulot d'étranglement pour de nombreuses applications distribuées **[1], [2], [3]**. Par exemple, les services offerts sur Internet sont souvent construits par la combinaison de plusieurs éléments, venant de différents nœuds, d'une manière ascendante et ces éléments sont traités par des nœuds ayant les compétences nécessaires **[2]**. La création de ces services peut être considérée comme un ensemble d'activités coopératives entre plusieurs agents ayant les compétences requises **[4], [5]**. La réalisation d'une tâche donnée par l'attribution de ses sous-tâches à des agents appropriés se réfère souvent au problème de formation de groupes dans un système multi-agents **[4], [6]**. Souvent, un service ne peut pas être fourni par un agent unique et même si c'est le cas, alors les performances qu'il présente ne sont pas acceptables, d'où l'idée de mettre les agents en groupes pour fournir des services efficients et efficaces **[7]**.

Ainsi, dans notre cas nous organisons un ensemble de robots sauveteurs en groupes pour effectuer des opérations de sauvetage et d'évacuation au profit des victimes piégées dans les décombres d'un environnement frappé par un désastre naturel. En outre, nous supposons que chaque robot sauveteur dispose d'un ensemble de compétences et de ressources pour aider et assister les victimes. En général, ces compétences et ressources sont souvent rares et insuffisantes, d'où la nécessité de les bien gérer et organiser les robots sauveteurs en groupes pour mener à bien les différentes opérations de sauvetage.

Alors, nous présentons une approche pour la formation de coalitions de robots sauveteurs dans un environnement très instable en utilisant des négociations basées sur deux valeurs d'utilité différentes. La première valeur est calculée en utilisant la théorie des arbres de décision (plus précisément l'algorithme C4.5 présenté dans **[8]**) et l'ensemble des compétences et ressources communes entre les robots sauveteurs et les victimes. La seconde valeur est calculée à l'aide d'un ensemble de connaissances stockées dans une base de connaissances (BC). Cette BC contient tous les traumatismes dont une victime peut avoir dans un environnement frappé par une catastrophe naturelle, et pour chaque traumatisme elle identifie toutes les compétences et ressources nécessaires pour l'assumer. Dans la section suivante, nous présentons quelques travaux connexes à notre approche. Ensuite, nous allons exposer notre problématique. Après, nous expliquerons notre approche et montrerons les résultats obtenus de sa simulation et enfin nous terminerons par une conclusion et quelques perspectives.

II. TRAVAUX CONNEXES

Dans les systèmes multi-agents, la formation de groupes et de coalitions a toujours attiré beaucoup d'attention de la part des chercheurs et il y en avait de nombreux travaux qui ont traité ce problème. Comme **Shehory** et **Karus [9]** qui ont proposé un algorithme pour la formation de coalitions en utilisant des solutions Pareto-optimales. En outre, **Yang** et **Luo [10]** utilisèrent les algorithmes génétiques pour trouver les meilleures correspondances entre un ensemble de tâches et des groupes d'agents. Néanmoins, leurs recherches supposent que l'état interne d'un agent et l'ensemble des tâches sont connus à l'avance, alors que cela n'est réellement pas possible dans la majorité des cas. En plus, ils supposent que les valeurs des coalitions sont connues à l'avance. Cependant, les recherches qui supposent que les tâches ne sont pas connues à l'avance ne sont pas nombreuses. Alors, comme **Sandholm** et al **[11]** qui ont montré que la génération d'une coalition optimale est un problème NP-difficile en utilisant une approche basée sur la théorie des jeux. En outre, **Rahwan** et al **[12]** ont proposé un algorithme anytime pour la génération des coalitions. **Abdallah** et **Lasser [13]** ont proposé aussi une méthode efficace pour la formation de groupes en utilisant l'apprentissage par renforcement dans un réseau d'agents structuré en arbre. Les groupes sont formés de manière

efficace car un agent dans le réseau apprend les politiques d'allocation de tâches d'après ses précédentes expériences et les diffuse ensuite aux agents dans le niveau inférieur de la hiérarchie. **Katayanagi** et **Sugawara [14]** ont étendu cette approche en rajoutant la réorganisation du réseau d'agents afin d'affecter plus de tâches aux agents.

III. PROBLEMATIQUE

Intuitivement, il est évident qu'après la survenue d'une catastrophe naturelle, comme les séismes ou les incendies, altérant un environnement que des opérations de sauvetage et d'évacuation devraient avoir lieu pour aider et assister les victimes qui s'y trouvent. Toutefois, avant de commencer ces opérations il est très utile d'avoir au moins une certaine connaissance sur l'état de chaque victime et déterminer, en termes de traumatismes et de compétences, les besoins de chacune et ce pour une meilleure prise en charge. Ainsi, dans le présent papier nous allons proposer une méthode pour la détermination des besoins de chaque victime coincée dans un tel environnement.

En plus, nous supposons que les victimes dans cet environnement sont mises en groupes en fonction de certains critères et après leur regroupement et la détermination de l'état de chacune, alors nous devons commencer les opérations d'évacuation et de sauvetage de ces victimes. En général, ces opérations sont assurées par des équipes humaines de recherche et de sauvetage. Cependant, il faut noter le cas où l'intervention de ces équipes est dangereuse pour ses membres ou même impossible. Par conséquent, nous recommandons l'utilisation de robots sauveteurs qui peuvent fournir une aide précieuse dans de telles situations. Donc, pour aider et évacuer les différents groupes de victimes dans un tel environnement, alors nous supposons que ces opérations de sauvetage sont effectuées par des robots sauveteurs qui doivent coopérer entre eux. Cette coopération peut avoir lieu entre un robot et un autre ou, plus généralement, entre un groupe de robots et un autre groupe. Comme la mise en groupes de ces robots sauveteurs peut fournir l'avantage du partage de leurs compétences et ressources, alors nous pouvons assumer le problème de manque de ces dernières. Mais la question qui se pose dans ce contexte est de comment organiser ces robots sauveteurs en groupes ou coalitions ?

Par définition, une coalition **[15]** est une alliance momentanément conclue entre des individus, des groupes ou des partis pour défendre des intérêts communs. Dans ce papier, nous allons proposer aussi une approche, basée sur des négociations entre les différents robots sauveteurs et ce en utilisant des valeurs d'utilité, pour la formation dynamique de coalitions de robots sauveteurs et ce dans un environnement très instable où les conditions et les circonstances changent de façon inattendue.

IV. APPROCHE PROPOSEE

Tout d'abord, nous supposons que nous avons un ensemble de victimes coincées dans un environnement altéré par une catastrophe naturelle. Ces victimes sont organisées en groupes en fonction de certains critères et nous devrons pour chacun former une coalition de robots sauveteurs afin d'effectuer les différentes opérations de sauvetage. Aussi, nous supposons que notre environnement est divisé en zones adjacentes où chacune est délimitée par quelques drones. En fait, ces drones coopèrent pour déterminer l'ensemble des groupes de victimes dans l'environnement et déterminer l'état et les besoins de chacune. Par conséquent, nous proposons une méthode pour déterminer l'état et les besoins de chaque victime et une approche pour la formation de coalitions de robots sauveteurs.

A. Méthode pour la détermination des besoins des victimes

Etant donné un environnement altéré par un désastre naturel, nous supposons que nous avons un ensemble de victimes $V = \{v_1, v_2, \cdots, v_n\}$, qui ont besoin d'être évacuées par des robots sauveteurs. Dans cet environnement une victime $v_i \in V$ est représentée par un quadruplet $(\mathrm{v_i}, \Psi^{\mathrm{v_i}}, \Omega^{\mathrm{v_i}}, \Pi^{\mathrm{v_i}})$, où l'ensemble $\Psi^{\mathrm{v_i}} = \{\psi_1^{\mathrm{v_i}}, \psi_2^{\mathrm{v_i}}, \cdots, \psi_{\mathrm{m}}^{\mathrm{v_i}}\}$ exprime tous les traumatismes qu'a la victime v_i, l'ensemble $\Omega^{\mathrm{v_i}} = \{\omega_1^{\mathrm{v_i}}, \omega_2^{\mathrm{v_i}}, \cdots, \omega_{\mathrm{p}}^{\mathrm{v_i}}\}$ dénote toutes les compétences requises pour évacuer la victime $\mathrm{v_i}$ et $\Pi^{\mathrm{v_i}}$ est un nombre réel qui représente la norme des sommes des quantités des ressources nécessaires pour chaque compétence $\omega_{\mathrm{p}}^{\mathrm{v_i}} \in \Omega^{\mathrm{v_i}}$. Dans les paragraphes suivants, nous expliquerons comment déterminer les éléments des ensemble $\Psi^{\mathrm{v_i}}$ et $\Omega^{\mathrm{v_i}}$ et calculer la norme $\Pi^{\mathrm{v_i}}$.

1. Détermination des éléments de l'ensemble Ψ^{v_i}

Pour faciliter la compréhension de notre méthode, nous supposons que tous les traumatismes qu'une victime peut subir sont connus à postériori et arrangés dans un ensemble $T = \{t_1, t_2, \cdots, t_p\}$. Aussi, nous rajoutons à cet ensemble l'élément $\perp$ qui dénote l'absence de tout traumatisme, c'est-à-dire que la victime est indemne. Alors l'ensemble T devient $T = \{\perp, t_1, t_2, \cdots, t_p\}$. Soit Ψ l'ensemble de toutes les partitions possibles de T, exceptant la partition vide, ainsi $\Psi = \{\{\perp\}, \{t_1\}, \{t_2\}, \cdots, \{t_1, t_2\}, \cdots\}$.

Intuitivement, nous supposons qu'une victime $v_i \in V$ peut avoir un ensemble de traumatismes Ψ_i qui doit être inclus dans Ψ, alors nous pouvons écrire :

$$\forall v_i \in V : has(v_i, \Psi_i) \Longrightarrow \Psi_i \subset \Psi$$

Où l'ensemble Ψ_i dénote une partition quelconque de Ψ. Pour déterminer l'ensemble Ψ_i , qui exprime tous les traumatismes que la victime $v_i \in V$ a subis, nous admettons aussi que nous avons un ensemble de capteurs $S = \{s_1, s_2, \cdots, s_q\}$. Ces capteurs sont sensibles à certains phénomènes prédéterminés et qui transforment une grandeur physique en signaux, souvent numériques, manipulables par des machines. Le papier **[16]** cite quelques capteurs ambiants utilisés dans le contexte médical. Dans notre cas, ces capteurs ont pour principal objectif le diagnostic d'une victime v_i afin de déterminer l'ensemble Ψ_i de ses traumatismes. Soient $\mathrm{X} = \{\{s_1\}, \{s_2\}, \cdots, \{s_1, s_2\}, \cdots\}$ l'ensemble de toutes les partitions possibles de S, exceptant la partition vide, et x une partition quelconque de X ($x \subset X$). Donc, nous définissons un prédicat $h(v_i, x)$, qui peut être écrit comme suit :

$$h(v_i, x) = h_1(v_i, x_1) \wedge h_2(v_i, x_2) \wedge \cdots \wedge h_j\,(v_i, x_j)$$

Où $(\mathrm{x_1}, \mathrm{x_2}, \cdots, \mathrm{x_j}) \in \mathrm{x}$ et $\mathrm{h_j}$ est une fonction booléenne bien définie, qui pour une victime v_i et un capteur $\mathrm{x_j}$ retourne une valeur booléenne. Plus formellement, nous pouvons écrire $\mathrm{h_j}\,(\mathrm{v_i}, \mathrm{x_j}) = \mathrm{y}$ et $\mathrm{y} \in \{0,1\}$. Maintenant, nous définissons une autre fonction f donnée par l'expression suivant :

$\forall x \subset X$:

$$\begin{cases}\{0,1\}^j \mapsto T \\ f\left(h(v_i,x)\right) = \\ f\left(h_1(v_i,x_1) \wedge h_2(v_i,x_2) \wedge \cdots \wedge h_j\left(v_i,x_j\right)\right) = t_p\end{cases}$$

Cette fonction nous donne, suivant la combinaison des valeurs booléennes des fonctions $h_j\left(v_i,x_j\right)$, un élément de T. Spécifiquement, la fonction f prend en entrée une victime v_i et une partition de capteurs $x \subset X$ et retourne une valeur $t_p \in T$. Cette fonction est calculée pour toutes les partitions possible de X afin de déterminer, exhaustivement, l'état de chaque victime v_i, c'est-à-dire déterminer les éléments de la partition Ψ_i qui est équivalente à l'ensemble Ψ^{v_i}. Autrement dit, nous calculons l'union:

$$\bigcup_{\forall x \subset X \wedge v_i \in V} f\left(h(v_i,x)\right) = \Psi_i = \Psi^{v_i}$$

2. *Détermination des éléments de l'ensemble Ω^{v_i} et la norme Π^{v_i}*

Une fois l'opération de détermination de l'ensemble des traumatismes Ψ^{v_i} s'est terminée avec succès, alors maintenant nous devons déterminer l'ensemble des compétences et ressources nécessaires pour prendre en charge chaque traumatisme, c'est-à-dire pour chaque traumatisme $\psi_m^{v_i} \in \Psi^{v_i}$ nous devons énumérer toutes les compétences et les ressources requises pour en assumer. Par définition, une compétence est la capacité de résoudre efficacement un problème dans un contexte donné **[17]**. Dans notre cas, nous supposons qu'un traumatisme $\psi_m^{v_i} \in \Psi^{v_i}$ peut être assumé au moins par une seule compétence, c'est-à-dire pour prendre en charge un traumatisme nous devons faire appel à au moins une compétence. D'une autre manière, soit $\omega_p^{v_i} \in \Omega^{v_i}$ une compétence donnée, qui exprime que la victime v_i a besoin de la compétence $\omega_p^{v_i}$ pour être prise en charge, définie par le couple $\left(\omega_p^{v_i}, h_{v_i}^p\right)$ où $h_{v_i}^p = \left[h_1^{[p,v_i]}, h_2^{[p,v_i]}, \cdots, h_d^{[p,v_i]}\right]$ est un vecteur d'entiers positifs qui exprime les quantités des ressources $h_d^{[p,v_i]}$ que requiert la compétence $\omega_p^{v_i}$ pour évacuer la victime v_i. Donc la norme Π^{v_i} peut être calculée comme suit :

$$\Pi^{v_i} = \left\| \sum\nolimits_{\omega_p^{v_i} \in \Omega^{v_i}} h_{v_i}^p \right\|$$

Dans notre méthode, nous supposons que toutes les compétences, pour prendre en charge tous les traumatismes, sont listées dans l'ensemble $C = \{\omega_1, \omega_2, \cdots, \omega_r\}$. Soit $\Omega = \{\{\omega_1\}, \{\omega_2\}, \cdots, \{\omega_r\}, \{\omega_1, \omega_2\}, \cdots\}$ dénotant l'ensemble de toutes les partitions possibles de C, exceptant la partition vide. Toutes les correspondances entre les compétences et les traumatismes sont listées dans notre BC. Donc, nous pouvons définir une autre fonction P qui pour un traumatisme donné $\psi_m^{v_i} \in \Psi^{v_i}$ retourne l'ensemble de toutes les compétences requises $\Omega_j \subset \Omega$ pour l'assumer. Donc, nous pouvons écrire :

$$\forall \psi_m^{v_i} \in \Psi^{v_i}: \begin{cases}\Psi^{v_i} \mapsto \Omega \\ P\left(\psi_m^{v_i}\right) = \Omega_j\end{cases}$$

L'union $\bigcup_{\psi_m^{v_i} \in \Psi^{v_i}} P\left(\psi_m^{v_i}\right)$ des ensembles Ω_j exprime l'ensemble de toutes les compétences nécessaires pour prendre en charge la victime v_i et cet ensemble est équivalent à l'ensemble Ω^{v_i}.

B. *Représentation des robots sauveteurs*

Soit $A = \{a_1, a_2, \cdots, a_n\}$ un ensemble de robots sauveteurs. Un robot $a_i \in A$ est représenté par un triplet $(a_i, \Omega^{a_i}, \Pi^{a_i})$, où $\Omega^{a_i} = \left\{\omega_1^{a_i}, \omega_2^{a_i}, \cdots, \omega_p^{a_i}\right\}$ dénote l'ensemble de ses compétences. Une compétence $\omega_p^{a_i} \in \Omega^{a_i}$ est représentée par un couple $\left(\omega_p^{a_i}, h_{a_i}^p\right)$, où $h_{a_i}^p = \left[h_1^{[p,a_i]}, h_2^{[p,a_i]}, \cdots, h_d^{[p,a_i]}\right]$ est un vecteur de valeurs positives exprimant les quantités de ressources $h_d^{[p,a_i]}$ que possède le robot a_i pour la compétence $\omega_p^{a_i}$. D'une façon similaire, le nombre Π^{a_i} exprime la norme des sommes des quantités des ressources nécessaires pour chaque compétence $\omega_p^{a_i} \in \Omega^{a_i}$ et elle est calculée comme suit :

$$\Pi^{a_i} = \left\| \sum\nolimits_{\omega_p^{a_i} \in \Omega^{a_i}} h_{a_i}^p \right\|$$

Donc, nous pouvons dire qu'un robot a_j peut prendre en charge une victime v_i lorsque l'inégalité suivante et vérifiée :

$$\forall\, p, a_j, v_i :\ h_{a_j}^p \geq h_{v_i}^p$$

Plus généralement, si nous avons un ensemble de victimes v_i qui appartiennent à un même groupe G et requiert chacune la compétence $\omega_p^{v_i}$ que possède le robot a_j, alors l'inégalité suivante doit être toujours vérifiées :

$$\forall\, p, a_j, v_i :\ h_{a_j}^p \geq \sum_{v_i \in G} h_{v_i}^p$$

C. *Formation de coalitions de robots sauveteurs*

L'approche pour la formation de coalitions de robots sauveteurs que nous avons proposée consiste en trois étapes principales qui sont le choix d'un leader pour chaque coalition de robots, son initialisation et enfin sa mise à jour en fonction de l'état de l'environnement. Ainsi, pour chaque groupe de victimes nous devons former une coalition de robots sauveteurs.

1. *Sélection d'un leader pour chaque coalition*

Etant donné une coalition de robots sauveteurs, nous supposons que chaque robot dans cette coalition peut jouer un rôle distinct, il peut être soit un leader soit un membre. Par conséquent, selon ce rôle nous pouvons déterminer les activités de chaque robot sauveteur dans sa coalition. Dans cette section, nous présentons un algorithme pour la spécification du rôle de chaque robot sauveteur.

Soit $V = \{v_1, v_2, \cdots, v_n\}$ un ensemble de victimes réparties sur un ensemble de groupes $G = \{G_1, G_2, \cdots, G_m\}$, où $\forall\, G_i, G_j \in G : G_i \cap G_j = \emptyset$. Alors, chaque victime $v_i \in V$ est définie par son quadruplet $(v_i, \Psi^{v_i}, \Omega^{v_i}, \Pi^{v_i})$. En plus, soit $A = \{a_1, a_2, \cdots, a_k\}$ un ensemble de robots sauveteurs, où chaque robot $a_i \in A$ est défini par son triplet $(a_i, \Omega^{a_i}, \Pi^{a_i})$. Par analogie, un groupe de victimes $G_j \in G$ est défini par son quadruplet $\left(G_j, \Psi^{G_j}, \Omega^{G_j}, \Pi^{G_j}\right)$, où les ensemble Ψ^{G_j} et Ω^{G_j} dénotent, respectivement, les unions des éléments des ensembles Ψ^{v_i} et Ω^{v_i} de toutes les victimes appartenant au groupe G_j et Π^{G_j} est la somme des normes des différentes victimes dans le groupe G_j. Donc, nous pouvons écrire :

$$\begin{cases}\Psi^{G_j} = \bigcup_{v_i \in G_j} \Psi^{v_i} \\ \Omega^{G_j} = \bigcup_{v_i \in G_j} \Omega^{v_i} \\ \Pi^{G_j} = \sum_{v_i \in G_j} \Pi^{v_i}\end{cases}$$

Nous supposons que chaque groupe $G_j \in G$ doit être dirigé par un robot leader. Donc, notre objectif sera la désignation de l'ensemble des robots sauveteurs ayant le rôle leader et

chacun doit diriger un et seulement un groupe de victimes. Un robot sauveteur s'il n'est pas leader alors il est membre.

Le processus d'élection des robots leaders est comme suit. D'abord, chaque robot sauveteur doit calculer pour chaque groupe de victimes $G_j \in G$ son gain en utilisant l'équation (1). Ensuite, il diffuse son résultat aux autres robots sauveteurs afin d'organiser une élection. L'algorithme 1 résume le principe de cette élection.

$$gain\left(a_i, G_j\right) = \left(\left|GainRatio\left(a_i, G_j\right)\right|\right)^{\alpha} + \left(\frac{card\left(\Omega^{a_i} \cap \Omega^{G_j}\right)}{card(\Omega^{a_i}) + card\left(\Omega^{G_j}\right)}\right)^{\beta} + \left(\frac{\Pi^{a_i}}{\Pi^{G_j}}\right)^{\gamma} \rightarrow (1)$$

Où $GainRatio\left(a_i, G_j\right)$ est le ratio de gain d'information qu'obtiendrait un robot sauveteur a_i pour le groupe G_j (se référer à **[8]**) et les trois paramètres balançant le gain sont α, β et γ qui contrôlent respectivement l'importance relative donnée au ratio de gain d'information, au nombre de compétences communes entre le robot sauveteur a_i et les victimes dans le groupe G_j et la pondération des normes Π^{a_i} et Π^{G_j}. Les leaders des groupes sont choisis en se basant sur leurs valeurs de gains, c'est-à-dire pour un groupe de victimes donné un leader est désigné si sa valeur de gain est la plus grande. D'une autre manière nous pouvons écrire :

$$\forall\, a_i \in A, G_j \in G: a_i \text{ est leader de } G_j \text{ alors}$$
$$a_i = \max_{i \in [1,\cdots,k]} \left\{gain\left(a_i, G_j\right)\right\}$$

2. *Initialisation des coalitions des robots sauveteurs*

Dans la section précédente, nous avons présenté un algorithme pour la sélection d'un robot leader pour chaque groupe de victimes. Alors, le leader d'un groupe jouera un rôle très important dans le processus d'initialisation des coalitions. De ce fait, chaque leader d'un groupe commence par énumérer toutes les compétences et les ressources que nécessitent les victimes dans son groupe puis il entame des négociations avec les robots sauveteurs ayant le rôle membre pour les faire adhérer à sa coalition.

Nous allons maintenant décrire le mécanisme d'initialisation des coalitions des robots sauveteurs. Nous supposons que nous avons un ensemble de groupes de victimes $G = \{G_1, G_2, \cdots, G_n\}$, où chaque groupe $G_j \in G$ est défini par son quadruplet $\left(G_j, \Psi^{G_j}, \Omega^{G_j}, \Pi^{G_j}\right)$. Aussi, nous avons un ensemble de robots sauveteurs $A = \{a_1, a_2, \cdots, a_m\}$, où chacun est défini par son triplet $(a_i, \Omega^{a_i}, \Pi^{a_i})$. Ces robots doivent être organisés en n coalitions afin d'assurer les différentes opérations de sauvetage et d'évacuation au profit des victimes appartenant aux différents groupes $G_j \in G$. En plus, pour chaque groupe de victimes nous devons désigner un robot leader en utilisant l'algorithme 1.

Aussi, nous supposons que le nombre de robots sauveteurs est supérieur au nombre de groupes de victimes, c'est-à-dire $m > n$. En plus, chaque robot leader $a_i \in A$ doit maintenir dans sa mémoire un vecteur de gain pour chaque groupe de victimes, c'est-à-dire il doit maintenir le vecteur défini par : $Gain_{a_i} = \left[\left(G_1, gain(a_i, G_1)\right), \cdots, \left(G_n, gain(a_i, G_n)\right)\right]$. Entre autres, un robot leader avec un rôle leader doit garder dans sa mémoire les informations concernant le groupe dont il dirige, c'est-à-dire si le robot sauveteur a_i est leader du groupe G_j alors il doit garder dans sa mémoire le quadruplet $\left(G_j, \Psi^{G_j}, \Omega^{G_j}, \Pi^{G_j}\right)$. Maintenant, le leader de chaque groupe de victime doit entamer un ensemble de négociations avec les robots sauveteurs ayant le rôle membre afin d'initialiser sa propre coalition. Les étapes de cette initialisation son résumées dans l'algorithme 2.

Algorithm 1: Algorithm of designation of leaders. Behaviour of an agent a_i

Input: Let $A = \{a_1, a_2, \ldots, a_n\}$ be the set of all rescue robots;
Input: Let $G = \{G_1, G_2, \ldots, G_m\}$ be the set of all groups of victims that we would designate leaders;
Input: Let L_1 be a list called list of gains;
Input: Let L_2 be a list called list of leaders;
Input: $temp_1 \leftarrow true$;$temp_2 \leftarrow false$;$p \leftarrow 0$;
Output: L_2;

```
while (true) do
    if (message − received() = true) then
        type ← content − of − received − message().type;
        if (type =' vote') then
            sender − name ← content − of − received − message().name;
            L3 ← content − of − received − message().list;
            for q ← 1 to L3.size() do
                if (L1.get(q) < L3.get(q)) then
                    L2.set(q, sender − name);
            p ← p + 1;
            if (p = (n − 1)) then
                temp2 ← true;
                p ← 0;
        if (type =' sort') then
            L3 ← content − of − received − message().list;
            for q ← 1 to L3.size() do
                calculate − the − occurrence − of − each − agent();
            p ← p + 1;
            if (p = n) then
                for − each − group − designate − the − agent − with − highest − occurrence − as − leader();
    if (temp1 = true) then
        foreach (Gj ∈ G) do
            L1.addLast(gain(ai, Gj));
            L2.addLast(ai);
        foreach (aj ∈ A) do
            if (aj ≠ ai) then
                content − message('vote', ai, L1);
                send − message(content − message, aj);
        temp1 ← false
    if (temp2 = true) then
        foreach (aj ∈ A) do
            content − message('sort', L2);
            send − message(content − message, aj);
        temp2 ← false
```

Algorithm 2: Initialization algorithm of coalitions of rescue robots. Behaviour of an agent a_i

Input: Let $G_j \in G$ be a group of victims that we would initialize its coalition of rescue robots;
Input: Let $a_i \in A$ be the rescue robot leader of the group G_j;
Input: Let σ_{a_i} be the coalition of rescue robots for the group G_j;
Input: $temp \leftarrow true$;
Output: σ_{a_i};

```
if (a_i.role =' Leader') then
    σ_{a_i} ← σ_{a_i} ∪ {a_i};
while (true) do
    if (message − received() = true) then
        if (a_i.role =' Leader') then
            a_p ← choose − the − member − with − the − biggest − proposed − gain();
            σ_{a_i} ← σ_{a_i} ∪ {a_p};
            a_p.state ← busy;
            temp ← true;
        if (a_i.role =' Member') then
            if (a_i.hasCompetence(δ)) ∧ (a_i.isFree()) then
                replay − the − leader − with − the − biggest − proposed − gain(gain(a_i, G_j));
    if (a_i.role =' Leader') ∧ (temp = true) then
        Ω^{G_j} ← extract − all − needed − competences − by − victims − in(G_j);
        Ω^{σ_{a_i}} ← extract − all − robots − competences − in(σ_{a_i});
        Δ^{G_j}_{σ_{a_i}} ← Ω^{G_j} − Ω^{σ_{a_i}} = {ω ∈ Ω^{G_j} ∧ ω ∉ Ω^{σ_{a_i}}};
        H^{Ω^{G_j}}_{G_j} ← extract − all − needed − resource − by − victims − in(G_j);
        H^{Ω^{σ_{a_i}}}_{a_i} ← extract − all − robots − resource − in(σ_{a_i});
        if (Δ^{G_j}_{σ_{a_i}} ≠ ∅) ∧ (H^{Ω^{σ_{a_i}}}_{a_i} ≥ H^{Ω^{G_j}}_{G_j}) then
            δ ← select − an − element − from(Δ^{G_j}_{σ_{a_i}});
            diffuse − a − message − to − members − to − ask − for(δ, gain(a_i, G_j));
            temp ← false;
```

3. Mise à jour des coalitions des robots sauveteurs

Quand les coalitions initiales des robots sauveteurs sont initialisées, alors chaque leader avec les membres de sa coalition doivent se déplacer vers les victimes pour effectuer les opérations de sauvetage. En fait, pour initialiser les coalitions des robots sauveteurs nous avons utilisé une fonction d'utilité de gain basée sur les compétences et les ressources requises par chaque victime. Cependant, dans un environnement très instable et qui change de façon inattendue nous devons également prendre en considération d'autres paramètres plus importants régissant la dynamique de l'environnement et qui peuvent grandement contribuer à la détermination du contenu des coalitions, comme la gravité des blessures, l'âge et le sexe de chaque victime. De même, nous pouvons prendre en compte le fait que d'autres victimes peuvent être découvertes entre temps et changer ainsi le contenu des coalitions. Bien entendu, d'autres critères peuvent être utilisés.

Dans ce papier, nous nous limiterons à ces paramètres et nous supposons que les robots sauveteurs, après avoir été déplacés sur les lieux où les victimes se trouvent, doivent faire un autre diagnostic plus profond afin de déterminer la gravité des blessures, l'âge et le sexe de chaque victime et ils doivent également faire des recherches périodiques en quête de nouvelles victimes. Par conséquent, avec ces paramètres, nous pouvons définir une autre fonction d'utilité qui détermine la priorité de chaque victime et cette fonction peut être exprimée comme suit :

$$priority(gravity, age, gender) = (gravity)^{\alpha} \times (age)^{\beta} \wedge \times (gender)^{\gamma}$$

Les paramètres α , β et γ contrôlent, respectivement, l'importance relative donnée à la gravité des blessures, à l'âge et au sexe de chaque victime. Pour définir l'âge approximatif et le sexe d'une victime, alors nous pouvons analyser son ADN ou son visage **[18], [19]**. Par analogie, nous pouvons attribuer une valeur numérique à chaque traumatisme qui définit la gravité d'une blessure par rapport à une autre et ce en utilisant notre BC. Notre BC est construite à partir des expertises de personnes spécialisées en traumatologie et des pompiers. De ce fait, cette BC est structurée dans un fichier XML partagé entre les robots sauveteurs et elle identifie la quasi-totalité des traumatismes qu'une victime peut avoir, leur gravité et les compétences et les ressources nécessaires pour les prendre en charge. Donc, si nous calculons la fonction d'utilité de priorité pour chaque victime, alors le contenu des coalitions de robots sauveteurs peut être modifié en fonction des besoins des victimes.

V. RESULTATS OBTENUS

Pour tester les algorithmes que nous avons proposés dans notre approche pour la formation de coalitions de robots sauveteurs dans un environnement dynamique, alors nous avons utilisé 40 victimes réparties sur 8 groupes. Le tableau I indique, pour chaque zone, comment nos victimes sont divisées en groupes.

TABLE I. LE CONTENU DE CHAQUE GROUPE DE VICTIMES

Zone	Nombre de groupes	Contenu des groupes
0	2	{v1,v25,v5,v6,v21,v23,v7,v37} {v29,v38}
1	1	{v34,v20,v14,v31,v9,v3,v27,v24,v32,v2,v15}
2	2	{v4,v28,v33,v18,v0,v13,v16,v35,v26} {v39}
3	3	{v30,v12,v10,v8,v36,v17,v19} {v22} {v11}
Total	8	

Pour déterminer les besoins de chaque victime, alors nous avons utilisé la méthode que nous avons proposée dans notre approche et un BK structurée dans un fichier XML. Le tableau II montre les compétences nécessaires par certaines victimes (dans notre simulation, nous avons utilisé 20 compétences différentes).

Algorithm 3: Updating Algorithm of coalitions of rescue robots. Behaviour of an agent a_i

Input: Let $G_j \in G$ be a group of victims that we would update its coalition of rescue robots;
Input: Let $a_i \in A$ be the rescue robot leader of the group G_j;
Input: Let σ_{a_i} be the coalition of rescue robots for the group G_j;
Input: $temp \leftarrow true$;
Output: σ_{a_i};

```
while (true) do
    if (message − received() = true) then
        if (a_i.role =' Leader') then
            a_p ← choose − the − member − with − the − biggest − proposed − gain();
            σ_{a_i} ← σ_{a_i} ∪ {a_p};
            if (a_p.state = free) then
                a_p.state ← busy;
            else
                leave − the − ex − coalition();
                a_p.state ← busy;
            temp ← true;
        if (a_i.role =' Member') then
            if (a_i.hasCompetence(δ)) then
                replay − the − leader − with − the − biggest − proposed − sum(gain(a_i, G_j));
    if (a_i.role =' Leader') ∧ (temp = true) then
        foreach (a_p ∈ σ_{a_i}) do
            look − for − new − victims();
            add − new − discovered − victims − to(G_j);
        foreach (v_p ∈ G_j) do
            L_1.addLast(priority(v_p));
        Ω^{G_j} ← extract − all − needed − competences − by − victims − in(G_j);
        Ω^{σ_{a_i}} ← extract − all − robots − competences − in(σ_{a_i});
        Δ^{G_j}_{σ_{a_i}} ← Ω^{G_j} − Ω^{σ_{a_i}} = {ω ∈ Ω^{G_j} ∧ ω ∉ Ω^{σ_{a_i}}};
        H^{Ω^{G_j}}_{G_j} ← extract − all − needed − resource − by − victims − in(G_j);
        H^{Ω^{σ_{a_i}}}_{a_i} ← extract − all − robots − resource − in(σ_{a_i});
        if (Δ^{G_j}_{σ_{a_i}} ≠ ∅) ∧ (H^{Ω^{σ_{a_i}}}_{a_i} ≥ H^{Ω^{G_j}}_{G_j}) then
            δ ← select − an − element − from(Δ^{G_j}_{σ_{a_i}});
            sum ← calculate − sum − of − priorities − of − victims − whose − need(δ, L_1);
            diffuse − a − message − to − members − to − ask − for(δ, sum);
            temp ← false;
```

TABLE II. LES BESOINS DE QUELQUES VICTIMES

Victimes	ω1	ω2	ω3	⋯	ω17	ω18	ω19	ω20
v0	1	0	0	⋯	0	1	1	1
v1	0	1	0	⋯	1	0	0	1
v2	1	0	1	⋯	0	1	1	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Dans le tableau II, une cellule contenant la valeur 1 signifie que la victime qui se trouve en ligne a besoin de la compétence qui est en colonne.

Dans notre cas de simulation, nous avons utilisé 40 robots sauveteurs qui doivent être organisés en 8 coalitions (le nombre de coalitions doit être égal au nombre de groupes de victimes). Par analogie, chaque robot sauveteur dispose d'un ensemble de compétences et le tableau III montre les compétences de certains d'entre eux.

TABLE III. LES COMPETENCES DE QUELQUES ROBOTS

Robots	ω1	ω2	ω3	⋯	ω17	ω18	ω19	ω20
a0	0	0	1	⋯	0	0	1	0
a1	1	0	0	⋯	1	1	0	0
a2	0	1	0	⋯	0	0	0	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Le tableau IV montre les besoins de chaque groupe de victimes, en termes de compétences, et l'identifiant de son leader obtenus par l'algorithme 1.

TABLE IV. LES LEADERS ET LES BESOINS DE CHAQUE GROUPE DE VICTIMES

Groups	Leaders	ω1	ω2	ω3	⋯	ω18	ω19	ω20
G0	a17	0	1	0	⋯	0	0	1
G1	a1	1	1	1	⋯	1	1	1
G2	a2	1	1	0	⋯	1	1	1
G3	a11	1	1	0	⋯	1	1	1
G4	a19	1	1	0	⋯	1	1	1
G5	a16	0	1	0	⋯	0	0	0
G6	a10	1	1	0	⋯	1	1	1
G7	a9	1	0	0	⋯	1	0	1

Le tableau V montre les coalitions initiales obtenues par l'algorithme 2.

TABLE V. COALITIONS INITIALES OBTENUES PAR L'ALGORITHME 2

Groups	**Leaders**	**Content of coalitions**
G0	a17	{a17,a29}
G1	a1	{a1}
G2	a2	{a2}
G3	a11	{a11,a3}
G4	a19	{a19,a21,a26}
G5	a16	{a16}
G6	a10	{a10}
G7	a9	{a9,a13,a8,a22,a27}

Le tableau VI montre trois itérations de l'algorithme 3 et comment le contenu de nos coalitions est modifié.

TABLE VI. DIFFERENTES COALITIONS OBTENUES PAR L'ALGORITHME 3

Groups	Leaders	Content of coalitions
G0	a17	{a17,a29,a14}
G1	a1	{a1,a4,a12,a20,a6}
G2	a2	{a2,a5}
G3	a11	{a11,a3}
G4	a19	{a19,a21,a26,a0,a24}
G5	a16	{a16}
G6	a10	{a10}
G7	a9	{a9,a13,a8,a22,a27}

Iteration 1

G0	a17	{a17,a29,a14}
G1	a1	{a1,a4,a12,a20,a6}
G2	a2	{a2,a5,a18,a25}
G3	a11	{a11,a3,a15,a23}
G4	a19	{a19,a21,a26,a0,a24}
G5	a16	{a16}
G6	a10	{a10}
G7	a9	{a9,a13,a8,a22,a27}

Iteration 2

G0	a17	{a17,a29,a14}
G1	a1	{a1,a4,a12,a20,a6}
G2	a2	{a2,a5,a18,a25}
G3	a11	{a11,a3,a15,a23}
G4	a19	{a19,a21,a26,a0,a24}
G5	a16	{a16}
G6	a10	{a10}
G7	a9	{a9,a13,a8,a22,a27}

Iteration 3

VI. CONCLUSION ET PERSPECTIVES

Dans ce papier, nous avons proposé une approche pour la formation de coalitions de robots sauveteurs. Tout d'abord, nous avons proposé une méthode pour la détermination des besoins, en termes de compétences, de chaque victime dans notre environnement. Ensuite, pour chaque groupe de victimes nous avons proposé un algorithme pour en désigner un leader. Enfin, nous avons proposé deux algorithmes de négociation fondés sur des valeurs d'utilité et ce pour l'initialisation et la modification du contenu de nos coalitions de robots sauveteurs.

Notre approche traite le problème de formation de coalitions dans un système multi-agents en utilisant un ensemble de compétences et nous avons supposé qu'aucune information sur les compétences des robots n'est partagée ce qui est en fait plus proche de la réalité. En outre, dans notre approche les valeurs d'utilité sont calculées dynamiquement ce qui nous permettra de prendre en compte la dynamique de l'environnement et modifier ainsi le contenu de chaque coalition en fonction de l'état de l'environnement.

Dans un futur travail, nous voulons tester l'applicabilité de notre approche sur d'autres études de cas et ce pour permettre de dresser un tableau comparatif de ses performances sur les différents champs d'application et ce afin de valider les résultats expérimentaux et mettre en avant l'adaptabilité de notre approche sur les différentes problématiques.

REFERENCES

[1] A. Galstyan, K. Czajkowski, and L. K., "Resource allocation in the grid with learning agents," J.GridComput, pp. 91–100, 2005.

