OCTA

6-8 February, 2020

ALECSO, Tunis -Tunisia

OCTA International Multi-Conference on:

Organization of Knowledge and Advanced Technologies

https://multiconference-octa.loria.fr/

Unifying the scientific contributions of the following conferences:

SIIE

8th. International Conference on:

Information Systems and Economic Intelligence

ISKO-Maghreb

Digital Sciences: impacts and challenges on Knowledge Organization

CITED

2nd. International Symposium on:

Advanced Technologies, Renewable Energies and Economic Development

TBMS

1st. international symposium on:

Big-Data-Analytics Technologies for Strategic Management: innovation and competitiveness

ORGANIZERS

Université de Tunis

ISKO-Maghreb

UNIVERSITÉ DE LORRAINE

IEEE Tunisia Section

www.ieee.tn

المنظمة العربية للتربية والثقافة والعلوم

SPONSORS

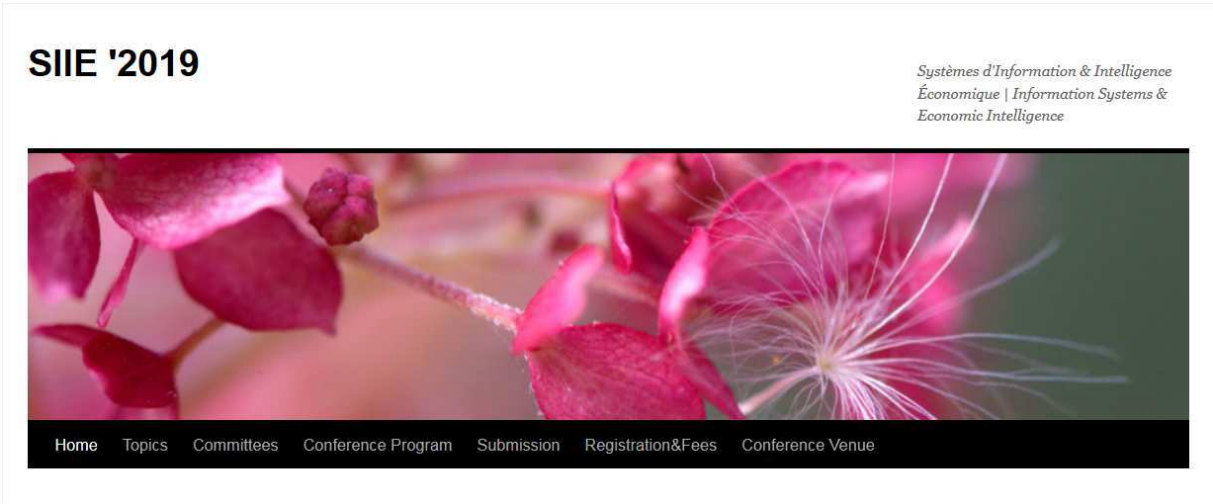
Tunisia China Friendship Association

ATLAS

SPECIAL ISSUES JOURNALS & BOOKS

isprs International Journal of Geo-Information

IEEE Xplore® Digital Library



PROCEEDINGS
8th. Edition of the International Conference on:

"Information Systems and Economic Intelligence".

February 6-7-8, 2020 Tunis (Tunisia)

Thematic Conferences & Tutoriels

www.siie.fr & https://siie2019.loria.fr

Organization in the event of the International Multi-Conference OCTA'2019 on: « Organization of Knowledge and Advanced Technologies »

(https://multiconference-octa.loria.fr/)

SIIE'2019 CFP (to download) & **Special issues in Journals & Books**

Description :

The SIIE conference aims to promote dialogue between experts and researchers from both the public and private sectors, on fundamental and experimental knowledge of Information Systems and Economic Intelligence (SIIE). This is to upgrade in a risk environment the technologies related to economic intelligence (IE). The dynamics of EI (ie. Competitive Intelligence) depend on mastering the knowledge and skills needed

to design the best strategies and to ensure that decision-makers make the right decisions.

SIIE will hold its eighth (8th.) edition in Tunis in November 2019, after the seven successful editions. This edition is organized by the University of Tunis and supported by the universities of the Maghreb countries, Europe and the sponsorship of the International Societies as: **IEEE Tunisia section** (in progress) and **ISKO-Maghreb Chapter** (confirmed).

The last editions and its proceedings since 2008, have allowed academic researchers and economic actors to come up with finalized projects. The goal of SIIE is to continue in this way by creating opportunities, innovative ideas and means to strengthen projects, and to build bridges between universities and industries on both shores of the Mediterranean.

SIIE'2019 will offer lectures, tutorials, authors' presentations and industrial panels that will be led by experts to identify new approaches and knowledge in economic intelligence, as well as applied research and feedback.

In a friendly and pleasant environment, as Tunisia knows how to do so, the SIIE conference promotes the creation of networks of trust between academics, industry and politicians, thus contributing to strengthening the alliances of the scientific community SIIE. The recommendations and opinions of the experts will help this community to find solutions to its many questions and problems.

Committees

Steering Committee OCTA'2019 :

Honorary President
Pr. Habib SIDHOM, President of the University of Tunis (Tunisia)

Presidents
Sahbi SIDHOM (LORIA – Université de Lorraine, France)
Amira KADDOUR (ENSTAB – Université de Carthage, Tunisia)
Anass EL HADDADI (ENSA Al-Hoceima, Morocco)
Mohamed ADDAM (ENSAH Al Hoceima, Morocco).
Abdelkrim MEZIANE (Centre de Recherche CERIST, Algeria)
Davy MONTICOLO (ENSGSI – Université de Lorraine, France)
Xi LU (Tsinghu University, China & Harvard University, Cambridge MA, USA)
Saoussen KRICHEN (Vice-President Université de Tunis, Tunisia)
Khaled KCHIR (Vice-President Université de Tunis, Tunisia)

Program committee / Comité scientifique :

Scientific Gerneral Chairs

Davy MONTICOLO (ENSGSI – Université de Lorraine, France)
Sahbi SIDHOM (LORIA – Université de Lorraine, France)

Members
ABOU EL MAJD Badr (Hassan II University, Morocco)
AGOSTINELLI Serge (LSIS UMR CNRS 6168 – Université Paul Cezanne Aix-Marseille III, France)
AïMEUR Esma (University of Montreal, Canada)
ALIMI Adel M. (IEEE Computer Society Tunisia & ENIS – University of Sfax, Tunisia)
AMAMI Mokhtar (Business Administration – Royal Military College of Canada, Kingston. Ontario, Canada)
BADACHE Nadjib (Directeur du CERIST, Algeria)
BEN AYED Mounir (Faculty of Sciences, University of Sfax, Tunisia)
BEN GHEZALA Henda (ENSI – Université de la Manouba , Tunisie)
BEN HAMADOU Abdelmajid (Lab. MIRACL Université de Sfax, Tunisie)
BENKARAACHE Taoufik (Hassan II University, Morocco)
BENSAOUD Othmane (IGA – Marrakech, Maroc)
BERRADA Ilham, ENSIAS, (Université Mohammed V de Rabat, Morocco)
BOYER Anne (LORIA– Université de Lorraine, France)
BIDAN Marc (Université de Nantes, France)
CALOF Jonathan (Telfer School of Management – University of Ottawa, Canada)
CHAFIK Khalid (ENCGT, Morocco)
CHAUDIRON Stéphane (Université de Lille 3, France)
CHICHTI Jameleddine (Montpellier Business School, France & Tunisie)
CUNG Van-Dat (Grenoble-INP, France)
DOUSSET Bernard (IRIT – Université de Toulouse, France)
EL AKKAOUI Zineb (INPT, Morocco)
EL GHAZI Hamid, (INPT, Morocco)
EL HACHANI Mabrouka, (ELICO – Université Jean Moulin Lyon3, France)
EL HADDADI Anass (ENSA Al-Hoceima, Maroc & Université de Toulouse, France)
FAUVET Marie-Christine (LIG – Université Joseph Fourier, France)
FEKI Jamel (Lab. Miracl, Université de Sfax, Tunisie)
FOHR Dominique (INRIA Grand Est & Université de Lorraine, France)
GAMMOUDI Mohamed Mohsen (Université de Carthage, Tunisie)
GLADY Marc (ESSEC Business School, France)
GRESSER Jean-Yves (Black Forest Group Inc. New York, USA & GELM Entreprises. Paris, France)
GRIVEL Luc (Université de Paris I Panthéon-Sorbonne, France)
GUAAYBESS Tourya (CREM – Université de Lorraine, France)
GUY Pascal (Oracle Toulouse, France)
HASSANALY Parina (IEP d'Aix-en-Provence, France)
HASSOUN Mohamed (ENSSIB de Villeurbanne, France)
HATON Jean-Paul (LORIA/INRIA – Université de Lorraine et Institut Universitaire de France, France)
ILLINA Irina (INRIA Grand Est & Université de Lorraine, France)
JAMOUSSI Salma (Lab. Miracl, Université de Sfax, Tunisie)