[2] J. Corchado, D. Tapia, and J. Bajo, "A multi-agent architecture for distributed services and applications," Int.J.Innov.Comput.Inf.Control, pp. 2453–2476, 2012.

[3] P. Papazoglou, D. Karras, and R. Papademetriou, "On multiagent based modeling and control in large scale wireless communication systems for improved resource allocation performance," Telecommun.Syst, pp. 1657– 1675, 2013.

[4] M. Huhns, M. Singh, M. Burstein, Decker, K., E. Durfee, T. Finin, L. Gasser, H. Goradia, N. Jennings, K. Lakkaraju, H. Nakashima, V. Parunak, J. Rosenschein, A. Ruvinsky, G. Sukthankar, S. Swarup, K. Sycara, M. Tambe, T. Wagner, and L. Zavala, "Research directions for service-oriented multiagent systems," IEEE Internet Comput, pp. 65–70, 2005.

[5] S. Stein, E. Gerding, and N. Jennings, "Optimal task migration in service-oriented systems : algorithms and mechanisms," in:Proceedings of the 19th European Conference on Artificial Intelligence, IOSPress, Amsterdam,TheNetherlands, pp. 73–78, 2010.

[6] S. Ramchurn, A. Farinelli, K. Macarthur, and N. Jennings, "Decentralized coordination in robocup rescue," Comput.J, pp. 1447–1461, 2010.

[7] S. Barrett and P. Stone, "An analysis framework for adhoc teamwork tasks," in:Proceedings ofthe 11th International Conference on Autonomous Agents and Multiagent Systems, pp. 357–364, 2012.

[8] S. Ruggieri, "Efficient c4.5 classification algorithm," Knowledge and Data Engineering, IEEE Transactions, pp. 438–444.

[9] O. Shehory and S. Kraus, "Feasible formation of coalitions among autonomous agents in non superadditive environments," Comput.Intell, pp. 218–251, 1999.

[10] J. Yang and Z. Luo, "Coalition formation mechanism in multi-agent systems based on genetic algorithms," Appl.SoftComput, pp. 561–568, 2007.

[11] T. Sandholm, K. Larson, M. Andersson, O. Shehory, and F. Tohme, "Coalition structure generation with worst case guarantees," Artif.Intell, pp. 209–238, 1999.

[12] T. Rahwan, S. Ramchurn, N. Jennings, and A. Giovannucci, "An anytime algorithm for optimal coalitions tructure generation," J.Artif.Intell.Res, pp. 521–567, 2009.

[13] S. Abdallah and V. Lesser, "Organization-based cooperative coalition formation," in:Proceedings ofthe IEEE/WIC/ACM International Conference on Intelligent Agent Technology,IAT'04,IEEE Computer Society, Washington,DC,USA, pp. 162–168, 2004.

[14] R. Katayanagi and T. Sugawara, "Efficient team formation based on learning and reorganization and influence of communication delay," in:Proceedings o fthe 11th IEEE International Conference on

Computer and Information Technology (CIT-2011), pp. 563–570, 2011.

[15] O. Shehory and S. Kraus, "Methods for task allocation via agent coalition formation," Artificial Intelligence, pp. 165–200, 1998.

[16] A. Giovanni, J. Diane, R. Parisa, and V. Athanasios, "A survey on ambient intelligence in health care," Proc IEEE Inst Electr Electron Eng, pp. 2470–2494, 2013.

[17] F. Jaques, "A meta-model for the analysis and desing of organization in multi-agent systems," Proceedings. International Conference on Multi-Agent Systems, pp. 3–7, 1998.

[18] S. Brooks and J. Suchey, "Skeletal age determination based on the os pubis : A comparison of the acsadi-nemeskeri and suchey-brooks methods," Springer Link, pp. 227–238, 1990.

[19] K. Sullivan, A. Mannucci, C. Kimpton, and P. Gill, "A rapid and quantitative dna sex test : fluorescence-based pcr analysis of x-y homologous gene amelogenin," Europe PubMed Central, 1993.

Les techniques de data mining dans les systèmes de recommandation

Abida Benahcene
Laboratoire des Sciences et Technologies de l'Information et de la Communication LabSTIC, Université 8 Mai 1945, Guelma
benahceneabida@gmail.com

Yacine Lafifi
Laboratoire des Sciences et Technologies de l'Information et de la Communication LabSTIC, Université 8 Mai 1945, Guelma
laf_yac@yahoo.fr

Résumé— ***Les dernières tendances dans le domaine des Environnements Informatiques d'Apprentissage Humain (EIAH) cherchent à personnaliser leurs dispositifs pour adapter le processus d'apprentissage à leurs utilisateurs. L'une des propositions est l'intégration d'un système de recommandation dans un EIAH. L'utilisation de cet outil permet de guider les utilisateurs pour atteindre leurs objectifs souhaitables dans une grande masse de données en se basant sur plusieurs dimensions (préférences, intérêts, besoins, …etc.). En plus, ces systèmes permettent de fournir une assistance dynamique ou statique lors de la prise d'une décision dans un moment critique. Ces systèmes ont rencontré des obstacles et des lacunes comme le démarrage à froid, l'insuffisance de données disponibles sur les utilisateurs pour faire des recommandations de qualité, …etc.***

Plusieurs méthodes et techniques ont été proposées comme remède pour résoudre ces problèmes. Parmi les plus utilisées celles de data mining dont le but est l'extraction des données pertinentes depuis les bases de données volumineuses, ainsi que la découverte des connaissances exploitées pour la prédiction des informations manquantes.

Dans ce papier, nous présentons un survol sur les systèmes de recommandation, ainsi que l'utilité des techniques de data mining pour améliorer le processus de la recommandation. Nous présentons aussi une architecture générale d'un système de recommandation basé sur l'une des techniques de data mining.

Mots clés— **EIAH , Système de recommandation, Data mining, Machine Learning, Personnalisation, Apprentissage humain.**

I. INTRODUCTION

Dernièrement, lors de l'accès à la plupart des sites web pour la recherche de l'information dans divers domaines, on rencontre une liste des items suggérés par la plateforme navigué qui sont similaires à notre recherche. Cette suggestion exprime une expérimentation d'une nouvelle filière dans le cadre de la communauté scientifique qu'on l'appelle « systèmes de recommandation ». Leur objectif principal est d'assurer une assistance utile en "bon moment" aux utilisateurs qui n'ont pas une expérience suffisante dans un domaine donné. L'assistance proposée par ces systèmes consiste à fournir une personnalisation des informations de la plateforme selon les intérêts, les préférences et les besoins de leurs utilisateurs.

Récemment, les systèmes de recommandation sont intégrés avec succès dans les applications mobiles (ex : les téléphones intelligents, tablettes, …etc.) dont le but est d'offrir un guide automatique, interactif et intelligent. L'avantage majeur d'utiliser cet outil est le filtrage d'une grande masse des informations téléchargées via le net. Dans ce cas, un système de recommandation est considéré comme un système de filtrage de l'information.

Malgré les avantages apportés par l'exploitation des systèmes de recommandation, ils souffrent de plusieurs lacunes, telles que la façon de réagir avec un utilisateur inconnu par le système (nouvel utilisateur), ou bien comment recommander un nouveau item ? Comment inférer les informations manquantes ? Sur quelle base on génère le top-N recommandation ? Comment choisir les voisins fiables à un utilisateur ? Toutes ces interrogations ont été abordées par la communauté des chercheurs en essayant de trouver les meilleures solutions. Ces lacunes influencent la performance totale du système et réduisent le niveau d'efficacité et de précision des prédictions qui causent l'insatisfaction de l'utilisateur, l'abondance de ce dernier ainsi que l'arrêt de l'utilisation du système de recommandation.

Les techniques de data mining peuvent être exploitées avec efficacité pour combler ces lacunes. L'objectif principal d'un processus de data mining est l'extraction des modèles, la découverte des relations entre les individus et les objets dans une base de données volumineuse. Ces découvertes connaissances sont exploitées dans le e-Learning pour différents objectifs tels que la prédiction de la performance, la prédiction du comportement des apprenants…etc. par l'analyse de leurs interactions avec le système.

L'objectif de ce travail est la présentation de l'utilité de l'utilisation des techniques de data mining dans les systèmes de recommandation à travers une étude comparative entre plusieurs approches.

Ce papier est structuré comme suit : section 2 présente les concepts de base des systèmes de recommandation. Section 3 aborde les différentes techniques de data mining investiguées dans les systèmes de recommandation pour résoudre leurs anomalies ainsi que l'amélioration de la qualité des recommandations. Section 4 discute le contexte de notre recherche où nous présentons une liste de contributions pour la conception d'un système de recommandation appliqué dans le e-Learning en améliorant le processus d'apprentissage humain et finalement, section 5 présente une conclusion générale.

II. SYSTEMES DE RECOMMANDATION

Dans une large quantité d'informations téléchargée depuis le net, l'intégration d'un système de recommandation est très utile dans ce cas. Un tel système aide les utilisateurs à atteindre leurs objectifs souhaitables en utilisant des techniques de filtrage de données selon les besoins, les intérêts, les préférences…etc. [Hsu, 2008a]. Son principe général est la collecte des données sur les utilisateurs et les items pour aboutir à des recommandations personnalisées [Capuano et al., 2013].

Les systèmes de recommandation sont intégrés en premier lieu dans les sites du commerce électronique, dont le but est la suggestion des produits à acheter aux clients, ainsi que l'offre d'un guide en ligne pour bien choisir le produit approprié. Dans le même contexte, ces systèmes sont exploités avec succès dans d'autres secteurs tels que le e-Learning, e-tourisme,…etc.

Dans le tableau suivant, on essaie de récapituler les différents objectifs de la recommandation dans quelques domaines.

TABLE I. OBJECTIFS DES SYSTEMES DE RECOMMANDATION DANS DIVERS DOMAINES

Domaine d'application	***Objectifs***	***Ressources recommandées***
E-commerce	-Fournir une assistance dynamique afin d'accomplir une tâche complexe. -Proposer des produits aux clients (préférences, besoins,…etc.). -Fournir des informations pour aider les consommateurs à déterminer quels produits à acheter. -Suggérer des consommateurs plus expérimentés pour les clients novices.	Films [Kim et al., 2011], vidéos, musique [Jawaheer,2010], …etc.
E-Learning	-Motivation des apprenants. -Création d'une meilleure combinaison des cours selon les préférences des apprenants. -Identification des points faibles des apprenants. -Aider les enseignants à améliorer leurs stratégies d'enseignement. -La recommandation des communautés pour la collaboration lors de l'apprentissage. -Optimiser le chemin lors de la recherche de l'information. -Améliorer la performance des apprenants dans différentes matières.	Cours[Hsu,2008a] Liens, ressources multimédia [Romero et al., 2007], Collaborateurs [Di Bitonto et al., 2011], Scénario d'activité [De Oliveira et al.,2013], …etc.
E-tourisme	-Recommander les endroits touristiques. -Faciliter la recherche de l'information pour l'hébergement ou le déplacement. -Offrir un guidage interactif (changement de destination, visualisation du trajet) prenant en compte le temps, le lieu et le moyen du transport. -Améliorer les services offerts par les agences de tourisme pour augmenter le chiffre d'affaire.	Lieux touristiques [Camara et al., 2010], Séjours [Picot-Clémente, 2011], Moyens de transport [Chelcea et al., 2004], …etc.

A. Mécanisme de fonctionnement d'un système de recommandation :

Si on considère le système de recommandation comme une boite noire, il a deux entrées : un ensemble des utilisateurs et un ensemble des items ; et comme sortie on a une liste des objets recommandés. Si on ouvre la boite noire, on va obtenir un ensemble d'algorithmes purement mathématiques dans la plupart du temps, pour prédire les données manquantes dans la matrice (utilisateur, item) en utilisant des mesures de similarité.

Le processus de recommandation nécessite en général le passage par trois étapes essentielles [http1] :

- La collecte des informations sur les utilisateurs et les items implicitement ou explicitement.
- L'élaboration de la matrice (utilisateur, item) selon un modèle d'utilisateur bien défini.
- L'établissement de la liste des objets recommandés.

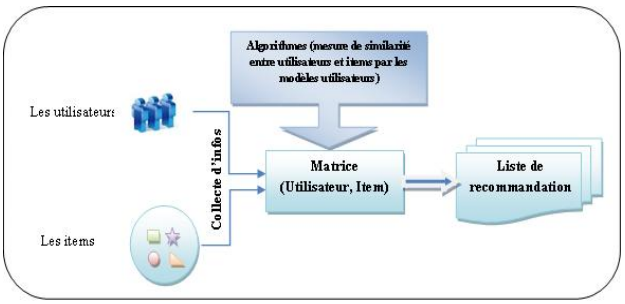


Fig. 1. Mécanisme de fonctionnement d'un système de recommandation.

B. Type des systèmes de recommandation :

Il existe trois types standards des systèmes de recommandation qui sont les plus utilisés et les plus connus [Zhuhadar et al., 2009]. Commençant par le système de recommandation basé sur le contenu. Le principe de base de ce type des systèmes est l'exploitation de l'historique de la navigation d'un utilisateur. Ses interactions sont représentées par les activités effectuées dans le passé. Elles sont stockées dans un profil utilisateur lors de la navigation dans la plateforme. Le système va suggérer des objets similaires aux objets déjà consultés antérieurement.

La recommandation à base de filtrage collaboratif est un autre genre de ces systèmes, qui se base sur la similarité entre les utilisateurs. Il cherche des objectifs communs et essaie de prédire les informations manquantes. Ce type de la recommandation est largement utilisé dans plusieurs domaines dans ce dernier temps.

Les deux types abordés précédemment sont réunis dans une approche dite hybride. Cette combinaison permet d'éviter certaines limitations de chaque type séparément.

L'utilisation de ces types n'a pas assuré des recommandations de qualité. L'intégration de data mining dans ce contexte est très efficace pour gagner la confiance de l'utilisateur avec un taux de satisfaction très élevé.

III. UTILITE DE DATA MINING DANS LES SYSTEMES DE RECOMMANDATION

A. Data mining :

Le data mining (*DM*) est l'une des étapes principales du processus de découverte des connaissances dans les bases de données de grande taille. Ce processus est constitué de trois phases essentielles : préparation de données, extraction des modèles et évaluation des modèles. La première phase représente un ensemble des prétraitements des données tels que la collecte, l'intégration, la transformation, le nettoyage, la réduction et la description de données. Dans la deuxième phase, plusieurs techniques de data mining sont utilisées pour l'extraction et la découverte des modèles tels que le clustering, les règles d'association,...etc. La dernière phase consiste à faire l'évaluation et l'interprétation des modèles extraits pour aboutir à la fin à des modèles, patterns valides, nouveaux, potentiellement utiles, et compréhensibles [http2] [Amatriain et al., 2011] [Liao et al., 2012].

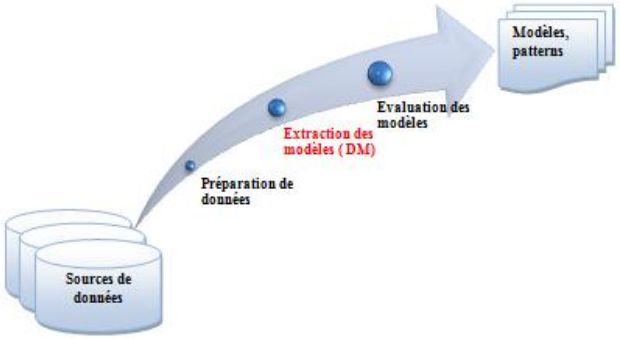


Fig. 2. Processus de la découverte des connaissances [http2].

Comme on a mentionné ci dessus, l'extraction des modèles se fait par l'application des techniques mathématiques en exploitant le processus de data mining. Ces techniques sont utilisées dans différents axes tels que les systèmes de recommandation.

B. Techniques de data mining dans les systèmes de recommandation :

Les techniques de data mining sont nombreuses et sont utilisées dans différents domaines pour divers objectifs. Dans ce cas, le choix d'une technique ou d'une autre est une tâche très difficile puisqu'il faut prendre en considération la problématique à traiter, le contexte applicatif, la disponibilité des données à l'entrée du système, ainsi que les avantages et les inconvénients de la technique choisie [Esslimani, 2010].

La technique de K plus proches voisins (K Nearest Neighbors : KNN) est largement utilisée par les systèmes de recommandation basés sur le filtrage collaboratif. Son principe est la découverte et l'extraction des K plus proches utilisateurs appelés voisins. Ces derniers partagent des intérêts communs avec l'utilisateur cible [Park et al., 2012].

Cette technique a prouvé un double avantage pour la recommandation des objets d'apprentissages aux apprenants. Le premier avantage est représenté par [De Oliveira et al., 2013] qui exprime l'utilité de l'algorithme "Multi Label-Nearest Neighbor" (ML-NN) pendant le regroupement des apprenants qui ont des difficultés d'apprentissage en utilisant leurs profils communs. Le deuxième avantage de cette technique est abordé dans le travail de [Di Bitonto et al., 2011] pour confronter le démarrage à froid d'un utilisateur. Pour cette raison, les auteurs ont utilisé "Shared-Nearest Neighbor clustering" (SNN). Ce dernier est appliqué dans un contexte social pour enrichir le processus d'apprentissage.

On déduit que la technique de KNN a été adoptée avec efficacité dans le processus de la recommandation dans le e-Learning. En plus, elle a résolu pas mal de problèmes rencontrés pendant le cycle d'apprentissage.

Malheureusement, la classification obtenue en appliquant cette technique est liée au choix initial du K. Ainsi, elle n'est pas flexible lors de l'évolutivité du nombre des utilisateurs et des objets recommandés.

Contrairement à la technique du KNN, les arbres de décision sont trop flexibles grâce à leur structure, qui est adaptable aux différents besoins du concepteur. Cette technique a été appliquée avec succès par [Anaya et al., 2013] pour améliorer le processus de la collaboration entre les apprenants pendant leur apprentissage. Leur approche vise à identifier les difficultés de la collaboration et la proposition des solutions alternatives aux apprenants sous forme de conseils. Le problème de démarrage à froid est encore discuté par [Lika et al., 2014]. Les auteurs ont appliqué un processus intelligent pour chercher les voisins qui partagent des caractéristiques communes avec le nouvel utilisateur. La classification de ce dernier se fait par l'application de l'algorithme C4.5.

L'inconvénient majeur rencontré par ces approches est la complexité dans la gestion d'un nombre important des utilisateurs et des objets, ce qui nécessite un temps considérable pendant la phase d'apprentissage ainsi que la phase de classification. Ce constat permet de diminuer la performance du système de recommandation en général.

Les règles d'association sont largement exploitées par les systèmes de recommandation grâce à leur structure symbolique pour expliquer le processus de la recommandation aux utilisateurs. Leur principe est la découverte des relations entre les groupes des apprenants qui partagent le même comportement et la liste des objets d'apprentissage sous forme d'un processus de personnalisation [Hsu, 2008a].

Les règles d'association sont utilisées pour la prédiction des actions des apprenants dans le futur. L'algorithme "Sequentiel Pattern Mining" (SPM) est exploité pour analyser la séquence d'accès aux matières effectuée par les apprenants dans le passé [Salehi et Kamalabadi, 2013].

TABLE II. DIFFERENTES TECHNIQUES DE DATA MINING UTILISEES DANS LES SYSTEMES DE RECOMMANDATION

Auteurs	*Objectifs de l'approche*	*Domaines d'application*	*Techniques de DM*	*Ressources recommandées*	*Utilisateur cible*
De Oliveira et al., 2013	Améliorer la performance des apprenants dans la programmation.	E-Learning	-Multi-Label K-nearest neighbor (ML-kNN)	Classes des activités	Apprenants
Di Bitonto et al., 2011	-Enrichir le processus d'apprentissage en exploitant les interactions sociales des apprenants. -Réduire le problème de démarrage à froid.	E-Learning	- The shared Near Neighbor (SNN) clustering algorithm	Groupes de discussion	Apprenants
Anaya et al., 2013	Détecter et résoudre les problèmes du processus de la collaboration.	E-Learning	-Influence diagram -Decision tree algorithm	Conseils en cas de problèmes de collaboration	Enseignants et apprenants
Lika et al., 2014	Éviter le problème de démarrage à froid en utilisant trois étapes: classification, similarité et prévision.	E-commerce	-Classification par les algorithmes C4.5 et Naïve Bayes	Films	Publique
Hsu, 2008a	Motiver les apprenants par la personnalisation des cours.	E-Learning	-Clustering algorithm -Association rules algorithm	Les cours	Apprenants
Kim et Yum, 2011	Amélioration de la qualité des recommandations en se basant sur le comportement des utilisateurs ainsi que leur historique de navigation.	E-commerce	-Association rule mining	Produits	Publique
Aher et Lobo, 2012	Perfectionner la précision de la recommandation.	E-Learning	-Simple k-means clustering. -Apriori association rule algorithm	Les cours	Apprenants
Romero et al., 2007	-Aider les étudiants à trouver un chemin optimal à travers les objets d'apprentissage. -Alléger le problème de démarrage à froid.	E-Learning	-Clustering algorithm -Sequential pattern mining algorithm	Lien optimal	Enseignants et apprenants
Chen et al., 2014	Recommandation des ressources d'apprentissage utiles pour guider les utilisateurs pendant le processus d'apprentissage.	E-Learning	-Sequential pattern mining	Ressources d'apprentissage	Apprenants
Kim et al., 2011	Améliorer la qualité des recommandations en exploitant les tags des utilisateurs comme des indicateurs de leurs préférences.	E-commerce	-Weighted Naïve Bayes	Films	Publique
Hsu, 2008b	Améliorer les stratégies d'enseignement en identifiant les points faibles des apprenants.	E-Learning	-Clustering algorithm	Stratégies d'enseignement	Enseignants et apprenants
Salehi et Kamalabadi, 2013	- Personnaliser la recommandation en utilisant les activités des apprenants. -Amélioration de la qualité des recommandations. - Résoudre le problème de la densité faible de données.	E-Learning	-Association rules -Pattern-tree -K-means clustering algorithm	Les matières	Publique
Wu et al., 2013	-Résoudre le problème de démarrage à froid dans les systèmes de recommandation collaborative. -Amélioration de l'exactitude des recommandations en se basant sur les utilisateurs actifs.	E-commerce, E-Learning	-K-means clustering algorithm	Films, Papiers académiques	Publique
Zhang et al., 2014	-Optimiser la matrice de factorisation par le regroupement des utilisateurs en utilisant leurs informations.	E-commerce	-K-means clustering algorithm	Films	Publique

Les règles d'association sont appliquées dans le e-commerce pour étudier le comportement des clients pendant leurs acquisitions des produits en ligne en analysant leur processus d'achat implicitement d'après leurs produits consultés, les produits achetés,…etc. et essayer de découvrir une relation d'association entre eux [Kim et Yum, 2011]. Un autre algorithme des règles d'association est "Apriori Association Rule" qui traite le cas des patterns qui se produisent souvent. Cette technique est utilisée par [Aher et Lobo, 2012] pour la prédiction d'une combinaison meilleure des cours aux apprenants selon leurs préférences. Elle se base sur l'historique de la navigation des utilisateurs pendant leurs interactions avec le système. Les traces d'interaction sont adaptées par cette technique selon les besoins des apprenants lors de la recherche des matières d'apprentissage [Romero et al., 2007]. En outre, Romero et ses collègues ont résolu le problème de démarrage à froid d'un nouvel utilisateur en se basant sur sa navigation actuelle, par la suggestion des cours selon cette situation. "Le Sequentiel Pattern Mining algorithm" est utilisé encore par [Chen et al., 2014] pour consolider la recommandation basée sur le filtrage collaboratif. Cette technique est appliquée sur les matières pour découvrir la relation entre elles.

L'exploitation de l'historique de la navigation nécessite un multiple scan de la base de données pour déterminer les patterns fréquents (temps coûteux), ce qui est considéré comme un inconvénient majeur des approches exploitant ce principe.

Les principes de la probabilité sont appliqués dans les systèmes de recommandation dont le but est la prédiction des préférences des utilisateurs. "Locally Weited Naïve Bayes" est une approche qui exploite le théorème de bayes dans une recommandation collaborative [Kim et al., 2011]. Cette approche utilise les étiquettes générées par les utilisateurs comme des indicateurs de préférence. Ce principe aide à confronter le problème de démarrage à froid.

Le clustering est une classification non-supervisée. Son objectif principal est le regroupement des objets similaires dans la même classe en maximisant la distance interclasses et en diminuant la distance intra-classes [Romero et al., 2007] [Zakrzewska, 2012]. La plupart des systèmes de recommandation utilisent un algorithme non hiérarchique très connu qui est le K means vu ses avantages majeurs pour combler les lacunes de la classification dans une base de données volumineuse, ainsi que sa simplicité et efficacité. L'utilité d'utiliser cette technique dans le e-Learning est abordée par [Hsu, 2008b] pour l'identification des points faibles des apprenants par un regroupement efficace, ainsi que pour fournir l'aide appropriée aux enseignants en leurs suggérant des nouvelles stratégies d'enseignement suite à la détection des difficultés de leurs apprenants. Par fois, cette technique de clustering est combinée avec d'autres techniques telles que les règles d'association et les arbres de décision afin d'améliorer la prédiction des recommandations [Salehi et Kamalabadi, 2013]. Cette combinaison entre différentes techniques de data mining peut être considérée comme une phase de prétraitement où on applique le clustering pour améliorer, réduire et nettoyer les données à exploiter [Aher et Lobo, 2012].

Dans la plupart des approches proposées, le clutering est appliqué soit sur les utilisateurs, soit sur les objets, et par fois sur les deux en même temps. Comme le cas dans [Wu et al., 2013] où les auteurs ont essayé d'améliorer la performance de la collaboration dans un réseau social et d'affronter le problème du démarrage à froid.

Enfin, le clustering est utilisé pour résoudre le problème de la dispersion de données (l'indisponibilité de données) en optimisant la matrice de factorisation standard par l'intégration des informations des utilisateurs [Zhang et al., 2014].

IV. CONTEXTE DE NOTRE RECHERCHE :

A travers l'énumération des travaux de recherche menés sur la recommandation, on constate que l'aspect crucial dans un moteur de recommandation est le moment d'intervention du système pour fournir l'assistance à l'utilisateur concerné en bon moment. Ce dernier est appelé moment critique ou situation d'impasse ou de blocage, qui peut être exprimé par un processus qui s'exécute en arrière-plan et lors de détection de ces situations le système de recommandation va se déclencher.

Le concept du moment critique est abordé par plusieurs approches dans divers contextes avec différentes notions. DRARS (Dynamic Risk-Aware Recommender System) est un système de recommandation dynamique qui est sensible au contexte [Bouneffouf, 2013]. Il prend en considération le niveau de risque de la situation de l'utilisateur dans un contexte industriel afin d'améliorer la performance de la recommandation. Bouneffouf (2013) a défini les situations critiques par l'ensemble des situations où l'utilisateur a besoin des informations importantes qui peuvent être recommandées par le système pour diminuer l'inquiétude de l'utilisateur dans un moment donné. [Zarka, 2013] a aussi abordé ces moments dans un contexte dynamique pour accomplir des tâches complexes. Il a représenté ces situations par des éléments observés qui sont utilisés par le paradigme de raisonnement à partir de cas.

De ce constat, on essaie de réaliser les contributions suivantes :

- La conception d'une approche qui permet la détection/prédiction des moments d'intervention en utilisant l'une des techniques de data mining/machine learning.
- L'identification des types d'intervention selon les moments d'intervention détectés (tutorat, collaboration, évaluation, apprentissage,…etc.).
- La conception d'un système de recommandation qui peut intervenir en bon moment pour suggérer une liste des ressources importantes (personne, activité pédagogique, scénario d'actions,…etc.)

L'ensemble des tâches abordées ci dessus est schématisé dans la figure suivante :

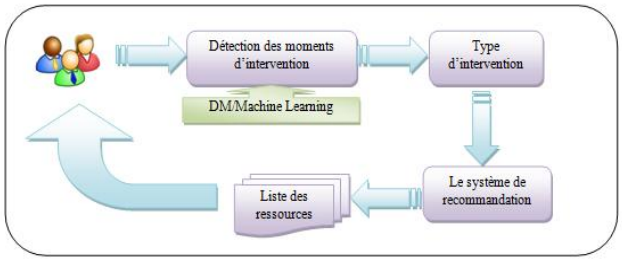


Fig. 3. Architecture générale de l'approche proposée.

Cette architecture est détaillée dans les phases suivantes :

- Phase de détection/prédiction des moments d'intervention : le choix d'une technique de DM/Machine Learning est indispensable dans cette phase pour la détection/prédiction des moments d'assistance. Elle peut incorporer une étape d'apprentissage automatique sur les comportements des utilisateurs. Nous proposons d'utiliser K-means dans un premier temps.

- Type d'intervention : d'après la phase précédente, le système va identifier le type d'intervention qui peut être exprimé par le type de ressource à recommander, par ailleurs, selon le contexte d'intervention on va recommander la bonne ressource.

- Système de recommandation : cette phase présente le fonctionnement du moteur de recommandation, qui est déclenché suite à la détection du moment d'assistance. Le système prend le type d'intervention comme entrée pour générer à l'utilisateur cible une liste de recommandations appropriées.

V. CONCLUSION

Aujourd'hui, l'intégration des systèmes de recommandation est devenue une nécessité dans la plupart des plateformes commerciales ainsi que dans d'autres domaines. Ces systèmes permettent une gestion intelligente des grandes bases de données d'une façon implicite dont le but est la satisfaction des utilisateurs. La publication des papiers académiques dans cette filière s'augmente d'une manière exponentielle chaque année, ce qui traduit l'attention des chercheurs envers cette discipline. L'évolution des dispositifs technologiques a influencé l'utilisation de la recommandation comme un système de filtrage de l'information, plus précisément avec les applications des Smartphones et des tablettes…etc.

D'après la littérature, on a constaté que les systèmes de recommandation ont été utilisés fréquemment par les sites du commerce électronique pour la recommandation des films (ex : EachMovie, Movielens, Netflix,…etc.).

Malgré que la maturité atteinte par les systèmes de recommandation, ils souffrent de plusieurs problèmes et lacunes. Parmi les solutions proposées l'utilisation des techniques de data mining.

D'après notre lecture, on a déduit que :

- Le clustering a été reconnu comme la technique la plus appliquée par la majorité des systèmes de recommandation, soit pour résoudre le problème de démarrage à froid, ou bien pour alléger le problème du manque de données.
- Le problème de démarrage à froid a été traité avec plusieurs techniques selon différents points de vue.
- La technique du KNN est largement utilisée par les systèmes de recommandation à base du filtrage collaboratif pour la recherche des voisins fiables, ainsi que la recommandation des ressources en se basant sur la similarité avec les autres utilisateurs.
- La classification à base de règles est intégrée avec succès dans les systèmes de recommandation pour inférer des nouvelles règles ainsi que la prédiction des informations manquantes.
- La combinaison entre plusieurs techniques du data mining a pour but de combler les lacunes de certains algorithmes par d'autres.

Enfin, cet état de l'art a prouvé l'efficacité de l'intégration de différentes techniques de data mining pour combler les lacunes rencontrées par les systèmes de recommandation plus précisément dans le domaine d'apprentissage humain. L'objectif est d'améliorer la qualité d'apprentissage des apprenants et faciliter leurs tâches pédagogiques.

L'union entre le data mining et les systèmes de recommandation a permis de fournir un terrain fertile pour expérimenter les idées de la communauté des chercheurs dans divers domaines.

REFERENCES

SB. Aher et L. Lobo, « Applicability of Data Mining Algorithms for Recommendation System in e-Learning », in Proceedings of the International Conference on Advances in Computing, Communications and Informatics, New York, NY, USA, p. 1034–1040, 2012.

X. Amatriain, A. Jaimes, N. Oliver, et JM. Pujol, « Data Mining Methods for Recommender Systems », in Recommender Systems Handbook, F. Ricci, L. Rokach, B. Shapira, et P. B. Kantor, Éd. Springer US, p. 39□71, 2011.

AR. Anaya, M. Luque, et T. García-Saiz, « Recommender system in collaborative learning environment using an influence diagram », Expert Syst. Appl., vol. 40, n° 18, p. 7193□7202, déc. 2013.

D. Bouneffouf, « DRARS, A Dynamic Risk-Aware Recommender System », thèse de doctorat, Institut National des Télécommunications, France, 2013.

F. Camara, R. Demumieux, G. Calvary, et N. Mandran, « Cocoon, un système de recommandation sensible au contexte: analyse de la valeur par une étude qualitative », in Proceedings of the Ergonomie et Informatique Avancee Conference, Biarritz, France, p. 211–218, 2010.

N. Capuano, R. Iannone, M. Gaeta, S. Miranda, P. Ritrovato, et S. Salerno, « A Recommender System for Learning Goals », in Information Systems, E-learning, and Knowledge Management Research, M. D. Lytras, D. Ruan, R. D. Tennyson, P. O. D. Pablos, F. J. G. Peñalvo, et L. Rusu, Éd. Springer Berlin Heidelberg, p. 515 521, 2013.

S. Chelcea, G. Gallais, et B. Trousse, « Recommandations personnalisées pour la recherche d'information facilitant les déplacements », Premières Journées Francophones: Mobilité et Ubiquité, p. 143-150. Nice, France, 2004.

W. Chen, Z. Niu, X. Zhao, et Y. Li, « A hybrid recommendation algorithm adapted in e-learning environments », World Wide Web, vol. 17, n° 2, p. 271–284, 2014.

MG. De Oliveira, PM Ciarelli, et E. Oliveira, « Recommendation of programming activities by multi-label classification for a formative assessment of students », Expert Syst. Appl., vol. 40, n° 16, p. 6641–6651, 2013.

P. Di Bitonto, T. Roselli, et V. Rossano, « Recommendation in e-learning social networks », in Advances in Web-Based Learning-ICWL 2011, Springer, p. 327–332, 2011.

I. Esslimani, « Vers une approche comportementale de recommandation: apport de l'analyse des usages dans un processus de personnalisation », thèse de doctorat, université Nancy II, France, 2010.

MH. Hsu, « A personalized English learning recommender system for ESL students », Expert Syst. Appl., vol. 34, n° 1, p. 683–688, 2008a.

MH. Hsu, « Proposing an ESL recommender teaching and learning system », Expert Syst. Appl., vol. 34, n° 3, p. 2102–2110, 2008b.

G. Jawaheer, M. Szomszor, et P. Kostkova, « Comparison of implicit and explicit feedback from an online music recommendation service », in proceedings of the 1st international workshop on information heterogeneity and fusion in recommender systems, Barcelona, Spain, p. 47–51, 2010.

HN. Kim, A. Alkhaldi, A. El Saddik, et GS. Jo, « Collaborative user modeling with user-generated tags for social recommender systems », Expert Syst. Appl., vol. 38, n° 7, p. 8488–8496, 2011.

YS. Kim et BJ. Yum, « Recommender system based on click stream data using association rule mining », Expert Syst. Appl., vol. 38, n° 10, p. 13320–13327, 2011.

SH. Liao, PH. Chu, et PY. Hsiao, « Data mining techniques and applications–A decade review from 2000 to 2011 », Expert Syst. Appl., vol. 39, n° 12, p. 11303–11311, 2012.

B. Lika, K. Kolomvatsos, et S. Hadjiefthymiades, « Facing the cold start problem in recommender systems », Expert Syst. Appl., vol. 41, n° 4, p. 2065–2073, 2014.

DH. Park, HK. Kim, IY. Choi, et JK. Kim, « A literature review and classification of recommender systems research », Expert Syst. Appl., vol. 39, n° 11, p. 10059–10072, 2012.

R. Picot-Clémente, « Une architecture générique de Systèmes de recommandation de combinaison d'items: application au domaine du tourisme », thèse de doctorat, université de Bourgogne, France, 2011.

C. Romero, S. Ventura, J. A. Delgado, et P. D. Bra, « Personalized Links Recommendation Based on Data Mining in Adaptive Educational Hypermedia Systems », in Creating New Learning Experiences on a Global Scale, E. Duval, R. Klamma, et M. Wolpers, Éd. Springer Berlin Heidelberg, p. 292 306, 2007.

M. Salehi et IN Kamalabadi, « Hybrid recommendation approach for learning material based on sequential pattern of the accessed material and the learner's preference tree », Knowl.-Based Syst., vol. 48, p. 57–69, 2013.

R. Zarka, « Trace-Based Reasoning for User Assistance and Recommendations », thèse de doctorat, université de Lyon, France, 2013.

CX. Zhang, ZK. Zhang, L. Yu, C. Liu, H. Liu, et XY. Yan, « Information filtering via collaborative user clustering modeling », Phys. Stat. Mech. Its Appl., vol. 396, p. 195–203, 2014.

L. Zhuhadar, O. Nasraoui, R. Wyatt, et E. Romero, « Multi-model ontology-based hybrid recommender system in e-learning domain », in Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology-Volume 03, p. 91–95, 2009.

D. Zakrzewska, « Building Group Recommendations in E-Learning Systems », in Transactions on Computational Collective Intelligence VII, N. T. Nguyen, Éd. Springer Berlin Heidelberg, p. 144 163, 2012.

[http1]http://www.podcastscience.fm/dossiers/2012/04/25/les-algorithmes-de-recommandation/ (Dernier accès le 21/11/2013).

[http2]http://www.memoireonline.com/04/10/3310/m_La-desambigusation-des-toponymes5.html (Dernier accès le 22/06/2014).

Fuzzy Rules based classification: Analysis of the effect of Membership functions and Rule Weights

Salah Halima, Nemissi Mohamed, Seridi Hamid
LabSTIC, Université 8 Mai 1945- Guelma- Algérie
{halima_sallah, nemissi_m, seridihamid}@yahoo.fr

***Abstract*—**
The aim of this paper is to examine the effect of certainty grades and membership functions on the performance of fuzzy rule-based classification systems. We present a theoretical study based on the analysis of the classification boundaries and the classification rate of synthetic examples. We carry out tests with different membership functions and with/without certainly grades.

***Index Terms*—Fuzzy rule based classification system, Membership function, Rule Weights.**

I. INTRODUCTION

Fuzzy approaches have become one of the well known solutions for classification problems due to their powerful capabilities to handle both imprecision and uncertainty concepts. In the fuzzy classification, a sample can be matched to numerous different classes with different degrees by applying a set of fuzzy rules of the form (If ... Then ...), based on the linguistic values of different attributes.

To build a fuzzy rule based classification system, the most difficult and important tasks are: generation, learning, selection rules, and the Learning of Rule Weights and membership functions. Numerous approaches have been proposed to accomplish these tasks: the bio inspired techniques have attracted significant research interest, such as Genetic Algorithm (GA) [9] [10] [11], Artificial Immune System (AIS) [12] [13], Ant Colony Optimization (ACO)[14][15] and Particle Swarm Optimization (PSO)[16][17].
In [4] a new approach was presented for the optimum determination of membership functions for fuzzy inference system (FIS) based on the use of extended Ant Colony Optimization. In [6] authors examined some heuristic methods for rule weight specification.

In this paper we focus on the effects of Rule Weights and membership functions on the performance of fuzzy rule based classification system.
Indeed, Nauck and Kruse [5] showed how the learning of rule weights can be equivalently replaced by the modification of the membership functions of antecedent or consequent fuzzy sets. They also showed that the use of rule weights for the learning in fuzzy rule-based systems is not necessary. In a similar focus Ishibuchi and Nakashima [7] examined the effect of rule weights in fuzzy rule-based classification systems. They showed that the learning of membership functions and the use of certainly grade have similar (but not the same) effects on the adjustment of classification boundaries.

This paper is organized as follows. A brief introduction on fuzzy logic is presented in section II. The general structure of the used Fuzzy rule-based classification system is described in section III. We discuss the effect of rule weight and membership function in section IV and V. Section VI concludes this paper.

II. FUZZY LOGIC

The term "fuzzy logic" was introduced in the 1965 by Lotfi A. Zadeh proposal of fuzzy set theory [1]. Fuzzy systems are very effective in expressing the natural ambiguity and subjectivity of human reasoning It contains four basic elements; 1) fuzzification of crisp information, 2) fuzzy knowledge-base derived from humans and the relations between the inputs and outputs, 3) fuzzy inference mechanism based on fuzzy logic, and 4) deffuzification of fuzzy outputs into crisp (non-fuzzy) outputs.
The four basic elements of a fuzzy system are briefly discussed in the following subsections [2].

A. Fuzzification:

The process of transforming crisp values into grades of membership corresponding to fuzzy sets expressing linguistic terms. Many types of curves can be used, but triangular or trapezoidal shaped membership functions are the most common.

B. Fuzzy Knowledge-Base (KB):

The knowledge-base comprised of the Rule Base (RB): a set of if-then rules generated by the expert or through consultations with the expert and the Data Base (DB) that describes the semantic of the Fuzzy subsets associated to the linguistic labels in the if-part of the rules.

C. Fuzzy Inference

The fuzzy inference machine is a decision making logic which employs fuzzy rules from the fuzzy rule base to determine fuzzy outputs of a fuzzy system corresponding to

the fuzzified input. The two commonly used methods to perform fuzzy logic inference are the max-min method and max-product method [8].

D. Defuzzification

The process of transforming the fuzzy outputs obtained in previous step into a single numerical value or crisp values. There are several defuzzification methods in the literature. One of the most popular defuzzification methods is the centroid.

III. FUZZY RULE BASED CLASSIFICATION SYSTEM

A fuzzy rule-based classification system (FRBCS) is one of the most popular approaches used in patterns classification problems, it is known for their transparency and ability of accounting for uncertainty.

We use the following form of Fuzzy if-then rules for an H-class pattern classification problem with n attributes :

$$\text{Rule } R_j\text{: If } x_1 \text{ is } A_{j1} \text{ and } \ldots \text{ and } x_n \text{ is } A_{jn} \text{ then Class } C = C_j \quad (1)$$

where:

- Rj is the label of the j-th fuzzy If Then rule,
- $x_n = (x_1 \ldots\ldots x_n,)$ is an n-dimensional pattern vector
- $A_{j1}, \ldots, A_{jn}$ are triangular fuzzy subsets on the unit interval [0, 1],
- C_j is the consequent or resultant class

We also use fuzzy if-then of the form:

$$\text{Rule } R_j\text{: If } x_1 \text{ is } A_{j1} \text{ and } \ldots \text{ and } x_n \text{ is } A_{jn} \text{ then Class } C=C_j \text{ with } CF=CF_j, \quad (2)$$

where CF is the certainty grade of the fuzzy IF–THEN rule R. Usually CF is a real number in the unit interval [0,1] (i.e.,$0 \leq Cf \leq 1$)

Let us assume that our pattern space is the unit square [0,1]X[0,1] . We also assume that $x_p=(x_{p1},x_{p2})$,p=1,2,….M are considered as training pattern from H-class. In this paper we use symmetric triangular membership function for A_i, i=1,2,…,K, where K is the number of linguistic variable in each axe. Each subspace has its own fuzzy rule.

The consequent part is determined from the given training patterns by the following procedure [3]:

Step1: For each subspace ij calculate the sum of the compatibility grade of all x_p in each class C to the fuzzy rule R_{ij} in(1) and (2)

$$\boldsymbol{\beta_{classC}} = \sum_{x_p \in \, class\ C} \boldsymbol{\mu_p(x_{p1}) \cdot \mu_p(x_{p2})} \quad (3)$$

Step2: Find class *C* for the rule R_{ij} with the maximum value of β class X

$$\boldsymbol{\beta_{class\ x} = \max\,(\beta_{classC})} \quad (4)$$

The consequent class C_{ij} is determined as class x

Step3: when we use fuzzy if-then rule as (2) CF_{ij} is determine as:

$$CF_{ij} = (\beta_{class\ C} - \beta) / \sum_{C=1}^{H} \beta_{classC} \quad (5)$$

where

$$\beta = \sum_{\substack{C=1 \\ C \neq x}}^{H} \beta_{classC} \ /(c-1) \quad (6)$$

After a rule set S is given an input pattern $x_p=(x_{p1},x_{p2})$ is classified by a single winner rule R_{ij} in S.

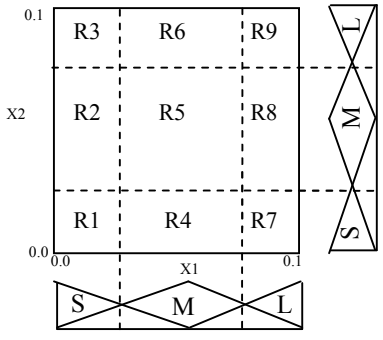
The winner rule for a new pattern is the fuzzy IF–THEN rule that has the maximum compatibility grade with the new pattern

$$\mu(x_p)=\max\{\ \mu_i(x_{p1}).\ \mu_j(x_{p2})|R_{ij}\} \quad (7)$$

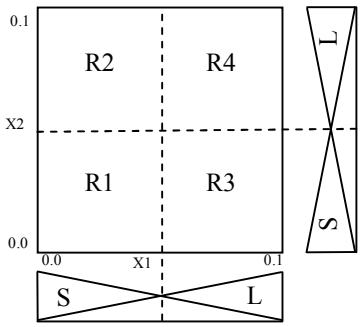
When we use fuzzy IF–THEN rules with certainty grades (i.e., rule weights), the winner is determined as the rule with the maximum product of the compatibility grade and the certainty grade.

$$\mu(x_p).CF_{ij}=\max\{\ \mu_i(x_{p1}).\ \mu_j(x_{p2}).CF_{ij}|R_{ij}\} \quad (8)$$

The generation of fuzzy rules for pattern classification problems consists of two phases: fuzzy partition of the pattern space and identification of a fuzzy rule for each fuzzy subspace



(a)



(b)

Fig. 1. Decision area of each fuzzy IF-THEN rule with: (a) K=3 and (b) K=2

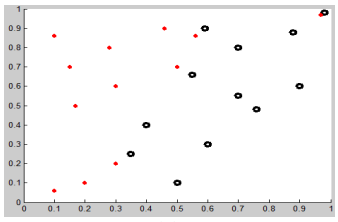
After we partitioned our pattern space into K^2 fuzzy subspaces we determined one rule for each fuzzy subspace. The generated fuzzy rules corresponding to subspaces presented in Fig 1.(b) can be written as :

R1: if x_1 is S and x_2 is S then C_1=class1 with CF_1
R2: if x_1 is S and x_2 is L then C_2=class2 with CF_2
R3: if x_1 is L and x_2 is S then C_3=class2 with CF_3
R4: if x_1 is L and x_2 is L then C_4=class1 with CF_4

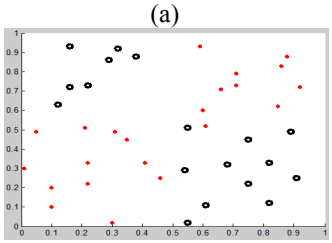
IV. INFLUENCE OF RULE WEIGHTS

In this section we examine the effect of rule weight on classification boundaries obtained using fuzzy IF–THEN rules with/without certainty grades.

We consider two classification problems (fig .2), these problems are bi-dimensional and they are 2-class (calss1 ‘x’, class2 ‘o’)



(a)



(b)

Fig. 2. (a) First Classification problem, (b) Second Classification problem

K	Classification rate	
	Without CF	With CF
K=2	75	86,36
K=3	72,72	81,81
K=4	81,81	95,45

	Classification rate	
	Without CF	With CF
K=2	92	67,5
K=3	67,5	90
K=4	95	97,5

TABLE I: The classification rate of the two classification problems (1,2) with various K.

TABLE I, illustrates the obtained classification rate. The use of certainty grade improved the classification rate except in the case of K=2 in the second problem.

Fig3., Fig4. illustrate the classification boundaries obtained using nine fuzzy if-then rules without and With CF successively.

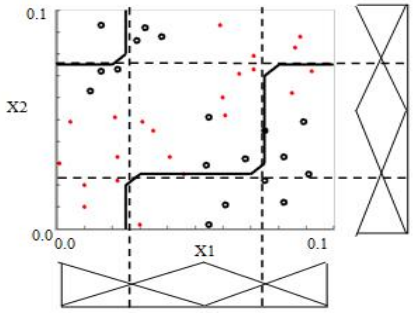


Fig. 3. Classification boundaries without CF, K=3.

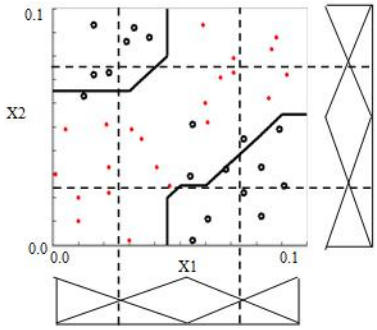


Fig.4. Classification boundaries with CF; K=3.

The fuzzy rules corresponding to the subspaces presented in Fig.4.are as follows:

R1: if x_1 is S and x_2 is S then C_1=1 with CF_1=1
R2: if x_1 is S and x_2 is M then C_2=1 with CF_2=0.3
R3: if x_1 is S and x_2 is L then C_3=2 with CF_3=0.37
R4: if x_1 is M and x_2 is S then C_4=2 with CF_4=0.16
R5: if x_1 is M and x_2 is M then C_5=1 with CF_5=0.17
R6: if x_1 is M and x_2 is L then C_6=1 with CF_6=0.015
R7: if x_1 is L and x_2 is S then C_7=2 with CF_7=1
R8: if x_1 is L and x_2 is M then C_8=2 with CF_8=0.07
R9: if x_1 is L and x_2 is L then C_9=1 with CF_9=1

From Fig 3. and Fig 4. We can note that the use of the certainty grade with fixed the membership function improve the classification boundaries of this example.

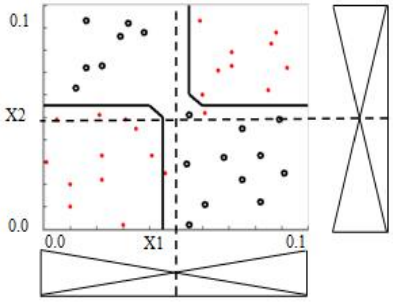


Fig 5. Classification boundaries with adjusted CF

	Classification rate		
	With CF	Without CF	Adjusted CF
K=2	67,5	95	97,5

TABLE II: classification rate with adjusted CF classification problem2

The obtained classification rates of this problem is 67,5% with heuristic certainty grades calculated in (5) and 95% without certainty grades. We can improve the classification performance (97,5%) and the classification boundaries by adjusting the certainty grade of each fuzzy IF–THEN rule, as shown in Fig 5 and *TABLE II* .

V. INFLUENCE OF MEMBERSHIP FUNCTION (MF)

In this section we use fuzzy rules (form 2) without CF and we modified the membership function on another classification problem 3.

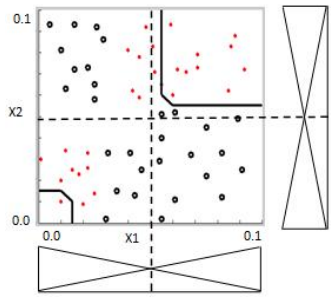


Fig. 6. Classification boundaries with CF, K=2.

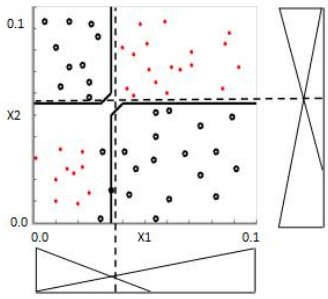
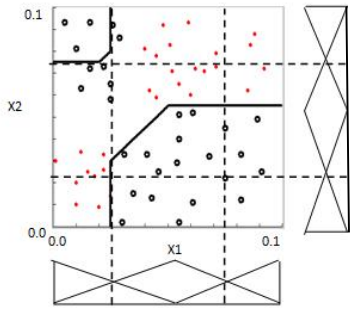
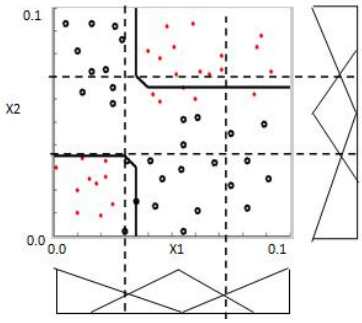


Fig. 7. Classification boundaries with modified membership function

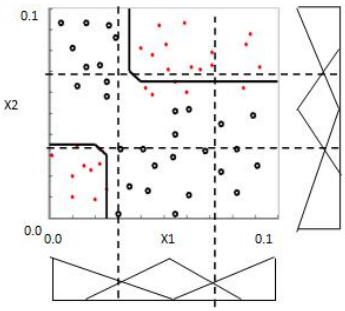
We remark that several patterns are not classified (Fig.6). but the modification of the membership function of each antecedent linguistic value permit improving classification boundaries performance even without certainty grade (Fig.7).



(a)



(b)



(c)

Fig.8. Classification boundaries: (a) with CF, (b) modified membership function without CF; (c) modified membership function with CF

	Classification rate		
	Non-modified MF	modified MF Without CF	Modified MF with CF
K=3	87,71	91,22	92,98

TABLE III: Classification rate of classification problem 3.

We show that the use of certainty grade improve the classification rate and classification boundaries also when we modified the membership function. But the use of both gives better results.

VI. Conclusion

In this paper, we analyzed the effect of membership function and certainty grade in fuzzy rules-based classification systems. The analysis was based on the examination of the classification rate and classification boundaries of synthetic bi-dimensional classification problems. For this task we performed tests, in one hand, with centered and modified membership function and, in the other hand, with and without certainly grade. The obtained results showed that: First, the use of adjusted membership function without rule weight has a significant effect on the classification performance. Second, the use of fixed membership function with the rule weights improves also the classification performances. Finally, adjusting both of them improve more the classification performances.

References

[1] L.A. Zadah , « Fuzzy sets » , Informationand Control, vol.8 , no3 ,p. 338-353 ,juin 1965 .

[2] A.M. Aziz, « Effects of Fuzzy Membership Function Shapes on Clustering Performance in Multisensor-Multitarget Data Fusion Systems », IEEE International Conference on Fuzzy Systems, Jeju , Island ,p.1839 – 1844, 2009.

[3] H. Ishibuchi, K. Nozaki, and H. Tanaka, «Distributed representation of fuzzy rules and its application to pattern classification ». Fuzzy Sets and System. vol. 52, p. 21-32. 1992.

[4] H. Jiang, H. Deng, Y. He, «Determination of Fuzzy Logic Membership Functions Using Extended Ant Colony Optimization Algorithm», Fifth International Conference on Fuzzy Systems and Knowledge Discovery, Shandong , IEEE, Vol.1, p. 581 - 585, 2008.

[5] D. Nauck and R. Kruse, «How the learning of rule weights affects the interpretability of fuzzy systems», in Proc. 7th IEEE Int. Conf. Fuzzy Systems, IEEE World Congress on Computational Intelligence, Anchorage, AK, vol.2, p. 1235–1240, May 4–9, 1998.

[6] H. Ishibuchi and T. Yamamoto, «Rule weight specification in fuzzy rule-based classification systems », IEEE Transaction on Fuzzy Systems, vol. 13, no. 4, p. 428-435, August 2005.

[7] H. Ishibuchi and T. Nakashima, «Effect of rule weights in fuzzy rule-based classification systems», IEEE Transaction on Fuzzy Systems, vol. 9, no. 4, p. 506-515, August 2001.

[8] D. Kaur, «Performance Enhancement of Digital communication Systems Using Fuzzy Logic», IJCSI International Journal of Computer Science Issues, Vol. 9, no. 3, p. 354-359, March 2012.

[9] O.Cordon, M.J. Del Jesus, F.Herrera, « Genetic Learning of Fuzzy Rule-Based Classification Systems Cooperating with Fuzzy Reasoning Methods», International Journal Of Intelligent Systems, vol. 13, 1025-1053, 1998.

[10]Y. Yuan, H. Zhuang, «A genetic algorithm for generating fuzzy classification rules», Fuzzy Sets and Systems, vol. 84, no. 25, p. 1-19,1996.

[11] J.J. Aguilera, M. Chica, M.J. Del Jesus and F. Herrera, « Niching genetic feature selection algorithms applied to the design of fuzzy rule-based classification systems», IEEE International Fuzzy Systems Conference, London, p. 1- 6, 23-26 July 2007.

[12] R.T. Alves, M.R. Delgado, H.S. Lopes, A.A. Freitas, «An Artificial Immune System for Fuzzy-Rule Induction in Data Mining», In proceeding of: Parallel Problem Solving from Nature - PPSN VIII, 8th International Conference, Birmingham, UK, vol. 3242, p. 1011-1020, 18-22 September 2004.

[13] Z. Lei, L. Ren-hou, «Designing of classifiers based on immune principles and fuzzy rules», Information Sciences, vol. 178, no. 7, p. 1836-1847, 1 April 2008.

[14] M.S. Saniee, J. Habibi and E. Soroush, «Induction of Fuzzy Classification Systems via Evolutionary ACO-Based Algorithms», International Journal of Simulation Systems, Science & Technology, vol. 9, no. 3, p. 1-8, September 2008.

[15] M.F. Ganji, M.S. Abadeh, «Using fuzzy Ant Colony Optimization for Diagnosis of Diabetes Disease», 18th Iranian Conference on Electrical Engineering (ICEE), Isfahan, p. 501- 505, 11-13 May 2010.

[16] C.C. Chen, «Design of PSO-based Fuzzy Classification Systems», Tamkang Journal of Science and Engineering, vol. 9, no. 1, p. 63-70, 2006.

[17] A. Kashanipour, N.S. Milani, A.R. Kashanipour, H.H. Eghrary, «Robust Color Classification Using Fuzzy Rule-Based Particle Swarm Optimization», IEEE, Congress on Image and Signal Processing (CISP '08), vol.2, Sanya, China, p. 110-114, 27-30 May 2008.

Efficient compute for complex illumination effects with virtual point light

[1]Khemliche sarra,[2] Babahenini Mohamed Chaouki, [3]Bahi Naima, [4]Zerari Abd El Mouméne

LESIA Laboratory
Department of computer science
Mohammed khider University
Biskra,Algeria

s.khemliche@univ-biskra.dz
chaouki.babahenini@gmail.com
naima.bahi@yahoo.fr
a.zerari@live.fr

Abstract— **Computing a global illumination solution in real-time is still a challenge in computer graphics. Global illumination algorithms are a fundamental part of the realistic image synthesis discipline. They take into consideration indirect illumination component to produce realistic images. On the other hand, the accurate evaluation of the indirect illumination tends to be the most time consuming part of the rendering process. There are plenty methods that focus on solving this problem most of them are not fast enough for interactive environments. Instant radiosity (IR) is a method that approximates the indirect lighting, as part of global illumination, by creating additional light sources.in this paper we present our model proposal wish based on IR algorithms that builds upon this algorithm, a technique to efficiently compute complex illumination effects to provide a fast and smoother global illumination.**

Index Terms— **global illumination (GI), virtual point light(VPL),visibility.**

I. INTRODUCTION

Interactive computation of global illumination is a major challenge in computer graphics research today. Effects such as multi-bounce indirect lighting, caustics, shadows, complex surface reflections and all other scenarios where the lighting behavior of one surface alters the lighting on another surface are important visual cues for the perception of photorealism in synthesized images. These effects not only consider the illumination coming directly from a light source (direct), but also reflected by other objects (indirect) the difference as illustrated in fig 1. This effect greatly enhances the realism of a scene. A fundamental difficulty in global illumination is the high computation cost acquired by indirect lighting, where all surfaces contribute illumination to the scene. Creating realistic looking images is an important goal in a lot of computer graphic domains. Although realism is important, there is another issue for computer graphics. It is the time we need to calculate the illumination and render our images. This two, the realism of an image and its time to render, are the major two issues we have to concern with. But unfortunately enhancing the one would worsen the other. So when we want both we have to make some restrictions, as static objects or for just one bounce light transport or for a specific king of materials.

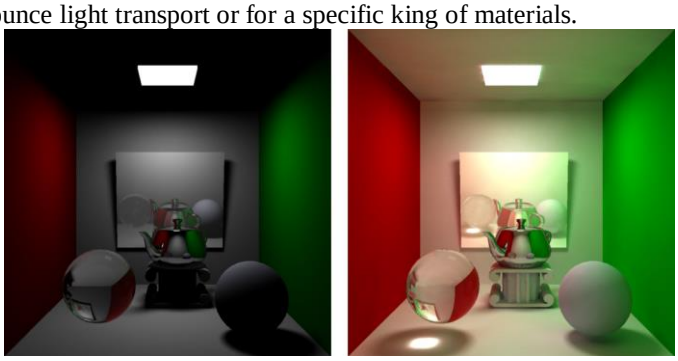


Fig1: *Image on the left is rendered with local illumination .Image on the right is rendered using a global illumination approach.*

Physically-based rendering aims to simulate light transport on a computer in a physically correct way wish is the task of global illumination. Recent research simulates most lighting effects (diffuse, direct and indirect lighting, shadow caustics ….etc.) with sufficient accuracy .These computations require extensive computations and are usually far from interactive. However High-quality global illumination at interactive speed is a difficult challenge, especially in complex and fully dynamic scenes. On the one hand, conventional methods for off-line rendering produce photo-realistic results, but do not allow interactivity. On the other hand, available interactive or real-time techniques often employ the computational power available in modern GPUs, but suffer from various limitations such as static geometry ,precomputed movement [7], low frequency lighting or low-glossy indirect illumination [8].

Our new system bring different contribution to satisfy a coherent functional rendering system to compute (multi-bounce) indirect illumination in large, dynamic scenes in real-time on modern GPUs.

The remaining of this paper is organized as follows. Section 2 introduces some solutions that mentioned in

literature and works that we are concentrate on our work .This later presented in section 3 as a model wish has parts that contribute in the process of image synthesizes .section 4 draws conclusion and some perspectives .

II. Related Work

State-of-the-art reference solutions of algorithms that used for realistic image synthesis usually rely on path tracing [2], photon mapping [5], ray tracing [10], or radiosity [3], these methods can perform the rendering in real-time and also creates realistic images, but unfortunately the fast rendering is gained by an expensive preprocessing step, an enormous amount of storage and a limitation to static scenes, and the realism just includes indirect illumination.

In contrast to those algorithms another method to do the rendering is the so-called **Instant radiosity** [1]. It is a popular technique that converts indirect illumination to a small set of virtual point lights (VPLs). This approach considerably reduces the computation cost, but is limited to a small number of VPLs and is only suitable for primarily diffuse materials(It is proven that point light- based global illumination to be a useful and efficient way to approximate diffuse inter-reflection in real time systems [2]). It constructs random paths from the light sources and creates new virtual point lights (VPLs) where these paths encounter a surface. When used together with shadow maps these point lights represent the full indirect illumination of the scene and are also optimized for usage with the current graphics hardware [6]. IR is a bidirectional method that first creates a set of virtual point lights (VPLs) **using random walks** that trace light paths starting from the primary light sources. Direct contribution of these VPLs then approximates the entire multi-bounce light transport in the scene.

Approaches based on instant radiosity and shadow mapping [1;4] compute single-bounce indirect lighting but ignore caustics and multibounce interreflections. Precomputed radiance transfer [9] incorporates many illumination effects but entails a large amount of precomputed data and is inefficient for dynamic geometry and materials.

Reflective shadow maps (RSMs) [4] treats shadow map pixels projected from the direct light source as indirect sources illuminating the scene. Determine the visible area from the light source.it provides information needed for the VPL (position, normal and outgoing color).we use it for combining pixels to reduce the number of VPL.

In our work we contribute two extensions to an Instant radiosity-type [1] interactive global illumination technique based on Reflective [4] and Imperfect Shadow Maps [8] (ISMs). The first extension improves the indirect lighting quality by identifying sources of indirect light that strongly contribute to the final image and by culling unnecessary ones. The second extension, improves the accuracy of indirect shadows by adapting the occlude representation to the viewpoint.in our work it represent an approximate representation of visibility to facilitate the computation of indirect illumination. Via these ameliorations, the technique achieves high-quality results, and runs offline at interactive frame rates, even for larger dynamic scenes

The use of VPLs is at the core of our rendering system; however, we extend it to simulate non-diffuse light paths as well. Traces paths from the primary light source and creates new virtual light points (VPLs). These lights sources are placed at each reflection point (where the light of the primary light source hits a reflector, e.g. a wall) and used to illuminate the scene. The final results are then placed into the scene when doing the rendering at the client side

III. SYSTEM OVERVIEW

The goal of our system is to provide a fast and smoother, global illumination solution as compared to approaches. This section provides an overview of our algorithm.

We focus on scenes where the direct light source is a point light, so the direct illumination can be computed in real-time using **GPU-based ray tracing**. In the following we focus on discussing the computation of indirect illumination.

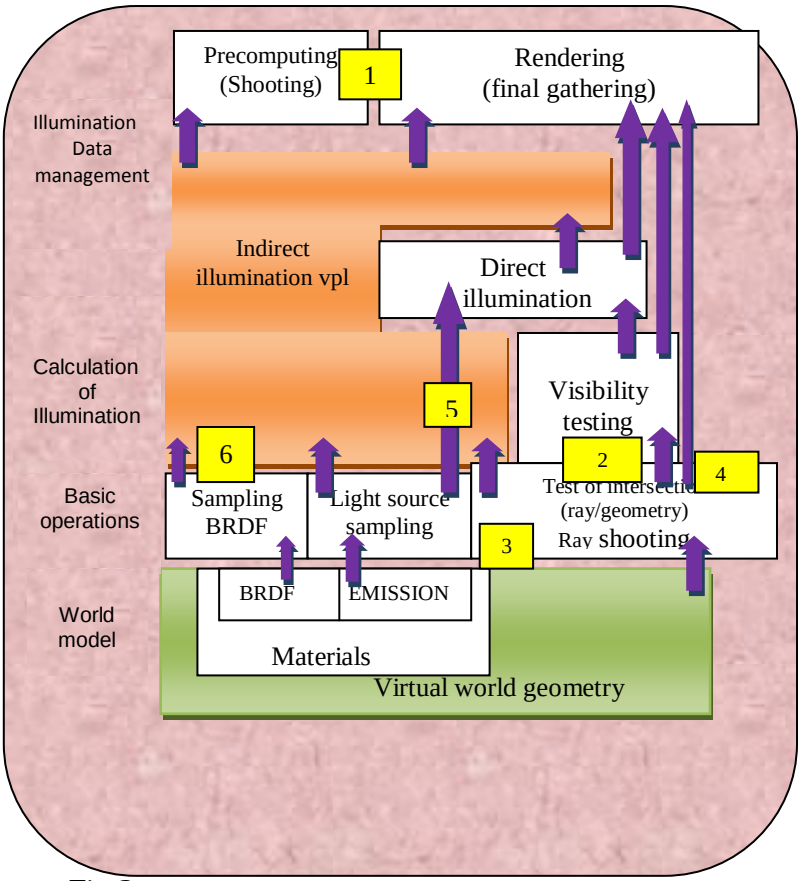


Fig 2: Modular diagram of image synthesis of our system with labels indicating our contributions.

Fig 2 Depicts the architecture of an image synthesis algorithm that is based virtual point light to achieve smoothly indirect illumination, **using random walks** that trace light paths starting from the primary light sources. Direct contribution of these VPLs then approximates the entire multi-bounce light transport in the scene (on random walks), The virtual light sources method is a perfect example of the approach depicted here -as we are mentioned before- .relies on ray tracing, and utilizes the idea of precomputation and final gathering phases.

All algorithms operate on a numerically defined virtual world, where the geometry elements have two sets of light transport parameters attached: **emission** and **reflection (BRDF)** properties. The basic operations the virtual world representation has to support are light source sampling, directional sampling according to the reflection properties **(BRDF sampling),** and the computation of the intersection of a light ray with the geometry **(ray shooting).**

A special case of ray shooting is when the visibility between two points has to be tested. In the virtual light sources method in particular, this is one of the most time critical tasks. The computation of illumination can be separated into a direct and an indirect part. Direct illumination requires light source sampling and then a visibility test for every sample has to be performed. Indirect illumination is computed by generating multi-bounce light paths using random walks. **Random walk generation** requires light source sampling, directional sampling and ray shooting. When next event estimation is used, computation of direct illumination is also required.

As indirect illumination is usually more expensive to evaluate, this is the part that is typically **precomputed**. The challenge is identifying those parts of the computation that are constant, or can be rapidly adapted to changes happening in the virtual world. The **final gathering** phase makes use of the precomputed data to render images. This is when direct illumination is typically evaluated. The gathering phase might also make use of random walks, or deterministic ray shooting. The main role of final gathering, however, is the validation and adaption of precomputed information. This typically means checking whether a moving object has interrupted a light path: that is, **visibility testing.**

The labels indicate where the points contribute to the image synthesis algorithm. Label (1) describes precomputation-based approaches for rendering animation. Label (2) proposes an approximation for visibility testing only, while label (3) improves ray shooting in general. The algorithm proposed in (4) is most effective for the ray tracing of visible reflective or refractive objects. Label (5) and (6) improves sampling in random walk algorithms and in direct illumination, respectively.

Together they provide a complete framework for the development of fast global illumination rendering algorithms. However, all of the above fields are open to improvement.

IV. Conclusion

In this paper we showed a detailed system that we are currently realize its principle contributions aiming on approach the various phenomena of interaction matter/light by inexpensive processing in real time.

References

[1] Alexander Keller. Instant radiosity(1997). In SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques.

[2] BENTHIN C., WALD I., SLUSALLEK P: A Scalable Approach to Interactive Global Illumination computerGraphics forum 22,3(2003),621-630 (Proceedings of Eurographics

[3] COHEN, M. F., WALLACE, J., AND HANRAHAN, P. 1993. Radiosity and realistic image synthesis. Academic Press Professional, Inc., San Diego, CA, USA

[4] DACHSBACHER C., STAMMINGER M.: Reflective shadow maps. In Proc. I3D (2005).

[5] JENSEN H. W.: *Realistic Image Synthesis Using Photon Mapping*. A. K. Peters, 2001.

[6] LAINE S., LEHTINEN J., KONTKANEN J.: Incremental Instant Radiosity for Real-Time Indirect Illumination In *Proc EGSR* (2007), pp. 4–8.

[7] MA V. C. H., MCCOOL M. D.: Low Latency Photon Mapping Using Block Hashing. In Proc. Graphics Hardware (2002).

[8] RITSCHEL T., GROSCH T., KIM M. H., SEIDEL H. P.,DACHSBACHER C., KAUTZ J.: Imperfect shadow maps for efficient computation of indirect illumination. ACM Trans. Graph.(Proc. SIGGRAPH Asia) 27, 5 (2008)

[9] SLOAN, P.-P., KAUTZ, J., AND SNYDER, J. 2002. Precomputed radiance transfer for real-time rendering in dynamic low frequency lighting environments ACM Trans. Graph. 21, 3, 527–536.

[10] WALD I., KOLLIG T., BENTHIN C., KELLER A SLUSALLEK P.: Interactive global illumination using fast raytracing In *Proc EGRW*(2002) pp15–24

Real-Time People Detection

Rouabhia Houssam Eddine
Département d'informatique,
Université 8 mai 1945 Guelma, B. P. 401, Algérie
LabSTIC, Université 8 mai 1945 Guelma, B. P. 401, Guelma 24000, Algérie
rouabhia.h@gmail.com

Farou Brahim
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

Seridi Hamid
Département d'informatique,
Université 8 mai 1945 Guelma, B. P. 401, Algérie
LabSTIC, Université 8 mai 1945 Guelma, B. P. 401, Guelma 24000, Algérie
seridihamid@yahoo.fr

***Abstract*— Despite the large number of promising approaches existing, it's still a hard work to obtain desirable results in the area of people detection, especially in terms of time. In this paper, we present a real-time and accurate human detector for video surveillance system, which includes background subtraction technique to extract candidate region, CENTRIST descriptor to encode the contour information and a linear SVM to make a decision, whether the selected region is a person or not. The used method achieves 25 to 30 fps speed and state-of-the-art detection accuracy.**

***Index* People detection, CENTRIST descriptor, background subtraction.**

I. INTRODUCTION

Detecting people in images and videos is one of the main tasks in many applications of computer vision , including human-robot interaction , image and video content management, assisted automotive system (which warns the driver if there is a person in the road) and video surveillance system which is our interest topic.

One of the main goals of video surveillance system is to monitor and identify people in public places such as airports and supermarkets or in sensitive places such as banks and ministries for security purposes. For these reasons, detecting and identifying people are very important phase in this system. At the same time, it is one of the most challenging problems, due to the difficulties related to the variations in people appearance and poses; these variations increase when it is combined with motion and clothes. The surrounding environment also plays an important role in the appearance (changing in illumination and cluttered background). In some cases, the position and orientation of the camera caused considerable difficulties. In the case of the presence of several people with a certain degree of interaction, it may cause an occlusion, which increases the difficulty of the task of detecting people.

It can be seen that the task of detecting people is related to the classification problem, which consists of three basic phases, the first one is the selection of a candidate region. The sliding window method [1] is the simplest method used in the stat of the art, to reduce the computational cost the background subtraction technique is used when the input data is a video from a fixed camera [2].

The second phase is the feature extraction. Many feature are used in the literature and the well-known are wavelet [3], edgelet [4],shape context [5],shapelet [6],LBP[7] and HOG [8] which is characterized by its efficiency and robustness when representing the human body. An extended version of HOG used a variable size block is proposed in [9] and they propose a rejection cascade to solve the problem of computational costs existing in HOG. In [10] a combination between HOG and LBP is used.

The classification is the last phase in detecting people, in this phase a decision is taken whether the vector of feature extracted from the candidate region represent a person or not. Most works in this field used Variants of Boosting [11][12] and SVMs [13][14] machine learning

Despite the large number of research in this area ,the current algorithms are characterized by their quality and efficiency and suffer from high computational cost or the contrary

in this paper, we build a system for detecting people in video surveillance system in real time, the system is characterized by its quality and speed since it use a fast and powerful discriminant descriptor CENTRIST [15],and background subtraction technique to reduce the computational cost in the phase of candidate selection.

II. OVERVIEW OF THE METHOD :

As illustrated in Fig1. This method is divided into three main processes: background subtraction technique to select region of interest (ROI); CENTRIST descriptor for describing the selected region; SVM learning machine to make a decision , whether the selected region is a person or not .

Fig1: The procedure of proposed method.

III. USING BACKGROUND SUBTRACTION TECHNIQUE TO SELECT CANDIDATE REGIONS :

The first step of our system shown in Fig.1 is background subtraction technique, with the purpose to localized regions of interest (ROIs) that contain person. We apply the algorithm from [16] to detect moving object and produce the foreground objects , then for every object detected a bounding box is extracted as illustrated in fig 2 a.

The bounding boxes extracted from the foreground image is then used to extract the candidate region from the current frame as illustrated in fig 2 b.

The reason for using the algorithm in [16] is that it used a recursive equations to constantly update the parameters of a Gaussian mixture model [17] and automatically select the needed number of components per pixel. These improvements making it fast and has more accuracy in the segmentation than the old GMM[17] which is one of the best algorithm for background subtraction.

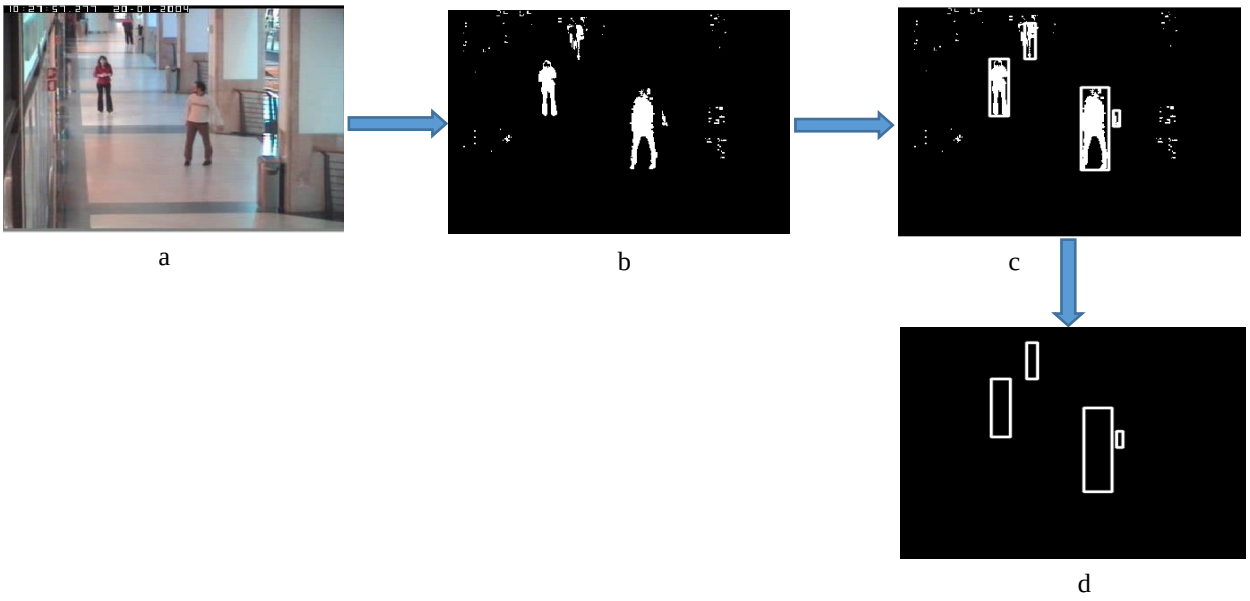


Fig. 2a. extract bounding box: where a: frame from a video, b: foreground image, c: extract object from the foreground , d: bonding box

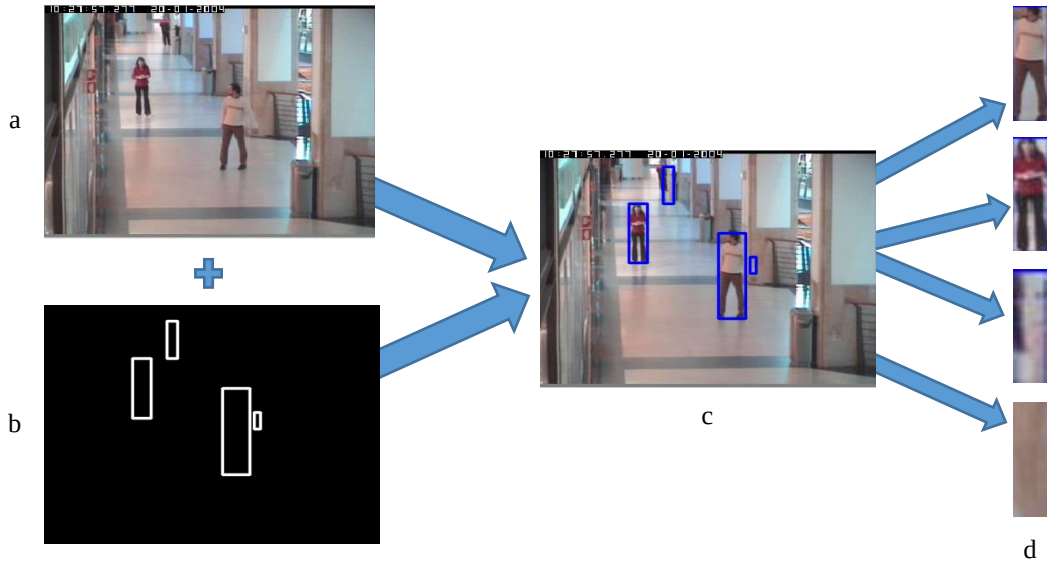


Fig 2b. extract the candidate region where a: the same frame, b:the same bonding box image, c: the fusion of the two images(a and b), d: candidate region.