KALLEL Ilhem (ISIMS, University of Sfax, Tunisia)
KAMMOUN Habib (ENIS, University of Sfax, Tunisia)
KASSOU Ismaïl (ENSIAS – Université Mohammed V Souissi de Rabat, Maroc)
KAZAR Okba (Département d'informatique – Université de Biskra, Algérie)
KHALIDI IDRISSI Mohamed, (Département de Génie Informatique, Université Mohamed V Rabat, Maroc)
LAMBERT Philippe (Université de Lorraine, France)
LARBAOUI Souad (Institut de Management Stratégique et d'Intelligence Economique, Algeria)
LAURENT Anne (LIRMM – Université Montpellier 2, France)
LELEU-MERVIEL Sylvie (Université de Valenciennes et du Hainaut-Cambrésis, France)
LÓPEZ-HUERTAS María (Universidad de Granada, Spain)
MAKHLOUF Mohamed (KEDGE Business School, France)
MEZIANE Abdelkrim (CERIST, Algeria)
MILLE Alain (LIRIS – Université Claude Bernard Lyon1, France)
MIRANDA Serge (Faculté des Sciences, Université de Nice Sophia Antipolis, France)
MOINET Nicolas (IAE Université de Poitiers, France)
NOURRISSIER Patrick (Director NetLor Concept, France)
NOWAKOWSKI Samuel (LORIA – Université de Lorraine, France)
OHLY Peter (President ISKO int. & GESIS – Leibniz-Institut für Sozialwissenschaften – Bonn, Germany)
PECQUET Pascal (Université de Montpellier 1, France)
PINON Jean-Marie (INSA de Lyon, France)
RAZMERITA Liana (Copenhagen Buisness School, Denmark)
ROUDIES Ounsa, (Ecole Mohammadia d'Ingénieurs, Université Mohammed V de Rabat, Morocco)
ROUSSANALY Azim, (LORIA – Université de Lorraine, France)
ROUSSET Marie-Christine (LIG – Université de Grenoble & Institut Universitaire de France, France)
SALEH Imad (Lab. Paragraphe, Université Paris 8, France)
SMAILI Kamel (LORIA – Université de Lorraine, France)
SOUISSI Omar, (INPT, Morocco)
SOLBERG SØILEN Klaus (Blekinge Institute of Technology, Sweden)
TAKI Mounaim (Mohamed V University, Morocco)
TANTI Marc (Centre d'épidémiologie et de sante publique des armées – UMR 912 – SESSTIM, France)
TARDIEU Hubert (ATOS, France)
TARI Kamel (Univeristé de Bejaia, Algeria)
THELWALL Michael (School of Technology, University of Wolverhampton, UK.)
YILDIZ Ustun (University of California – Davis, USA)
ZREIK Khaldoun (CITU-Paragraphe & Université de Paris 8, France)

Results of multi-agent system and ontology to manage ideas and represent knowledge in a challenge of creativity

Pedro Chávez Barrios
Escuela de Bachilleres, Universidad Autónoma de Querétaro
Moises Solana s/n, 76070 Querétaro, México
pedro.chavez@uaq.mx

Davy Monticolo
ERPI Laboratory, University of Lorraine
8 rue Bastien Lepage, 54000 Nancy, France
davy.monticolo@univ-lorraine.fr

Sahbi Sidhom
LORIA (Kiwi) Laboratory, University of Lorraine
F-54506 Vandœuvre lès Nancy, France
sahbi.sidhom@loria.fr

Abstract—**This article is about an intelligent system to support ideas management as a result of a multi-agent system used in a distributed system with heterogeneous information as ideas and knowledge, after the results about an ontology to describe the meaning of these ideas. The intelligent system assists participants of the creativity workshop to manage their ideas and consequently proposing an ontology dedicated to ideas. During the creative workshop many creative activities and collaborative creative methods are used by roles immersed in this creativity workshop event where they share knowledge. The collaboration of these roles is physically distant, their interactions might be synchrony or asynchrony, and the information of the ideas are heterogeneous, so we can say that the process is distributed. Those ideas are writing in natural language by participants which have a role and the ideas are heterogeneous since some of them are described by schema, text or scenario of use. This paper presents first, our MAS and second our Ontology design.**

Keywords: MAS Multi-agent system, Ontology, Intelligent system, Knowledge, Creativity and Idea;

I. INTRODUCTION

The University of Lorraine organizes every year a creativity workshop called "48 hours generating ideas" (48H). We have observed that more than 1200 idea cards (IdC) were generated during the last 48H creativity workshop in 2017 [1]. **In order to manage these ideas a multi-agent system is studied and proposed since the multi-agent system has been proved to be efficient in a distributed process and to propose an ontology to represent knowledge**. The concept of multi-agent system appears at the end of last century. The multi-agent system has two forms of vision the interaction among agents and the interaction among humans, *the first, as an artificial intelligence (AI) concept attributed to Nils Nilsson "all AI is distributed-1980" and the second as artificial life (Alife) based in the complex adaptive behaviors of communities of humans* [2]. B*y relating an individual to a program, it is possible to simulate an artificial world populated of interacting processes* [3]. The individual is an agent that interacts according to his environment which is clearly defined with respect to the reality. These interactions among agents and their environments are an important aspect in the MAS. In the beginning of the century XXI, *an initial model tools used to create generic multi-agent platforms based on an organizational mode based in the core model agent-role-group* [4], and also *multi-agent model involving some agents to hundreds focusing in break down a problem therefore the agent can solve a simple problem* [5]. At present, the multi-agent systems have been used to improve energy efficiency [6]. However, thinking in our intelligent system based in agents, the interactions among the actors in the creativity workshop 48H is complicated and we have to help them in their individual and collaborative activities inside this organization where *the organization is defined by a group of roles that interact among them* [7]. Several design methodologies of multi-agent system exist such as GAIA [8], [9] and DOCK [10] are examples of these design. The multi-agent systems have two principal methodologies, [11] the *methodologies oriented to agents and the methodologies oriented to organizations* that *base in organizational unit, service, environment and norm* [12]. Due to the uncountable times that agent is mention and the interaction of agent in a multi-agent system, we have to write some definitions about the concept of agent, it takes some primordial functions during the creation of our intelligent system. Since the last century and initial years of this century the concept of agent and its characteristics appears. *There are several definitions about agent, description of agent's requirements, uses of agent* [13] *and description about agent's evolution* [14]. In a software design an agent represents state components, which are structured aggregations *(sequences, sets, multisets, etc.) of elements such as events, actions, beliefs, plans and tasks* [15]. An actual definition, "Agent" is a system whose behavior is neither casual nor strictly causal, but teleonomic [16], "goal-oriented" toward a certain state of the world . an agent is specified as *an active communicating entity which plays roles inside groups*[4]. *An agent is a computer system **situated** in some **environment**, and that is capable of autonomous **action** in this environment in order to meet its design objectives where **autonomous** means to act without interventions of humans of other agents* [17].

Having proposed the MAS next step is to define ontology. With respect to the ontology, it is dedicated to ideas and specifically to assist participants in the idea generation during the creativity workshop. In addition, our ontology represents knowledge from this CWS like ideas, processes, activities, actors, roles, methods, idea cards and possible solutions; this ontology is used to annotate ideas and to facilitate the ideas management. As initial definition of ontology, the etymology of ontology comes from ancient

Greeks, but the concept of ontology appeared in the century XVII in the work "Ogdoas Scholastic" by Jacob Lorhard, he provided a useful key to the understanding of the Protestant Europe in a grammar text book [18]. Also, last century, the concept of ontology focus in the definition of objects, concepts, entities, relationships among them in a defined area [19], [20], and ontology works as database with information, properties, relationships about concepts that exist in the world or domain [21].

The objective of this article is to present the results about our multi-agent and ontology proposals. Initially, we present the context, the problem and the methodology; the section two is about the state of art about MAS and Ontology approaches; the third section, our approaches. Finally, last section is dedicated to results and conclusions.

II. Multi-agent System (MAS) and Ontology

A. MAS and Ontology, state of the art

Multi-agent system

Inside the multi-agent system, the concept of agent is vital. The definition of agent in a general and complete AI idea: *an agent is anything that can be viewed as perceiving its environment through sensors and acting upon that that environment through effectors. A human agent has eyes, ears, and other organs for sensors, and hands, legs, mouth, and other body parts for effectors. A robotic agent substitutes cameras and infrared range finders for the sensors and various motors for the effectors* [22].

There are some varieties of agents but in a robotic sense: *Autonomous agents require be reactive to changes in the environment, it must be able to predict, incompatible goal management and adaptivity (prediction)* [23]. The agents, in our environment of creativity workshop 48H, is part of teams and can play one or more roles. Another definition, an agent is a physical or virtual entity that have several properties. These properties are the capacity to interact with its environment, the capacity of communication with other agents, the necessity to achieve an objective, the capacity to manage its resources, the capacity to perceive the environment, the capacity to represent partially or totally the environment and eventually the capacity of reproduction [24]**.** With a different Wooldridge's conceptions, an agent is an informatic system in a specific environment, with autonomous actions to achieve its objectives [25]. For a language, agent could be a mental state consisting in beliefs, desires and intentions [8]. A definition based in software, agent are coarse-grained computational system, each making use of computational resources, they are heterogeneous [8]. A Final definitions according to a multi agent system where agents interact among them to achieve a global objective, there exist two kind of agents, cognitive and reactive agents[24]. The artificial intelligence (AI) is an important discipline that defines agent in different ways. *AI borrows concepts (states, actions and rational agents) and techniques for autonomic computing. The definition of rational agent: is any entity that perceives and acts upon its environment, selecting actions that, on the basis of information from built-in knowledge, are expected to maximize de agent's objective* [26]. The most simple definition about agent, *A software agent has encoded bit strings as its percepts and actions* [22].

Ontology

In the last century, the concept of ontology focus in the definition of objects, concepts, entities, relationships among them in a defined area [19], [20], and ontology works as database with information, properties, relationships about concepts that exist in the world or domain [21].

Berners-Lee proposes to use the ontologies in the context of the Internet in order to bring a semantic dimension of the Web. He explains, "The Semantic Web will enable machines to comprehend semantic documents and data, semantic web uses collections of information called ontologies that is a component of the semantic web. Artificial-intelligence and Web researchers have co-opted the term for their own jargon, and for them an ontology is a document or file that formally defines the relationships among terms [27]". The semantic Web allow us to build ontologies by using a set of languages as RDF, RDFs and OWL to structure knowledge resources.

The creation of a domain ontology need to define in detail the concepts, the procedures, the activities and the relationships that belong specially to this domain or field trying to eliminate ambiguity and doubts due to the communication among web researches and machines (computers) using software applications, as Staab and Studer explain in (Staab and Studer, 2004).

Mathieu d'Aquin defines "Ontologies represent the essential technology that enables and facilitates interoperability at the semantic level, providing a formal conceptualization of the data which can be shared, reused, and aligned" [28]".

Elbassadi in [29] complete those definitions by explaining that the Ontologies provide a semantic representation of a common language to foster interoperability, declaratively, and intelligent services between tools and to support the innovation life cycle.

There are several existing libraries dedicated to ontology. An ontology library is a distributed data space where users and software agents can publish and access information from many different sources, the format RDF guarantees the interoperability making it possible for applications to reuse data and to link diverse data [30]. BioPortal is a library of biomedical ontologies developed by the National Center for Biomedical Ontologies, it provides essential domain knowledge to drive data integration, information retrieval, data annotation, natural-language processing and decision support [31].

B. MAS Methodologies

There are some methodologies to design a multi-agent system (MAS). These methodologies involve mainly roles, agents, interactions among agents and the environment. First methodologies were developed at the end of the century based on interaction of roles [25] but in this century

several methodologies appears such as ICTAM [32], DOCK methodology [33], MOBMAS [34], ADELFE [35], GAIA [36], [9].

Wooldridge's GAIA Methodology

The GAIA methodology is for agent-oriented analysis and design, in that it is applicable to a wide range of multi-agent systems, and comprehensive, in that it deals with both the macro-level (societal) and the micro-level (agent) aspects of systems. Gaia is founded on the view of a multi-agent system as a computational organization consisting of various interacting roles. Gaia has been specifically tailored to the analysis and design of agent-based systems [36].

DOCK Methodology

DOCK [37] helps to model a multi-agent system based in knowledge management. This methodology describes an intelligent knowledge system; it uses human organizations, roles, collaborations, skill, goals and knowledge. The methodology DOCK defines four elements: the organizational structure to identify agents (roles), the process model, the activity model and the role model. The model defines three stages: human organization, agent organization and interactions.

MOBMAS Methodology

This ontology-based methodology, used for the analysis and design of multi-agent systems [34]; MOBMAS is the first methodology that explicitly identifies and implements the various ways in which ontologies can be used in the MAS development process and integrated into the MAS model definitions. Conforming to the definition of a software engineering methodology [38], MOBMAS is comprised of a software engineering process that contains activities and associated steps to conduct the system development, techniques to assist the process steps and a definition of the models.

GAIA methodology with abstractions

This multi-agent system paradigm introduces a number of new abstractions and design/development issues. Gaia exploits the organizational abstractions to provide clear guidelines for the analysis and design of complex and open software systems [9].

SABPO Methodology

The SABPO methodology use an organizational metaphor, in which each agent plays a specific role to achieve the global goals of the organization, in addition, the methodology introduces a new interaction pattern. This approach puts the FIPA (the foundation for Intelligent Physical Agent) abstract architecture specification to the center of the methodology as a basic organizational structure and tries to create a concrete FIPA architecture that satisfies system requirements.

ADELFE Methodology

ADELFE is another methodology focus on the adaptive nature of the environment. It introduces the concepts of Non Cooperative Situations (NCS) that defines the behavior of local agents when they encounter an unpredictable situation in a dynamic environment. Rational Unified Process leads ADELFE that is devoted to software engineering adaptive MAS. ADELFE guarantees that the software is developed according to the adaptive multi-agent systems AMAS theory. The AMAS theory provides a solution to build complex systems for which classical algorithmic solutions cannot be applied, these systems are open and complex. All the interactions the system may have with its environment cannot be enumerated[35]. ADELFE is an agent-oriented methodology for designing Adaptive Multi-Agent Systems (AMAS), it is a French acronym for "Atelier de Développement de Logiciels à Fonctionalité Emergente" [39].

C. Methodologies to design ontologies

Ontology requires a well-defined process to represent the reality. Several methodologies already exist to build ontologies.

The construction of ontologies is very much an art rather than a science. This situation needs to be changed and will be changed only through an understanding of how to go about constructing ontologies. We will present in the next section some approaches to build ontologies.

Enterprise Approach

Uschold and King propose a methodology for ontology construction in [40]. This methodology according to Jones [41] bases on four steps:

- Identify the purpose determines the level of formality at which the ontology should be described.
- Identify the scope: a "Specification" is produced which fully outlines the range of information that the ontology must characterize. This may be done using motivating scenarios and informal competency questions, as in TOVE or by "brainstorming and trimming" i.e. produce a list of potentially relevant concepts and delete irrelevant entries and synonyms.
- Formalization: create the "Code", formal definitions and axioms of terms in the Specification.
- Formal evaluation: the criteria used may be general; this stage may cause a revision of the outputs of stages 2 identify the scope and 3 formalizations.

Methontology

The goal of this methodology, [42], is to clarify to readers interested in building ontologies, the activities they should perform and in which order, as well as the set of techniques to be used in each phase of the methodology. He thinks ontology is an art and tries to transform in an engineering. Here, the steps:

(1) Specification: identify the purpose of the ontology, including the intended users, scenarios of use, the degree of formality required, etc., and the scope of the ontology including the set of terms to be represented, their

characteristics and the required granularity. The output of this phase is a natural-language ontology specification document.

(2) Knowledge acquisition: this occurs largely in parallel with stage (1). It is non-prescriptive as any type of knowledge source and any elicitation method can be used, although the roles of expert interviews and analyses of texts are specifically discussed.

(3) Conceptualization: domain terms are identified as concepts, instances, verbs relationships or properties and each are represented using an applicable informal representation.

(4) Integration: in order to obtain some uniformity across ontologies, definitions from other ontologies, e.g. Ontolingua standard units ontology, should be incorporated.

(5) Implementation: the ontology is formally represented in a language, such as Ontolingua.

6) Evaluation: much emphasis is placed on this stage in METHONTOLOGY. The techniques used are largely based on those used in the validation and verification of KBSs. A set of guidelines is given on how to look for incompletenesses, inconsistencies and redundancies.

(7) Documentation: collation of documents that result from other activities.

KBSI IDEF5

This method is devoted to assist in the creation, modification and maintenance of ontologies according to Jones in [41] and [43], the process of IDEF5:

(1) Organizing and scoping establishes the purpose, viewpoint, and context for the ontology development project. The purpose statement provides a set of "completion criteria" for the ontology, including objectives and requirements. The scope defines the boundaries of the ontology and specifies parts of the systems that must be included or excluded.

(2) Data collection: the raw data needed for ontology development is acquired using typical KA techniques, such as protocol analysis and expert interview.

(3) Data analysis: the ontology is extracted from the results of data collection. First, the objects of interest in the domain are listed, followed by identification of objects on the boundaries of the ontology. Next, internal systems within the boundary of the description can be identified.

(4) Initial ontology development: a preliminary ontology is developed, which contains proto-concepts i.e. initial descriptions of kinds, relationships and properties.

(5) Ontology refinement and validation: the proto-concepts are iteratively refined and tested. This is essentially a deductive validation procedure as ontology structures are "instantiated" with actual data, and the result of the instantiation is compared with the ontology structure.

Methodology for ontology Ontolingua

Ontolingua is mechanism for writing ontologies in a canonical format, such that they can be easily translated into a variety of representation and reasoning systems. This allows one to maintain the ontology in a single, machine-readable form while using it in systems with different syntax and reasoning capabilities. The syntax and semantics are based on the KIF knowledge interchange format [44]. Ontolingua extends KIF with standard primitives for defining classes, relationships, and organizing knowledge in object-centered hierarchies with inheritance.

The methodology used to design Ontolingua was:

- A well-defined, declarative semantics for all statements in the language.
- A mechanism that allows operational use of the ontologies within in a variety of implemented representation systems.
- A syntax that facilitates the modular definition of terms in an ontology, and the modular packaging of ontologies.
- A means for capturing conventions in knowledge representation and organization, such as class hierarchies and domain and range constraints on relationships, in a system independent, declarative form without sacrificing the efficient implementation of these conventions by various representation systems.
- An architecture and support library that makes it easy to write additional KIF translators.

Methodology to design ontologies from organizational models

The phases and activities applied to creativity workshops that represent the ontology process (cf. figure IV.37), it describes an ontology to model knowledge in creativity workshop, its description:

Phase 1: Definition

- Definition of domain, the Scope and Purpose.
- Definition of the questions-skills of the ontology (aptitudes).

Phase 2: Conceptualization

- Conceptualization and acquisition.
- The reuse of existing ontology concepts

Phase 3 Development

- The development of the ontology (programming, formalization).
- Population of the ontology

Phase 4: Validation/Evaluation

- Evaluation

III. MAS AND ONTOLOGY OUR APPROACH PROPOSALS

A. Aplying MAS Methodology GAIA

The design processes of GAIA have three models, agent model, services model and acquaintance model that help us to understand the roles and interactions described before in the analysis phase. *Gaia is concerned with how a society of agents cooperate to realize the system-level goals, and what is required of each individual agent in order to do this. Actually how an agent realizes its services is beyond*

the scope of Gaia, and will depend on the particular application domain [36].

The models in GAIA are in two phases, first the analysis phase with the role model and interaction models and the second phase, the design phase with the agent model, service model and acquaintance model (cf. Figure 1).

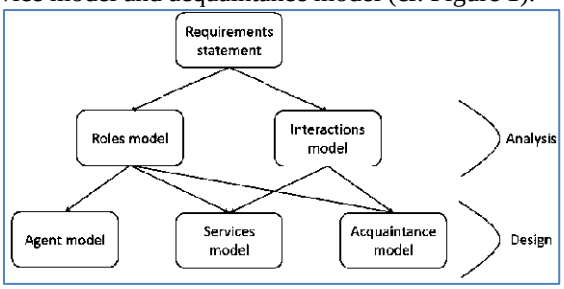


Figure 1: Relations between the GAIA's models[36] *p. 3.*

The objectives of a multi-agent system are to manage idea cards, to take decisions, and to enhance the creative techniques by means of the reactive and cognitive agents in the environment of 48H creativity workshop. The most abstract concept is the system. The organization is a collection of roles and interactions among them (Figure cf. 2). The analysis moves from abstract to concrete concepts.