IV. CENTRISTE DESCRIPTOR :

HOG is the most popular and one of the best descriptor in the field of people detection, but it cannot be used because its high computational cost in video surveillance.

CENTRIST[15] descriptor use contour information as HOG and use the signs of comparisons among neighboring pixels to encode this information, in addition, it does not need any image preprocessing or feature vector normalization, which make it better then HOG in the computational cost with almost the same accuracy [18]

CENTRIST means CENsus TRansform hIStogram, the census transform(CT) has developed by [19] for establishing correspondence between local patches .

the CT value of a pixel is obtained by comparing its intensity with those of its neighbors using Eq.1

$$CT_i = \sum_{N=0}^{7} \left[2^n \times S(P_i - P_n)\right], S(a) = \begin{cases} 1, a \geq 0 \\ 0, a < 0 \end{cases} \ldots (1)$$

Where P_N (N=0,1…..7) is the Nth neighbors of P_i pixel, and P_i is the center pixel in a scanning window of 3 * 3 range

CENTRISTE visual descriptor is a histogram of census transform value. An image is described as a histogram sequence by the following procedure:

- Compute the well know Sobel gradient of an input image.
- Compute the Census Transform image using the Sobel filter result
- The Census Transform image (CT) is divided into (N*M) rectangle regions with specific size(h*w), and a histogram of CT values is computed for each 2*2 adjacent region.
- The CT histograms of all regions are concatenated to form the final vector feature used for training the SVM machine learning ,
- In the phase of testing we don't need to extract the vector feature ,because the feature extraction component is seamlessly embedded into classifier evaluation[18] and a patch is classified as containing person if :

$$\sum_{i=0}^{N-2} \sum_{j=0}^{M-2} \sum_{x=1}^{2h-1} \sum_{y=1}^{2w-1} W_{i,j}^{CT((i-1)\times h+x,(j-1)\times w+y)} > \theta \ldots (2)$$

Where W is the linear SVM classifier and θ is a threshold.

V. EXPERIMENTS :

To evaluate our algorithm, we use the INRIA database that contain 614 positive images and 1218 negative images for training a linear SVM machine learning, we extract person from the positive images and choose randomly patches sample from the negative images. In order to be used by the SVM, both positives and negatives patches are resized; because the SVM classifier cannot be adapted to a variant patches size.

In the testing phase we use CAVIAR database , the evaluation is based on two aspects , the speed and the detection accuracy.

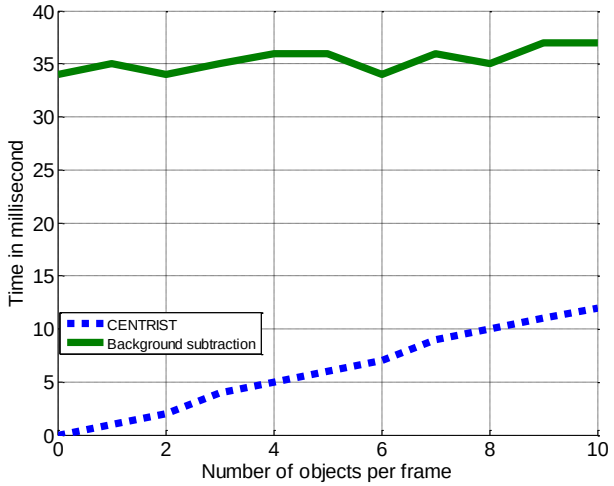


Fig 3 : Time spent by the background subtraction and CENTRIST versus number of objects per frame .

Figure 3 shows curves of the time spent by CENTRIST descriptor and the background subtraction technique relative to the number of found objects (ROI). From the figure, we find that the time of the background subtraction technique is not affected by the number of found object, unlike the time of the CENTRIST which increases with the rise in the number of objects, but the amount of this increase is about(1 millisecond per object). Since the time consumed by CENTRIST is small compared to background subtraction time, the total time is not much affected by the increasing of the number of objects , which lead to achieve 25 to 30 fps speed using only 1 processing core of a 2.0 GHz.

To evaluate the detection accuracy we use the Precision-recall curve.

$$\text{Recall} = \frac{TruePositives}{AnnotatedHumans} \ldots (3)$$

$$\text{Precesion} = \frac{TruePositives}{TruePositives + FalsePositives} \ldots (4)$$

A bonding box(BB) is counted as true positives when it is classified as human by the algorithm and match the ground-truth. But BB that is classified as human and not matched to any annotated human are counted as false positives, the matching is determined if the following inequality is satisfied [20].

$$\frac{area\left(BB_{gt} \cap BB_{dt}\right)}{area\left(BB_{gt} \cup BB_{dt}\right)} \ldots (5)$$

When BB_{gt} is the ground truth bonding box and BB_{dt} is the detected bonding box .

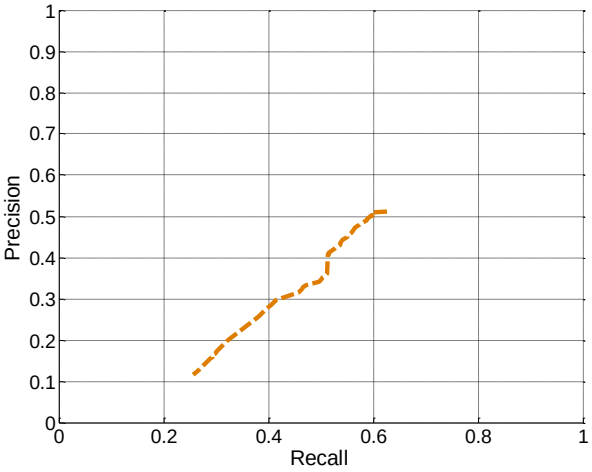


Fig4: Precision-Recall curve.

From figure 4, we find that recall is proportional to the precision; the interpretation of this relationship is that the used method does not produce a lot of false positives detection, which leads to good accuracy. On the other hand, this method is limited by detecting only moving people, because the background subtraction technique detects only moving object and that lead to produce false negative detection (people not detected).

VI. Conclusion :

The main work in this paper is consist of: experimentally verify the effectiveness and efficiency of the CENTRIST descriptor for people detection on the CAVIAR database and using the background subtraction technique, to reduce the computational cost and improve the detection accuracy. In the future, we will use another method to select the candidate region to resolve the problem of standing people.

References

[1] C. Papageorgiou et T. Poggio, « A Trainable System for Object Detection », Int. J. Comput. Vis., vol. 38, no 1, p. 15-33, juin 2000.

[2] O. Javed et M. Shah, « Tracking and Object Classification for Automated Surveillance », in Computer Vision — ECCV 2002, A. Heyden, G. Sparr, M. Nielsen, et P. Johansen, Éd. Springer Berlin Heidelberg, 2002, p. 343-357.

[3] P. Viola, M. J. Jones, et D. Snow, « Detecting Pedestrians Using Patterns of Motion and Appearance », Int. J. Comput. Vis., vol. 63, no 2, p. 153-161, juill. 2005.

[4] B. Wu et R. Nevatia, « Detection of multiple, partially occluded humans in a single image by Bayesian combination of edgelet part detectors », in Tenth IEEE International Conference on Computer Vision, 2005. ICCV 2005, 2005, vol. 1, p. 90-97 Vol. 1.

[5] E. Seemann, B. Leibe, K. Mikolajczyk, et B. Schiele, « An evaluation of local shape-based features for pedestrian detection », in In Proc. BMVC, 2005.

[6] P. Sabzmeydani et G. Mori, « Detecting Pedestrians by Learning Shapelet Features », in IEEE Conference on Computer Vision and Pattern Recognition, 2007. CVPR '07, 2007, p. 1-8.

[7] Y. Mu, S. Yan, Y. Liu, T. Huang, et B. Zhou, « Discriminative local binary patterns for human detection in personal album », in IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008, 2008, p. 1-8.

[8] N. Dalal et B. Triggs, « Histograms of oriented gradients for human detection », in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005, 2005, vol. 1, p. 886-893 vol. 1.

[9] Q. Zhu, M.-C. Yeh, K.-T. Cheng, et S. Avidan, « Fast Human Detection Using a Cascade of Histograms of Oriented Gradients », in 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006, vol. 2, p. 1491-1498.

[10] X. Wang, T. X. Han, et S. Yan, « An HOG-LBP human detector with partial occlusion handling », in 2009 IEEE 12th International Conference on Computer Vision, 2009, p. 32-39.

[11] H.-X. Jia et Y.-J. Zhang, « Fast Human Detection by Boosting Histograms of Oriented Gradients », in Fourth International Conference on Image and Graphics, 2007. ICIG 2007, 2007, p. 683-688.

[12] M. J. Jones et D. Snow, « Pedestrian detection using boosted features over many frames », in 19th International Conference on Pattern Recognition, 2008. ICPR 2008, 2008, p. 1-4.

[13] S. Zhou, Q. Liu, J. Guo, et Y. Jiang, « ROI-HOG and LBP Based Human Detection via Shape Part-Templates Matching », in Neural Information Processing, T. Huang, Z. Zeng, C. Li, et C. S. Leung, Éd. Springer Berlin Heidelberg, 2012, p. 109-115.

[14] H. Bin et Z. Chunxia, « An HOG-LGBPHS human detector with dimensionality reduction by PLS », Pattern Recognit. Image Anal., vol. 24, no 1, p. 36-40, janv. 2014.

[15] J. Wu et J. M. Rehg, « CENTRIST: A Visual Descriptor for Scene Categorization », IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no 8, p. 1489-1501, août 2011.

[16] Z. Zivkovic et F. van der Heijden, « Efficient adaptive density estimation per image pixel for the task of background subtraction », Pattern Recognit. Lett., vol. 27, no 7, p. 773-780, mai 2006.

[17] C. Stauffer et W. E. L. Grimson, « Adaptive background mixture models for real-time tracking », in Computer Vision

and Pattern Recognition, 1999. IEEE Computer Society Conference on., 1999, vol. 2, p. -252 Vol. 2.

[18] J. Wu, C. Geyer, et J. M. Rehg, « Real-time human detection using contour cues », in 2011 IEEE International Conference on Robotics and Automation (ICRA), 2011, p. 860-867.

[19] R. Zabih et J. Woodfill, « Non-parametric local transforms for computing visual correspondence », in Computer Vision — ECCV '94, J.-O. Eklundh, Éd. Springer Berlin Heidelberg, 1994, p. 151-158.

[20] P. Dollar, C. Wojek, B. Schiele, et P. Perona, « Pedestrian detection: A benchmark », in IEEE Conference on Computer Vision and Pattern Recognition, 2009. CVPR 2009, 2009, p. 304-311.

Un modèle pour la génération de situations d'apprentissage de pédagogie par projet collectif pour un environnement d'apprentissage collaboratif.

Hana BENDJEDID
LRI, Laboratoire de Recherche en Informatique
Groupe de Recherche e-Learning (GReLearn)
Annaba, Algérie
Bendjedid-hana@hotmail.fr

Tahar BOUHADADA
LRI, Laboratoire de Recherche en Informatique
Groupe de Recherche e-Learning (GReLearn)
Annaba, Algérie
bouhadada.tahar@univ-annaba.org

***Résumé*— L'ingénierie des environnements informatiques pour l'apprentissage humain (EIAH) est devenue un domaine de recherche actif, ces dernières années ; Afin d'assurer une articulation entre l'aspect informatique et l'aspect humain. Par ailleurs, l'environnement collectif favorise une meilleure productivité, car c'est la lumière sur laquelle un processus d'apprentissage se développe. Dans cet article, notre principale contribution est la génération de situations d'apprentissage dans un contexte de pédagogie de projet collectif. Notre travail s'inscrit au confluent de deux axes, les situations d'apprentissage (formulées par l'enseignant) et les projets collectifs (réalisés par les apprenants), dont l'objectif est la construction de la connaissance. Le système permet de générer des situations d'apprentissage sous forme de projet (les entrées susceptibles d'amener les apprenants à la construction de leurs connaissances) destiné à un ou plusieurs groupes d'apprenants qu'ils ont été créés par eux-mêmes ou par d'autres enseignants. En outre, La méthode que nous avons adoptée consiste à modéliser un environnement social à distance (échanger, interagir, confronter les points de vue, ..) en s'inspirant de la réalité dans le but de bâtir des projets communs.**

***Mots clés*— projet pédagogique, apprentissage, socialisation des apprenants, tâche collective, EIAH.**)

1) CONTEXTE DE L'ETUDE ET PROBLEMATIQUE

Les nouvelles technologies de l'information de la communication (NTIC) progressent rapidement et possèdent une place considérable dans différents domaines. Leur utilisation quotidienne est devenue incontournable, Quant à leur développement, il a permis au domaine de l'enseignement de bénéficier pleinement de l'apport de ce cadre technologique, dans le but de se délivrer des contraintes liées à l'éloignement ou à l'enclavement des apprenants cibles.

L'intégration des NTIC dans les EIAH transforme d'une part la façon d'apprendre et d'autre part le moment, la nature et le lieu d'apprentissage.

Or un apprenant de par sa nature a besoin d'un environnement social où il peut échanger, interagir et communiquer ses idées, un réel besoin de communication sera crier entre ces derniers via des outils de communication, le simple fait de fournir des outils classiques de communication n'est pas suffisant pour faire émerger des interactions entre apprenants, un seul outil ne peut être le déclencheur d'une dynamique collective; Afin de favoriser les interactions, l'enseignant formule des situations d'apprentissage en se basant sur les lacunes repérées chez les apprenants. La principale difficulté est alors d'énoncer une situation d'apprentissage de pédagogie par projet collectif pour favoriser la construction collective de connaissances. Dans ce contexte, notre travail vise à concevoir un environnement informatique pour l'apprentissage collectif par pédagogie de projet, permettant une meilleure utilisation des différents acteurs.

La réalisation d'un tel système pose divers problématiques : l'énoncé de la problématique, la durée du projet, les buts poursuivis, les ressources pédagogiques nécessaires au projet (la documentation, les liens, …), les acteurs concernés, l'affectation des rôles et des tâches, …etc. Les questions de recherche que nous abordons dans ce cadre portent sur : l'énoncé de la situation d'apprentissage de pédagogie par projet collectif en se focalisant sur les difficultés repérées chez l'apprenant. Effectivement, la distance peut représenter un inconvénient lors de l'utilisation des plateformes, en l'occurrence l'isolement de l'apprenant qui peut engendrer l'ennui, et puis l'abondant de ce dernier ; c'est pour cela qu'il faut se poser les bonnes questions :

- Comment fournir un soutien informatique à un enseignant-concepteur pour la mise en place d'une situation d'apprentissage (y compris la gestion de l'énoncé) de type projet ?
- Qui et comment créer les groupes, la taille de ces derniers, répartir les tâches et choisir les membres idéaux au projet pédagogique.
- Comment apporter un soutien aux apprenants pendant le déroulement de leurs activités collectives selon un contexte a distance ?

- Comment assister les apprenants pendant le déroulement collectif de la situation d'apprentissage ?

C'est à travers l'ensemble de ces obstacles que nous pouvons mesurer l'ampleur des difficultés à concevoir un environnement informatique pour l'apprentissage collectif par pédagogie de projet.

2) SITUATIONS D'APPRENTISSAGE DANS UN CONTEXTE DE PEDAGOGIE PAR PROJET COLLECTIF

Tchounikine définit un EIAH comme un environnement informatique conçu dans le but de favoriser l'apprentissage humain [1], Selon le même auteur, le travail sur la réalisation d'un EIAH, ne s'agit pas d'étudier l'éducation (Science humaine) puis la technologie (informatique) ; Réciproquement, ni d'étudier la technologie puis l'éducation, l'un ne complète pas l'autre, mais c'est la recherche en ingénierie d'EIAH [2]. Donc La réalisation d'un tel environnement nécessite une véritable collaboration entre différents acteurs ; C'est le résultat d'une convergence entre plusieurs disciplines (informatique, pédagogie, psychologie, etc).

Pour Faerber, les enseignants qui sont parties prenantes dans les enseignements à distance, sont amenés à formuler les situations d'apprentissage [3]. Une situation d'apprentissage est une situation pédagogique (amener l'apprenant à déployer une activité adéquate à l'atteinte d'un ou de plusieurs objectifs précis) conçue afin de susciter ou accompagner un apprentissage cible, c'est-à-dire, la construction des connaissances cibles chez l'apprenant [4]. Elle est donc initiée d'une situation réelle de la vie. Les deux principaux acteurs dans une situation d'apprentissage sont l'enseignant et les apprenants, D'après George il ressort deux autres acteurs bien distincts : le concepteur pédagogique et le tuteur [5]. Pour le bon fonctionnement de l'environnement collectif (Collaboratif et coopératif), notamment la motivation et l'engagement des apprenants, la taille et la composition du groupe d'apprentissage sont des facteurs considérables, plusieurs travaux de recherche s'intéressent à ces deux derniers. On trouve par exemple que Faerber estime qu'un groupe devra contenir entre 6 et 13 membres [6]. Quant à Leburn, la taille du groupe idéal se situe entre 5 et 10 personnes [7]. Selon Henry et Lundren-Cayrol, préconisent la constitution de petits groupes de 3 à 5 apprenants [8].

Divers types de situations d'apprentissage existent, la pédagogie par projet collectif (PPC) s'avère la plus adéquate aux situations d'apprentissage, Selon Leburn cité dans [9], La PPC est une méthode de la pédagogie active, rassemblant deux méthodes : Pédagogie par Projet (PP) et Apprentissage Collectif (AC).

La plus part de la communauté de recherche recommande des environnements où l'on tolère aux acteurs pédagogiques à travers *Educational Modelling Language* (EML) de devenir concepteur de leurs propres scénarios d'apprentissage, ces derniers formulés par le biais d'un langage formel de description de scénario (EML) sont par la suite inclus dans une plate-forme qui déroulerait ces scénarios. Dans l'approche citée auparavant, IMS *Global Learning Consortium* ont fait une proposition majeure, *Instructional Mana-gement Systems-Learning Design* (IMS-LD) [10]. Cependant, il y'a lieu de se poser une question : est ce que cette approche a une limite ? Car ce mode de pédagogie est parfois complexe et difficile à maitriser dans les enseignements. D'après [11], l'approche est limitée du fait du peu de familiarisation des enseignants avec certains concepts. Donc elle nécessite une compréhension approfondie du langage.

Plusieurs autres outils sont mis à la disposition des enseignants-concepteurs comme *Reload* [12], cet outil auteur permet la création de contenus e-learning. On trouve même des plates-formes comme *NetUniversité* [16] où les enseignants-concepteurs énoncent leurs situations d'apprentissage sans même avoir une connaissance préalable du langage. D'autres plateformes sont *open source*, hélas elles ne s'adaptent avec la PPC [13]. Il existe une autre solution, c'est de permettre aux enseignants de définir dans un méta-modèle (spécifique au domaine de l'enseignant) les concepts de base du modèle pédagogique qu'ils souhaitent [14].

En outre, Abdallah propose une méta-modèle qui réside donc, à guider l'enseignant afin d'implémenter son scénario PPC dans la plate-forme cible choisie (*Moodle*) qui est utilisée à l'IUT de Laval, plutôt que développer une nouvelle plate-forme [9]. Sa démarche consiste à transformer les modèles depuis le méta-modèle de la situation d'apprentissage de la PPC à celui de la plate-forme cible (*Moodle*) suivant une approche qui met au centre met au centre la technique de transformation de modèles de type MDA (*Model Driven Architecture*). Deux lacunes se présentent : la première est due aux problèmes de réutilisation des plates-formes, qui sont d'ailleurs une problématique centrale dans les activités du génie logiciel [15] ; Quant à la seconde, l'auteur cible les enseignants, sans donner trop d'importance aux activités des apprenants (auto-organisation), qui dépendent du groupe, son aptitude à s'organiser et des compétences informatiques de chacun de ses membres et des caractéristiques dynamiques du groupe (par exemple, des conflits entre ses membres ou l'apparition d'un leader).

Notre recherche en EIAH nous a permis de constater une carence des langages de scénarisation pédagogique dans un contexte de Pédagogie par Projet Collectif (PPC). Egalement, la plus part des chercheurs adoptent des pédagogies traditionnelles plutôt que des pédagogies ouvertes (la PPC appartient au courant socio-constructivisme).

Notre démarche consiste à réaliser un système d'apprentissage, a travers le quel l'enseignant-concepteur progresse dans une attitude pédagogique qui lui est propre (Générer des situations d'apprentissage) ; et qui permet aux apprenants de communiquer et d'interagir en synchrone et à distance dans le cas d'une situation d'apprentissage collective supportée par ordinateur.

3) ARCHITECTURE DU SYSTEME CLB PROJECTS

CLB PROJECTS (Collective Learning by Projects ou apprentissage collectif en réalisant des projets) est composé de plusieurs modules (figure I).

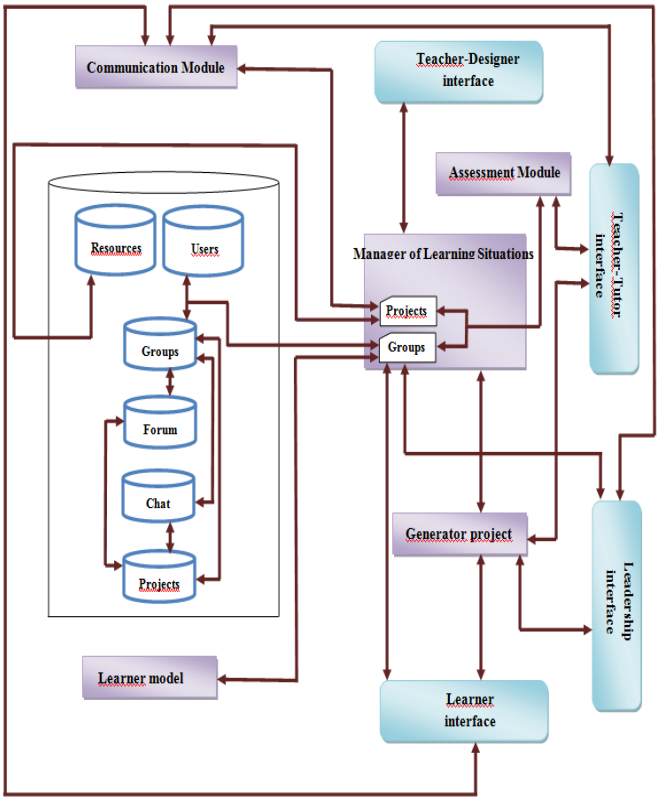


Figure 1. Architecture du système CLB PROJECTS

Un des objectifs principaux du système est l'apprentissage collectif à distance mettant en œuvre une pédagogie de projet ; afin d'améliorer les compétences en favorisant l'apprentissage. Pour atteindre ces objectifs nous nous sommes orientés vers l'utilisation d'une architecture distribuée qui permet un apprentissage à distance (les membres d'un groupe travaillent collectivement). Notre proposition est fondée sur la réalité (source de notre inspiration). Afin de permettre aux membres d'un groupe de travailler collectivement, il est indispensable de mettre à leur disposition des outils de communication ou de discussion tels que les Forums et le Chat. Certes ils ne sont pas spécifiques à l'éducation mais des bénéfices pédagogiques y sont attendus.

D'après des études [KRU, 99], on retient ***80%*** de ce qu'on enseigne aux autres, à contrario, on retient entre ***5*** et ***50*** % pour le cas d'un apprentissage individuel. La collaboration s'avère donc efficace pour la construction de la connaissance.

Pendant la session d'apprentissage, l'environnement permet aux apprenants de coopérer ou de collaborer pour réaliser le projet au sein de petits groupes qui sont créés au départ soit par l'enseignant-concepteur, soit par eux même (Auto-organisation). L'enseignant-concepteur détenteur des connaissances met en œuvre des situations d'apprentissage (Type projet), destiné à un groupe d'apprenants pour qu'ils travaillent collectivement (en coopération ou en collaboration ou les deux).

La réalisation du modèle ainsi que la conception de l'environnement permettent de structurer et surveiller les situations d'apprentissages collaboratives afin que de fortes interactions se produisent, afin de parer, autant que possible, aux difficultés d'apprentissage synchrone et à distance.

A. Gestionnaire de situations d'apprentissage

C'est le noyau du système, il interagit avec les différents autres modules, sa principale fonction est vis-à-vis à l'interface enseignant-concepteur et il permet dans ce cas la gestion de la conception des énoncés des situations d'apprentissage (projets) ainsi que la création des groupes d'apprenants.

Ce module permet la création d'un projet et sa gestion (modification et suppression) , il permet également la consultation des projets conçus par d'autres enseignants-concepteurs et facilite leur réutilisation, cela devra permettre l'échange de projets entre enseignants . Lors de la création d'un nouveau projet, le Gestionnaire de situations d'apprentissage est l'intermédiaire entre l'interface enseignant-concepteur et la base de données, c'est lui qui effectue tous les traitements nécessaires.

Cependant, la création d'un projet nécessite l'existence d'un groupe destinataire, le GSA gère la création des groupes, leur stockage dans la base de données, il propose également à l'enseignant concepteur d'accéder aux profils de tous les apprenants et cela en interagissant avec le modèle apprenant. Il est également responsable de la modification ou la suppression d'un groupe. Le GSA interagit directement avec l'interface apprenant et l'interface Leadership et cela pour lui présenter les groupes auxquels il appartient.

B. Générateur de projets

C'est le module intermédiaire entre le gestionnaire de situations d'apprentissage et les interfaces de l'apprenant, de l'enseignant-tuteur et du leadership. Le gestionnaire de situations d'apprentissage modélise la fiche technique du projet de chacun des destinataires (apprenant, leader ou tuteur) et puis vient le rôle du générateur de projets qui est de présenter l'énoncé aux différentes interfaces.

C. Modèle apprenant

Le modèle de l'apprenant permet au système de s'adapter à l'apprenant qui interagit avec lui. C'est-à-dire, il présente des connaissances compréhensibles.

Dans notre environnement on vise d'une part à repérer les lacunes des apprenants afin de mieux formuler les groupes d'apprenant, c'est le rôle principal attribué à ce modèle. D'autre part s'intéresse à qui concerne le collectif, et ceci ne peut être atteint que par la connaissance du profil.

D. Module de communication

Il interagit avec le gestionnaire de situations d'apprentissage, mais également avec les interfaces enseignant-tuteur, apprenant et leadership. Il consiste à permettre aux différents acteurs (apprenant et enseignant) d'accéder aux différentes fonctionnalités offertes. Pour l'apprenant, l'apprentissage et pour l'enseignant-tuteur superviser l'apprentissage. Lors de l'élaboration d'un projet c'est le module qui gère la communication entre les membres d'un groupe ainsi qu'avec l'enseignant tuteur. Il existe différents

outils de communication : Synchrones et asynchrones, Nous nous sommes focalisés sur les deux types (Le chat pour les étudiants présents, et les forums pour ceux qui manquent ou ceux qui veulent aider les autres groupes à résoudre les différents problèmes rencontrés).

E. Groupes

Cependant pour favoriser le sentiment d'appartenance et faciliter les interactions entre les membres du groupe, il est préférable que le nombre d'individu dans un groupe ne soit pas trop élevé. C'est pourquoi l'idéal pour un manager est de diriger un groupe de 5 ou 6 personnes.

Lors de la conception de l'énoncé d'une situation d'apprentissage (projet), il sera demandé au concepteur de choisir un groupe destinataire, s'il n'existe aucun groupe l'enseignant-concepteur doit créer un à partir des apprenants inscrits. Un apprenant peut appartenir à un ou plusieurs groupes. Tout projet destiné à un groupe donné est automatiquement attribué à tous les membres de ce dernier.

Une fois le groupe formé, l'enseignant-concepteur accède aux profils de tous les apprenants.

Dans chaque groupe, il y'a deux acteurs différents : Un leader et des simples apprenants.

Leader : Le leader c'est un utilisateur qui mène les activités du groupe, il validera le groupe, lancera le chat pour discuter, choir le mode de résolution et validera la solution proposées par son groupe.

Simple apprenant : L'utilisateur s'identifiera, choisira la façon de s'intégrer au groupe et attendra les instructions du leader de son groupe afin de communiquer et interagir.

En outre, la tâche (ce qui doit être réalisé) est proposée par l'enseignant équipé d'une intention pédagogique qui a sa propre interprétation. Il choisira aussi l'intimité des activités du groupe (Publique ou privées).

L'enseignant-concepteur peut modifier un groupe qu'il a créé s'il désir changer un ou plusieurs de ses membres. Si un apprenant n'appartient plus à un groupe, il sera également dissocié des projets affectés à ce dernier. Il peut également supprimer un groupe, dans ce cas tous les membres se verront dissociés des projets attribués au groupe en question.

F. Projets

Nous nous sommes référé à la caractérisation de l'énoncé d'une situation d'apprentissage à destination des apprenants dans un contexte collectif proposée par R. Faerber dans [Faerber04]. Donc le système permet au concepteur pédagogique la création d'un projet ainsi que sa gestion (modification et suppression).

Le système propose la création mais aussi la consultation des projets déjà conçus ou encore leur réutilisation par un autre enseignant.

1) *Création du projet :* Pour la création d'un projet, trois possibilités se présentent à l'enseignant-concepteur ; il pourra :

a) Créer un nouveau projet qui n'existe pas dans la base de données du système (c'est lui l'auteur) : cette première étape est préliminaire, c'est sur la base des informations fournies dans celle ci que la structure du projet va se dessiner. L'enseignant-concepteur donne un titre au projet, et dans le cas ou c'est lui qui a formé les groupes, il choisit le type d'interactivité entre les membres et le leader, le groupe destinataire, l'enseignant-tuteur, le nombre d'objectifs poursuivis, le type et le nombre des ressources pédagogiques ; après avoir validé la première étape, et tenant compte des informations fournies auparavant, le système propose à l'enseignant-concepteur d'introduire l'énoncé de la situation d'apprentissage, l'énoncé est subdivisé en trois parties (Les informations générales sur le projet, énoncé et consigne et ressources pédagogiques).

b) Consulter un projet déjà conçu par un autre concepteur (c'est-à-dire qui existe) pour avoir des idées pour son propre projet : Le système permet à l'enseignant-concepteur de consulter les énoncés des projets qu'il a déjà conçus mais aussi ceux des autres enseignants. À partir de cette consultation il pourra créer un nouveau projet.

c) Réutiliser un projet existant en apportant des modifications selon ses besoins pédagogiques, techniques, etc : Si un enseignant projette de réutiliser un projet, il voudra certainement apporter quelques modifications ne serait-ce pour le groupe destinataire. Le système permet à l'enseignant de modifier les informations qu'il veut (sans pour autant changer le projet originel). Après validation le projet sera envoyé aux membres du groupe destinataire et il sera considéré comme auteur du nouveau projet.

2) Modification du projet : Il arrive qu'un enseignant-concepteur veut apporter des modifications à un projet donné et cela pour de multiples raisons (Techniques, Pédagogiques,…). Le système permet à l'enseignant de modifier uniquement les projets dont il est le concepteur, ces modifications sont permanentes, c'est-à-dire que le projet n'existera plus sous sa forme initiale.

Un projet peut être réutilisé par un autre groupe, si l'enseignant-concepteur désir affecté un projet qui existe déjà à un autre groupe, il n'aura pas à répéter les étapes de la création, le système lui propose simplement de modifier le groupe destinataire.

3) Suppression de projets : L'enseignant-concepteur peut supprimer seulement les projets dont il est l'auteur. Les apprenants ne seront plus concernés et ne peuvent plus travailler sur le projet supprimé. Pour l'environnement destiné aux apprenants, ils pourront accéder à leurs projets, aux groupes auxquels ils appartiennent, aux groupes auxquels il n'appartiennent pas (Groupes publics) et aux différents outils de communication. Chaque apprenant a l'accès de visualiser les membres de chaque des groupes et d'y accéder au profil des membres des groupes auxquels il appartient.

G. Evaluation

Permet le suivi des apprenants pour ainsi connaître leurs progressions individuelles et collectives dans leur projet et ensuite d'évaluer la production finale. L'enseignant-tuteur est chargé d'effectuer l'évaluation et retourner le résultat à l'enseignant-concepteur.

4)CONCLUSION

Le système de génération de situations d'apprentissage de pédagogie par projet collectif assure un environnement d'apprentissage (Interactions, échange, débats,...). Il ne peut donc pas être considéré comme la simple réunion ou le partage des résolutions individuelles ; mais un processus cognitif sollicitant une réflexion et une communication entre les membres du groupe, l'enseignant-concepteur et le l'enseignant-tuteur.

Un prototype est en cours de réalisation, actuellement notre intérêt est de finaliser les situations d'apprentissage dans un contexte de Pédagogie par Projet Collectif, afin d'améliorer l'apprentissage en favorisant la collaboration.

REFERENCES

[1] P. Tchounikine. « Pour une ingénierie des Environnements Informatiques pour l'Apprentissage Humain ». 2002.

[2] Tchounikine P. Directions to acknowledge learners' self organization in CSCL macro-scripts. In: Groupware: Design, Implementation and Use, LNCS n°4715, Haake J.M., Ochoa S.F., Cechich A. (eds), Springer Berlin / Heidelberg, p. 247-254. (2007).

[3] R. Faerber, « Caractérisation des situations d'apprentissage en groupe », Sciences et Technologies de l´Information et de la Communication pour l´Éducation et la Formation, Volume 11, 2004.

[4] P. Tchounikine, « Précis de recherche en ingénierie des EIAH », 2009.

[5] S. George, « Interactions et communications contextuelles dans les environnements informatiques pour l'apprentissage humain », Habilitation à Diriger des Recherches présentée devant L'Institut National des Sciences Appliquées de Lyon et l'Université Claude Bernard Lyon I, Laboratoire d'Informatique pour l'Entreprise et les Systèmes de Production, 2010.

[6] R. Faerber, « Le groupe d'apprentissage en formation à distance : ses caractéristiques dans un environnement virtuel », in : KARSENTY, LAROSE, La place des TIC en formation initiale et continue : Bilan et perspectives. Chapitre 4, Edition de CRP, Université de Sherbrooke.

[7] M. Lebrun, Des méthodes actives pour une utilisation effective des technologies. Techno et Pédagogie : de la méthode www.ipm.ucl.ac.be/marcel/METHODES.PDF, 19 septembre 2001.

[8] M. Arnaud « Les limites actuelles de l'apprentissage collaboratif en ligne », Sciences et Technologies de l´Information et de la Communication pour l´Éducation et la Formation, Volume 10, 2003.

[9] F. Abdallah, « Méta-modélisation pour décrire et instrumenter une situation d'apprentissage de Pédagogie par Projet Collectif », Thèse de doctorat de l'Université du Maine, 2009.

[10] V-S. Kumar. computer-supported collaborative learning : Issues for research. In Eight Annual Graduate Symposium on Computer Science, University of Saskatchewan, 1996.

[11] P. Gounon and P. Leroux. Ediscéne, un editeur de scénarios d'apprentissage et d'encadrement. TICE'2006, 2006.

[12] Reload. Reusable eLearning Object Authoring Delivery. http ://www.reload.ac.uk/, 2006.

[13] S. George. Apprentissage collectif à distance. SPLACH : un environnementinformatique support d'une pédagogie de projet. PhD thesis, Thèse de doctorat, Université du Maine, 2001.

[14] P.A. Caron, F. Hoogstoel, X. Le Pallec, and B. Warin. Construire des dispositifs sur la plateforme moodle - application de l'ingénierie bricoles. In MoodleMoot-2007, Castres, France, 14 - 15 Juin 2007.

[15] Swebok. Guide to the software Engineering of body of knowledge. IEEE Computer Society, 2004.

NetUniversité. NetUniversité. http ://www.cepiahhds.utc.fr :8080/CEPIAHV3/web/index-1.jsp.

Diff- HEEP: Hybrid Energy Efficiency Protocol with differentiated services for Wireless Sensor Network

Bourmada Amal ,
Dept of computer science UHL Batna
abourmada@yahoo.fr

Bilami Azeddine
Dept of computer science UHL Batna
abilami@yahoo.fr

ABSTRACT— In these recent years Wireless Sensor Networks (WSN) has been widely used in many types of applications such as environment monitoring or asset tracking. In these applications, each sensor node may collect different types of data with different levels of importance. Each type of traffic requires a different treatment from the network to meet the objective of QoS (Quality of Service). We argue that the sensor networks should be willing to spend more resources in disseminating packets that carry more important information. In this paper we propose a new protocol with differentiated services for WSNs called Diff-HEEP. The proposed protocol is an enhancement of HEEP protocol that can provide requested quality of service for high priority real time classes and distinguish high priority real time traffic from the low priority non-real time traffic, and input traffic streams are then served based on their priorities.

KEYWORDS— Wireless sensor network; Quality of service; Routing Protocol; Service differentiation, HEEP.

I. INTRODUCTION

A Wireless Sensor Network (WSN) is a type of wireless network that consists of distributed autonomous devices, referred to as nodes or motes, equipped with sensors that can monitor environmental or physical conditions, such as humidity, pressure, etc. These devices are randomly deployed in an area of interest. In recent years, Wireless Sensor Networks (WSNs) have gained attention in the research community due to their application in a multitude of real world situations. Invariably, the application domain imposes a restriction on the size and the cost of a single sensor node Thus a WSN will have the capability to transmit different kinds of data, such as still pictures, stream video, voice, animal sounds and monitoring data. Each of these kinds of data has their own unique requirements making QoS provision and service differentiation a necessity. In these new areas, there is real-time and non-real-time traffic in the sensor networks. Each type of traffic requires a different treatment from the network to meet the objective QoS (Quality of Service).