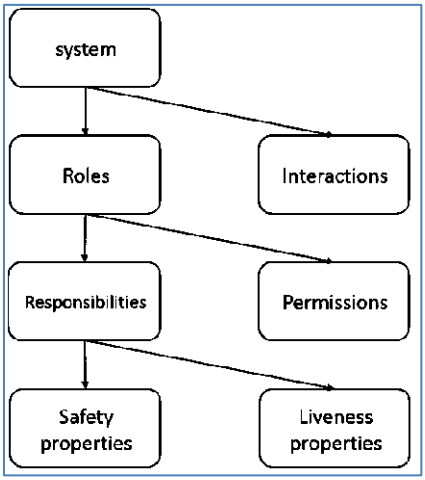


Figure 2: Analysis of concepts [36] *p. 4.*

B. Applying the ontology Uschold ontologies

The proposed methodology, to build our ontology, must help to represents the evolution of the ideas, the individual ideas, after idea cards and finally possible solutions, all this evolution in an organizational model called 48H. This methodology follows a process based primordially on building the ontology. Initially, it expresses the meaning of the organization after it works on the design of ontology, finally the judgement and documentation.

The methodology chosed is the ontology of [45]. It has four phases:

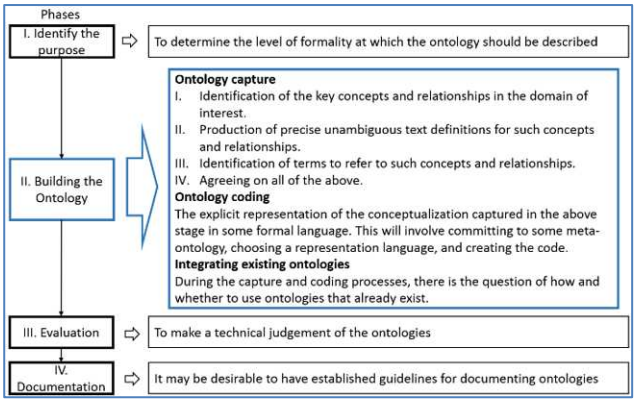


Figure 3: Ontology Uschold phases.

With this ontology, we can define easily the phases and identify the steps with the finality to do several iterations to correct our job. The Uschold ontology (cf. Figure 3) focus mainly in building the ontology that is something that we appreciate, for us the capture, coding and to integrate existing ontologies are vital steps without forget the iteration to improve.

IV. Results and Conclusion

A. Multi-agent system results

Analysis – Roles Model

We will describe the roles of the agents also we have chosen to give the same role for the agents than for the participants of the creativity workshop i.e. "creative expert", "Industrial manager", "Organizer", "solver participant" and "technical expert". The objective of the agents is to assist the participants to achieve their activities and to manage their ideas during the creativity workshop process.

The schema of solver participant (cf. Table 1) details the production of ideas an idea cards. There are fifteen protocols (showed in the table 1 at the row of Protocols and activities) where solver participant acts. The permissions are producing an individual idea using individual activities, to produce at least two idea cards using a collaborative creative method, to have the same problem in WorkIdeaCards at the time of sharing with colleague team, to evaluate idea cards from the same problem except your idea card.

Role Schema	**Name**
Description and objective	The role solver participant: To produce ideas in an individual way using activities. To produce idea cards by mean of collaborative creative methods.
Protocols and Activities	RequirementsInscription (Name, Last name, Institution), GiveRequirements (Name, Last name, Institution) Assignation (Assign_InstToWork, Assigned_ind, assigned_rol), Provides (part_team, problem) Offer_activity, SelectActivity, WorkIdeas Offer_method, SelectMethod, WorkIdeaCards, Improve, CompareIdeas, SendingIdeaCards, ReceivingPossibleSolutions, WatchingPossibleSolutions, AwardsEnd
Permissions	The actor must be register as a Solver Participant; To produce at least 1 individual idea using individual activities.

	To produce 2 Idea Cards using a collaborative creative method. To have the same problem in WorkIdeasCards at the time of sharing with colleague team. To evaluate idea cards from the same problem except your idea cards and the idea cards from your team partner.
	Responsibilities
Liveness	Solverparticipant = (RequirementsInscription.GiveRequirements)+ · (Assignation)+ · (Provides)+ · (Offer_activity.SelectActivity.WorkIdeas)+ · (Offer_method.SelectMethod.WorkIdeaCards.Improve)+ ·(CompareIdeas)+ · (SendingIdeaCards.ReceivingPossibleSolutions)+ · (WatchingPossibleSolutions.AwardsEnd)+
Safety	Idea > 0 Idea Card = 2 by team.

Table 1: Role Solver Participant.

Analysis – Interaction Model

This second model describes the communication's protocols for each agent. The agent's protocol has some elements (cf. Figure 4) that help us to improve the explanation about the protocol's description (Colleman et al. 1996) in the interaction of the agents.

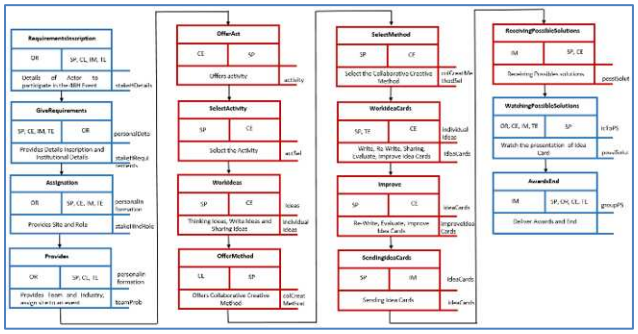


Figure 4: Definition of protocols associated with Role Solver Participant.

Design – Agent Model

In the agent model (cf. Figure 5), we identify seven agents and their instances that will make up the system. During the creativity workshop, it identifies easily the five roles according to the agent model proposed. The role creative expert and Technical Expert will form an agent called Creative Technological Expert Agent (CTEAgent-CTEA), the number of CTEA agents are 1 or more. The rest of the roles have their agents, making note that the agent Organizer (ORAgent-ORA) has one or more instances. However, we add three agent, semantic model knowledge agent (SMKA), width semantic distance agent (WSDA) and the comparative similarity agent (CSA). These agents help us to order ideas according to semantic distance, width-density and comparative similarity of Idea Cards. The definitions of the agents are:

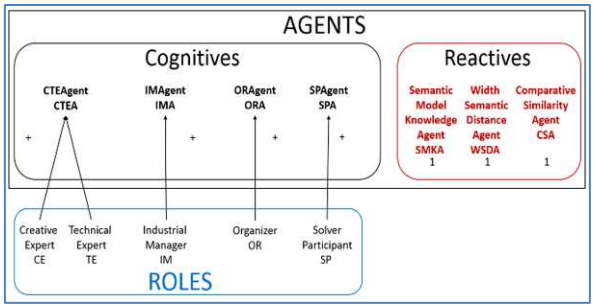


Figure 5: The Agent Model.

Design – Service Model

The services model identifies the main services that are required to realize the agent's role. These services (cf. Table 2) are functions that the agents have to execute according to the protocols described before.

Service	Inputs	Outputs	Pre-condition	Post-condition
Obtain information of actors and assignation of roles	Actor details Name, last name, institution, sex, date of birth	Actor requirements	Event=1	Institution>=1 Industry>=1 Role>=2 Team>=2 Problem>=1
Selection and application of activity for ideas	Group, Creative Technical Expert, Activities	Ideas	Ideas per participant at least in mind	Idea>0
Selection and Application of Methods for idea Cards	Thousands of Ideas, many methods	Idea Cards	2 Idea Cards per group =2	Idea Cards >2
Evaluation by partners and improving idea card as a goal	Two ideas per group	Idea Cards	2 Idea Cards per group	Idea Cards > 2
Classification of Idea Cards	n Idea Cards	n Idea Cards	At least 2 Idea Cards	Idea Cards >n
Sending Possible Solutions	Idea Cards	Possible Solutions	2 possible solution per group	Possible solution >=2

Table 2: Service Model.

Design – Acquaintance Model

The acquaintance model, (cf. Figure 6), documents the lines of communication among the different agents. The agents Creative Expert CTEA, Organizational ORA, Solver Participant SPA and Industrial Manager have communication among them during the entire creativity

workshop, the agents CTEA and SPA has relation with the agents Semantic Distance SDA, Width Density WDA and Comparative Similarity CSA. The agents SDA, WDA and CSA take action with the purpose of classify the ideas at the end of the creation, evaluation (among partner group and the rest of the groups) and improving.

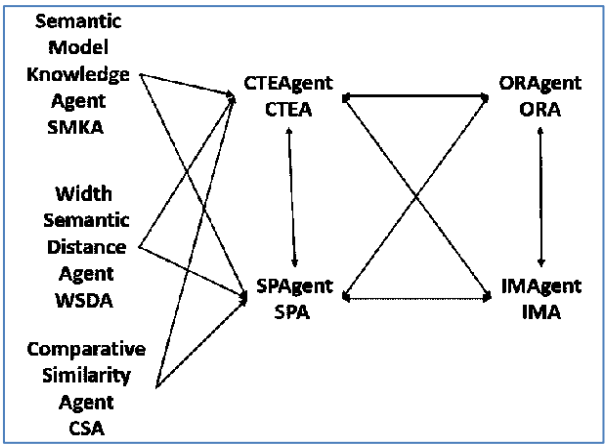


Figure 6: The Acquaintance Model.

B. Ontology results

The Domain is an ontology dedicated to the creqtivity and we called it "The Collaborative Creative Ideas Ontology " CCIDEAS.

The scope describes, specifies and represents all the concepts relate to the creativity workshop 48 hours challenge. These concepts are identified inside of an organizational model.