Quality of service can be provided by generously over-provisioning a network so that interior links are considerably faster than access links. But the over-provisioning mechanisms waste the network resources. Similar to services, such as real time voice and video conference and file transfer, which have varying needs for packet delay, packet loss, bandwidth, availability, etc., different classes of services also require varying levels of preferential traffic treatment. Consequently, organizations are seeking ways to guarantee that the most important traffic is always given the highest priority throughout the network. The Internet Engineering Task Force (IETF) has defined architecture for Differentiated Services (DiffServ) for delivering quality of service.

The protocol HEEP [1] is a routing protocol for wireless sensor networks (Hybrid Energy Efficiency Protocol) that provides more improvement in term of energy efficiency. The proposed protocol is considered as an enhancement of HEEP protocol that can provide requested quality of service for high priority real time classes and distinguish high priority real time traffic from the low priority non-real time traffic.

II. RELATED WORK

A. An overview of HEEP (Hybrid Energy Efficient Protocol)

In this protocol, nodes of the same cluster are organized as a chain which improves energy consumption, and minimizes the overhead on the CH (cluster-head). Indeed, the nodes can not communicate directly with the CH but only with their closest neighbors, which limits power consumption and provides better utilization of bandwidth. Data aggregation of each node in a chain reduces the amount of exchanging data between nodes and their CH, to preserve node energy reserves [1].

Figure 1 Shows how the nodes are organized within clusters, each node N transmits its data to its nearest neighbor which in turn aggregates the received data with its own and sends it to its other neighbor until reaching the CH which transmits directly to the BS.

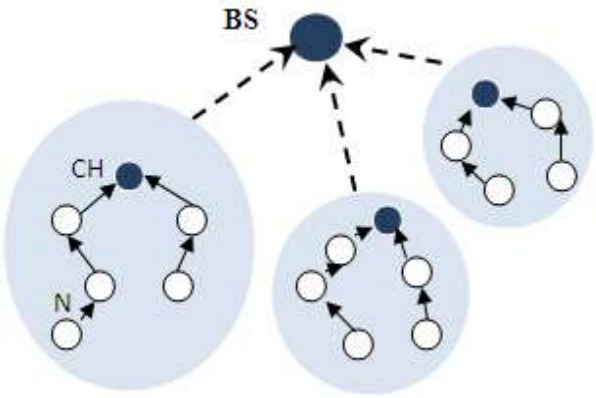


Figure 1: Chains cluster organization in HEEP.

This protocol is inspired by two protocols LEACH [4.5] and PEGASIS [11]. But unlike LEACH, the number of nodes that communicate with the CH is considerably reduced. That implies better energy saving and prolongs CHs lifetime, the HEEP protocol adopts the concept of random rotation of CH role proposed by LEACH, which regulates the energy dissipation and prevents that the selected nodes as CHs do not die faster. However, HEEP uses the concept of PEGASIS to organize the cluster nodes as a chain. CHs establish a transmission plan (TDMA schedule) that attributes to each node the exact time which he must transmit data collected. HEEP has two essential phases: the initialization phase that forms chained clusters, and elects CHs, and transmission phases where collected data is transmitted.

B. Service differentiation in wireless sensor network

In a wireless sensor network each sensor node may have different types of sensor data and with different levels of importance. Some applications of WSNs require the transmission of real-time traffic by sensor nodes. This real time traffic requires a small latency and high reliability.

Service differentiation is the main technique used in wireless sensor networks to support QoS. However, used alone, it is not sufficient to achieve deterministic guarantees. QoS in wireless sensor networks is largely best- effort.
The first part in the differentiation of service is the traffic classification, i.e. the specification of the priority of each packet flowing through the network. The choice can be static or dynamic. In the first case, a priority is assigned to the packet since its inception. It does not change during its passage through the network [7]. The allocation may be based on the traffic type of the data collected or any order specified by the network designers. In the second case, the priority varies during the traverse of the packet. It depends, for example, the number of hops traversed, the number of hops remaining, the remaining energy of a node. In wireless sensor networks, it is mainly based on changes in the parameters of the protocol based on the priority of traffic. The service differentiation is one of the basic mechanisms of QoS. It is more important when the network require managing multiple traffic classes, with different priority. The main classes defined in the DiffServ model are the follows:

Expedited Forwarding (EF): that provides low packet loss, low latency, low jitter and end-to-end bandwidth service.
Assured Forwarding (AF): This class is intended for applications sensitive to packet loss. Packets from this class must be removed as often as possible in case of congestion.
The DiffServ architecture allows up to 64 classes of traffic. The two examples above are not the only politic.

III. MODEL OF DIFFERENTIATION SERVICE FOR HEEP

The model that we adopt for supporting QoS in a wireless sensor network is that presented by M. Yaghmaee, and D. Adjeroh [2]. The model has been implemented into HEEP protocol. The modified protocol is called diff-HEEP.

This model can support two basic traffic classes, namely, expedited forwarding (EF) class which is assigned to real time traffic, and Assured forwarding (AF) for non real-time traffic class. The non real-time traffics (AF) can be further divided into different levels of importance based on their requirement for network resources. We consider four types of traffic class: real time traffic (EF class), high priority non real time traffic (AF1 class), medium priority non real time traffic (AF2 class) and low priority non real time traffic (AF3 class). Each traffic class is assigned a different priority. Since guaranteeing a low delay bound is an important issue for real time traffic, this model guaranties a low delay bound for this type of traffic. Real time traffic is buffered in a separate queue with low buffer size, while non real-time traffic is managed using the active queue management algorithm.
Figure 2 shows an organization of the WSN using HEEP protocol in order to guarantee energy saving. Each node has different traffic classes.

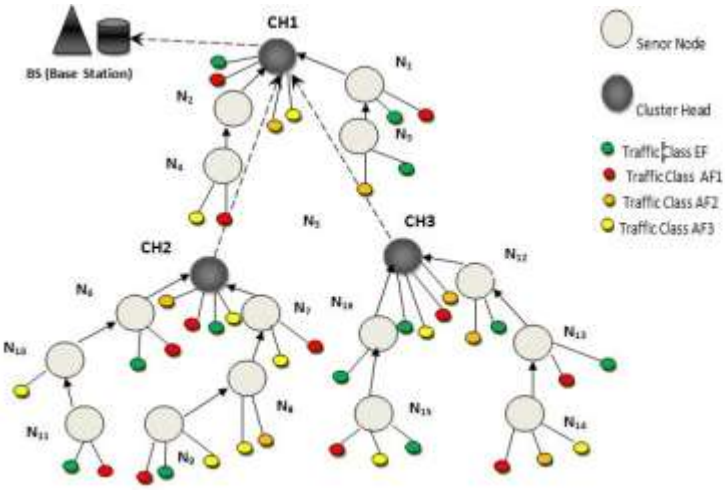


Figure 2: WSN using HEEP protocol with different traffic classes

This model uses separate queue for each type of traffic class. The queuing model of each node is shown in

Figure 3. To discriminate between traffic classes, each wireless node adds an identifier traffic class for each packet and puts them in the appropriate queue.
This identifier represents the traffic class of each packet. In each intermediate wireless node, arriving packets are sent to different queues according to their traffic class.

.

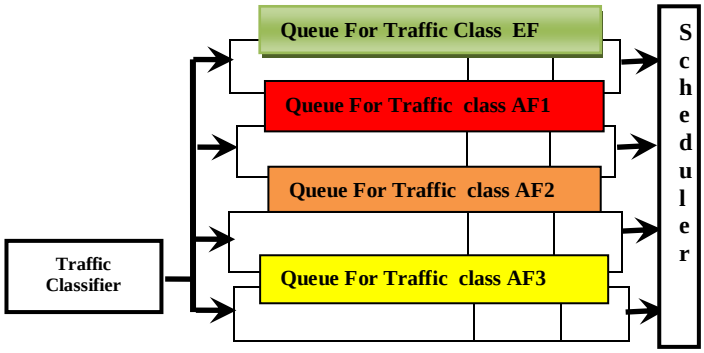


Figure 3: Traffic classification at each node N

For AF classes the RED protocol [3] is used to achieve service differentiation within sensor nodes. To offer quality of service for high priority EF flows in a WSN, we use the mechanism priority queuing (PQ) which prioritizes the packet transmission process at each node. *Figure 4* resume the mechanism of differentiated service in the WSN

Each sensor node i has different traffic class priority as showing in *Figure 2*.

SP_j (i) represents the priority of the traffic source j in sensor node i. The value of SPj (i) could be set manually to achieve service differentiation. For high priority traffic the value of $SP_j(i)$ should be high for discriminate it from the other low priority traffics.

The priority of the node i NP (i) is equal to the sum of source priority $SP_j(i)$ and is calculated as follows:

$$NP(i) = \sum_j SP_j(i)$$

Where:

J: is the traffic class, jє {EF, AF1, AF2, and AF3}.

The global priority of each node i, GP_i is calculated as:

$$GP(i) = \sum_{k \in c(i)} GP(k) + NP(i)$$

Where c (i): the set of the i's child nodes.

As each sensor node may have different traffic classes. The global priority for each traffic class can be calculated as follows:

$$GP_{EF}(i) = \sum_{k \epsilon\, c(i)} GP_{EF}(K) + SP_{EF}(i)$$
$$GP_{AF1}(i) = \sum_{k \epsilon\, c(i)} GP_{AF1}(K) + SP_{AF1}(i)$$
$$GP_{AF2}(i) = \sum_{k \epsilon\, c(i)} GP_{AF2}(K) + SP_{AF2}(i)$$
$$GP_{AF3}(i) = \sum_{k \epsilon\, c(i)} GP_{AF3}(K) + SP_{AF3}(i)$$

T_s (i): The service time of the current packet in the node i.

$$\bar{T}_s(i) = (1-\alpha).\, \bar{T}_s(i) + \alpha.\, T_s(i)$$

Where: α is a constant

$\bar{T}_s$ (i) : : Is the time taken to successfully transmit a data packet over MAC layer and measured starting from the time when the network layer first sends the packet to the MAC layer to the time when the MAC layer notifies the network layer that the packet has been transmitted.

The node rate ri, is calculated as follow:

$$ri = \frac{1}{\bar{T}s(i)}$$

Let $\bar{T}_\alpha(i)$ denote the average inter arrival time of each packet to sensor node i .this value is updated periodically whenever there are M new packets arriving β is another constant, and T_M is the time interval over which the measurements are performed, and within which The M new packets arrive. As follow:

$$\bar{T}_\alpha(i) = (1-\beta).\, \bar{T}_\alpha(i) + \beta.\, \frac{T_M}{M}$$

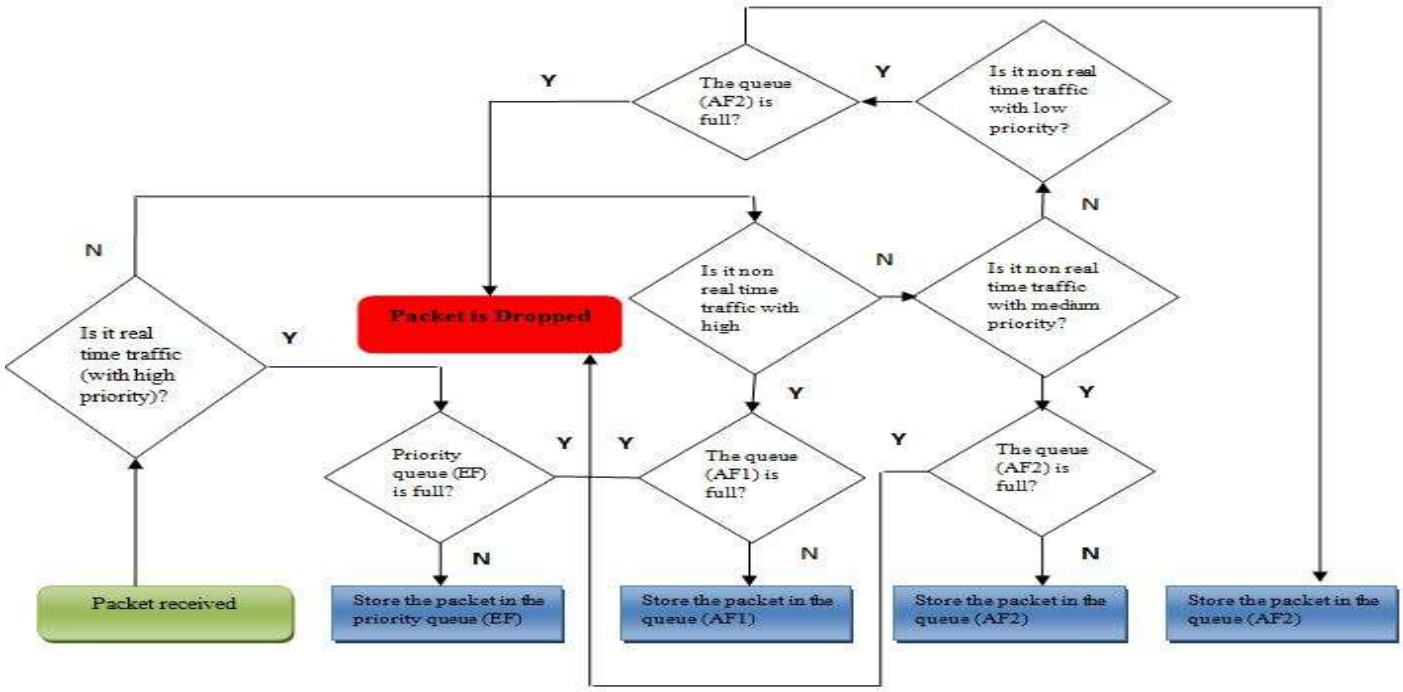


Figure 4: Traffic Classification according to priority diagram

IV. QUEUING MODELS

Regardless of its numerous limitations, First In, First Out (FIFO) is the default queuing algorithm in several topologies that require no configuration. Most importantly, FIFO queuing makes no decision about packet priority. FIFO queuing involves storing packets and forwarding them in order of arrival. Explode sources can cause extended delays in delivering time-sensitive application traffic and, potentially, in networking control and signaling messages. Although FIFO queuing was an effective network traffic controller before, more recent intellectual networks need more sophisticated algorithms [9]. Furthermore, in FIFO, a full queue will cause dropping of packets – even high-priority packets. In fact, the sensor node cannot prevent this undesirable packet dropping, as it has no room in its queue. Furthermore, FIFO cannot differentiate between a high priority and low-priority packet [8].

Different types of queuing tools, such as Priority Queuing (PQ), Custom Queuing (CQ) [6], Weighted Fair Queuing (WFQ), Weighted Round-Robin (WRR), Class-Based Weighted Fair Queuing (CBWFQ) [7] and Low Latency Queuing (LLQ), Deficit Weighted Round Robin (DWRR) have been developed to provide different services to heterogeneous traffic classes. The sensor nodes need elegant and uncomplicated queuing techniques, since they usually work as small routers. The implementation of Priority Queuing (PQ) in a sensor node is exploited to calculate the queuing delay for two different kinds of traffic in a sensor node [10].

Priority Queuing scheduler

To overcome the limitations of the FIFO queuing discipline, Priority Queuing (PQ) is suggested as one of the applicable solutions to meet the desired QoS for real time traffic. It is used in a wide variety of applications including operating systems, real-time systems, and discrete event simulations. In this work, Four queues in a sensor node are considered: one for real time traffic (EF class), the others for high priority non real time traffic (AF1 class), medium priority and low-priority. Each element is ordered by its associated priority. The packet with the highest priority goes first, regardless of the order of arrival. A priority queuing mechanism provides low delay bound for high priority traffic. Low priority traffics are all queued in the lower queues, while high priority traffic, is queued in the upper queue. The queue policy determines how the queues will be emptied when all queues have packets waiting. In the simple priority policy, high priority packets are serviced first, and then the remaining packets in the low priority queues. If additional high-priority packets arrive while the lower queues are being emptied, service is immediately switched to the high priority queue. So Priority Queuing (PQ) scheduling is very suitable for providing quality of service for real time traffic.

V. CONCLUSION

New applications made possible by the rapid developments and miniaturization in hardware has

motivated recent developments in Wireless Sensor Networks (WSNs). For some applications, there is a need to send real time traffic toward the sink node with low latency and high reliability so that immediate remedial and defensive actions can be taken, as appropriate. Further, when an important event occurs in the system, the sensor node that detected the event should send some alarm message to the cluster head. Usually this kind of high priority traffics are bursty. This means that high priority traffic are generated only for a short period of time while low priority traffics usually exist in the network and produce thousands of packets generated periodically. For such environments, service differentiation in wireless sensor networks becomes an important problem. To provide service differentiation in WSNs, it is necessary to consider a different priority for each traffic source. In this paper we integrate a model for service differentiation in the protocol HEEP. The proposed protocol can support four different traffic classes including: real time traffic class (EF class), high priority, non real-time traffic class (AF1 class), medium priority, non real-time traffic class (AF2 class) and low priority, non real-time traffic class (AF3 class).

VI. REFERENCES

[1] D. Boubiche, A. Bilami, "HEEP (Hybrid Energy Efficiency Protocol) based on chain clustering" International Journal of Sensor Networks archive volume 10 Issue 1/2 , june 2011 pages 25-35.

[2] M. Hossien Yaghmaee, D. Adjeroh, "A Model for Differentiated Service Support in Wireless Multimedia Sensor Networks" IEEE Xplore February 5, 2009

[3] S. Floyd, V. Jacobson, Random early detection gateways for congestion avoidance, IEEE/ACM Transactions on Network, August 1993

[4] W.R. Heinzelman, A. Chandrakasan, et H.Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Micro sensor Networks," IEEE Transactions on the wireless communications, Vol. 1, no. 4, pp. 660-670, Oct 2002.

[5] W.R. Heinzelman, A. Chandrakasan, H. Balakrishnan "Energy-efficient communication protocol for wireless micro sensor networks," IEEE Hawaii International Conference on System Sciences, 2000.

[6] Aalsalem, M. Y., Zomaya, A., Iftikhar, M., and Taheri, J., "A QoS framework for wireless sensor networks," Journal of IEEE Transactions on Mobile Computing, 2008.

[7] Eui-Jik Kim, Meejoung Kim, Sung-Kwan Youm, Seokhoon Choi and Chul-Hee Kang. Priority-based service differentiation scheme for IEEE 802.15.4 sensor networks. AEU – International Journal of Electronics and Communications. Volume 61, Issue 2, Pages 69-81, February 2007.

[8] RAJSHREE S. DUBEY1, " Challenges for Quality of Service (QoS) in Wireless Sensor Networks" International Journal of Engineering Science and Technology Vol. 2 (12), 2010

[9] Wee Lum Tan; OnChing Yue; Wing Cheong Lau; "Performance Evaluation of Differentiated Services Mechanisms Over Wireless Sensor Networks», Vehicular Technology Conference, 2006. VTC-2006 fall. 2006 IEEE 64 thSept. 2006 Page(s):1 – 5

[10] Rakesh Kumar Pradhan, B.Tech. "Priority Based Energy Efficient Access Networks" JANUARY 2013

[11] S. Lindsey, C. Raghavendra, "PEGASIS: Power-Efficient Gathering in Sensor Information Systems," IEEE Aerospace Conference Proceedings, Vol. 3, pp. 1125-1130, 2002.

Routing protocol for energy efficiency and fault tolerance in Wireless sensor network

Yasmine Djebaili, Azeddine Bilami
Department of computer science
laSTIC laboratory
University of Batna, Algeria
d-nimsay@hotmail.com, abilami@yahoo.fr

Abstract— Sensors in a Wireless Sensor Network (WSN) are prone to failure, due to energy depletion, hardware failures, etc. Fault tolerance is one of the critical issues in WSNs. The existing fault tolerant mechanisms either consume significant extra energy to detect and recover from the failures or need to use additional hardware and software resource. In this paper, we propose a novel energy-aware fault tolerance protocol for WSN.

Keywords: Wireless sensor network, Fault tolerance, fault tolerant routing protocol

I. INTRODUCTION

Wireless sensor networks (WSN) are self-organized networks that typically consist of a large number of low-cost and low-powered sensor devices, called sensor nodes, which can be deployed over a geographical area for monitoring physical phenomena like temperature, humidity, vibrations, seismic events, and so on [1]. Now, WSN are permeating a variety of application domains such as avionics, environmental monitoring, structural sensing, telemedicine, space exploration, and command and control.

WSN should have a long lifetime to accomplish the application requirements. However, In addition to resource constraints in WSN, the failure of sensor nodes is almost unavoidable due to energy depletion since they have been usually deployed in hostile environments and their batteries cannot be recharged or replaced, hardware failure, communication link errors, and so on.

Therefore, fault tolerance has become as important as other performance metrics such as energy efficiency, latency and accuracy. In general, the consequence of these failures is that a node becomes unreachable, violates certain conditions that are essential for providing a service or returns false readings which could cause a disaster especially in critical applications.

Furthermore, the above fault scenarios are worsened by the multihop communication nature of WSN. It often takes several hops to deliver data from a sensor node to the remote base station; therefore, failure of a single node or link may lead to missing reports from the entire region of WSN. Therefore, since sensors are prone to failure, fault tolerance should be seriously considered in many sensor network applications. Recently, several studies have dealt with fault tolerance in WSN, particularly in the routing process. Moreover, these works focus on the detection and recovery of failures in WSN.

In this article, we present a new method for fault management of nodes and CH's for HEEP [2]. The rest of this article is organized as follows: First, we will give a brief introduction to wireless sensor networks. Then we present some fault-tolerant routing protocols, that have been proposed to manage failures in this type of network .thereafter, we present the mechanisms of detection and recovery of faults that we have we proposed to HEEP, and Finally, we conclude our paper

II. RELATED WORK

Several fault tolerant routing protocols have been proposed. These can be classified into three groups based on the methods used to find the path: proactive routing in which all paths are calculated and maintained in advance and stored in a routing table, reactive routing where all paths are created on demand, and hybrid routing combining both types of previous routing.

The research proposed in [3], present a proactive fault-tolerant routing protocol in which each node has two paths leading to the BS (a main path and a backup path). Whenever a node wants to send data to the BS, he chooses the path with the highest degree of energy. In this protocol, the communication between nodes is achieved by messages, where the sender node sends a "Hello!" to the destination node. This type of message is used to check if the neighbor is reachable. In addition, functions to report errors are incorporated in this protocol which includes different types of messages (failed links, critical battery message, destination unreachable, etc). This protocol has some limitations such as power consumption is quite large when the distribution of error reports. In addition the robustness of this solution is limited by the number of nodes registered (failure of the backup node).

In [4], Another fault tolerant routing protocol have been proposed, in which each node send two types of packets to the

CH during its time slot, a data packet or a packet to inform the CH that is not out of order but it has no data to send. At the end of the round, the data sent to the CH are aggregated and sent to the BS, if it does not receive any packet from a CH, she sent him a "Hello packet!". If it receives no ACK, she declares it out of order and goes to the election of another CH for the next round, in this case the data collected at the cluster where the CH is down will be lost.

ENFAT-AODV (The Enhanced Fault-Tolerance Mechanism of AODV Routing Protocol) is a reactive fault tolerant routing protocol [5], in which two paths are used for data delivery to the base station: a main path and a backup path, during the transmission phase, if the node has an invalid main path or receive a data packet from a node for which it doesn't have a main path, it immediately uses it's backup path to deliver the coming data packets without interruption. After that, the node launches the process of discovery of a new main path.

IHR [6] is a fault tolerant protocol that has been created based on DHR, in which each CH choose among the nodes of the group who has the greatest level of energy and select it as a BCH, whose role is to route data where the CH fails. At each period the BCH send a "Beacon message" to the CH, if it is not broken he replied else the BCH waits for 3 rounds, if after that time he don't receive any response, he declares the node as faulty and send a message to the nodes of its cluster to ask them to send their data.

CFS (Cluster based Fault Tolerant Scheme) [8] is a fault tolerant protocol in which two CH's are elected: a primary CH (CH_p) and vice CH (CH_v), both CH's are elected based on the weight of the sensor , this parameter is calculated at the beginning of each round in order to prevent rapid energy depletion, when the node detects an event, it sends a data packet to the (CH_p), if he don't reply with un ACK, the (CH_v) waits a period of time then he send un ACK to the node and declare the (CH_p) as faulty . When all the nodes send their data they turn off their radio in order to conserve their energy.

FTCP-MWSN [10] is a fault tolerant routing protocol designed for mobile sensor networks which assumes that all nodes are homogeneous in terms of mobility and move with constant speed. In this protocol each node generates two types of packets: a data packet or a small packet to inform the CH that is not broken and does not move outside the cluster. Once the data packet or the small packet received, the CH sends an ACK. If it receives nothing he declares the node as out of the cluster to the BS and removes it from the list of cluster members and also removes its time slot TDMA. Similarly, when the node doesn't receive an ACK he deduces that it is not attached to CH due to mobility, and sends "JOIN-REQUEST" message to all CH's. Although this protocol seems good results compared to other protocols, but it is incomplete as it did not address the problem of failure of CH's something that can cause a loss of data packets. In addition, it seems expensive in terms of energy given that there's a lot of message exchange between nodes. This protocol seems expensive also in terms of cost since it requires each time the location of all nodes.

IFTMRP (Inter-Cluster Fault Tolerant Multilevel Routing Protocol) [11] is a fault tolerant routing protocol whose objective is to consider the fault of CH when transmitting and receiving and ensuring inter-cluster fault tolerance. To make this, every CH choose two neighboring CH instead of one in order to be able to recovered data when the first CH fails. When the CH send data to its neighbor CH, his neighbor must send an "ACK" packet, if the sending CH doesn't receive any "ACK", he transmits data to the second CH. The limit of this protocol is that the data collected in the cluster where the CH is out of order are lost.

IIFTMRP (Inter-Intra cluster Fault Tolerant Multilevel Routing Protocol) [11] is a protocol that was developed to solve the problem of data loss within the cluster. In IIFTMRP, the CH chooses from the members of its cluster that has the greatest degree of energy as vice CH , this vice CH is responsible for sending data when the CH fails. During the slot reserved for the vice CH, if the CH doesn't receive data, it sends an error message to the vice CH notify it down. After receiving the message, the vice CH is responsible for sending data. This solution has solved the problem but it is costly in terms of energy due to the large number of messages sent and received by nodes.

III. PROPOSED PROTOCOL

In this paper we propose a new protocol to manage faults. This new protocol is an extension of HEEP (Hybrid Energy Efficient Protocol) [2], in which the nodes in the same cluster are organized like a chain, this organization can improve the energy consumption. Indeed, the nodes cannot communicate directly with the CH but only with their close neighbors, which saves energy and provides a better use of the bandwidth. The aggregation of data at each node in the chain reduces the amount of data exchanged between the nodes and CH, which has the effect of preserving the energy. Figure 1 shows how the nodes are organized within the clusters.

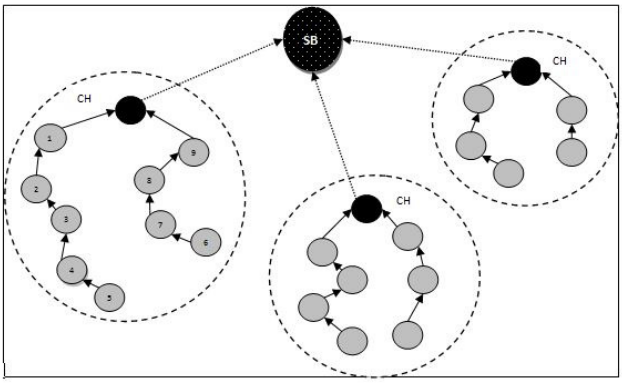


Figure1.Organization of the nodes according to HEEP.

Contrary to HEEP protocol which introduces no mechanism for fault tolerance, our new protocol provides a mechanism for fault tolerance for nodes as well as for CH's; this new mechanism is based on the following three rules: Rapid detection of the fault; Instant recovery of the failure; Continuation of the mission.

A. DETECTION OF THE FAILURE NODES AND CH'S

Fault detection is the first phase in the fault management, where an unexpected failure should be correctly identified. It can be done with preventive or corrective way.

1. PREVENTIVE APPROACH

The objective of this first approach is to avoid failures due to the energy problem [3]. During the initialization phase, and after sending the ID, location and the energy level of each node to the BS, it checks the energy level of each node, if the reserve of energy has reached a low level; it removes it from the list of nodes that will participate in the routing of information and also removes its plan TDMA. Else it allocates a slot during which it will submit its data to the neighboring node.

2. CORRECTIVE APPROACH

During the transmission phase and before data communication, each node including the CH must generate a "Hello packet!" During his slot TDMA, this will allow its nearest neighbors in the chain to know their status [9]. If the node generates the packet it means that it is still alive else the node who has submitted data declares it down to the CH (send her ID) and pass to the recovery procedure. If this node is itself a CH the last node in the chain will declare it faulty to the BS. To do this we must extend the period of time allocated to each node as follows:

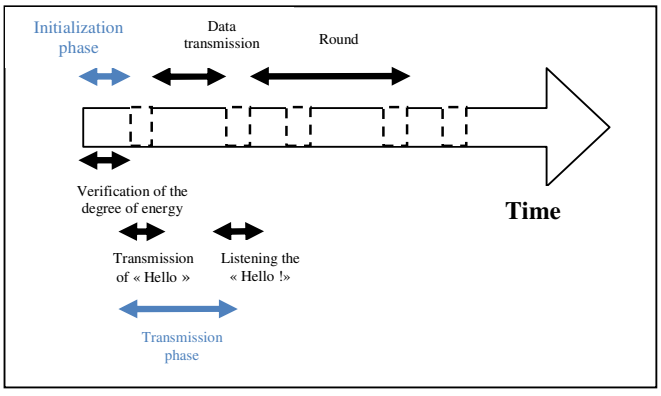


Figure 2. Execution steps of the protocol

B. THE FAILURE RECOVERY

The failure recovery is the step in which the sensor network is reconfigured and reconstituted in such a way that failures do not affect network performance.

1. RECOVERY OF THE FAILURE OF NODES

After the chain formation, each node choosing the neighbor's of its neighbor node as backup node if ever his neighbor fails the backup node is responsible for routing data through the chain until they reach the CH.

For example, when Node 4 fails, the node 5 uses its backup node (node 3) to route data through the chain until they reach the CH.

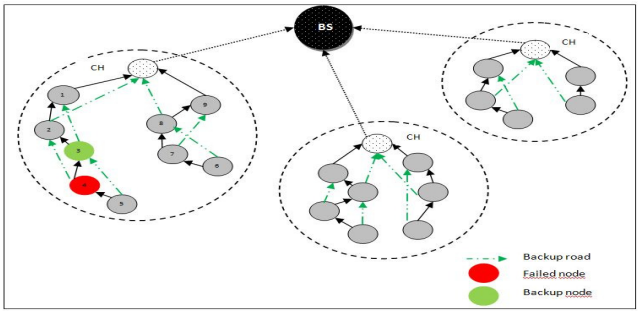


Figure 3. Scheme recovery of the failure of nodes

2. RECOVERY OF THE FAILURE OF CH'S

When the CH fails, the node that has the greatest degree of energy and located in one hop from him, in our case the node 1/9 for example is responsible for routing data directly to the BS, as shown in the figure:

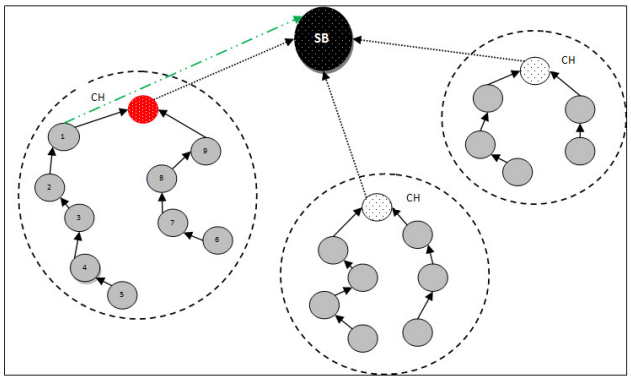


Figure 4. Scheme recovery of the failure of CH's

IV. CONCLUSION

In this paper, we proposed a solution for the detection and recovery of failures of nodes and CH's for HEEP. The idea was to check the level of energy of each node in the initialization phase and generate small packets during the transmission phase, to ensure instantaneous fault detection.
As for the idea of recovery, it is to choose the neighbor of its neighbor as backup node in order to adjust the failure of nodes and select the node that is located in one hop from the CH and that have the greater degree of energy to route data if the CH fails.

Finally, after testing our solution under NS2 simulator, we plan to evaluate the performance of our protocol and compare them to those of fault-tolerant protocols in WSN such as IHR, CFS, etc…

REFERENCES

[1] Akyildiz F, Weilian Su, Sankarasubramaniam Y. et Cayirci E, "A survey on sensor networks", *IEEE Communications*, August, 2002.
[2] D.Boubiche et A.Bilami, " HEEP (Hybrid Energy Efficiency Protocol) Based on Chain Clustering", *Int. J. Sensor Networks*, 2011.
[3] K. Kulothungan, J. Angel Arul Jothi et A. Kannan, "An Adaptive Fault Tolerant Routing Protocol with Error Reporting Scheme for Wireless Sensor Networks", European Journal of Scientific Research, ISSN 1450-216X, Vol.60, No.1, pp.19-32, 2011
[4] Lutful Karim, Nidal Nasser et Tarek Sheltami, "A Fault Tolerant Dynamic Clustering Protocol of Wireless Sensor Networks", in proceedings of the IEEE "GLOBECOM", 2009
[5] Zamree Che-Aron, Wajdi Al-Khateeb et Farhat Anwar, "The Enhanced Fault-Tolerance Mechanism of AODV Routing Protocol for Wireless Sensor Network", IJCSNS International Journal of Computer Science and Network Security, VOL.10, No.6, June 2010.
[6] Meikang Qiu, Jianning Liu, Jiayin Li, Zongming Fei, Zhong Ming et Edwin H. M. Sha, "A Novel Energy-Aware Fault Tolerance Mechanism for Wireless Sensor Networks", IEEE/ACM International Conference on Green Computing and Communications, 2011.
[7] N. Jain, V. M. Vokkarane, et J. P. Wang, "Performance analysis of dual-homed fault-tolerant routing in wireless sensor networks", IEEE Conference on Technologies for Homeland Security, pp. 474–479, May 2008
[8]Mohamed LEHSAINI et Chifaa TABET HELLEL, "A Novel Cluster-based Fault-tolerant Scheme for Wireless Sensor Networks", in Proceedings of the 24th International Conference on Microelectronics (ICM), 2012
[9] Charles E Perkins et Elizabeth M Royer, "Adhoc On Demand Distance Vector Routing ".
[10] Lutful Karim et Nidal Nasser. " Energy Efficient and Fault Tolerant Routing Protocol for Mobile Sensor Network", the proceedings of IEEE ICC, 2011
[11] Zibouda Aliouat et Makhlouf Aliouat, "Improving Wireless Sensor Networks Robustness through Multi-level Fault Tolerant Routing Protocol", Springer, 2013

An Overview of Question-Answering Systems

Benaissa Safa
Superior national school of computer science,
Algiers, Algeria
s_benaissa@esi.fr

Omar Nouali
Theory Division & Computer Systems Engineering,
CE.R.I.S.T., 3 brothers street, Aïssiou, Ben Aknoun,
Algiers, Algeria
onouali@cerist.dz

***Abstract*— The question-answering systems (QAS) can be defined as systems advanced information retrieval which goal to return a specific answer, or a fragment of text (paragraph, sentence, ...) containing the answer to a user query, as opposed to research engines that return a standard set of documents considered pertinent. This query is a question expressed by the user in natural language without any constraint vocabulary. The question asked undergone a process of analysis and processing to automatically extract from a documentary base, a response is considered correct at the question. A difference a traditional research systems, the answer extraction does not require human intervention. This paper presents an overview of the existing question-answering systems, and a study of different mechanisms and techniques used for the development of such systems.**

***Keywords*— question-answering systems, question analysis, information research, answer extraction.**

I. Introduction

The information volume available on the Internet creates a problem of an important information overload. It therefore becomes necessary to develop tools to filter and to better target relevant answers provided to users as possible and eliminate unnecessary information.

The question-answering systems go beyond information retrieval where information retrieval systems do not return to the user that the documents deemed relevant, this relevance is measured by the value of the correspondence of keywords of the documents to those of the user's request. The question-answering systems aim to provide a specific answer to a question, not just the documents that might contain the answer. [24] The answer must be extracted or reconstructed from documents.

Today, there are an important number of question-answering systems. They are distinguished by several factors such as the questions processed category (definition, factual ...), the type of data processed (structured data, semi-structured, unstructured or multimedia ...), and the use of one or more languages at the same time...

This article presents an overview of the existing question-answer systems, and a study of the various mechanisms and techniques used for the development of such systems. It is organized as follows, primarily, the first section provides an introduction, Section 2 describes the global architecture of a question-answering system, Section 3 discusses the evaluation of such systems, Section 4 presents related work, and section 5 presents a summary of previous work. Finally, a conclusion is presented in Section 6, which gives a brief overview on the future directions of research in the domain.

II. Architecture of Question-Answering Systems

In order to achieve a final answer, precise and concise, it is not a simple task, ordinary or trivial. Briefly, this operation requires a sequence of process chained as a process of understanding and filtering in order to reduce the research field: from the corpus to a reduced documents collection that contain the desired answer, then, from this collection to passages considered relevant, and finally, from these passages to a specific and precise answer [7].

Generally, a question-answer system can be described by the sequence of three main process, namely a question analysis module, a documents processing module includes that two sub-modules, one for information retrieval and the other for the detection and selection of the relevant passages, and an answer extraction module [10] (Fig. 1).

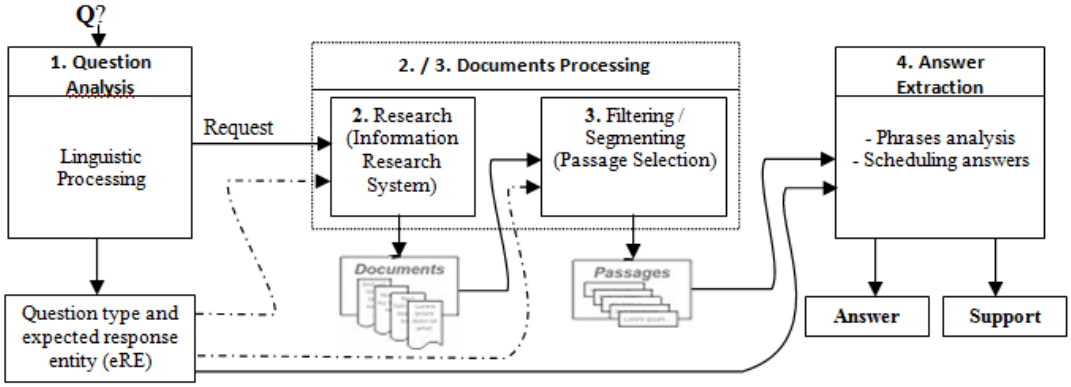


Fig. 1 Generic architecture of question-answering system

A brief detail about modules of Question-Answering System is presented below:

A. Question analysis module

This step is very important in the process of a question-answer system, because valuable processing as possible to the question, is applying the research strategy most appropriate.