The purposes of this ontology are to represent knowledge and ideas and to understand the creativity workshop. The ontology will be used to compare ideas in the intelligent system.

Creation of concepts

These concepts and relationships determine the ontology. The definition of every concept that are part during this creativity workshop (*cf.* Table IV.16); this definition will be used to create triplets among concepts.

Name of concept	Definition
1 Activity	The action(s) that actor follows to produce ideas (I) in the phase of divergence. The activities of divergence used to produce individual ideas. Type: String
2 Actor	The person that will participate in the event and can take a role to solve problems. The concept will indicate the role or several roles to assume during a CWS. Type: String
3 Collaborative creative method (CCM)	Set of instructions applied by solver participants to generate idea cards (IC). The methods of convergence are used to produce Idea cards. Type: String.
4 Event	The name of the CWS and its edition. Type: String
5 Idea	The individual ideas (I) are produced in the phase of divergence; they are created using combinational, exploratory and transformational techniques with individual activities. The actor captures the initial individual idea. Type: String
6 IdeaDesc	Idea's description. Type: String
7 IdeaCard	The result of the use of ideas (I) and collaborative creative methods (CCM). An actor with the role of solver participant and from a team creates idea cards using CCM.
8 ICDesc	Idea Card's description. Type: String.
9 ICTitle	Idea Card's title. Type: String.
10 ICScenery	Idea Card's scenery. Type: String.
11 ICPrioCli	Idea Card's priority client. Type: String.
12 ICAdvant	Idea Card's advantage. Type: String.
13 ICRisk	Idea Card's risk. Type: String.
14 Industry	The name of the industry, this concept has the problem. The industrial manager proposes the industry that contains the problem. Type: String
15 Problem	The reasons why the organization creates the creativity workshop CWS events every year. Industries like Assystem, Bostik, CEA Tech, Decathlon, GRDF, ICM, MSA Safety, Muller, Normande Aerospace, Pierre Fabre, and Scarabée Biocop participate in those events. The problem is assigned to a team by mean of the industry. Type: String
16 Role	The type of character that the actor takes. (Organizer, solver participant, creative expert, technical expert, Industrial manager). Type: String
17 Site	The place where actors will work (ASU BAHRAIN, CESI NANTERRE, ENSGSI NANCY, etcetera). The site given to an actor. Type: String
18 Team	The set of actors with the same role (solver participant), event, problem, site, name of team and colour. (Str_Ass_1, Lyo_Assy_1, Uca1, Str_Ass_2, etcetera). An actor takes part of a team. Type: String
19 Vocabulary	The vocabulary is formed by ADJECTIVE, ADVERB, NOUN, VERB, ARTICLE, PRONOUN, PREPOSITION, CONJUNCTION AND INTERJECTION; some fields of the idea card's concept use these concepts such as title, description, priority client, name, scenery, advantages and risk. The vocabulary is part of idea's description. Type: String
20 Organizer	The actor who takes the role in activities of assignation of roles, industries, team and event. Initially, this role create the event and he is asking for information to the actors with the purpose to do the inscription. Type: String

21 Solver Participant	This concept is part of the roles and part of the team. The site is assigned to the solver participant; other relationships are: Select activity, Create ideas, Send possible solutions. Type: String
22 Creative Expert	The creative expert offers activities and collaborative creative methods to the team of solver participants. Type: String
23 Technical Expert	Technical expert helps teams to improve idea cards. Type: String
24 Industrial Manager	The industrial manager proposes an industry to the creativity workshop. Type: String
25 Possible Solutions	The concept possible solutions are the idea cards with best score according to the semantic skills like width, semantic distance and similarity.

Table 3: Definition of concepts.

Creation of relationship proposed

The creation of relationships uses the format subject-verb-object with the purpose of create sets of three elements; *the term of relationships the concepts (Bachimont 2000) where the relationship represents the verb* (*cf.* Table 4 in the Annexe). Finally, we present the global ontology figure 7 in the annexe too.

V. CONCLUSION

A. Multi-agent system and ontology

Contributions

Our main contribution with the MAS and ontology is to assist participants of the creativity workshop to manage their ideas and knowledge and to propose an intelligent system based in multi-agent and an to develop an annotation system (ontology).

The agents inside MAS manage ideas but probably in the future other agents could be focused to assist participants in the selection of a better activity or collaborative creative method.

References:

[1] "Ecole d'été RRI 2017, les 28-29 août 2017: ' L'innovation agile : Quels défis pour les individus, les organisations et les territoires ?' - soutenue par IDEFI - InnovENT-E Nancy, Grande Région Est, France. [Online]. Available: https://rni2017.event.univ-l."

[2] D. Weyns, V. D. H. Parurak, and F. Michel, *Environments for Multi-Agent Systems, Lecture Notes in Artificial Intelligence*. 2004.

[3] A. Drogoul, J. Ferber, and C. Cambier, "Multi-agent Simulation as a Tool for Analysing Emergent Processes in Societies," pp. 49–63, 1992.

[4] O. Gutknecht and J. Ferber, "MadKit: A generic multi-agent platform," *Proc. fourth Int. Conf. Auton. agents - AGENTS '00*, pp. 78–79, 2000.

[5] O. Simonin and J. Ferber, "Un modèle multi-agent de résolutio collective de problèmes situés multi-échelles," pp. 1–13, 2003.

[6] W. Zhang, W. Liu, X. Wang, L. Liu, and F. Ferrese, "Distributed multiple agent system based online optimal reactive power control for smart grids," *IEEE Trans. Smart Grid*, vol. 5, no. 5, pp. 2421–2431, 2014.

[7] J. Ferber, "Coopération réactive et émergence," pp. 19–52, 1994.

[8] M. Wooldridge, M. Fisher, M. Huget, and S. Parsons, "Model Checking Multi-Agent Systems with MABLE ∗," pp. 952–959, 2002.

[9] F. Zambonelli, N. R. Jennings, and M. Wooldridge, "Developing Multiagent Systems: The Gaia Methodology," *ACM Trans. Softw. Eng. Methodol.*, vol. 12, no. 3, pp. 317–370, 2003.

[10] J. Girodon, D. Monticolo, E. Bonjour, and M. Perrier, "An organizational approach to designing an intelligent knowledge-based system: Application to the decision-making process in design projects," *Adv. Eng. Informatics*, vol. 29, no. 3, pp. 696–713, 2015.

[11] S. Esparcia, E. Argente, and V. Botti, "An agent-oriented software engineering methodology to develop adaptive virtual organizations," *IJCAI Int. Jt. Conf. Artif. Intell.*, no. i, pp. 2796–2797, 2011.

[12] E. Argente, V. Julian, and V. Botti, "MAS Modeling Based on Organizations," *Agent-Oriented Softw. Eng. IX 9th Int. Work.*, pp. 16–30, 2009.

[13] W. Shen and D. H. Norrie, "Agent-Based Systems for Intelligent Manufacturing: A State-of-the-Art Survey," *Knowl. Inf. Syst.*, vol. 1, no. 2, pp. 129–156, 1999.

[14] L. Vanhée, J. Ferber, and F. Dignum, "Agent-Based Evolving Societies (Extended Abstract) Categories and Subject Descriptors," pp. 1241–1242, 2013.

[15] D. Kinny, "A Visual Programming Language for Plan Execution Systems," *Proc. First Int. Jt. Conf. Auton. Agents Multi-Agent Syst. (AAMAS-2002, Featur. 6th AGENTS, 5th ICMAS 9th ATAL), 15--19 July, Bol. Italy*, pp. 721–728, 2002.

[16] C. Castelfranchi, "Guarantees for autonomy in cognitive agent architecture," pp. 56–70, 2012.

[17] N. R. Jennings and M. Wooldridge, *Applications of Inteligent Agents. Agent Technology: Formations, Applications and Markets.* 1998.

[18] P. Øhrstrøm, H. Shârphe, and S. Uckelman, "Jacob Lorhard's Ontology: A 17th Century Hypertext on the Reality and Temporality of the World of Intelligibles," in *Conceptual Structures: Knowledge, Visualization and Reasoning.Lecture Notes in Artificial Intelligence.*, 2008, pp. 74–87.

[19] T. R. Gruber, "Toward Principles for the Design of Ontologies," *International journal of human-computer studies*, vol. 43, no. 5. pp. 907–928, 1995.

[20] M. R. Genesereth and N. J. Nilsson, “Logical Foundations of Artificial Intelligence: Nomotonic reasoning,” *Logical Foundations of Artificial Intelligence*. 1987.

[21] K. Mahesh, “Ontology development for machine translation: Ideology and methodology,” *Comput. Res. Lab. New Mex. State Univ. MCCS-96-292*, p. 87, 1996.

[22] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach.*, 2nd. Ed. 2003.

[23] V. Decugis and J. Ferber, “Action selection in an autonomous agent with a hierarchical distributed reactive planning architecture,” *AGENTS ’98 Proc. Second Int. Conf. Auton. agents*, pp. 354–361, 1998.

[24] J. Ferber, *Les systèmes Multi-Agents vers une Intelligence Collective*, InterEditi. 1995.

[25] M. Wooldridge, N. R. Jennings, and D. Kinny, “A Methodology for Agent-Oriented Analysis and Design: A Conceptual Framework,” *Proc. Third Int. Conf. Auton. Agents*, pp. 69–76, 1999.

[26] J. O. Kephart and W. E. Walsh, “An artificial intelligence perspective on autonomic computing policies,” *Proc. - Fifth IEEE Int. Work. Policies Distrib. Syst. Networks, POLICY 2004*, pp. 3–12, 2004.

[27] T. Berners-Lee and J. Hendler, “The Semantic Web,” *Sci. Am.*, vol. 21, 2002.

[28] M. D’Aquin and N. F. Noy, “Where to publish and find ontologies? A survey of ontology libraries,” *J. Web Semant.*, vol. 11, pp. 96–111, 2012.

[29] L. Elbassiti and R. Ajhoun, “Semantic Representation of Innovation, Generic Ontology for Idea Management,” *J. Adv. Manag. Sci.*, vol. 2, no. 1, pp. 128–134, 2014.

[30] M. D’Aquin and N. F. Noy, “Where to publish and find ontologies? A survey of ontology libraries,” *J. Web Semant.*, vol. 11, pp. 96–111, 2012.

[31] N. F. Noy *et al.*, “BioPortal : ontologies and integrated data resources at the click of a mouse,” vol. 37, no. May, pp. 170–173, 2009.

[32] S. Elsawah, J. H. A. Guillaume, T. Filatova, J. Rook, and A. J. Jakeman, “A methodology for eliciting, representing, and analysing stakeholder knowledge for decision making on complex socio-ecological systems: From cognitive maps to agent-based models,” *J. Environ. Manage.*, vol. 151, pp. 500–516, 2015.

[33] J. Girodon, D. Monticolo, and E. Bonjour, “How To Design a Multi Agent System Dedicated to Knowledge Management ; the DOCK Approach,” pp. 113–121, 2015.

[34] Q. N. N. Tran and G. Low, “MOBMAS: A methodology for ontology-based multi-agent systems development,” *Inf. Softw. Technol.*, vol. 50, no. 7–8, pp. 697–722, 2008.

[35] G. Picard and M. Gleizes, “THE ADELFE METHODOLOGY,” in *Methodologies and Software Engineering for Agent Systems - The Agent-Oriented Software Engineering Handbook*, 2006, pp. 1–2.

[36] M. Wooldridge, N. R. Jennings, and D. C. N.-W.-2000-01 Kinny, “The Gaia Methodology for Agent-Oriented Analysis and Design,” *J. Auton. Agents Multi-Agent Syst.*, vol. 3, no. 3, pp. 285–312, 2000.

[37] A. Gabriel, “Gestion des connaissances lors d’un processus collaboratif de créativité. Université de Lorraine, ERPI Lab. Décembre 2016.,” 2016.

[38] B. Henderson-Sellers, A. Simmons, and H. Younessi, *The OPEN Toolbox of Techniques*. UK, 1998.

[39] S. Rougemaille, J. P. Arcangeli, M. P. Gleizes, and F. Migeon, “ADELFE design, AMAS-ML in action: A case study,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5485 LNAI, pp. 105–120, 2009.

[40] M. Uschold and M. King, “Towards a Methodology for Building Ontologies,” *Methodology*, vol. 80, no. July, pp. 275–280, 1995.

[41] D. Jones, T. Bench-Capon, and P. Visser, “Methodologies for Ontology Development,” no. November 2012, 1998.

[42] M. Fernandez, A. Gomez-Pérez, and N. Juristo, “METHONTOLOGY : From Ontological Art Towards Ontological Engineering,” p. 40, 1997.

[43] B. Peraketh *et al.*, “Ontology Capture Method (IDEF5),” 1994.

[44] M. R. Genesereth, R. E. Fikes, and T. Gruber, “Knowledge Interchange Format,” *Interchange*, no. January, 1992.

[45] M. Uschold and M. King, “Towards a Methodology for Building Ontologies,” *Methodology*, vol. 80, no. July, pp. 275–280, 1995.

Annexe

Relation name proposed	Domains (Concepts)	Range (Concepts)	Triplet and/or Definition
1Select	Solver Participant	Activity. Examples: Brainstorming, write storming, Bend it and Shape it, Brain borrow, Copy cat, …	Solver Participant selects Activity. The property indicates that the Solver Participant select an Activity to create individual ideas during the phase of divergence
2 Offers	Creative Expert	Activity. Examples: Brainstorming, write storming, Bend it and Shape it, Brain borrow, Copy cat, …	Creative Expert offers Activity. The property indicates that Creative offers an activity.
3 Plays	Actor	Role. Examples: Creative Expert, Technical Expert, Industrial Manager, Solver Participant and Organizer	Actor plays a Role.
4 Assign	Organizer	Role. Examples: Creative Expert, Technical Expert, Industrial Manager and Solver Participant.	Organizer assign Role. The property indicates that Organizer assigns all the roles in the creativity workshop.
5 Propose	Industrial Manager	Industry. Examples: Decathlon, ICM, Bostik, etc.	Industrial Manager proposes an industry. The property indicates that Industrial Manager proposes an industry.
6 Create	Organizer	Event. Examples: 48h InnovENT-Edition 2016, Operation 2015 InnovENT-E 48 hours to bring ideas to life.	Organizer creates an Event.
7 Assign	Organizer	Site. Examples: INSA LYON, ENSGSI, UCA MARRAKECH, etc.	Organizer assigns Site.
8 Assign	Organizer	Industry; Examples: Examples: Decathlon, ICM, Bostik, etc.	Organizer assigns Industry.
9 Assign	Organizer	Team. Examples: Nan_Dec_1, Nan_Dec2, Str_Ass_2, etc.	Organizer assigns Teams.
10 Requires	Organizer	Actor. Examples: Any institutional, educative or industrial person interested in creativity and solving problems.	Organizer requires Actor.
11 Help	Technical Expert	Team. Examples: Nan_Dec_1, Nan_Dec2, Str_Ass_2, etc.	Technical Expert helps Team. The property indicates that Technical experts helps teams.
12 IsAssignedTo	Industry	Team. Examples: Nan_Dec_1, Nan_Dec2, Str_Ass_2, etc.	Industry is assigned to Team.
13 Receive	Industrial Manager	Possible Solutions	Industrial Manager receives Possible Solutions.

			The property indicates that Industrial Manager receives the possible Solutions.
14 IsPartOf	Actor	Team. Examples: Nan_Dec_1, Nan_Dec2, Str_Ass_2, etc.	Actor is part of Team.
15 IsAssignedTo	Site	Event. Examples: 48h InnovENT-Edition 2016, Operation 2015 InnovENT-E 48 hours to bring ideas to life.	Site is assigned to an Event.
16 Send	Solver Participant	Possible Solutions	Solver Participant sends Possible Solutions. The property indicates that Solver Participant sends the possible Solutions.
17 Present	Team	Possible Solutions	Team presents Possible Solutions. The property indicates that Team presents the possible Solutions.
18 IsAssignedTo	Site	Role. Range: Technical Expert, Solver Participant and Creative Expert.	Site is assigned to Role.
19 IsAssignedTo	Team	Role. Range: Technical Expert, Solver Participant and Creative Expert.	Team is Assigned to Role.
20 Create	Team	Idea Card	Team creates Idea Card. The property indicates that Team creates the Idea Cards.
21 Improve	Team	Idea Card	Team improves Idea Card. The property indicates that Team improves the Idea Cards.
22 Select	Team	CCM. Examples: Six hats of thinking, The shirt off your back, Puzzle pieces, Organizational brainstorms, Best off, Rice storm, …	Team select CCM. The property indicates that Team selects the Collaborative Creative Method.
23 Use	CCM	Idea.	CCM uses Ideas.
24 Form	Idea	Idea Cards	Ideas form Idea Card.
25 Use	Idea Card	CCM. Examples: Six hats of thinking, The shirt off your back, Puzzle pieces, Organizational brainstorms, Best off, Rice storm, …	Idea Card uses CCM.
26 Offer	Creative Expert	CCM. Examples: Six hats of thinking, The shirt off your back, Puzzle pieces, Organizational brainstorms, Best off, Rice storm, …	Creative Expert offers CCM.
27 IsPartOf	IdeaDesc	Idea	IdeaDesc is part of Idea. The property indicates that Idea Description (IdeaDesc) is part of the Idea.
28 Create	Solver Participant	Idea	Solver Participant creates Idea. The property indicate that Solver Participant creates ideas.
29 IsPartOf1	ICDesc	Idea Card	ICDesc is part1 of idea card. The property indicates that the field Idea Card Description (ICDesc) is part of the Idea Card.
30 IsPartOf2	ICTitle	Idea Card	ICTitle is part2 of idea card. The property indicates that the field Idea Card Title (ICTitle) is part of the Idea Card.
31 IsPartOf3	ICScenery	Idea Card	ICScenery is part3 of idea card.

			The property indicates that the field Idea Card Scenery (ICScenery) is part of the Idea Card.
32 IsPartOf4	ICPrioCli	Idea Card	ICPrioCli is part4 of idea card. The property indicates that the field Idea Card Priority Clients (ICPrioCli) is part of the Idea Card.
33 IsPartOf5	ICAdvant	Idea Card	ICAdvant is part5 of idea card. The property indicates that the field Idea Card Advantage (ICAdvant) is part of the Idea Card.
34 IsPartOf6	ICRisk	Idea Card	ICRisk is part6 of idea card. The property indicates that the field Idea Card Risk (ICRisk) is part of the Idea Card.