The analysis of the question is, on the one hand, a linguistic processing of the sentence and extracts the key words, and secondly to determine the type of waited answer. More specifically, this step is to identify the informative elements of a question (type of question, subject "focus" to which the question based, the type of response waited and important terms) to find the answer in the corpus. So, this module must determine:

- The categorization of the question: to assign a type to the question according to a predefined classification (definition, factual, Boolean ...). A question as a factual « What is the capital of Japan? » And another type definition such as « What is the definition of management? »;
- The focus of the question: is the entity sought by the question;
- Named entities in question: this step determines all named entities make up the question. For example, there are two named entities in the following question: « When Jacques Chirac preside France? » A person entity « Jacques Chirac » and another type place « France »;
- The type of waited answer: generally determined according to the form of a type of named entity (person, date, place ...) or a broader type of entity (disease, treatment, event ...). For example, in the question « Who is the president of France? », the type of waited answer is an entity named person;
- The theme of the question (topic): represents the topic to which the question. For example, in the question « In which city is born President of the United States? », « the city » is the focus and the topic is « President of the United States ».

The important terms to define the context of the research and formulate a query used by the second module. In the following question: « What is the date of birth of President of the United States? », interesting terms: « date, birth, President, United States » and the desired answer is an entity named date.

B. Documents processing module

The objective of this module is the screening of the relevant passages. Firstly, a search engine has questioned the existence of relevant documents responsive to the request from the keywords extracted in the first module. Then, the selected documents undergo a filtering / segmentation in order to detect the best parts of texts or sentences that may contain the desired response based on key terms. The initial request may be enriched by the contributions of techniques of the natural language processing (NLP) and therefore, by semantic variations (synonyms, hyponyms or semantic links between nouns and verbs ...).

C. Answer extraction module

The task of seeking an answer is to search phrases or passages considered relevant returned by the second module to locate and retrieve the response is considered correct. Generally, this module is divided into two stages, the first is used to analyze sentences or passages candidates to retrieve a set of responses deemed correct and second is to order the set of answers extracted and select the final answer.

The extraction of a response is strongly dependent on the type of the expected response, if the response type is named entity; a technique applied is to identify all named entities corresponding to the semantic type of the desired response in the passages. However, if the expected response is not a named entity, the systems use predefined extraction patterns. Another approach to exploit external knowledge sources such as the Wikipedia encyclopedia as possible answers bases.

To order the answers, the system uses a scoring function based on a sorting criterion, a rate of correspondence between the type of the candidate response and that awaited response or frequency of occurrence of the candidate answer in the selected documents. There are various forms to submit the response returned to the user, is the most widely used form of an ordered list of candidate answers the most common.

III. EVALUATION OF THE QUESTION-ANSWERING SYSTEMS

The evaluation is to provide participants with a collection of textual documents, and a set of questions that their systems must return responses accompanied by the document of justification. The answers returned are judged by human experts according to predefined guidelines.

The first evaluation of question answering systems was marked in 1999 as part of the campaign cycle assessment TREC-8 in the United States [21]. For this edition, the assessment task has become habitual and by such annual campaign, which focuses exclusively on the English language. The following versions of TREC take into account changes in some parameters such as the size of the document collection, the complexity and the number of questions (200 to 700), the maximum length and the number authorized to provide answers (between 1 and 5) [22].

Since 2000, a series of annual evaluation campaigns appeared as the QA@CLEF [1] is interested in European languages, and is derived from TREC. The existence of other evaluation campaigns organized since 2003 as NTCIR[2] for Asian languages (Japanese in particular) and, more recently, in 2004, the birth of the companion EQueR[3] in France for the French language. [5]

[1] CLEF : Cross-Language Evaluation Forum ; see website http://clef.iei.pi.cnr.it/

[2] NTCIR : NACSIS Test Collection for IR Systems ; see website http://research.nii.ac.jp/ntcir/

[3] EQueR : Evaluation des systèmes de Question-Réponse.

IV. Related Work

Below, a brief overview of some key approaches that have achieved the best results in recent campaigns TREC and CLEF evaluation [5].

A. Architecture multiple-Stream

The effectiveness of question-answering systems is reflected by the rate of recall (proportion of relevant documents found compared to those that are relevant in the corpus). Indeed, the formulation of the question and that of the documents probably containing the answer is frequently different, implying that a portion of the relevant documents are not returned, because there are a variety of words or phrases can express the same concept. A question-answering system was developed by the University of Amsterdam [4] based on architecture with multiple streams, each stream has its own way to determine the answer candidates. These flows can be divided into four classes:

- Pattern Matching: to achieve matching, this flow is based on the generation of regular expressions from the questions;
- Look up: a flow uses a knowledge base built from prior to collection;
- Dutch Tequesta : (TExtual QUESTion Answering system) a textual question answering system for Dutch [13] incorporating a named entity recognizer, a tagging system morphosyntactic POS (Part-Of-Speech) and based on the notion of proximity to select candidates responses;
- n-gram: responses to search candidates, this flow is based the technique of n-grams.

The example below shows decomposition into n-grams of words in the sentence « This document may contain the exact answer »:

- 1-grams: This document may contain the exact answer.
- bi-grams: This-document document-may may-contain contain-the the-exact exact-answer.
- tri-grams: This-document-may document-may-contain may-contain-the contain-the- exact the-exact-answer.

Firstly, the responses determined by each of these streams are filtered to merge the responses deemed relevant in a single list and remove those deemed improper [19]. Similarly, eliminating answers that are from the web and having no arguments in the corpus. Finally, a voting system to differentiate the remaining responses to select the best response that the system will present the user. This system showed the best performance at the QA@CLEF-2004 [14].

B. Semantic networks

A Question-Answering System based on the technique of semantic networks has been proposed by the University of Hagen (Germany) German named InSicht. Firstly, a syntactic-semantic analyzer performs of questions individually analyze and the documents. Then a representation is developed as a semantic network [3].

In this system, different representations of the same question can be generated by extension of the query. After that, the system performs a comparison of each question representation with those of candidate documents. Configuration as compared guaranteeing correspondence can generate a candidate answer. To select the best answer, the system is based on two preferred features, such as the length and frequency of occurrence of answers in the collection [16].

C. Logical Reasoning

PowerAnswer is a QAS developed by the research team of the LCC[4] with an architecture based on logical reasoning [9]. In his analysis module of the question, the system determines the category of the question, the terms and descriptors that compose the type of awaited answer. In this analysis, the system will apply a named entity extractor. It uses a hierarchy of concepts se-mantic originated from WordNet [6] to define the named entities present in the question. Then, the system is based on the question under significant components to detect interesting passages in the document collection. Finally, the system uses the same module named entity recognition to identify and extract named entities corresponding to the type of awaited answer in selected passages. If the module fails to find an answer, the system uses a demonstration program automatic response based on a set of axioms mainly from WordNet to construct a logical reasoning to achieve matching between the logical representation of the question and that of the answers candidates. The enrichment of the axioms base done dynamically using a set of axioms product automatically between the key terms of the question and those answers. Therefore, the system can justify the answer extracted by referring to the user and the relevant passages considered reasoning formula linking the question and the answer.

D. Extraction Patterns

A Question-Answering System has been developed by the team InsightSoft-M (Moscow) based entirely on the application of extraction patterns as regular expressions and use a base of factual knowledge [18]. The team members make use of patterns describing the surface structures of strings may contain some semantic information [17]. The analysis of the question is carried out to determine the type and select patterns corresponding to implement. Then, each pattern characterized by a general semantics is applied to all candidate passages. Finally, matching is done on the basis of the pattern applied and the characteristic elements of the knowledge base, such as a name of a country, measures or currencies ... This system is ranked among the first in TREC-2002 [23], with 77% of correct answers. [2]

E. QUALIFIER

The team from the School of Computing at the University of Singapore conducted a QAS called QUALIFIER[5] [26]. In a first step, the system performs an analysis of the question to

[4] LCC : Language Computer Corporation.

[5] QUALIFIER : QUestion Answering by LexIcal FabrIc and External Resources

determine the category and the type of awaited answer. Named entities extracted in this step are used for the construction of ontology of classes of questions. Then the question is extended through the web and lexical resources such as WordNet [6]. To find the desired response, the system QUALIFY uses the information retrieval system Managing Gigabytes (MG) [25] to identify and track the responses candidates based on keywords constituents extent the question. Then, sentences considered relevant returned documents are classified according to association rules. Finally, the desired response is extracted by applying the named entity recognition, anaphora resolution and justification of response [26].

F. Diogene

Diogene is a name of a multilingual question-answering system developed by the center of scientific research and technological ITC-irst Italian [15]. At first, this system uses a POS tagger (Part-Of-Speech) to analyze the questions and identify the type of awaited answer. And thereafter to detect candidate documents, a standard information retrieval system called Managing Gigabytes (MG) [25] is used. Then, the system extract the response with a named entity recognizer [11] and linguistics patterns [20] using regular expressions. Finally, and with the tool AllTheWeb [6], the answer is validated by searching on the Web [12].

G. FQA System

A question-answering system for French was proposed by Laura Perret [5] in his doctoral thesis. This system is based on the combination of a probabilistic model of information retrieval standard for the detection of interesting documents with an approach based on linguistic parsing texts for matching. The research model probabilistic Okapi chosen does not document that contains all the keywords of the query to be selected. It is capable of tracking a document that contains a portion of the query terms by giving more weight to rare terms in the collection. It exploits syntactic analyzer (FIPS French Interactive Parsing System) developed in LATL[7] of University of Geneva [8]. This analyzer capable of associating with each sentence a syntactic structure reflecting the main syntactic properties of the sentence. It is based on grammars chomskyennes inspired theories [1].

V. SUMMARY

TABLE I

COMPARATIVE STUDY FOR SOME QAS FROM THE QUESTION ANALYSIS MODULE

QAS	Question Analysis
Multi-stream architecture	using a knowledge base, POS tagger or a named entity recognizer.
Semantic networks InSicht	syntactic analysis and semantic issues and documents.
Logical reasoning: Logical form of question and answer	Type and terms of question, response type: identified by a named entity recognizer or semantic hierarchy of concepts in WordNet.
Extraction patterns: as expressions regular	The type of question and select patterns corresponding to implement.
QUALIFIER	Extract named entities from question to build an ontology of classes. Question extended by Web and lexical resources ex. WordNet.
Diogene	POS analyzer and type of expected response identified.
FQA System	Eliminate words (prepositions, auxiliaries, conjunctions, articles). enracineur uses to remove suffixes. Parsing to extract the relevant terms and determine the type of response expected.

TABLE II

COMPARATIVE STUDY FOR SOME QAS FROM THE DOCUMENTS PROCESSING MODULE

QAS	Documents Processing	
	Documents Research	**Passages selection**
Multi-stream architecture	Matching with the generation of regular expressions.	
Semantic networks InSicht	question has several forms, each form comparative forms of candidate documents	That match: a response generated from a semantic network of the document.
Logical reasoning: Logical form of question and answer	Uses terms of the question to detect candidate passages.	
Extraction patterns: as expressions regular	Each boss has a general semantics applied to passages. Matching (pattern and applied elements of the basis of factual knowledge).	
QUALIFIER	IRS: Managing Gigabytes to detect responses using terms of the question extended.	Relevant sentences selected documents are classified according to association rules
Diogene	IRS: Managing Gigabytes (Boolean model).	
FQA System	Decompose the document into paragraphs. Use the Okapi probabilistic model (by selecting the first 10 documents).	Parser applied (FIPS) Break the paragraph into sentences.

TABLE III

COMPARATIVE STUDY FOR SOME QAS FROM THE ANSWER EXTRACTION MODULE

QAS	Answer Extraction
Multi-stream architecture	Selection of answer based on proximity or technical n-grams.

[6] Voir le site http://www.alltheweb.com/

[7] LATL : Laboratoire d'Analyse et de Technologie du Langage ; http://www.latl.unige.ch/

Semantic networks InSicht	Two factors: length and frequency of occurrence of the response in the collection.
Logical reasoning: Logical form of question and answer	applying a named entity extractor, or logical theorem (axioms) for representation of the join and the question of the response
Extraction patterns: as expressions regular	If match: a response generated (corresponding element).
QUALIFIER	Answer extracted named entity recognition, anaphora resolution and justification of the answer.
Diogene	Named entity recognizer and linguistic pattern using regular expressions.
FQA System	Decomposing fragments words. Calculating the frequency of occurrence of each word and ranking in descending order of frequency.

VI. CONCLUSION

The main objective of a question-answering system is to return a precise answer to a question formulated in natural language rather than complete documents as most information retrieval systems. This paper presents a global description of question-answering systems, as well as some approaches used in existing systems, such as based on a multi-stream architecture, semantic networks and logical reasoning, ... It also presents a comparison between these different types of systems on three levels, namely: question analysis module, document processing and extraction at the relevant answer. In perspective, a more detailed question analysis can lead to improved performance of these systems such as the introduction of constraints depending on the question to clarify and limit the area of the better looking of the relevant answer.

REFERENCES

[1] Chomsky, N.: The Minimalist Program, MIT Press, Cambridge (1995).

[2] Embarek, M.: Un système de question-réponse dans le domaine médical : le système Esculape. Thèse doctorat, université Paris-Est (2008).

[3] Helbig, H., Carsten, G.: Multilayered extended semantic networks as a language for meaning representation in NLP systems, In Computational Linguistics and Intelligent Text Processing (CICLing 2003), vol. 2276, pp. 69-85. Springer, (2002).

[4] Jijkoun, V., Mishne, G., de Rijke, M.: How frogs built the Berlin Wall, In Proceedings CLEF 2003, LNCS, Springer, (2003).

[5] Laura, Perret. : Extraction automatique d'information : génération de résumé et question-réponse. Thèse doctorat, université de Neuchâtel, Suisse (2005).

[6] Leacock, C., Chodorow, M.: WordNet : An Electronic Lexical Database and Some of its Applications, pp. 265-283. MIT Press, Cambridge (1998).

[7] Laurent, G., Patrice, B., Marc, El-Bèze. : Quelles combinaisons de scores et de critères numériques pour un système de Questions-Réponses ?, TALN 2008, Avignon (2008).

[8] Laenzlinger, C., Wehrli, E.: FIPS : Un analyseur interactif pour le français, In TA Informations, 32(2), pp. 35-49. (1991).

[9] Moldovan, D., Clark, C., Harbagiu, S., Maiorano, S.: Cogex : A logic prover for question answering, In Proceedings of the Human Language Technology and North American Chapter of the Association for Computational Linguistics Conference, pp. 87-93. (2003).

[10] Mitkov, R.: The Oxford Handbook of Computational Linguistics, Oxford University Press, (2003).

[11] Magnini, B., Negri, M., Prevete, R., Tanev, H.: A WordNet-Based Approach to Named Entities Recognition, In Proceedings of SemaNet02, COLING Workshop on Building and Using Semantic Networks, pp. 38-44. (2002).

[12] Magnini, B., Negri, M., Prevete, R., Tanev, H.: Comparing Statistical and Annual Content-Based Techniques for Answer Validation, In Proceedings of the 40th Meeting of the Association for Computational Linguistics (ACL-2002), pp. 425-432. (2002).

[13] Monz, C., de Rijke, M.: Tequesta: The University of Amsterdam's textual question answering system, In Proceedings of The Tenth Text REtrieval Conference (TREC-2001), NIST Special Publication 500-250, pp. 519-528. (2002).

[14] Magnini, B., Vallin, A., Ayache, C., Erbach, G., Penas, A., de Rijke, M., Rocha, P., Simov, K., Sutcliffe, R.: Overview of the CLEF 2004 Multilingual Question Answering Track, In Working Notes for the CLEF 2004 Workshop, pp. 281-294. Pisa, Italy (2004).

[15] Negri, M., Tanev, H., Magnini, B.: Bridging Languages for Q-A: DIOGENE at CLEF 2003, In Working Notes for the CLEF 2003 Workshop, pp. 321-329. Pisa, Italy (2003).

[16] Neumann, G., Xu, F.: Mining answers in German web pages, In Proceedings of the International Conference on Web Intelligence (WI-2003), pp. 125-131. (2003).

[17] Ravichandran, D., Hovy, E.: Learning surface text patterns a Question for Answering System, In Proceedings of the 40th Annual Meeting of the ACL-2002, pp. 41-47. (2002).

[18] Soubbotin, M. M., Soubbotin, S. M.: Use of patterns for detection of likely answer strings: A systematic approach, In Proceedings of the Eleventh Text REtrieval Conference (TREC-2002), NIST Special publication # 500-251, pp. 325-331. (2003).

[19] Scholbach, S., Olsthoorn, M., de Rijke, M.: Type checking in open-domain Q-A, In Proceedings of the 16th European Conference on AI (ECAI 2004), pp. 398-402. (2004).

[20] Tanev, H., Kouylekov, M., Negri, M., Coppola, B., Magnini, B.: Multilingual pattern libraries for Question Answering: A case study for definition questions, In proceedings of the 4th International Conference on Language Resources and Evaluation, Lisbon (2004).

[21] Voorhees, E. M., Haman, D.: TREC Experiment and Evaluation in Information Retrieval. Chapter 10. pp. 233-257. MIT Press, Cambridge (2005).

[22] Voorhees, E. M.: Overview of the TREC 2001 Question Answering Track, In Proceedings of the Tenth Text REtrieval Conference (TREC-2001), NIST Special Publication # 500-250, pp. 42-51. (2002).

[23] Voorhees, E. M.: Overview of the TREC 2002 Question Answering Track, In Proceedings of the Eleventh Text REtrieval Conference (TREC-2002), NIST Special Publication # 500-251, pp. 57-68. (2003).

[24] Voorhees, E. M.: Overview of the TREC 2003 Question Answering Track, In Proceedings of the Twelfth Text REtrieval Conference (TREC-2003), NIST Special Publication # 500-255, pp. 54-68. (2004).

[25] Witten, I. H., Moffat, A., Bell, T.: Managing Gigabytes: compressing and Indexing Documents and Images (second ed.), Morgan Kaufmann Publishers, New York (1999).

[26] Yang, H., Cui, H., Maslennikov, M., Qiu, L., Kan, M. Y., Chua, T. S. : QUALIFIER In TREC-12 Q-A Main Task, In Proceedings of the Twelfth Text REtrieval Conference (TREC-2003), NIST Special Publication # 500-255, pp. 480-488. (2004).

Ombres douces avec source de lumière surfacique

Abd El Mouméne Zerari[1], Mohamed Chaouki Babahenini[2], Naima Bahi[3], Sarra Khemliche[4]

[1,2,3,4] *Département d'informatique, Laboratoire LESIA, Université Med Khider Biskra, 07000, Algérie*

[1]a.zerari@live.fr

[2]chaouki.babahenini@gmail.com

[3]naima.bahi@yahoo.fr

[4]sarakh19@gmail.com

***Résumé*— Dans cet article nous allons présenter l'implémentation des ombres douces, qui sont généré par une source de lumière surfacique. Une approche proposé sur la subdivision adaptative de source de lumière avec l'utilisation de la technique de Cascaded Shadow MapsCSM [1] pour diminuer le temps de calcul. L'idée principale est de commencer avec quelques échantillons sur la zone de la lumière, et on ajoutant plus de points d'échantillonnage si nécessaire. L'utilisation des shadow Maps [2] ou des volumes d'ombre aboutissent à la génération d'ombres en temps réel, bien que les volumes d'ombre exigent des modèles polygonales fermés et convexes, c'est-à-dire que la scène doit être bien modélisées préférables, par contre les shadows maps cartes d'ombre fonctionnes toujours avec n'importe quelles géométrie scène quelconque, donc l'utilisation des shadow maps est le plus exigé en temps réel surtout pour la génération d'ombres douces. Donc nous choisissons d'utilisé la méthode des cartes d'ombres pour générer les ombres. Les avantages de cette méthode sont de pouvoir facilement générer des ombres douces en temps réel.**

***Mots clés* — Ombre douce, Filtrage, rendu 3D réaliste, GPU.**

I. INTRODUCTION

Les ombres et l'ombrage représentent un domaine particulièrement actif dans la 3D. L'ombrage est au cœur du réalisme d'une scène 3D. Les ombres résultent d'une interaction entre la lumière et les objets, et nous permettent de capter la profondeur de la scène et surtout confèrent un poids aux objets 3D. Les ombres se classent en deux grandes catégories: les ombres douces et les ombres dures.

Dans la réalité, les ombres dures n'existent pas, seules les ombres douces ont une existence réelle. Produire des ombres douces est un sujet traité depuis longtemps en synthèse d'image, elles se caractérisent essentiellement par leur réalisme mais sont relativement complexes à reproduire en 3D temps réel. L'ajout de la contrainte temps réel est relativement récent, il est en fait directement lié aux possibilités de programmation des cartes graphiques de dernières générations.

L'objectif de notre travail est de présenter les effets que peut produire une source lumineuse surfacique, nous devons ensuite mettre en avant les problèmes liés aux algorithmes d'ombres douces temps réel, pour ce fait en utilisant les GPUs.

II. PROBLEMATIQUE GENERALE ET MOTIVATIONS

Les techniques de génération des ombres douces en temps réel introduisent un certain nombre de difficultés par rapport à leur visualisation. Le problème principal réside dans l'interactivité car il faut calculer en temps réel le rendu de plusieurs images. En général, ces calculs dépassent les capacités de traitement des cartes graphiques actuelles dans le cas des scènes complexes qui contiennent un nombre important de polygone. Les méthodes ainsi basées sur des maillages produisent des résultats assez satisfaisants, en termes de qualité et de performances, mais se montrent globalement complexes et très lourdes à mettre en place, en ayant, par exemple, recours à des longues phases de pré-calculs et calculs implémenté sur le CPU, comme pour la génération de plusieurs shadow map afin de générer l'ombre douce. Pour résoudre ce problème, il est possible de faire appel aux récentes innovations dans le domaine du matériel graphique afin d'optimiser le calcul des d'ombres douces. Un autre problème réside dans la qualité des ombres et s'ils sont physiquement correct ou non. Des ombres douces physiquement corrects sont générer à partir des sources de lumières surfacique.

Pour ce fait, il faut diminuer le nombre d'échantillon de la source de lumière pour engendrer des ombres douces physiquement corrects, et utiliser la technique de CSM pour réduit le temps de calcul. En s'appuyant sur la nouvelle génération des GPU, qui permettent, pour chaque pixel affiché à l'écran, de réaliser des opérations programmables, accroissant ainsi le réalisme en minimisant le coût.

III. MODELE PROPOSE

Dans ce travail, nous proposons une nouvelle approche basée sur l'idée d'échantillonner la source lumineuse plusieurs fois afin d'obtenir les ombres douces physiquement correcte et en utilisant la technique de *Cascaded Shadow Maps* pour accélérer le temps de calcul. Il faut optimiser le nombre de points d'échantillonnage nécessaire pour avoir des résultats satisfaisants en commençant par seulement très peu de points d'échantillonnage et en ajoutant de manière adaptative de plus en plus d'elles, selon, que la densité de d'échantillonnage soit assez élevé ou pas [3].

A. Estimation des ombres douces avec l'échantillonnage de la source lumineuse

Une source lumineuse surfacique peut être approximer par n différents échantillons de source de lumière ponctuelle, donc il faut produire n shadow map pour générer une ombre douce[3].

$$\tau_i(x,y) = \begin{cases} 0 & : \text{éclairé par la lumière ponctuelle i} \\ 1 & : \text{Dans l'ombre de la lumière ponctuelle i} \end{cases}$$

τ_i x,y est le résultat du test de l'ombre dur pour la carte d'ombre i pour le fragment de l'espace écran à la position x, y. En supposant que l'échantillonnage du point sur la source de lumière surfacique est assez dense c'est à dire, n est suffisamment élevée, le résultat de l'ombre douce Ψ psi) peut être estimée par la proportion Ψn d'échantillons d'ombre [3].

$$\hat{\psi}_n(x,y) = \frac{1}{n}\sum_{i=1}^{n} \tau_i(x,y). \qquad 1)$$

B. Évaluation de subdivision

La comparaison de quatre cartes d'ombre voisines dans l'espace de la caméra est faite dans un fragment shader en appliquant une stratégie de 2 passages:

1 *Dans le premier passe:* Les valeurs qui représentent la profondeur des quatre cartes d'ombre sont évaluées comme dans l'algorithme original de carte d'ombre : Pour chaque fragment de l'espace d'écran, les 4 valeurs correspondantes d'ombre sont calculées et additionnées c. -à-d. chaque fragment obtient une valeur de nombre entier entre 0 et 4, et enregistré dans la cible du rendu de comparaison.

2 Dans le deuxième passe: Les cumuls stockés d'ombre sont calculer et afficher sur écran.

C. Produire des points d'échantillonnage additionnels

Si l'évaluation de subdivision propose de créer encore un autre niveau de raffinement sur la source lumineuse surfacique, de nouveaux points d'échantillonnages et les cartes d'ombre correspondantes doivent être créés. Dans le cas d'une source lumineuse surfacique rectangulaire bidimensionnel, nous proposons d'employer un quadtree comme structure: Si les étapes de comparaison rendent une subdivision nécessaire, le rectangle est coupé en 4 secondaire-quadruples, et de nouvelles cartes d'ombre sont produites sur tous les nouveaux coins voir figure 1.

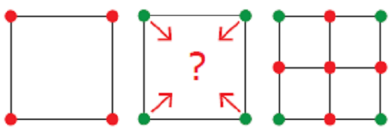


Fig. 1 Subdivision d'une source lumineuse surfacique rectangulaire

D. Attribuer des poids de contribution de carte d'ombre

Si toutes les cartes d'ombre produites avec notre stratégie de raffinement contribuent à la solution finale d'ombre douce avec le même poids, l'obscurité de la pénombre peut parfois varier un peu de la solution exacte, si la distribution des points d'échantillonnages adaptative choisis varie considérablement. Nous mettons donc des poids sur les points échantillonnés : Dans les surface trop subdivision, les différents échantillons sont attribuées à un plus petit poids, et ne contribueront pas autant à l'obscurité de la pénombre que celui avec un grand poids.

En cas de source lumineuse surfacique qui est subdivisée comme proposé dans la section B., le poids ωi attribués à la carte d'ombre i^{éme} est calculé avec [3]:

$$\omega_i = \frac{1}{(2^d+1)^2}, \qquad 2)$$

où **d** est la profondeur de subdivision. La somme de tous les poids est 1 si tous les échantillons atteignent la même profondeur de subdivision **d**. Autrement, les poids doivent être normalisés pour s'assurer que les cumuls finaux d'ombre se trouvent entre 0 entièrement visible et 1 entièrement ombr é.

E. Cascaded Shadow Maps

Le concept de base de Cascaded Shadow Maps CSM est facile à comprendre. Différentes régions de la pyramide de vue de la camera exigent des cartes d'ombres avec des résolutions différentes. Les Objets les plus proches de l'œil nécessitent une résolution plus élevée que ne le font des objets plus éloignés [1].

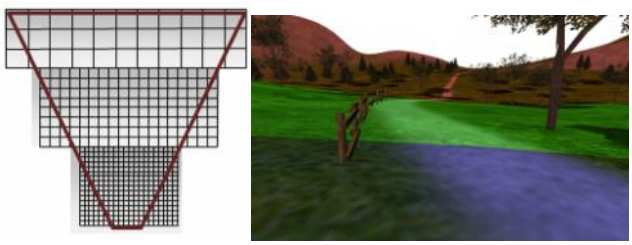


Fig. 2 Qualité du CSM

L'idée de base de CSM est de partitionner la pyramide de vue frustum en multiples sous -frustum. Une carte d'ombre est rendu pour chaque sous-frustum; le fragment shader puis des échantillons de la carte qui correspond le mieux à la résolution souhaitée figure 2.

F. Shadow map rapide

Nous avons pu optimiser l'algorithme du shadow map, en utilisant le FBO Frame Buffer Object) [4] pour éviter le transfert GPU/CPU. Les FBOs ont l'avantage de permettre de conserver et d'utiliser les informations calculées par le GPU dans la mémoire du GPU. En effet, avec les anciennes méthodes, il fallait faire des allers-retours entre le programme OpenGL et le GPU. Maintenant, les données restent sur le GPU, les programmes ne subissent plus des latences à cause du bus de données.

Le shadow mapping est une technique permettant de calculer les ombres pour les objets d'une scène. La technique est décomposée en deux passes: le calcul des ombres puis le rendu proprement dit. Pour le calcul des ombres, il faut placer la caméra à l'emplacement de la lumière, l'orienter avec la direction de la lumière, faire un rendu de la scène et ne garder que la profondeur. Celle-ci est conservée dans l'espace homogène. Grâce aux FBO ce résultat est stocké dans une texture, appelée shadow map. Nous devons créer la shadow

map pour chaque source, afin de connaître les ombres pour correspondante à chaque lumière.

Dans la seconde passe, la scène est rendue normalement, sauf que dans son fragment shader, on calculera si la distance entre le point actuel et la lampe est supérieur à la valeur retournée par la shadow map pour la même position. Si la valeur de la texture est inférieure, cela indique qu'il y a un objet entre la lampe et ce pixel, donc, qu'il doit être ombré.

G. La problématique de l'éclairage en temps réel

Dans les jeux modernes, il est très courant que les objets 3D soient éclairés par plusieurs lumières à la fois. Sauf qu'avec un rendu classique Forward Rendering, l'éclairage multiple est assez lourd au niveau des ressources.

Le rendu classique Forward Rendering : Le but est de dessiner un objet 3D à l'écran, celui-ci étant illuminé par plusieurs sources lumineuses. Il y a deux façons d'effectuer cette tâche avec le Forward Rendering:

La première façon est le Single-Pass Lighting, ou éclairage en simple passe. Ici, chaque objet est dessiné, et toutes les lumières affectant cet objet sont calculés dans un seul shader. Mais ce n'est pas sans inconvénients: vu que tout se passe dans un seul shader, on devient vite limité en nombre d'instructions et donc en nombre de lumières qui peuvent affecter cet objet 8 lumières maximum avec la version 3.0 du Shader Model). On utilise le Single-Pass Lighting dans des environnements nécessitant peu de lumières.

La seconde méthode est le Multiple-Pass Lighting. Ici, on prend un peu le problème à l'envers. Pour chaque lumière et pour chaque objet de notre scène. On a résolu le problème du nombre d'instruction et du nombre de lumières, puisque théoriquement on peut en dessiner à l'infini, sauf qu'un autre inconvénient majeur apparaît: le nombre de batch explose. Dans le pire des cas, ce nombre sera égal à N objets x N lumières, ce qui est inadmissible dans un jeu professionnel, sans parler des performances qui risquent d'être pitoyables.

C'est là que les *Texture Arrays* [5] rentrent en jeu. L'introduction de ce qu'on appelle *Texture Arrays* dans de plus nouveaux api de graphiques permet d'envoyer jusqu'à 2048 textures avec les mêmes tailles et le format au shader, où elles peuvent être échantillonnées. Cette fonctionnalité est parfaitement adaptée à nos buts, car elle nous permet d'échantillonner en même cas de nombreuses cartes d'ombre dans le Fragment shader. La valeur de l'occlusion du fragment courant peut donc être obtenue sans avoir besoin de passes supplémentaires, et plus de problème de limite en nombre d'instructions dans le même Shader.

IV. RESULTATS

La figure 3 à gauche illustre le rendu du shadow map avec 4 sources de lumière, à droite avec 289 sources de lu mière. La figure 4 démontre la variation du nombre de carte d'ombre en utilisant relation 2^d+1 où d représente la profondeur de subdivision, la valeur du d dans notre cas varie de 0 jusqu'à 5. Le graphique de la figure 5 montre les valeurs des FPS en fonction de nombre de triangles dans la scène. Cette même scène est rendue en utilisant deux méthodes différentes, la première avec un rendu classique du shadow Map, la deuxième avec un rendu en utilisant les FBOs, donc l'introduction des FBOs accélère le rendu.

Fig. 3 Ombre douce avec OGRE 3D.

Fig. 4 Ombre douce en variant le nombre de carte d'ombre.

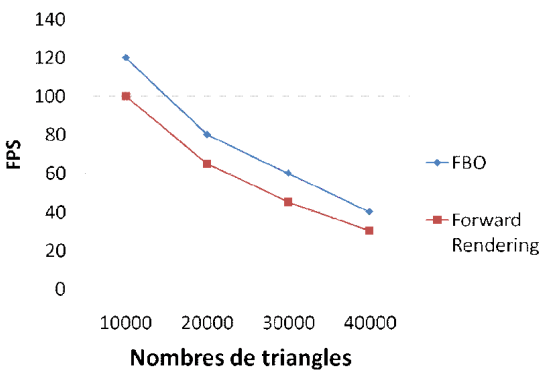


Fig. 5 Rendu du shadow map

V. CONCLUSION

Nous avons présente une amélioration d'une méthode d'ombres douces avec source de lumière surfacique basée sur le GPU, capable d'effectuer des rendus en temps réel.

L'intégration des FBOs pour accélérer le calcul a montré son efficacité dans l'optimisation du temps de calcul, ce qui nous a permis de générer des scènes en temps réel. La contribution que nous avons présentée dans ce papier nous a permis d'éviter le transfert GPU/CPU, donc effectuer les traitements dans la mémoire du GPU au lieu de les effecteurs dans la RAM.

VI. PERSPECTIVES

Comme futures travaux, nous envisageons d'utiliser les Textures Array pour diminuer le nombre de passe de rendu, et implémenter notre stratégie en utilisant la technique de CSM, dans le but de minimiser le temps de calcul. Afin d'évaluer rapidement le besoin de subdivision, nous exploitons la fonctionnalité du *hardware occlusion queries* [7, 8], qui sont généralement utilisées pour évaluer la visibilité en comptant le nombre de Pixel dessinés sur l'écran. En rejetant tous les fragments pour lesquels le niveau de subdivision est suffisant, les fragments restants peuvent être efficacement calculés. Si au moins un Pixel est produit, la source lumineuse surfacique a besoin davantage de d'échantillonnage dans ce frame.

Si peu de cartes d'ombre sont produites, les bandes d'artéfacts vont devenir visibles. Afin d'améliorer la douceur des transitions entre les cartes d'ombre, Nous devons proposer donc de les échantillonner en utilisant un petit filtre PCF Percentage Closer Filtering [6] ou autre filtre, dans de telles situations. Le filtrage de PCF adoucie les frontières d'ombre.

REFERENCES

[1] ENGEL, W. 2007. Cascaded shadow maps. In ShaderX5: Advanced Rendering Techniques,W. Engel, Ed. Charles River Media, 197–206.

[2] Lance Williams: Casting curved shadows on curved surfaces. In Proc. SIGGRAPH 78, volume 12, pages 270–274, août 1978.

[3] Michael Schwärzler, Oliver Mattausch, Daniel Scherzer, and Michael Wimmer. " Fast Accurate Soft Shadows with Adaptive Light Source Sampling". Workshop on Vision, Modeling, and Visualization, page 39-46. Eurographics Association 2012.

[4] Opengl Working Group: OpenGL FrameBuffer Object Extension. https://www.opengl.org/registry/specs/ARB/framebuffer_object.txt. Septembre 23, 2013.

[5] OPENGL WORKING GROUP: OpenGL Texture Array Extension. www.opengl.org/registry/specs/EXT/texture_array.txt, September 2008.

[6] REEVES W. T., SALESIN D. H., COOK R. L.: Rendering Antialiased Shadows with Depth Maps. In Proceedings of the 14th annual conference on Computer graphics and interactive techniques 1987, ACM Press, pp. 283–291. 2.

[7] BARTZ D., MEISSNER M., HÜTTNER T.: Extending Graphics Hardware for Occlusion Queries in OpenGL. In Proceedings of the ACM SIGGRAPH/EUROGRAPHICS workshop on Graphics hardware New York, NY, USA, 1998, HWWS '98, ACM, pp. 97 –ff.

[8] OPENGL WORKING GROUP: OpenGL Occlusion Query Extension. http://www.opengl.org/registry/specs/ARB/occlusion_query.txt, Apr. 2007.

Un Nouveau Modèle De Négociation Adaptatif Des Systèmes Multi-Agents, Application Dans Le E-Commerce

TOUMI Chahrazad
Département des TLSI, *Faculté des NTIC*
Université Constantine 2
Guelma, Algerie
Toumi.chz@gmail.com

MAAMRI Ramdane
Département des TLSI, *Faculté des NTIC*
Université Constantine 2
Constantine, Algerie

***Résumé* : La négociation est une forme importante d'interaction et de résolution de conflit dans les systèmes multi agents. Ce pendant la majorité des modèles présents sont plus au moins spécifique et non adaptatif aux différentes situations de négociation qu'on peut trouver. La complexité du processus de négociation augmente avec le nombre de participants impliqués et avec la complexité de l'objet de négociation (en termes de nombre et de nature). Le modèle proposé garantit une négociation entre multiples parties avec multiples objets à multiples attributs. On considère que les agents clients appartiennent au même ensemble (groupe, coalition, SMA) et que les agents serveurs peuvent appartenir à des groupes différents.**

***Mots clef* : Systèmes multi-agents, négociation automatique, E-Commerce, modèle de négociation.**

INTRODUCTION

La négociation est un domaine d'étude vaste et plus au moins ancien et abordée dans plusieurs disciplines telles que : les sciences de l'aide à la décision, l'intelligence artificielle distribuée (IAD), …etc. On peut la définir, selon [3], comme : une discussion dans laquelle des individus intéressés échangent des informations et arrivent à un accord commun. En informatique, et dans les systèmes multi agents, la négociation est automatique entre les agents, et est une forme fréquente d'interaction entre eux. On peut dire que la négociation est un processus qui permet de résoudre les conflits entre deux ou plusieurs agents. Elle consiste à aboutir à un accord, le plus commun qui satisfait aux plus les membres en conflits sur un objet (ressource, service, bien ...). Les agents d'un même SMA ou une coalition, coopèrent (généralement) vu qu'ils partagent tous un intérêt commun, il leur arrive de négocier une ressource d'une façon compétitive tout en prenant en considération cet objectif. Entre différents systèmes, les agents ne coopèrent pas et ils n'ont aucun intérêt commun, la négociation entre eux est compétitive.

Dans toute négociation (commerciale ou non) il y a toujours un demandeur dit "client", un fournisseur dit "serveur", "l'objet" sur quoi porte la négociation et des "règles" organisant les interactions entre eux. Dans un système multi-agent, la négociation prend en compte des aspects principaux suivants : l'objet, qui peut être simple ou composé (ensemble d'objets), où chaque objet peut avoir un seul ou plusieurs attributs. En fonction de l'objet en trouve des négociations "mono attribut" c'est généralement le prix est l'attribut le plus souvent négocié ou "multi attribut" où plusieurs attributs sont associés à un objet comme la qualité, la quantité, le délai, couleur, les pénalités...etc. Un protocole qui définit les participants possibles dans la négociation, ainsi que les règles qui gèrent les interactions entre eux. En fonction du nombre des participants à la négociation, on trouve : une négociation entre deux parties (bilatérale), une négociation un à plusieurs, une négociation plusieurs à plusieurs ou multilatérale et enfin une négociation multi-bilatérale. Et un processus de décision qui revient à répondre à la question : "que dois-je faire maintenant ?" qui définit quelle stratégie appliquée et à quel moment de la négociation.