Table 4: Definition of relationships

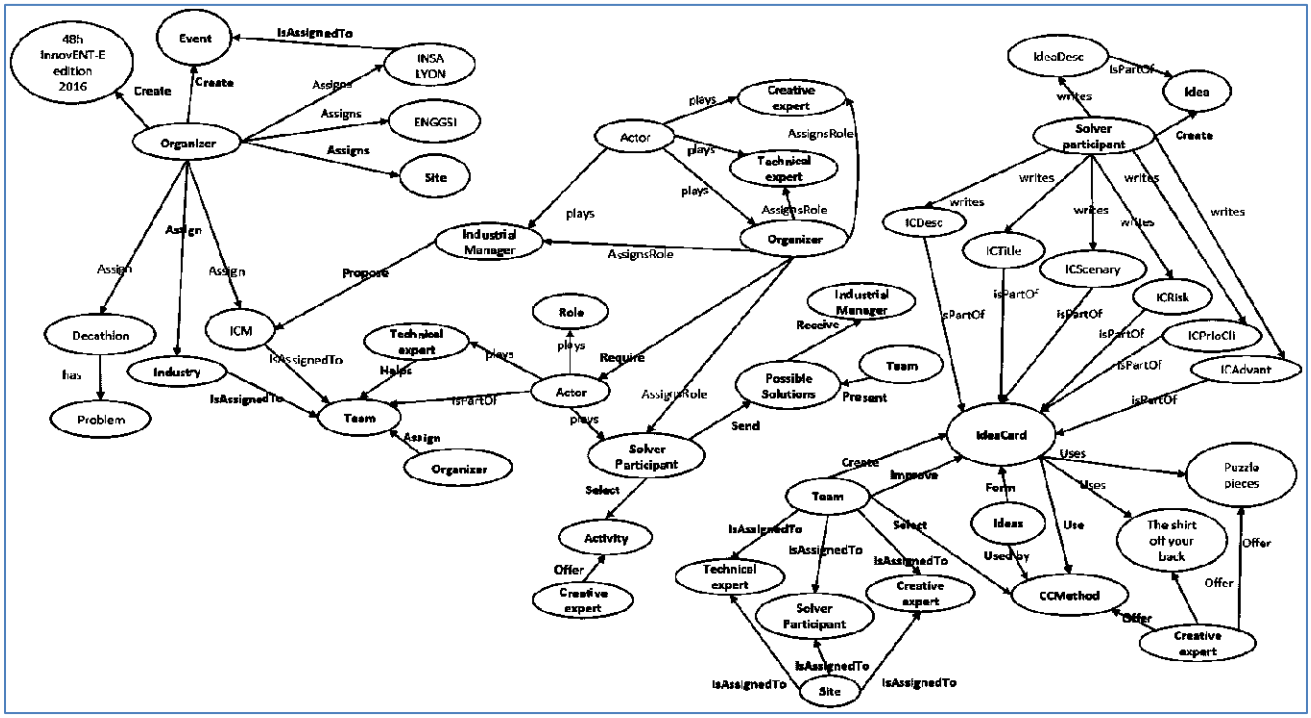


Figure 7: Global Ontology.

Social networking application, visual communication system for the protection of personal information

Marilou Kordahi

Abstract— We contribute to the field of Information Systems by attempting to develop a first innovative social networking application, the “SignaComm”. The SignaComm's objective is to enable multilingual communication between users worldwide for the protection of personal data on the Web. We design this application while relying on the theory of patterns as well as the principles of ontologies and signage system. The theory of patterns presents good practices for creating a model, which describes the characteristics of a generic solution to a specific problem. It permits the reuse and remodelling of patterns to serve as resources for software development and problem solving. Ontologies describe a structured set of concepts and objects by giving a meaning to an information system in a specific area, and allow the construction of relationships between these concepts and objects. The signage system is a visual communication system with an international vocation where, the “signagram” is the writing unit. When creating the SignaComm we use an automatic translation software of key-phrases into signagrams. The social networking application is written with the PHP and Javascript programming languages and then tested technically. We hope that users from any culture, social environment or with disabilities could use it.

Index Terms— Communications applications, picture/image generation, information resource management, information systems, social networking

1 Introduction

In this paper, we contribute to the field of Information Systems by analysing the connections between Social Networking Sites (SNS) and artificial communication systems (e.g., visual communication systems). Several Web applications, namely Social Networking Sites and virtual communities, are currently using artificial visual communication systems to facilitate interactions and knowledge exchange between different users and members worldwide. This may be justified by the emergence of global social developments [1] as well as an available international audience. For example, one can notice the presence of visual communication systems in the Web applications' user-interface and dictionary of emoticons [2].

In the following paragraphs we will introduce the SNS as well as artificial communication systems.

Social Networking Sites: An SNS is an online information system for building social relationships between individuals (or organisations) sharing interests, activities, contacts in real life [4], [5], [6], [7], [8], [9]. The growth of these social ties (strong or weak ties) can only take place if these individuals (or organisations) have become members of the SNS [8], [10]. The information exchange may be done through instant messaging, emailing, voice recording, posting.

In this paper, we will attempt to design the “SignaComm”, a first SNS with an internationally oriented communication system for the protection of personal data on the Web. The SignaComm will be informative [8], [11]. It will execute two functions dynamically and in real time. Firstly, it will translate the member's input text into “signagrams” and deliver the result to another member. Secondly, it will display the history of instant messages in the chat room page. Our SNS would be used to deliver information to be understood and used quickly by its members. We hope that users from any culture, social environment or with disabilities could use it. The protection of personal data is defined by laws and regulations prohibiting the processing, storage or sharing of certain types of information about individuals without their knowledge and consent (e.g., analysing user's behaviour on a Website) [12].

Artificial communication systems: A number of artificial communication systems have been developed to improve the management of information, regardless of a specific natural language (e.g., Universal Playground, Istotype) [13], [14], [15]. We have been interested in the signage system, an artificial visual communication system with an international vocation where, the “signagram” is the writing unit [16]. The signagram's type is figurative as it is created from a direct representation of the object that evokes the object or situation to be represented [17]. Each signagram is made of an “external shape” (including the contours) and an “internal shape” [16] (Fig. 1).

The signage system and the signagram (the signage's unit) [16] will be integrated in our SNS, the SignaComm, to enable internationally oriented communication.

This paper's goal is to present the preliminary results of work in progress on the creation of the “SignaComm”. This SNS would support multilingual communication between users worldwide for the protection of personal data on the Web.

We design the SignaComm while relying on a theory and two principles: the theory of patterns [18], [19], [20], [21], as well as the principles of ontologies [22], [23], [24] and signage system [16]. At the core of Alexander's theory, a pattern describes the characteristics of a generic solution to a specific problem (e.g., the communication in real time between users worldwide). The theory of patterns permits the reuse and remodelling of patterns to serve as resources for software development and problem solving. According to Alexander [18, p. 313], “each pattern sits at the centre of a network of connections which connect it to certain other patterns that help to complete it”. The network of these relationships between small and large patterns creates the pattern. The ontology describes a structured set of concepts and objects by giving a meaning to an information system in a specific area (e.g., the user profile), and allows the construction of relationships between these concepts and objects [22], [23], [24].

The SignaComm could be implemented in the structure of a company's or public organisation's information system. Many fields may be interested in this SNS, for example, the cybersecurity, serious games, online learning. In our case, we are interested in the field of administrative authorities, namely the National Commission for Informatics and Liberty (in French, *Commission nationale de l'informatique et des libertés (CNIL)*) [25]. The CNIL is responsible for monitoring the data protection of professionals and individuals. We will explain the approach followed to develop the SignaComm for the protection of personal data when there may be a breach of privacy rights (e.g., the email advertising).

Our work consists of six sections. In section 2, we will present previously published works. In section 3, we will explain the SNS' characteristics and then design its pattern. In section 4, we will design the pattern for the automatic translation of text phrases into signagrams for the protection of personal data. In section 5, we will develop and test the prototype application that executes the SignaComm and communicates in visual messages, using the signage system and translation software of key-phrases into signagrams. In section 6, we will discuss the overall approach and finally conclude our work.

2 Related works

To our best knowledge, research projects addressing both topics, the SNS for multilingual communication and the protection of personal data, are limited. However, research projects are conducted on SNS combined with instant messaging and translation and, automatic translation of text phrases into signagrams. We will use our studies over these related works to fine-tune this research project and create the SignaComm.

In their published works, Yang and Lin [26] and Seme [27] have respectively developed a system and patent to automatically translate and send instant messages between members who communicate in different languages. Members, engaged in a session of instant messages, could send a message in a source language that could be translated automatically and received in a target language. The translation process has followed the Natural Language Processing (NLP) approach.

We have published works regarding a social networking site for crisis communication [28]. The objectives have been to translate in real time a sequence of syntagms into a series of signagrams, and to facilitate communication between members around the world. This SNS has translated automatically a source text into a target text (e.g., a message from the French language to the signage system) and has displayed the results in the SNS. The SNS has been based on the principles of the signage system, modular architecture and ontologies.

We have designed a software to automatically translate an input text into a sequence of signagrams [29]. We have relied on the semantic transfer method [30] with the linguistic rules and dictionaries for the source language and target communication system. The input has been the source text and has been written in the user's preferred language. The output has been the target text and has been written in the visual communication system, signage.

We rely on our works [16], [31] to show an example of signagram representing the syntagm *identify partners and data recipient* [25] (Fig. 1).

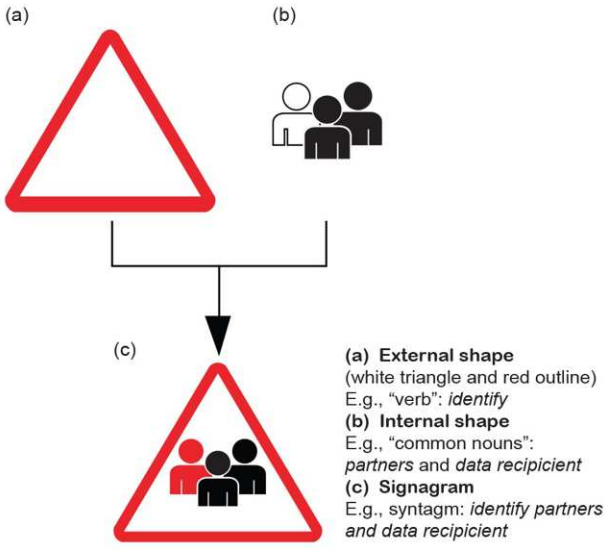


Fig. 1. Example of signagram.

3 Pattern for the SignaComm, first approach

We rely on the writings of Alexander [21] to create a pattern for the SignaComm [18]. This pattern is a first approach. The section consists of two main paragraphs, the descriptions of the SNS' context and design.