La complexité du processus de négociation augmente avec le nombre de participants impliqués et avec la complexité de l'objet de négociations (en termes du nombre et de nature). Aussi, on veut prendre en considération le cas où un ensemble d'agents (groupe, coalition, SMA) partagent un même intérêt veulent négocier le même objet de négociations. Et plus le processus est complexe, plus il est lourd. De ce point de vue, un modèle de négociations avec « multiples parties - multiples objets - multiples attributs » tout en respectant la nature de l'objet et de ses attributs est très complexe et difficile à mettre en œuvre. Cependant, pour réaliser un modèle de négociation adaptatif, il faut prendre en considération tous les cas possibles de la négociation et essayer de faire en sorte que le modèle s'y adapte au plus nombre de cas possibles. L'objectif de notre

travail est de présenter un tel modèle tout en essayant de le simplifier au maximum.

Dans la section suivante on présente un état de l'art sur quelques travaux existants dans le domaine. La section qui suit, on présente notre approche proposée : modèle de négociation : le protocole, l'objet de négociation et le processus de décision. La section trois présente un exemple d'application du modèle. Et enfin, dans la dernière section une conclusion sur notre travail.

TRAVAUX CONNEXE

Plusieurs modèles de négociation et mécanisme de décision ont été proposés pour la négociation automatique des systèmes multiagents. La théorie des jeux est utilisée comme stratégie dans, elle qui vise à trouver une solution optimale entre les participants en analysant les interactions entre eux et en trouvant un équilibre [5] et [6]. Le protocole « *Contract net* » proposé par [3] une solution au problème d'allocation de tâches qui reprend le fonctionnement du marché. Plusieurs domaines utilisent la négociation automatique entre agents, cependant le E-Commerce est le domaine où la négociation est le plus abordé. L'on trouve différent forme comme B2B (Business to Business), B2C (Business to Consumer) et C2C (Consumer to Consumer). Les enchères électroniques sont le type de négociation la plus utilisé pour le E-Commerce et surtout pour le B2C, les formes les plus connues sont les catalogues électroniques, la vente en ligne *ebay* [4]. Il existe des systèmes ou les agents assistés pour le E-Commerce comme *Kasbah* [8], the Michigan Internet AuctionBot [7] et MAGNET [9]. Le nombre de participants et celui des attributs de l'objet de négociation sont à aborder, [1] propose une approche basée heuristique avec multi partie et multi attributs. La stratégie de négociation comporte trois familles de tactiques basées sur : le temps, les ressources et le comportement en prennent en considération un seul critère à la foi.

Depuis, d'autres modèles bilatéraux et multilatéraux avec plusieurs attributs pour l'E-Commerce ont été proposés comme [10], [11] et [12]. Cependant ces modèles proposés sont spécifiques, ont des critères bien définis et ne peuvent pas s'adapter au dynamisme du changement de la négociation. [2] dans son travail propose un modèle de négociation adaptatif pour le E-Commerce, le modèle est utilisable et adaptable aux trois formes (B2B, B2C et C2C), cependant ce modèle et plus au moins spécifique pour le E-Commerce et supporte l'interaction entre plusieurs participants et plusieurs attributs, ce qui ne couvre pas les cas où l'objet est multiple.

PROPOSITION DE L'APPROCHE

Le choix de « multiple » dans notre modèle est pour garantir une flexibilité du modèle et pour couvrir au mieux les différentes situations de la négociation. La négociation se porte sur un objet pouvant être multiple avec plusieurs attributs sans aucune restriction sur leur nature. Pour la décision, on utilise une approche de négociation à base d'heuristique proposée par [1] et [2] en l'adaptant à notre modèle, avec une stratégie de négociation comportant plusieurs tactiques et prenant en compte plusieurs critères. On peut arriver à définir les éléments du processus de négociation dans notre comme suite :

Définition 1 **:** Un processus de négociation est composé d'un ensemble fini d'Agent « *Ag* », un objet de négociation «*Object Need* », des préférences relatives à l'objet de négociation « *IM* », un temps limite de négociation « *timef* », un ensemble de propositions « *P* » et un ensemble de contre-propositions « *CP* » :

$$\boldsymbol{Ag} = \{AgC, AgS, Co, AgN, ag\}$$

timef : définit le temps maximum que peut prendre la négociation, si l'agent n'a pas de contrainte sur le temps timef = 0 ;

$$\boldsymbol{Object\ Need} = \{Ob_1, Ob_2, \dots, Ob_f\} \backslash\ f \geq 1$$

$$\boldsymbol{P} = \{\Pr(ob_1), \Pr(ob_2), \dots, \Pr(ob_f)\}$$

$$\boldsymbol{CP} = \{\mathrm{CPr}(ob_1), \mathrm{CPr}(ob_2), \dots, \mathrm{CPr}(ob_f)\}$$

P ou CP est une suite des valeurs pour les attributs des objets à négocier noté « *Object Need* ».

Dans ce qui suit on détaillera notre modèle proposé, protocole, l'objet et le processus de décision.

Protocole de négociation

De la formule (1) définition précédente 'définition 1' on a cinq types d'agent

$\boldsymbol{Ag} = \{AgC, AgS, Co, AgN, ag\}$ où :

- **Co (Agent Coordinateur)** son rôle est de : i) définir l'objet de négociation Object Need ; ii) initialiser la négociation ; iii) décomposer ou regrouper les sous-parties négociables de l'Object Need ; iv) lancer l'appel à l'offre AO ; v) prendre la décision finale de la négociation ;
- **AgC (Agent Client) :** l'ensemble d'agent Client qui demande un service.

$$\mathrm{AgC} = \{\mathrm{AgC}^1, \mathrm{AgC}^2, \dots, \mathrm{AgC}^m\} \backslash\ \mathrm{m} \geq 1$$

- **AgS (Agent Serveur) :** l'ensemble des agents contractants ou participants à la négociation en réponse d'un appel à l'offre ;

$$\mathrm{AgS} = \{\mathrm{AgS}^1, \mathrm{AgS}^2, \dots, \mathrm{AgS}^n\} \backslash\ \mathrm{n} \geq 1$$

- **AgN (Agent négociateur) :** les agents négociateurs qui se chargent de la négociation avec le serveur. ;

$$\mathrm{AgN} = \{\mathrm{AgN}^1, \mathrm{AgN}^2, \dots, \mathrm{AgN}^n\}$$

- **ag :** l'ensemble des agents non spécifiques (agent ne participant pas à la négociation) ;

$$\mathrm{ag} = \{\mathrm{a}, \mathrm{d}, \ldots, \mathrm{c}\}$$

Dans notre protocole, on a huit (8) types de messages échangés entre les agents du système qui sont les suivants : appel d'offres AO (Call for proposal), Information sur besoin NEED, proposition, contre-proposition, confirmation, refus, acceptation, fin de négociation. La communication entre les agents s'effectue dans les deux sens, mais à tour de rôle entre le contractant et l'agent négociateur.

La négociation s'effectue principalement entre l'agent négociateur et l'agent serveur le déroulement de notre protocole est comme suite :

a) Tout agent communique ses besoins au coordinateur ;
b) Le coordinateur les analyse, définit l'objet à négocier «Object Need», il détermine si l'Object Need est décomposable où non et si décomposable, il définit en quelles parties.
c) Initialise la négociation et lance un appel à l'offre AO ;
d) Les agents souhaitant participer à la négociation font des propositions, soit pour l'ensemble de Object Need ou des parties de l'Object Need ; Si après un délai d'attente le coordinateur ne reçoit pas de proposition il revient à (3) ;
e) Après réception des propositions, le coordinateur les analyse ;
f) Si la proposition comporte tout Object Need alors il crée des instances de l'agent négociateur de telle sorte que chaque participant « serveur » négocie avec un agent négociateur (négociation bilatérale) et commence la négociation ;
g) Dans le cas contraire, le coordinateur vérifie s'il y a d'autres participants qui proposent la sous-partie restante (le but est de négocier la totalité de l'Object Need) ; si oui le coordinateur fait comme dans le premier cas et initialise un agent négociateur pour chaque participant, puis commence la négociation.
h) La négociation entre un agent négociateur et un serveur peut : ne pas aboutir à un accord dans ce cas, il termine la négociation et communique le résultat (par échec) au coordinateur ; ou aboutir à un accord qui sera communiqué au coordinateur ;
i) Tous résultats est enregistrée dans l'historique ;
j) Le coordinateur, à la réception les résultats des agents négociateurs, les analyse et choisi la meilleure (compte tenu des préférences initiales) ;
k) Le processus de négociation prend fin lorsque un contrat est accepté, tous les contrats sont refuser ou si le temps limite de la négociation se termine ;

L'objet de négociation

Si on prend l'exemple de la négociation dans le E-Commerce le prix d'une marchandise mise en vente est l'attribut le plus négocié, parfois même le seul attribut. Le temps de délivrance et la quantité peuvent être aussi des attributs importants pour le client, cela dépend des situations de la négociation. Vu que l'objet de négociation peut être selon sa nature : une ressource, un plan, un service, une tâche, un bien...etc. Et peut avoir des attributs qui peuvent être quantitatifs ou qualitatifs. En plus, c'est la valeur de ces attributs qui est négociée à travers des propositions ou contre-propositions. On a choisi dans notre travail de couvrir l'aspect nombre et nature de l'objet et de ces attributs et on a des hypothèses sur l'objet de négociation :

Hypothèse sur l'objet de négociation

- Tous les agents clients (*AgC*) veulent négocier le même objet *Object Need* ;
- L'*Object Need* peut être composé de 1 à plusieurs objets,
- Chaque objet, noté *Ob* de l'*Object Need* contient les mêmes attributs pour tous les agents ;
- Chaque *Ob* peut contenir de 1 à plusieurs attributs ;
- Le nombre d'attributs peut-être différent d'un *Ob* à l'autre ;
- Les préférences des *Ob* peuvent être différentes d'un agent à l'autre, il en est de même pour les préférences entre les attributs ;
- Les attributs peuvent être de nature qualitative ou quantitative ;

Object Need est formé à la réception des besoins des clients qu'on note *OBC*. Soit OBC^i l'ensemble des besoins *d'*$AgC^i \in AgC$. OBC^i est composé par un ensemble d'objets. Chaque objet *Obc* est composé à son tour par un ensemble d'attributs *att*. On modélise les besoins des agents clients *OBC* comme suite :

$$OBC^i = \{obc_1^i, obc_2^i, \ldots, obc_f^i\}$$

$$obc_f^i = \{att_{f1}^i, att_{f2}^i, \ldots, att_{fz}^i\}$$

Pour modéliser les préférences des agents par rapport un objet de négociation *ob*, une valeur numérique lui est attribuée pour déterminer son importance notée *Im* et celles qui sont entre les attributs d'un objet notée par *Pr*. *Im* et *Pr* sont modélisés comme suite :

$$Im^i = \{im_1^i, im_2^i, \ldots, im_f^i\}$$
$$Pr^i = \{pr_1^i, pr_2^i, \ldots, pr_z^i\}$$

Pour des raisons de normalisation, la somme des préférences est égale à 1

$$\sum_{1 \le j \le f} Im_j^i = 1 \;\; ; \;\; \sum_{1 \le l \le fz} Im_{jl}^i = 1$$

Le coordinateur forme *Object Need* suite aux OBC reçue. Les préférences entre les objets de. Pour un objet de l'*Object Need* son importance notée *IM* est la moyenne de la somme des importances des agents pour le même objet. Il est de même pour les préférences des attributs.

Les valeurs des attributs dans l'*Object Need* son déduit à partir des valeurs des attributs des *OBC*. Dans le cas d'attributs quantitatifs :

$$val(Att_z) \in \{MIN_z , MAX_z\}$$

Tel que : $MIN_z = \{\sum min_z\}$ et $MAX_z = \{\sum max_z\}\}$ Dans le cas d'attributs descriptifs, c'est l'ensemble de toutes les valeurs admissent pour tous les agents par rapport au même attribut :

$$val(Att_l) \in \{q_{l1}^1, q_{l2}^1, \dots, q_{ls}^z\}$$

Stratégie de négociation

La négociation s'effectue par échange mutuel de proposition *P* et de contre-proposition *CP*. Une proposition ou contre-proposition est une suite de valeurs des attributs pour les objets de la négociation. Comme déjà vu dans la *définition 1 P* et *CP* sont modélisés comme suite :

$$P = \{\mathrm{SPr}(Ob_1), \mathrm{SPr}(Ob_2), \dots, S\mathrm{Pr}(Ob_f)\}$$

$$CP = \{\mathrm{SCPr}(Ob_1), \mathrm{SCPr}(Ob_2), \dots, S\mathrm{CPr}(Ob_f)\}$$

$\mathrm{SPr}(Ob_1)\ et\ \mathrm{SCPr}(Ob_1)$ représente une sous-proposition et une sous contre-proposition pour l'objet Ob_1. Cette sous-proposition est une suite de valeurs des attributs de l'objet Ob_1. Leur succession à tour de rôle nous donne le vecteur de séquence de négociation à travers le temps « *t* » jusqu'à *timef*.

$$SPr(Ob) = \{\mathrm{Val}(Att_1), \mathrm{Val}(Att_2), \dots, \mathrm{Val}(Att_z)\}$$

$$\mathrm{SCPr}(Ob) = \{\mathrm{Val}(Att_1), \mathrm{Val}(Att_2), \dots, \mathrm{Val}(Att_z)\}$$

$$\boldsymbol{VS} = \{P_t, CP_{t+1}, P_{t+2}, \dots\}$$

Pour évaluer une proposition on utilise une fonction *Scoring function* présenter dans [2] qui attribue un score *V* valeur entre 0 et 1 à une valeur proposée *val*, V(val) → [0,1]. On calcule l'utilité *U* de chaque objet de cette proposition (contre-proposition) en tenant compte des préférences. Ensuite, on calcule l'utilité globale *UG* de la proposition.

$$\mathrm{UG\ (P)} = \sum_{1 \le i \le m} \mathrm{IM}^i \times (\mathrm{U}(Ob^i)/\mathrm{m})$$

$$\mathrm{U}(Ob^{i\prime}) = \sum_{1 \le i\prime \le m} PR^{i\prime} \times V(x^{i\prime})$$

$$\boldsymbol{VS} = \{P^t, CP^{t+1}, P^{t+2}, \dots, \boldsymbol{X}\}$$

Soit l'agent $a \in AgC$ et l'agent $b \in AgS$, on note $P_{a\to b}^t$ la proposition faite dans un temps « t » par l'agent 'a' à l'agent 'b'. Pour la contre-proposition qui suit dans le temps « t+1 » est $CP_{b\to a}^{t+1}$. $\boldsymbol{VS_{a\leftrightarrow b}}$ est le vecteur de séquence entre les agents « a » et « b ».

La fin du vecteur de séquence on la représente avec *X* tel que $X \in \{accepter, refuser\}$.

ETUDE DE CAS

Dans cet étude de cas, on suppose qu'on a deux agents client, que l'on note respectivement AgC1 et AgC2, qui appartient à un même groupe d'agents. En plus, on suppose que ces deux agents veulent satisfaire des besoins que l'on note respectivement OBC^1 et OBC^2. De ce fait, chacun de ces deux agents commence par communiquer : ses besoins, ses priorités ainsi que le temps limite durant lequel il attend un contrat au coordinateur, que l'on note CO, pour que ce dernier puisse former le besoin global que l'on notre *Object Need* ainsi que les priorités et les importances globales pour pouvoir effectivement commencer la négociation.

Dans la table suivante on va présenter les éléments de la négociation que possède chaque agent :

TABLE I. ELEMENTS DE NEGOCIATION POUR CHAQUE AGENT

Les agents	**Composante de la négociation**		
	Les objets	***Les priorités***	***Le temps***
AgC 1	OBC^1= {table (prix[10,20],quantité[5,8]), armoire(prix[20,30],quantité[2,3])}	Pr1={(0.7, 0.3),(0.8,0.2)} I^1={0.6 ,0.4}	*Timef1 = 15*
AgC 2	OBC2= {table (prix[8,18],quantité[5,8]), armoire(prix[20,30],quantité[3,4])}	Pr^2={(0.4,0.6) , (0.7,0.3)} I2={0.8, 0.2}	*Timef2 = 15*
CO	Object Need = { table(prix[18,38],quantité[10,16]), armoire(prix[40,60],quantité[5,7]}	PR={(0.55,0.45),(0.75,0.25)} IM ={0.7,0.6}	*Timef= 15*

Après formalisation de *l'Object Need* et les préférences, le coordinateur lance un appel à l'offre $AO = \{(18,16), (40,7)\}$. Ensuite, le coordinateur reçoit deux propositions P^1 et P^2 des participants AgS1 et AgS2, où P^1={(20 , 13),(45, 5)} et P^2={(20 , 10),(60, 7)} . A la réception le coordinateur initialise deux agents négociateurs AgN1 et AgN2. L'AgN1 négocie avec AgS1 et AgN2 négocie avec AgS2. Les agents négociateurs évaluent la proposition avant de la contre-proposer. L'évaluation des propositions et leur interprétation par les agents négociateurs ainsi que les contre-propositions faites sont présentées dans le tableau suivant :

TABLE II. PROPOSITIONS ET LEURS EVALUATIONS

Les agents	***Evaluation***	***L'interprétation***	***La contre-proposition***
AgS1	U^1(table) = 0.27 U^1(armoire) =0.44 $UG^1(P^1)$=0.7*0.27+0.3*0.44 =0.32	I(P1)=CP^1 (cotre proposé)	CP^1={(20,16),(43,7)}
AgS2	U^2(table)= 0.05 U^2 (armoire) =0,12 $UG^2(P^2)$=0.7*0.05+0.3*0.12 =0.03	I(P1)=CP^2 (cotre proposé)	CP^2={(23,16),(42, 7)}

Après 13 unités de temps les agents négociateurs reçoivent les propositions suivantes P^1={(27 , 16),(52, 7)} et P^2={(30 , 16),(46, 7)}. Après leur interprétation respective par les agents AgS1 et AgS2, les deux propositions sont à accepter. De ce fait,

ils envoient les contrats au coordinateur qui doit décider quel contrat accepter. Le coordinateur calcule les utilités globales des deux contrats gagnants le résultat est :

$UG(P^1_{13}) = (0,95*0,7) + (0,69*0,3) = 0,91$

$UG(P^2_{13}) = (0,78*0,7) + (0,8*0,3) = 0,78$

Après la comparaison entre $UG(P^1_{13})$ et $UG(P^2_{13})$, l'agent coordinateur choisit la proposition P^1_{13} de la part de l'agent AgS1. De ce fait il envoie un message de refus, et donc fin de négociation, à l'agent AgS2 et un message d'acceptation à l'agent AgS1.

CONCLUSION

Ce papier présente un modèle de négociation des systèmes multi-agents adaptative. Le protocole proposé garanti la négociation entre plusieurs agents. L'objet de négociation peut être un seul ou multiple, les attributs peuvent varier d'un seul à plusieurs attributs en ce qui concerne leur nombre et pour la nature il n'y a pas de restriction. Dans un travail futur on envisage d'utiliser les réseaux de neurones dans le processus de décision.

REFERENCES

*** *** *** *** *** *** *** *** ***

[1] H. Raiffa. The Art and Science of Negotiation. Harvard University Press, 1982.

[2] Wang, Gong, T. N. Wong, and Chunxia Yu. "A computational model for multi-agent E-commerce negotiations with adaptive negotiation behaviors." Journal of Computational Science 4.3 (2013).

[3] R.G. Smith, "The Contract Net Protocol: high-level communication and control in a distributed problem solver". IEEE Transactions on computers, C-29(12):1104-1113.

[4] eBay: http://pages.ebay.com.

[5] S. Kraus, Strategic Negotiation in Multiagent Environments, MIT Press, Cambridge, MA, 2001.

[6] M. Oprea, An adaptive negotiation model for agent-based electronic commerce, Studies in Informatics and Control 11 (2002) 271–279.

[7] P.R. Wurman, M.P. Wellman, W.E. Walsh, The Michigan Internet Auction-Bot: a configurable auction server for human and software agents, in:Second International Conference on Autonomous Agents, Minneapolis, 1998,pp. 301–308.

[8] A. Chavez, P. Maes, Kasbah: an agent marketplace for buying and selling goods, in: First International Conference on the Practical Application of Intelligent Agents and Multi-Agent Technology, London, 1996, pp. 75–90.

[9] J. Collins, W. Ketter, M. Gini, A multi-agent negotiation testbed for contracting tasks with temporal and precedence constraints, International Journal of Electronic Commerce 7 (2002) 35–57.

[10] A. Chavez, P. Maes, Kasbah: an agent marketplace for buying and selling goods, in: First International Conference on the Practical Application of Intelligent Agents and Multi-Agent Technology, London, 1996, pp. 75–90.

[11] M. Louta, I. Roussaki, L. Pechlivanos, An intelligent agent negotiation strategy in the electronic marketplace environment, European Journal of Operational Research 187 (2008) 1327–1345.

[12] Chen, Y. M., & Huang, P. N. (2009). Agent-based bilateral multi-issue negotiation scheme for e-market transactions. Applied Soft Computing,9(3), 1057-1067.

.

Building Adaptive Mobile Applications Using Composition Mechanism

Afrah Djeddar
LAMIS Laboratory
University of Tebessa, Algeria
afrah-djeddar@hotmail.fr

Hakim Bendjenna
LAMIS Laboratory,
University of Tebessa, Algeria
hbendjenna@yahoo.fr

Abdelkrim Amirat
LIM Laboratory,
University of Souk Ahras, Algeria
abdelkrim.amirat@yahoo.com

***Abstract*—Despite the plethora of available mobile applications (apps) with theirs different implementations forms (e.g. component, service, or app) the user's needs differ from user to another. Moreover, the mobile devices are characterized by heterogeneous software and hardware configurations. Thus, an important challenge in the development of mobile apps is their deployment in the heterogeneous devices available on the market. To tackle these challenges, there is a need for a composition process to reuse existing heterogeneous entities for developing mobile apps according to user's requirements and that the behavior of the desired apps can be customized according to theirs various context information. In this paper we address this issue by proposing context-aware mobile apps composition process based-on existing heterogeneous software entities. We follow a metamodeling approach by taking in consideration the quality of the composite apps and the cost of composition. This approach aims to make future internet provide better composed mobile apps for users with context-aware contents adapted to theirs requirements.**

***Index Terms*— Mobile apps, Composition process, Metamodeling, Heterogeneous software entities, Context-aware, Composition cost.**

INTRODUCTION

Mobile apps usage and development is experiencing exponential growth. According to Gartner, by 2016 more than 300 billion apps will be downloaded annually. The mobile domain presents new challenges to software engineering [1]. Mobile platforms are rapidly changing, including diverse capabilities as GPS, sensors, and input modes. For this, mobile apps need to be able to specialize their behavior according to the context information and thus must be Omni-Channel and work on all platforms. Moreover, despite the large number of available mobile apps, user's needs in the daily-life differ from each other. These issues give some need to a composition mechanism of mobile apps in order to achieve the desired functionalities while considering the different context information (i.e. software and hardware characteristics of mobile device). By using the composition mechanism we try to fulfill a key software engineering challenge: the reusability of existing entities [2].

To the best of our knowledge, there is no such composition mechanism of mobile apps which designed to consider various factors like: software entities kind, mobile devices characteristics, current state of mobile devices, quality of the generated app, and composition cost. Thus taking advantage from existing works, we present in this paper a composition process for composing context-aware heterogeneous mobile apps while considering the quality and the cost of composition. Thereby, the main objective of this paper is threefold: First, compose mobile apps using existing heterogeneous software entities by providing *heterogeneous composition process* in order to maximize the benefits from the existing functionalities and thus satisfy user's requirements. Second, consider the context information of the mobile device during the composition in order to develop mobile apps that can sense and adapt to their run-time environment. Third, composing mobile apps according to the minimum cost of composition while ensuring the efficiency and the good quality of the desired app. This work provides a first step towards a composition process of context-aware apps based-on heterogeneous entities by a metamodeling approach in mobile environment. The remainder of this paper is organized as follows: section 2 provides some related works. After, in section 3 we introduce a motivating scenario. While section 4 presents an overview of our composition process and describes its steps. Finally, section 5 draws some conclusions and perspectives of this research work.

RELATED WORKS

In order to meet the needs of users by reusing the existence entities, a lot of composition mechanisms were developed. In this paper, we consider that the composition mechanism can be classified in *Macro* and *Micro* composition. Where the first one concerns the composition of classical apps whereas the other one concerns the composition of mobile apps which need to be submitted and adapted to several constraints of execution environment. The notion of *Micro composition* refers to the obligation of composition with respect to limited resources offered by mobile devices and its execution state variability. In contrary, the notion of *Macro* reflects the composition in a wide sense (i.e. without any constraints).

In this section we describe several approaches that are focused on composition mechanism. *Chakraborty and al.* [3] describe some issues related to services composition in

mobile environment. They presented a distributed *Service Composition Protocol* for mobile environments that take into consideration mobility, dynamic changing service topology and device resources. They mainly concentrated on a distributed architecture to facilitate the task of composition but they haven't focused on the application layer nor on its adaptation capabilities. In [4], authors presented an automatic approach for web service composition, while addressing the problem of process heterogeneities and data heterogeneities by using a planner and a data mediator. On the contrary, we were intended in [5] to reify the relevant notions of existing mechanisms of composition in a composite service meta-model. This meta-model defines all interlaced features and provides a global and explicit vision of the service composition. Moreover, this approach allows the specification of the auto-composition process: the composite's ability to dynamically modify its architecture and its composition logics according to the environmental context. In [6], authors created a software infrastructure called AppSpotter that makes possible the dynamic and automated selection and composition of software components for building mobile apps. They proposed mechanism for selection of software components is based-on mobile device features. In [7], authors presented a model to design context-aware services that can be exploited as a flexible domain to automatically generate context-aware compositions by means a specific tool. Additional, the tool exploits an extension of the problem definition that includes also the context representation for extending the services. These multitude of approaches with theirs features often specialized does not have a global vision of mobile apps composition. Our research intends to clearly express the relevant notions of these existing mechanisms in *heterogeneous composition process* for mobile apps using meta-modeling approach. However, we find that the proposed approach will have some advantages with respect to the existing solutions in the area of composition mechanisms for mobile environment. This modest work represents the first attempt to support the development of context-aware mobile apps by composing heterogeneous software entities at architectural level. So, this composition subjected to the following specific constraints at architectural level: 1) Respect the limited resources offered by mobile devices in order to build adaptive mobile apps. 2) Reuse and compose any kind of software entities in order to take profit of existing entities regardless their implementation type. Therefore, our approach aims to address and solve the heterogeneity problem from the architectural perspective which proves the need to distinguish between: app, component, service. Moreover, another advantage of the proposed process is to compose mobile apps with a lower cost of composition while respecting the way of composition to ensure the reliability and efficiency of the future generated apps.

MOTIVATING EXAMPLE

In this paper, we based on the example of ShopReview App discussed in the work of Gianpaolo in [8] to explain the motivations behind our research work and the main encountered problems. ShopReview App provides the users with alternatives nearby places where the same product is sold at a more with convenient price when a user publishes the price of a product she found in a certain shop. Thus, it allows users to share their opinion (i.e. various information concerning the products and the shops) using social network such as Twitter. To perform this app following to the composition mechanism: Firstly, it is necessary to establish the list of the different desired functionalities and after to choose for each of them the existing concert software entity that will be used to implement it (cf Table 2).

Desired functionalities	Description	Selected concrete entities
Fun 1: Acquire Photo	Photo acquisition	Component1:(photo acquisition from mobile camera)
Fun 2: Read Barcode	Acquire the barcode of the product	Component2:(local barcode recognition)
Fun 3: Get product Name	Translate the barcode into product name	Service3:(HTTP:/www.searchupc.com/)
Fun 4: Input Price	Collects the product's price from the user	Component4:(textual input from the user)
Fun 5: Search the Web	Retrieves more convenient price offered on e- commerce sit	Service5:(HTTP:/www.kelkoo.it/)
Fun 6: Get Position	Retrieves the current user location	Component6:(user localization via GPS
Fun 7: Search the Neighborhood	Retrieves other nearby shops which offer the product at a more convenient price.	Service7:(HTTP:/Shopsavvy.mobi/)
Fun 8: Share Price	Found on a given shop on twitter.	Application8:(Twicca:HTTP://twicca.r246.jp./)

Table 2- Desired functionalities to compose *ShopReview* mobile app.

Then, to compose theses selected entities to obtain the desired composite mobile app as illustrated in Fig.1. Following to the selected concrete entities, the composite *ShopReview* mobile app needs to: GPS for executing the Fun 6 and Wi-Fi for Fun 5, Fun 7, and Fun 8. This example illustrates a first problem which is the definition of the collaboration between heterogeneous reused entities (e.g. compose *Component4* with *Service5*). None of existent composition approaches address this part. Typically, they define a composite app as a collection of software entities of the same kind using for example: SOA approach [9] or component-based-approach [10] or other paradigms [15]. Therefore, it does not such composition

process aims to specify this heterogeneous coordination. In the mobile environment, composition mechanism is intended to respect some constraints. It must consider the different settings of the mobile devices where the users will run the desired app. However, another problem that can arise is the adaptation with the mobile devices context information. Thus, all reused entities must be adaptable to the context information to ensure the correct functioning of the composed app. In our example, the composed app does not work with the mobile device which its courant context is: Has-GPS is GPS-disabled, Has-Wi-Fi is Wi-Fi-enabled, Has-Fixed-Focus-Camera, Battery Level is high, and Free Storage capacity is: 500 MB. Thus, the choice of *Component 2*, *Component 6* as concrete entities for implement the Fun 2, Fun 6 (i.e. without the respect of the courant context) limits the usability of the composed app. *Component 2* not allow a correct barcode recognition of an image acquired with fixed-focus camera and *Component 6* cannot be executed while the state of GPS is disabled.

In order to represent the desired composite mobile app (CMA) at architectural level we propose in Fig.4 the corresponding metamodel that provide a high-level description of the CMA architecture. It is based on the reification of the different properties and functionalities of existing composition mechanisms. We associate to each role or information a specific architectural element used to compose the desired mobile app. This metamodel aims at defining how the architectural elements of the composite app are related each other. While defining the CMA architecture, we focus firstly on the different functionalities of the desired app (i.e. customer requirements) where each of them refers to an abstract software entity, ignoring how they will be implemented (e.g. components, services, apps). Secondly, we focus on the detailed description of the desired app. We will be based on the choice of the implementation type of each abstract software entity (i.e. each functionality described in the first architecture) to construct this detailed architecture. This choice is delayed to the time when concrete software entities are selected and defined in step 4 of the proposed process. For this purpose, we need to the two following architectures for describe the CMA (cf Fig.4.b): 1) ***CMA abstract functional architecture:*** that represents a high-level description of the desired functionalities and theirs interconnections used to accomplish the composite's goal. 2) ***CMA abstract detailed architecture:*** that represents a high-level description of the different concrete software entities that will be used to implement the desired functionalities and the different *Mediators* that represents the collaboration between reused concrete entities and they will be used to eliminate the heterogeneity among these entities.

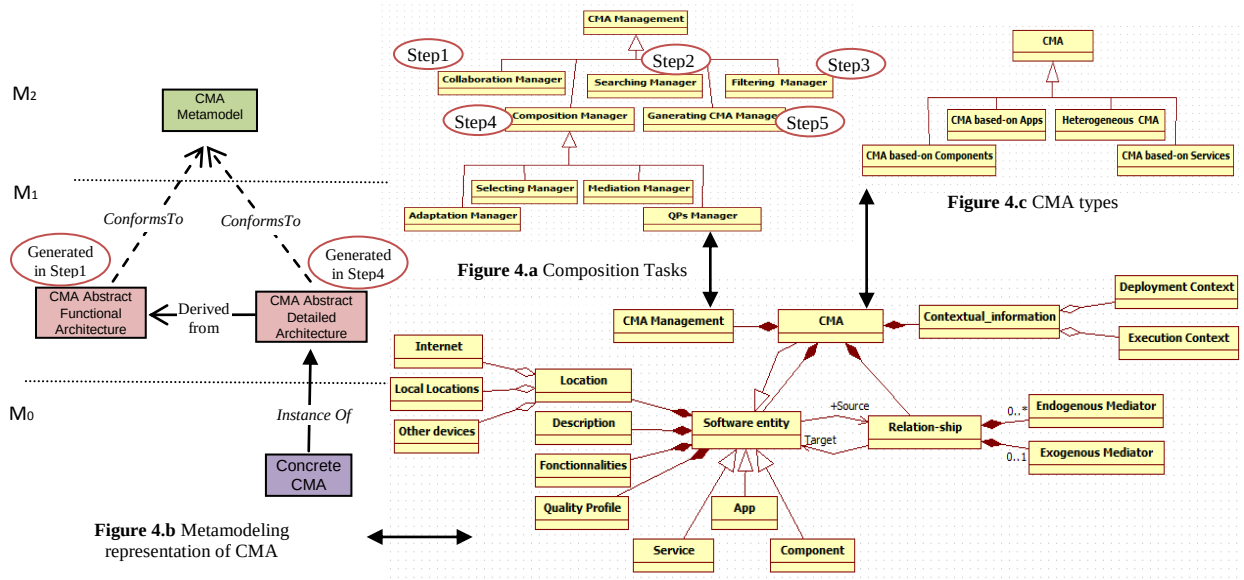


Figure 4.b Metamodeling representation of CMA

Figure 4.a Composition Tasks

Figure 4.c CMA types

Figure 4- Composite Mobile App Metamodel Overview.

The proposed process aims, as a first step, to define abstractly the desired mobile app. Thereby, it must firstly focus on the definition of the different functionalities which are required to compose the future mobile app and provide the dependencies between these desired functionalities (cf Fig.5). We associate this ability with the collaboration manager (cf Fig.4.a). *Collaboration manager* specifies *workflow* and *dataflow* between identified functionalities. *Workflow* schedules the invocations of functionalities (i.e. establishes precedence links) and *dataflow* expresses the data exchanges, inputs to output, between them (i.e. establishes use links).

The collaboration schema of the desired CMA presented in the motivating example of this paper will be as following:

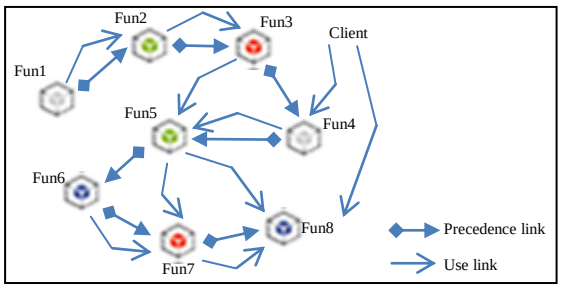


Figure 5- CMA abstract functional architecture.

Other than traditional software systems, the development of apps on mobile devices (e.g. smartphones, tablets, etc…) is constrained by theirs limited resources such as: small memory, a battery powered computing environment, and availability of some devices (e.g. Wi-Fi, GPS, auto-focus camera, etc…) [11]. For this purpose, our objective is not only to create mobile apps through the composition of existing software entities according to user needs but also to compose mobile apps that are sensitive to their contextual information (i.e. adaptive to their run-time environment). Thus, it is necessary to identify the context information of the deployment environment. It represents all software and hardware characteristics of the mobile device that will be used to run the desired mobile app. We model this context information in three categories: a) ***Deployment context***: represents hardware characteristics of mobile devices. b) ***Execution context*:** represents the current state of available devices. c) ***Execution platform:*** represents mobile device operating system, for example: iOS, android. This different contextual information may be inferred by the system of the mobile device or identified by the developer. We formalize the proposed categories of the mobile device in which we will run the desired CMA discussed in the motivating example as following:

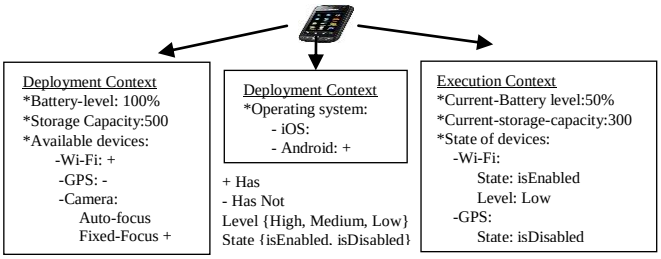


Figure 6- Context information.

Step 2: Discovering suitable concrete software entities

After setting the desired mobile app in abstract level (i.e. according to the architectural representation defined in the first step of the process), it is necessary to connect each abstract entity defined in the *CMA abstract functional architecture* with theirs corresponding existing concrete entities. We associate this role with the *Discovery manager* illustrated in the Fig.4.a. Discovery manager searches and chooses to have the corresponding concrete entities for each desired functionality by downloading them through *Internet* where some of them are free while others are paying (e.g. line stores such as Google Play, App Store or the Windows Phone Store), or by bringing them from *local locations* or *other devices* (e.g. laptops, mobile devices, etc...). Thus, the *Discovery manager* exploits the abstract functional representation of the desired app to find the different concrete entities that can be better-turned to the user's requirements described in this representation. Inspired from existing works in the area of services discovery ([12] [13] to access on available services, and on [14] to identify services adequate to needs), the *Discovery manager* will carry out his role as illustrated in *concrete software entities discovery* pseudo code. The result of this discovery task is a set of suitable concrete entities for each abstract entity (i.e. Suitable-CE-Fun i). The concrete entities mapped to the same abstract entity are functionally equivalent, but may vary in several non-functional aspects. In order to build a functioning mobile app, of good quality and adapted to the mobile device that will support it; we propose to associate each entity with a specific *Quality Profile* (*QP*) which contains all non-functional aspects that represent the necessary conditions for their execution. Thus, the different selected concrete entities must have a *QP* (cf Fig.7). We assume that these *QPs* are already defined (i.e. defined by developers of concrete entities) and we obtain them with the selected concrete entities during the execution of discovery task. *QPs* contain various metrics associated with values as illustrated in Fig.7. These metrics are the criteria that we propose to ensure the efficiency and the proper deployment and to measure the quality of the desired CMA in order to build it with a good quality as well as adaptive to their context information. In particular, it is necessary to notice that, using the described motivating example we are not presented the real *QPs* values of each concrete entity but only their relative difference. Following our motivating example, the functionality *Read Barcode* reflects an abstract software entity which aims to read the barcode written in the image of the product. The *Discovery manager* provides for these tow functionalities the following result: Suitable CE-Fun $_2 \leftarrow \{CE_{21}, CE_{22}, CE_{23}\}$.

```
Needed-Fun= {Fun1..... Fun n}; /*A set of desired functionalities*/
Available-CE= {CE1.....CEm}; /*A set of available concrete entities*/
Suitable-CE- Fun i= { }; /*A set of suitable concrete entities of Fun i*/
Sub-Fun i= { }; /* A set of sub-abstract entities of Fun i*/
Discovery 1-1: one concrete entity is needed to implement the same Fun.
Discovery 1-N: several concrete entities are needed to implement the same Fun.
is-correspond-to→ CEj adequate to the need Fun i.
1:  Begin
2:  For each Fun i ∈ Needed- Fun do /* discovery 1-1*/
3:      Access to Available-CE; /* access to locations of CE */
4:      For each CEj ∈ Available-CE do
5:          if CEj is-correspond-to (Fun i) then
6:              Add CEij (Suitable-CE- Fun i){
7:              Suitable-CE- Fun i←{Suitable-CE- Fun i + CEij}
8:                  }/* add CEj to a set of suitable CE*/
9           end_if
10:     end-for;
11:     If Suitable-CE- Fun i-is-empty then /* discovery 1-N*/
12:         Consult developers for decompose Fun i into k of sub- Fun;
13:         For each sub- Fun ik do
14:             Add sub- Fun ik (Sub- Fun i) {
15:                 Sub- Fun i← {Sub- Fun i + sub- Fun ik}
16:                 } /* thus: Sub- Fun i={ sub- Fun i1.... sub- Fun ik}*/
17:         end for
18:         For each sub- Fun ik ∈ Sub- Fun i do
19:             repeat 3, 4, 9.
20:         end-for;
21:     end_if;
22:  end-for;
23: End.
```

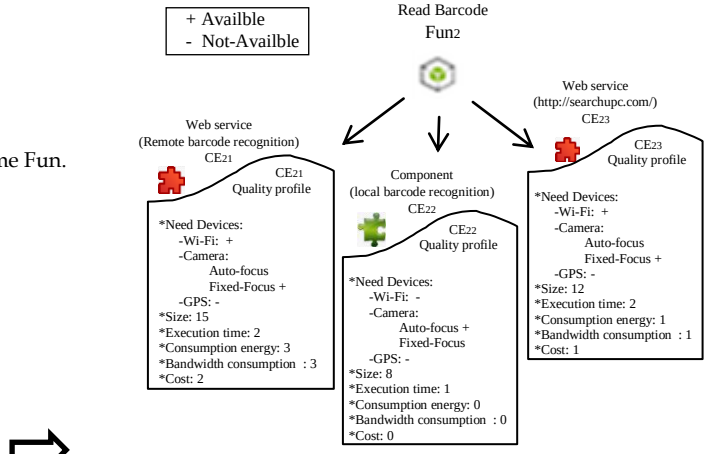


Figure 7- Concrete entities Quality Profile.

Step 3: Selecting context aware concrete software entities

The 3rd step of this process aims to solve the challenge of mobile devices heterogeneity that can arise when using such mobile device to deploy the desired CMA. Thereby, to ensure the correct deployment and the good functioning of the composed mobile app it is first necessary to assure that their composed concrete entities are adaptable to the current context of mobile device. This step corresponds to a filtering operation which aims to select among all concretes entities of each abstract entity those that are best suited with respect to the deployment and execution context of the mobile device where the composed mobile app will be deployed. We associate this filtering capability with the *Filtering manager* (cf Fig.4.a). It based on *QP* of each concrete entity and the different characteristics of the mobile device to perform the comparison in order to select those that are adaptive to these characteristics as illustrated in Fig.8. A concrete entity is conforms to mobile device context information if comparisons of all quality metrics with all mobile device characteristic are satisfied. Thereby, the result of this step is a set of concrete entities which are suitable to the deployment and execution context at the same time. According to the presented example in this paper and based on the research result of step 2 (cf Fig.7); the concrete entity CE_{22} that will be used to implement the functionality *read barcode* is not adaptive to the deployment context. It needs to Wi-Fi device to can perform its task. In this case, the comparison A *compared with* D is not satisfied. However, the result for this functionality will be: Suitable CE-Fun $_2$←{CE_{21}, CE_{23}}. The selection of the adaptive composed concrete entities to the mobile device characteristics only does not guarantee the correct deployment and the efficiency of the CMA but it is also necessary to satisfy the both following properties: 1) The overall size of composite app does not exceed the remaining storage capacity in the mobile device. 2) Handling the composed app does not consume an energy which exceeds the current energy level of the mobile device. These two properties can be checked during the collaboration of the selected composed concrete entities in the next step of the proposed process.

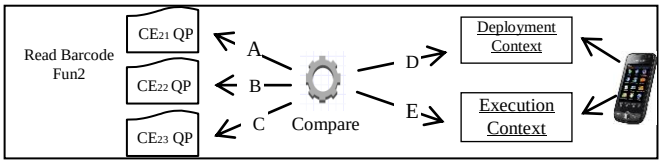


Figure 8- Filtering task.

Step 4: Composing Mobile app

After connecting each abstract entity with its corresponding context-aware concrete entities, the proposed process in this step designed to build the final app architecture (i.e. CMA *abstract detailed architecture*) by composing these concrete entities according to the coupling that generates the desired app with optimal composition cost. This process tries to overcome the heterogeneous composition limitation by providing a description representing the CMA architecture with: services, components, apps separately, or with heterogeneous entities (i.e. with several software entities types). Thus, our process allows building mobile apps as exogenous or endogenous composition (cf Fig.4.c) where each of their constituents' software entities will provide a specific service. However, heterogeneity issue in the composition mechanism can present two kinds of heterogeneity problems: 1) *Heterogeneous nature of entities:* the composed entities cannot directly communicate because the data which are exchanged between these entities are not understandable (e.g. microphone provides an audio stream and the jukebox needs to string input to perform its task). 2) *Heterogeneous type of entities*: represent the coordination of two different types of software entities (e.g. component connected with service). The proposed metamodel aims to address these heterogeneity issues by proposing two kinds of mediators (cf Fig.4) where: a) *Endogenous mediator*: that overcomes the heterogeneity between two entities of different nature; exchange data can require some transformation to be understandable. *Endogenous mediator* represents the *Mediation services* that are selected to ensure these data transformations. b) *Exogenous mediator*: this kind of mediators is intended to eliminate the heterogeneity between two entities of different kinds. These entities cannot directly communicate owing to their different type. Endogenous mediator aims to encapsulate related heterogeneous entities, it builds well-formed interface for each of them in order to take advantage of their services but just with manipulating necessary inputs and outputs independently from language implementations of these constituents entities. At this purpose, this composition process treats four type of composition as illustrated in the following table:

Composition Type	Heterogeneity problems	Proposed Mediators
Heterogeneous Exogenous	Heterogeneous Nature of entities Heterogeneous Type of entities	Endogenous Mediators Exogenous mediators
Homogenous Exogenous	Heterogeneous Type of entities	Exogenous mediators
Heterogeneous Endogenous	Heterogeneous Nature of entities	Endogenous Mediators
Homogenous Endogenous	None heterogeneity problems	None needed mediators

Table 2- Composition types.

Homogeneous endogenous composition allows to compose software entities with the same type and the same nature by a simple relationship and therefore without adaptation. In this case, the composed entities will be regrouped in a single software entity (e.g. composing two services of the same nature in a composite service). *Mediation manager* presented in Fig. 4.a is the responsible for this heterogeneous and endogenous composition, it manages *use links* which represent the *dataflow* between desired functionalities and identifies if it is necessary to associate this relationship with services mediation (i.e. during the composition of entities of different natures) or/and to encapsulate the composed entities (i.e. during the composition of entities of different types). The composition of adaptable concrete entities is not sufficient to ensure that the composite app itself will be adaptable to the context of the mobile device. *Adaptation manager* has the role to verify, after each composition step, if the actual free storage capacity is sufficient for deploy the composed app in this mobile device. Thus, if the consumed energy to handle the composed app does not exceeds the current battery level. If one of these two constraints is not satisfied, *Adaptation manager* triggers the recomposition of the app using alternative concrete entities that are elected in step 3. This leads to the conclusion that *Adaptation manager* aims to ensure the proper deployment and the correct functioning of the overall CMA. Last but least, it is necessary to generate for each composed mobile app its own *QP* and identify its own properties (i.e. location, description, and functionalities). This task is based on the different *QP*s and properties of the constituent entities of the composite app. We associate this role to *QFs manager* (cf Fig.4.a); it has the potential to fulfill CMA *QP* according to the different characteristics of their constituents.

Composition manager aims to handle CMA abstract functional architecture in order to derive the CMA abstract detailed architecture (cf Fig.4.b) by replacing each functionality by its appropriate context-aware concrete entity that can implement it with respect to the cost of its composition. This task is performed with *Selection manager* in parallel with *Mediation manager*, *Adaptation manager*, and *QFs manager*. *Selection manager* is based on composition type to calculate the composition cost of each concrete entity. Thus, it has the role of selecting those that are best suited which allow to compose the desired mobile app with the lower composition cost. The Fig.9 represents the generated abstract detailed architecture of the desired CMA and its *QP*.

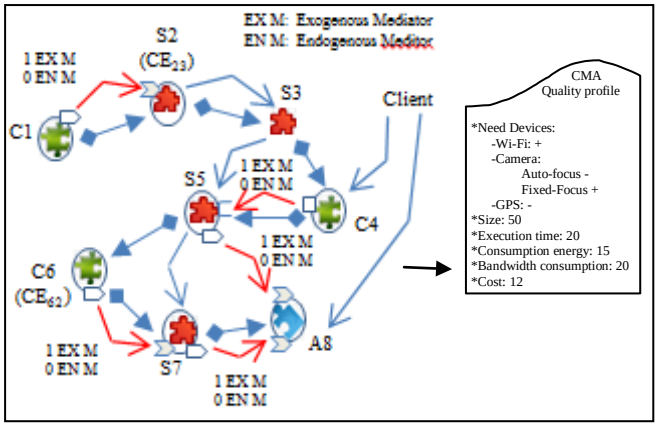


Figure 9- CMA abstract detailed architecture and its *QP*.

Step 5: Generating executable app

Mobile apps are composed visually with the proposed architectural representation without the need to write any lines of code. The executable model will be generated from the CMA description after searching, filtering, and selecting the most appropriate concrete entities as illustrates the instantiation relation presented in the Fig.4.b. Thus, our approach copes with the heterogeneity of mobile platforms. It is able to support heterogeneous target environments ranging from *CMA* architecture to mobile platforms. From this architectural model, the generation of the concrete CMA towards a specific platform is based on MDD mechanisms using transformations: *model to code*. This step improves that our process has the potential to deploy and migrate the same CMA between different mobile platforms (e.g. android, iOS). To perform this task, it is necessary to provide language syntax (e.g. using EMF/GMF technologies) in order to describe the architectural models of the desired CMA and to perform transformation rules using transformation languages (e.g. using ATL language in eclipse platform) in order to obtain the desired application code. The generation of specific code from architectural model by the execution of the defined transformation rules is the responsibility of *Generating CMA manager* (cf Fig.4.a). This step is still in progress.

5 Conclusions

In this paper we have presented conceptual framework for our proposed idea to provide *heterogeneous composition process* for mobile apps. This process is defined at architectural level. It is aimed to meet the needs of users and compose mobile apps that are sensitive and adaptive to the contextual information of the mobile device in which they will be installed, and also to facilitate the task of mobile apps composition while ensuring the efficiency and the good quality of the generated app. We have presented a conceptual framework of a composition process in mobile environment and we have introduced the motivations behind this process, its goals and its workability to build context-aware composite mobile apps using heterogeneous software entities. Moreover, this work is the cornerstone toward the definition of auto-composition and self-adaptation in a CMA. Thus, as a future work, we will aim to integrate and to clear specify these tow mechanisms in our composition process.

References

1. Jeffrey S. Hammond and Julie A. Ask , The Future Of Mobile Application Developmentn, January 17, 2013.
2. Hock-Koon, A. and M. Oussalah. Expliciting a composite service by a metamodeling approach. in Research Challenges in Information Science (RCIS), 2010 Fourth International Conference on. 2010. IEEE.
3. Chakraborty, D., et al., Service composition for mobile environments. Mobile Networks and Applications, 2005. 10(4): p. 435-451.
4. Wu, Z., et al., Automatic composition of semantic web services using process and data mediation. 2007.
5. Hock-Koon, A. and M. Oussalah, Composite service metamodel and auto composition. Journal of Computational Methods in Science and Engineering, 2010. 10: p. 215-229.
6. Rosa, R.E.V. and V.F. Lucena Jr. Smart composition of reusable software components in mobile application product lines. 2011. ACM.
7. Furno, A. and E. Zimeo, Context-aware Composition of Semantic Web Services. Mobile Networks and Applications, 2014. 19(2): p. 235-248.
8. Cugola, G., et al., Selfmotion: A declarative approach for adaptive service-oriented mobile applications. Journal of Systems and Software, 2013.
9. as Erl, T., Service-Oriented Architecture Concepts, Technology and Design. 2005.
10. Jifeng, H., X. Li, and Z. Liu, Component-based software engineering, in Theoretical Aspects of Computing–ICTAC 2005. 2005, Springer. p. 70-95.
11. Zhang, X., et al., Towards an elastic application model for augmenting the computing capabilities of mobile devices with cloud computing. Mobile Networks and Applications, 2011. 16(3): p. 270-284.
12. Verma, K., et al., The METEOR-S approach for configuring and executing dynamic web processes. 2005.
13. Kalasapur, S., M. Kumar, and B. Shirazi, Dynamic service composition in pervasive computing. IEEE Transactions on, 2007. 18(7): p. 907-918.
14. Patel, P. and S. Chaudhary. Context aware semantic service discovery. in Services-II, 2009. SERVICES-2'09. World Conference on. 2009. IEEE.
15. Abdelkrim Amirat, Hock-Koon Anthony, and Mourad Chabane Oussalah, Chapter 1: Object-Oriented, Component-Based, Agent-Oriented and Service-Oriented Paradigms, In M. C. Oussalah, editor, Software Architectures 1, Wiley-ISTE, April 2014.

InsMT: Instance Matching System for Ontology Alignment

Abderrahmane Khiat
LITIO Laboratory
University of Oran Es-Senia
B.P 1524 El M'Naouar 31000, Oran, Algeria
abderrahmane_khiat@yahoo.com

***Abstract*— Instance-based ontology alignment represents a very promising technique to identify semantic correspondences between entities of different ontologies when they contain many instances. In this paper, we propose a new approach for managing ontologies which contains disjoint instances. This approach aims to annotate the instances in order to enrich them with terminological information then we apply different terminological matchers with a local filter, for calculate the similarities and identify similar instances. We present the official results of our approach on Instance Matching[1] track of OAEI 2014[2]. The InsMT system provides good result in terms of recall compared to other systems participated in this track.**

Keyword: Instance Matching, Instance Annotation, Ontology Alignment, Ontology Matching, Semantic Interoperability, Semantic Web.

I. INTRODUCTION

THE SEMANTIC WEB community relies on ontologies to overcome the crucial problem of semantic heterogeneity. However, these ontologies are themselves heterogeneous. This heterogeneity may occur at syntactical, terminological, conceptual and semiotic levels [12]. Ontology alignment, is defined as the process of identification of semantic correspondences between entities of different ontologies to be aligned [1], is prerequisite and necessary for a better understanding and interactions between not only people but also software agents [2]. It likewise represents the solution to the problem of semantic interoperability between different sources of distributed information.

As one can easily imagine, this cannot be done manually beyond a certain complexity (due to the size and the number of ontologies). Automatic techniques or at least semi-automatic should be developed to reduce the burden of manual creation and maintenance of ontology alignment [2]. However the automatic identification of correspondences between ontologies is very difficult due to their conceptual divergence [7]. Much ontologies have been developed in recent years in different domains and several methods have been elaborated to semantically align their entities. These methods are generally based on the calculation of the similarity of their names, relationships and instances. These methods can be terminological, structural, semantic or extensional respectively, and recently modern systems combine these approaches [1].

The work proposed in this article is in the context of extensional alignment i.e. instance-based ontology alignment. In fact, when ontologies contain a lot of instances there is a very good opportunity to use the extensions of concepts in order to align them using instance-based techniques. In practice, two situations may arise in terms of sharing instances between ontologies to align [4]. When they have common instances, it is recommended to use metrics, such as JACCARD metric [1], which evaluates the overlap of instances between concepts, and these concepts are considered similar when their overlap is important [16]. In contrast, when ontologies do not share or share few instances, one generally proceeds to aggregate the information of instances of the concepts to align.

Our approach [13] aims to annotate instances with the name and label of concept that contain these instances also we add the name and label of properties related these concepts.

The advantage of our approach relative to others is that we enrich instances with terminological information obtained from the ontology, especially when the instances lack from these information and they are disjoint.

The rest of the paper is organized as follows. First, we present background in Section 2, and related work is presented in Section 3 on existing approaches for instance-based ontology alignment. Section 4 describes our contribution by giving a detailed account of our approach, and Section 5 presents the aspects related to the implementation of our system also we present the official results of our approach and the Section 6 contains concluding remarks and sets directions for future work.

[1] http://islab.di.unimi.it/im_oaei_2014/

[2] OAEI (**O**ntology **A**lignment **E**valuation **I**nitiative) organizes evaluation campaigns aiming at evaluating ontology matching technologies. http://oaei.ontologymatching.org/

II. BACKGROUND

In this section, we present the basic notions of ontology alignment and correspondence.

2.1 Ontology Alignment

The *alignment* is a process that starts from two representations O and O' and produces a set of correspondences between pairs of (simple or complex) entities <e, e'> belonging to O and O' respectively [2].
There are some other parameters that can be used in the alignment process [1]:

- A an input alignment, which is to be completed by the process.
- p the alignment parameters, e.g., weights, thresholds.
- r external resource used by the alignment process, e.g., common knowledge and domain specific thesauri.

The *Ontology Alignment* is a set of correspondences between two or more (in case of multi-alignment) ontologies.

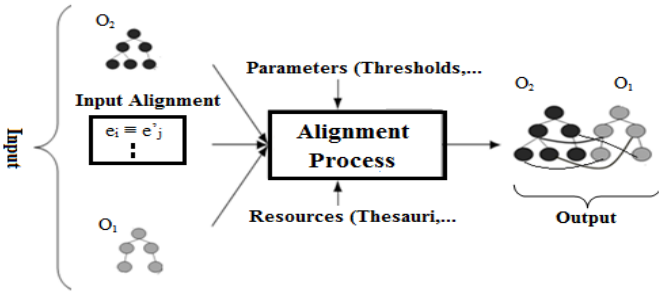


Fig. 1. Ontology Alignment Process.

2.2 Definition of Correspondence

Correspondence is a formal expression that states the semantic relationship between two entities belonging to different ontologies [1].

Given two ontologies O_1 and O_2, to align one ontology with another means that for each entity (concept C, relation R, or instance I) in ontology O_1, we try to find a corresponding entity, which has the same intended meaning, in ontology O_2.

In general a correspondence can be described as a quadruple: <e, e', n, R>:

- e and e' are the entities between which a relation is asserted by the correspondence.
- n is a degree of trust (confidence).
- R is the relation associated to a correspondence, where R identifies the relation holding between e and e', which can be:

  - simple set-theoretic relation.
  - a fuzzy relation.
  - a similarity measure.

III. RELATED WORK

We present and discuss in this section major works relevant to instance-based ontology alignment.

3.1 Instance-Based Ontology Alignment

Recently, instance-based alignment has become a promising family of solutions to find semantic relations among concepts of ontologies. In this section, we present some approaches relevant to this category.

Instance-based alignment compares the extensions of concepts, i.e. their instances rather than their interpretation. It provides notably the following advantages [3]:

- It focuses on the active part of ontologies, i.e. the instances, which reflects what those concepts really refer to in practice.
- It is less subjected to lexical ambiguity.
- It is not very sensitive to errors due to manual annotations which are generally inevitable.

There are two main categories [4] for instance-based ontology alignment.

3.2 Approaches in the Presence of Common Instances

These approaches are the most common of instance-based ontology alignment and use measures such as the JACCARD metric to calculate the similarity between entities of ontologies to align [16].

3.3 Approaches in the Absence of Common Instances

We mention below some systems of instance-based ontology alignment in the absence of common corpora between the ontologies to align:

1) RiMOM [5, 6]: is an acronym of Risk Minimization based Ontology Mapping. The principle of RiMOM is to construct a document from the comments, property values and instance names.

This approach uses cosine-similarity to compare documents. However it does not work well when instances do not use the same vocabulary to describe entities. Another disadvantage of this approach, resides potentially in insufficiency of information describing ontologies to be aligned.

2) GLUE [9]: The principle of GLUE[3] is to use multiple machine learning to construct for each concept of the ontology source a classifier based on names and textual content of instances (web pages URI), then the classifier decides whether or not an instance of a concept to align of the target ontology can belong to the concept of source ontology.

[3] GLUE: The semantic correspondences are in effect the "glue" that hold the ontologies to-gether into a "web of semantics".

This approach uses naive bayes algorithm on instances. It does not work well when instances are syntactically different or insufficient.

3) FCA-merge [10]: is an acronym of Formal Concept Analysis-Merge. The principle of FCA-merge is to populate ontologies with instances extracted from a set of documents that are relevant to both ontologies.

This approach uses NLP[4] techniques to extract instances from documents. It does not work well when instances are syntactically different or if there is a lack of documents that are relevant to both ontologies to be aligned.

4) IBOMbIE [3]: is an acronym of Instance-Based Ontology Alignment by Instance Enrichment. The principle of IBOMbIE is to align the instances and enriching instances of the source ontology by annotations of the most similar instances of the target ontology by creating a doubly annotated instance.

This approach uses cosine-similarity to compare instances of ontologies to be aligned. It does not work well when instances describing same concepts do not have the same features or if there is an insufficiency of instances.

5) Other Approaches

There are several other instance-based approaches like **MAFRA, LSD, COMA++** and the works proposed in [11] and [12] which are not covered by this paper due to minor importance for our approach (most them turn these instances to documents). An overview and a short description of several instance-based matching systems can be found in [1] [2].

* When ontologies contains disjoint instances, most of the ontology alignment approaches proceed generally, to aggregate the information of the instances by building a document containing these information, using machine learning techniques, methods of calculating similarities or using an external resource such as documents, dictionaries.

* Contrary to our approach, the approaches which aggregate only instances information do not work well especially when instances are disjoint and lack from terminological information.

* In our approach we aim to annotate instances with terminological information, and then we use different String-based methods to calculate similarities between these annotated instances. The advantage of our approach relative to others is that we enrich instances with terminological information obtained from the ontology, especially when the instances lack from these information and they are disjoint.

IV. OUR INSTANCE-BASED ONTOLOGY MATCHING APPROACH

The approach that we propose in this paper is situated in **Terminological-based methods** to resolve the problem of instance matching for ontology alignment when instances of the ontologies to align lack form terminological information and they are disjoint.

Our approach aims first to annotate instances of a concept with the name and label of the associated concept also with name and label of proprieties related this concept.

Then we use different string-based matchers in order to calculate similarities between these annotated instances with a local filter.

Finally, we select the instances similar using a threshold S. if the similarity value is greater than the threshold then the two instances is considered similar.

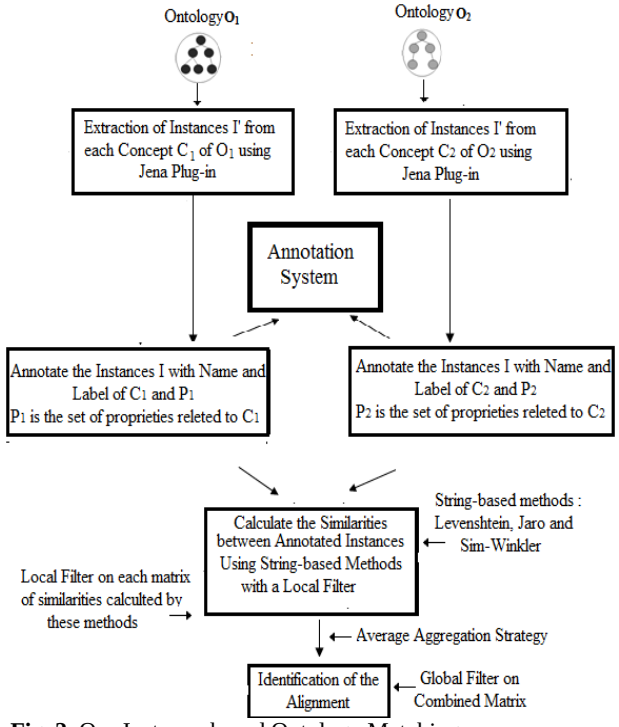


Fig. 2. Our Instance-based Ontology Matching.

We summarize the process of our approach in Fig.2 to provide a general idea of proposed solution. It consists in the following successive phases:

4.1 Phase 1: Annotation of Instances

In this phase, our system annotates the instances with the name and label of the concept also with property name. The purpose of this annotation is to enrich the instances with terminological information. This step is very import especially when instances do contain terminological information.

4.2 Phase 2: Calculation of Similarities

In this phase, our system calculates the similarities between instances, annotated in previous phase, using various string-based matching algorithms. The calculations of similarities by each string matching algorithm are represented in matrix.

[4] NLP: **N**atural **L**anguage **P**rocessing.

4.3 Phase 3: Local Filter

In this phase, our system applies a local filter on each matrix i.e. we choose for each string-based matching algorithm a threshold to realize a filter. We consider that: the similarities which are less than the threshold are set to 0. Our intuition behind this local filter is that the similarities which are less than the threshold can influence the strategy of the average aggregation.

4.4 Phase 4: Identification of Alignment

In this phase, our system combines the similarities of each matrix (after we have applied a local filter) using the average aggregation method and the result of the aggregation is represented in a matrix.

4.5 Phase 5: Global Filter and Identification of Alignment

In this final phase, our system applies a second filter on the combined matrix (result of the previous step) in order to select the correspondences found using the maximum strategy with a threshold.

V. Implementation, Evaluation and Experiment Results

A. Implementation

In order to test our approach we have implemented a platform in JAVA. The extraction of the instances of the two ontologies to align was produced with Jena Plug-in of the Protégé platform. The tasks of calculating the similarity and alignment extraction were performed using the different string-based methods namely: Levenshtein-distance, Jaro, and Slim-Winkler with a local filter, strategy of average aggregation and global filter for selection of semantic correspondence between instances.

B. Evaluation and Experiment Results

We present in this section the evaluation results of our InsMT System on Instance Matching track of OAEI 2014. The protocol of experimentation in this paper is the same taken from the OAEI instance matching track website due to different evaluation used in this track.

5.1 The Official Evaluation

We have implemented our InsMT system in order to participate in OAEI 2014 evaluation campaign, which represents as the most known evaluation campaign in the field of Ontology Matching. To participate in this campaign we must wrap and send our system in order to test the system on the SEAL platform which evaluation different ontology alignment systems. The alignment file generated from our InsMT system is analyzed and the results can be found on OAEI website.

5.2 Experimental Protocol

Our contribution is situated in instance-based ontology alignment for that the InsMT system is designed especially for instance matching track of OAEI evaluation campaign. The alignment file generated from InsMT contains the correspondences between instances.

The instance matching track aims at evaluating tools able to identify similar instances among different RDF and OWL ontologies. The instance matching edition for OAEI 2014 is articulated in two sub-tasks for identity recognition and similarity recognition. The evaluation was executed with support of SEALS. The evaluation was blind as described in details on the instance matching page.

6.2.1 Identity Recognition (*id-rec task*)

The goal of the id-rec task is to determine when two OWL instances describe the same real-world entity.

The datasets of the id-rec task have been produced by altering a set of original data with the aim to generate multiple descriptions of the same real-world entities where different languages and representation formats are employed.

The task goal is to produce a set of mappings between the pairs of matching instances that are found to refer to the same real-world entity. For example A book instance in the source Abox can have none, one, or more than one matching counterparts in the target Abox.

C. Evaluation Strategy

The mappings produced by our system were compared against a ground truth where an instance i in the source dataset is associated with all the instances in the target dataset that represent an altered description of i. Evaluation was performed through precision, recall, and F-measure [15]. These measures of are defined as follows:

$$\text{Precision} = P(A,R) = \frac{|R \cap A|}{|A|} \quad \text{Recall} = R(A,R) = \frac{|R \cap A|}{|R|}$$

$$\text{F-measure} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}$$

Where |R| is the number of correspondences of the instances reference alignment and |A| is the number of correspondences of instances found by our approach.

D. Experiment Results

1) *Analysis by Precision, Recall, and F-measure*

Systems	Precision	Recall	F-measure
InsMT	0.0008	0.7785	0.0015
InsMTL	0.0008	0.7785	0.0015
LogMap	0.6031	0.0540	0.0991
LogMap-C	0.6421	0.0417	0.0783
RiMOM-IM	0.6491	0.4894	0.5581

Tab. 1. The Evaluation Results of Different Systems

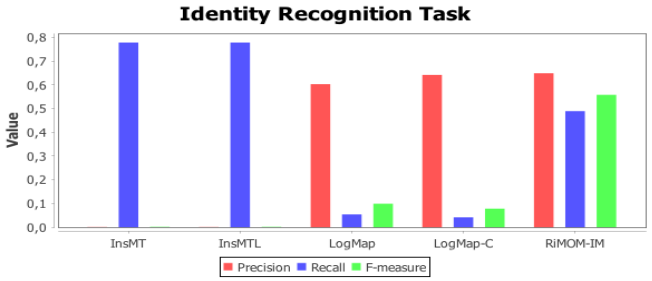


Fig. 3. The Evaluation Results of Different Systems on id-rec.

2) *Analysis by Increasing Values of Matching Threshold*

For each matching tool participate in this sub-track, the precision, recall, and F-measure have been observed when increasing values of matching threshold are specified. The threshold defines the minimum matching value between a pair of instances that is required for considering them as referring to the same real-world entity. We present above the analysis of our InsMT system, the results complete with other systems can be found in OAEI website in Instance Matching track.

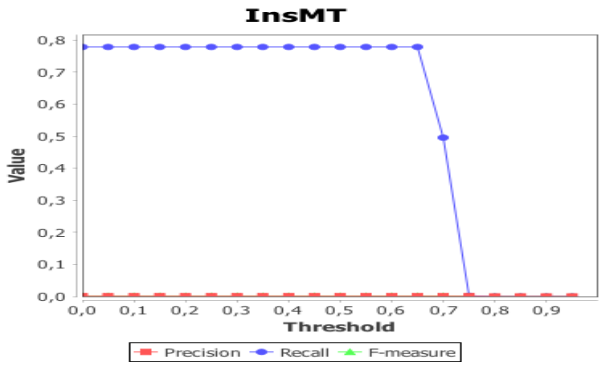


Fig. 4. The Evaluation Results of InsMT System on id-rec.

E. Comment on Obtained Results

The results of our InsMT system are better than all other systems in terms of recall but in terms of F-measure are not very good. The RiMOM-IM is considered as the best system in id-rec task. As we have mentioned before our InsMT system uses terminological information for annotation and when these ontologies do not contain these information our system fails.

6.2.2 Similarity Recognition (*sim-rec task*)

Contrary to id-red task, the goal of the sim-rec task is to evaluate the degree of similarity between two OWL instances, even when the two instances describe different real-world entities.

The datasets of the sim-rec task have been produced through crowdsourcing by employing the Argo system (Italian language). More than 250 workers have been involved in the crowdsourcing process to evaluate the degree of similarity between pairs of instances describing real books. Crowdsourcing activities have been organized into a set of HITs (Human Intelligent Task) assigned to workers for execution. A HIT is a question where the worker is asked to evaluate the degree of similarity of two given instances. The worker exploits the instances (i.e., book descriptions) "at a glance" and she/he has to specify her/his own perceived similarity by assigning a degree in the range [0, 1].

The task goal is to produce a complete set of mappings between any pair of instances. For example if the source Abox contains 173 book instances and the target Abox contains 172 book instances, then our system must provide 173x172 = 29756 mappings, each one featured by a degree of similarity in the range [0,1].

A. Evaluation Strategy

The mappings produced by our system were compared against the mappings obtained through crowdsourcing. Given a mapping m, the evaluation results is to compare the similarity degree assigned to m by the matching tool against the corresponding similarity degree assigned by the crowdsourcing workers. The evaluation was performed through the Euclidean distance.

B. Experiments Results

1) *Mappings Analysis*

The reference alignment is the mapping set obtained through crowdsourcing, where each mapping is associated with a similarity degree σ defined by the workers involved in the crowdsourcing activities. The cardinality of the reference alignment is `4.104` mappings.

In the analysis, for each mapping, the similarity degree of the reference alignment has been compared against the similarity degree calculated by the matching tools of the participants.

In the diagram, for a mapping `m` (i.e., the mapping between a certain instance `i` of the source dataset and a certain instance `j` of the target dataset):

- The similarity degree σ of the mapping m in the reference alignment (red line);
- The distance between σ and the similarity degree calculated by the matching tool (blue line for InsMT; green line for RiMOM-IM).

For the sake of readability, the mappings of the reference alignment are sorted according to the associated similarity degree.

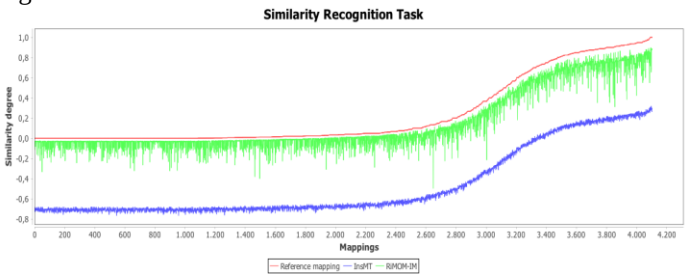


Fig. 5. The Evaluation Results of RiMOM-IM and InsMT Systems on sim-rec.

2) *Analysis by Range of Distance*

In this evaluation the range [0, 1] split of possible similarity degrees into ten smaller ranges of size 0.1 called range of distance. A range of distance r is populated with

those mappings whose distance from the reference alignment is in the range of r. Consider a mapping m:

- σ is the similarity degree of the mapping m in the reference alignment;
- τ is the similarity degree of the mapping m calculated by a considered matching tool.

The mapping m is placed in the range of distance corresponding to the value |σ - τ|. The similarity recognition task - analysis by range of distance - As a further, synthetic analysis indicator, for each matching tool, the Euclidean distance has been calculated with respect to the reference alignment. Consider:

- M is the overall number of mappings in the reference alignment based on crowdsourcing (i.e., 4.104 mappings);
- σ is the similarity degree of a mapping in the reference alignment;
- τ is the similarity degree of a mapping calculated by the considered matching tool.

The Euclidean distance (1) is calculated as follows:

$$d = \sqrt{\sum_{i=1}^{M} (\sigma_i - \tau_i)^2} \qquad (1)$$

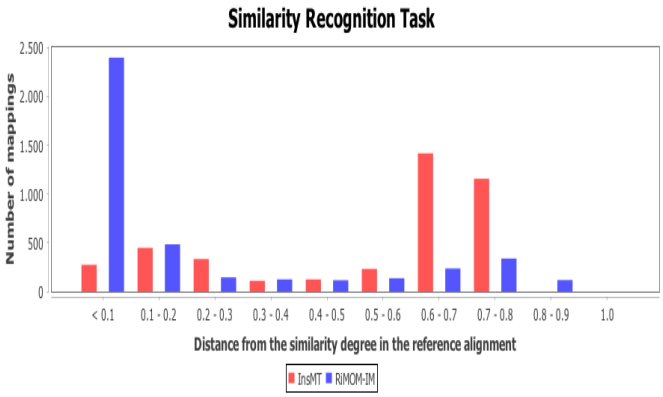


Fig. 6. The Evaluation Results of RiMOM and InsMT Systems on sim-rec distance from similarity degree.

Systems	Euclidean Distance
InsMT	d(InsMT) = 37.03
RiMOM-IM	d(RiMOM) = 21.83

Tab. 2. The Euclidean Distance of RiMOM-IM and InsMT Systems.

C. Comment on Obtained Results

Only two systems have been participated in this sub-task (sim-rec) namely RiMOM-IM and InsMT. The results of our InsMT system still not achieve good results.

V. CONCLUSION

In this article, we have introduced a new approach of instance-based ontology alignment. Our solution is very useful to identify the semantic correspondences between concepts of different ontologies when the instances are disjoint and lack from terminological information.

This is the first time the InsMT system has participated in SEAL platform and OAEI campaign. InsMT was designed especially for instance matching track of OAEI 2014 evaluation campaign.

The important part of InsMT system is the annotation of instances with concept and property names then our system combines the various string-based matching algorithms in order to calculate similarities with average aggregation method. Then we have applied a filter on the combined matrix for the selection of semantic correspondences between instances of ontologies to be aligned.

The purpose of this annotation is to enrich the instances with terminological information.

Finally the results show that our system provide good results in terms of recall but doesn't provide good results in terms of F-measure. The objective behind the implementation of our InsMT system is to find the best strategy of annotation of instances with local filter as we have proposed in section 4.1 and 4.3.

We have proposed method of annotation and we envision in the future to improve our InsMT system in order to get better results in terms of F-measure and propose other annotation (of instances) methods.

REFERENCES

[1] J. Euzenat and P. Shvaiko "Ontology Matching", Springer-Verlag, Heidelberg, 2013.

[2] M. Ehrig "Ontology Alignment: Bridging the Semantic Gap", Springer, 2007.

[3] B. Schopman, S. Wang, A. Isaac and S. Schlobach "Instance-Based Ontology Alignment by Instance Enrichment", Springer, Vrije Universiteit Amsterdam, Netherlands, 2012.

[4] E. Rahm "Towards large-scale schema and ontology Alignment", ReCALL, 2011.

[5] Z. Wang, X. Zhang, L. Hou,Y. Zhao, J. Li, Y. Qi and J. Tang "Rimom: a dynamic multi-strategy ontology alignment framework", OAEI, 2010.

[6] J. Li, J. Tang, Y. Li and Q. Luo "Rimom: a dynamic multistrategy ontology alignment framework", IEEE Trans Knowl, 2009.

[7] P. Bouquet, J. Euzenat, E. Franconi, L. Serafini, G. Stamou and S. Tessaris "Specification of a common framework for characterizing alignment", 2004.

[8] A. Maedche, B. Motik, N. Silva and R. Volz "Mafra¡Xa mappingframework for distributed ontologies", Springer, Benjamins VR (eds) EKAW, Berlin, vol 2473, pp 235,V250, 2002.

[9] A. Doan, J. Madhavan, P. Domingos and A. Halevy "Ontology Alignment: a machine learning approach", Springer, Berlin, 2004.

[10] G. Stumme, A. Maedche "Fca-merge: bottom-up merging of ontologies", Proceedings of the 17th international conference on artificial intelligence (IJCAI), Seattle, 2001.

[11] KS. Zaiss "Instance-based ontology Alignment and the evaluation of Alignment systems", Ph.D. thesis, Heinrich Heine Universitat Dusseldorf, 2010.

[12] K. Todorov, P. Geibel, KU. Kuhnberger "Mining concept similarities for heterogeneous ontologies", Springer, Berlin, ICDM, vol 6171. , 2010.

[13] A. Khiat "InsMT results for OAEI 2014 instance matching", OAEI 2014.

[14] A. Khiat and M. Benaissa "InsMTL results for OAEI 2014 instance matching", OAEI 2014.

[15] Euzénat J., "Semantic Precision and Recall for Ontology Alignment Evaluation", IJCAI, 2007.

[16] P. Jaccard, "Distribution de la flore alpine dans le bassin des dranses et dans quelques r´egions voisines", Bulletin de la societe Société vaudoise des sciences naturelles, 1901.

View publication stats