3.1 Description of the SignaComm's context

In general terms, the SNS interface follows the universal design principles of simplicity, flexibility and accessibility of use [32]. In addition, an SNS interface is graphical and contextual [33], [34]. Its graphical nature is based on a template that meets already defined and precise rules to ensure homogeneous and uniform results. These rules are the following: a simple and figurative content, uniqueness of graphical representations and uniqueness of colour contents [16], [28]. As for the dictionary of emoticons, it contains emotion symbols that are used worldwide.

So far, we haven't found published works regarding the standardisation of visual communication systems for SNS. A number of companies have developed their own communication system to integrate it into the SNS interface (e.g., Graphical User Interface (GUI) of WhatsApp) and the new technology tools (e.g., GUI of Apple iPhone). The company's (or organisation's) aim may be to intuitively guide users in their actions in entirely different and various contexts [33]. Each company (or organisation) chooses to adapt the charter of its visual communication system and its corresponding tools (e.g., SNS and new applications) according to the targeted countries. This adaptation approach may include the countries' laws, cultures, customs and traditions. Social media applications (GUI and emoticon dictionaries included) are essentially altered for two reasons. Firstly, to be compatible with international standards and regulations defined by every country government. And secondly, to meet the universal design principles depending on users' cultures.

To fulfil its objective, the SignaComm for the protection of personal data should utilise the signage system [16] as well as the graphical and contextual interface [33], [34]. The latter

should be meeting the universal design principles [32]. We have chosen both criteria to ensure the SignaComm's perception and spontaneous understanding worldwide. Furthermore, the SignaComm's development is in relation to three main concepts: the signage system [16], the SNS new technology tools [7] and user's adaptation. This interrelation makes the SignaComm dependent on these social, technical and human environments. For now, we will not include the emotional aspect as this SNS is an informative one.

3.2 Description of the SignaComm's pattern

The pattern for the SignaComm holds a network of connections between large and small patterns. In this work, we will present twelve large and two small patterns (in total 14 patterns) [19], [21]. The description is divided into several stages. A diagram will follow the explanations (Fig. 2).

We start with pattern 1 (Larger environments) that refers to many environments influencing the growth of SNS, such as information and communication technologies as well as social environments [6], [7].

Pattern 2 (Virtual communities environment) is contained inside pattern 1. Pattern 2 holds pattern 10 (SignaComm community) [6], [7].

Pattern 10 contains and describes the SignaComm functionalities (pattern 11), information technology administration (IT administration) (pattern 40) and interface (pattern 100) [3], [8].

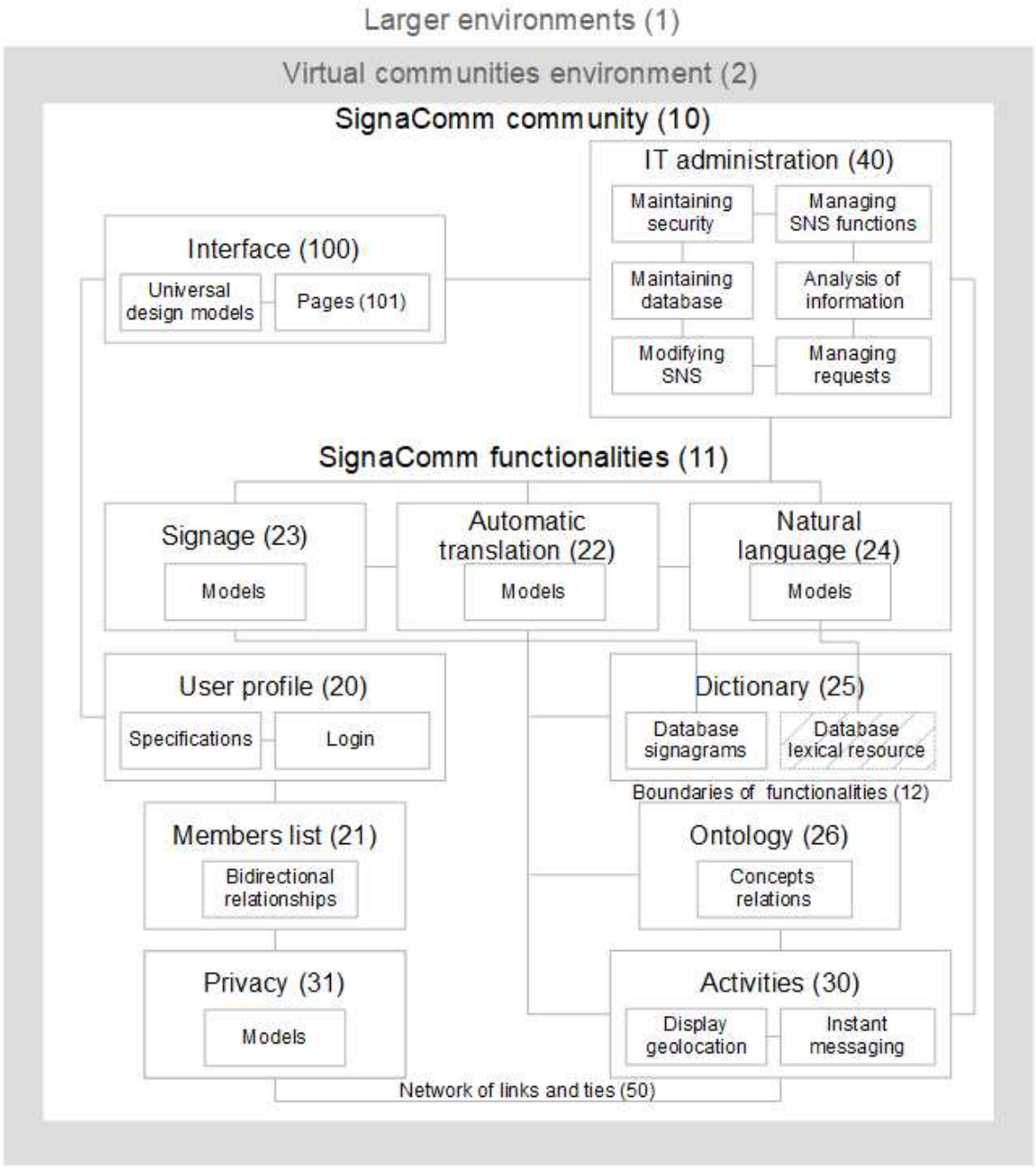


Fig. 2. Diagram for the SignaComm pattern.

Pattern 11 has various functionalities, listed as follow: the automatic translation of syntagms into signagrams (pattern 22), signage and signagrams models (pattern 23), natural languages and linguistic rules (pattern 24), dictionary (pattern 25), ontology (pattern 26), user profile characteristics and members' list (patterns 20 and 21), activities (pattern 30), privacy (pattern 31) [3], [8], [16], [30]. Every functionality has its own programming functions.

The SignaComm community (pattern 10) requires that all functionalities (pattern 11) execute their tasks to ensure the smooth running of the SNS. Pattern 12 (Boundaries of SignaComm's funtionalities) establishes boundaries to each functionality allowing it to perform its assigned tasks. It avoids the overlap with other functionalities.

Pattern 20 relates to user's profile characteristics [35], [36]. The SignaComm community encourages the diversity of members in order to enrich its growth [3]. Therefore, the growth of the SNS depends on a well balanced and represented community of members. This community would be able to support the interactions (pattern 30) between its members. For example, the interactions would help a member to solve a situation [3].

Pattern 20 has links with pattern 21 (Members list). The latter pattern specifies the SignaComm target audience (e.g., the bidirectional relationships) [37]. Members would belong to different cultures and social classes as well as different age groups [3].

Pattern 22 (Automatic translation) is a central functionality as it facilitates the communication between SignaComm members (pattern 21). Pattern 22 relies on the signage and signagrams, natural languages and linguistic rules, ontology as well as dictionary to translate members' requests (pattern 30).

Pattern 30 (Activities) is mostly linked to patterns 20, 21, 22 and 31 to create nodes of activities thus allowing members or groups to engage in various ways [7]. These activities may include invitations to join the SignaComm, instant messaging. Here, members have the opportunity to make acquaintances and connections, as well as to chat with members and groups of their choice [7]. Depending on the proximity of members, some ties are strong while others are weak [8], [10].

Pattern 31 (Privacy) is mainly for patterns 2, 20, 21, 22, 30 and 40. This pattern allows every member to set her/ his data sharing options with the IT administration, members and SNS environment [9], [12], [35], [36], [37], [38]. We provide the following example: a member chooses not to publicly display her/ his profile and then, not to share her/ his geographical position with the SignaComm and its environment. The SignaComm community (pattern 10) must respect the member's choice [12].

Pattern 40 (IT administration) is connected to both patterns 11 (SignaComm functionalities) and 100 (Interface). To make the interface and functionalities real, it is necessary to set up an IT administration. The latter manages the database and security, modifies the SNS, analyses the generated information and answers to members' requests.

Pattern 50 (Network of links and ties) creates and manages the network of relationships between all the patterns [7], [8]. It allows the information to circulate instantly and correctly in the SignaComm community.

Pattern 100 (Interface) gives an overview of the SignaComm interface, with an emphasis on the space of exchange between SignaComm members. The SNS' functionalities and IT administration contribute to its design [33], [34]. It includes the universal design principles [33].

Pattern 101 (Pages) is the continuation of pattern 100. The SignaComm is created with a reduced number of pages, such as the registration, members and chat room pages. This design is followed to quickly access information, provide flexibility in use and initiate intuitive interactions [32].

Fig. 2 shows an overall view of the SignaComm's pattern. It includes the fourteen patterns. We show the main links between the patterns to simplify the diagram's representation.

4 From text phrases to signagrams for the protection of personal data, first approach

Once we have designed the SignaComm pattern, we start developing pattern 22 (Automatic translation). As a reminder, the latter is a central functionality to achieve the SignaComm's objective. We rely on Emele et al. works and ours [16], [28], [29], [30] to accomplish this task. We will explain the methodology of work for developing both, the software and dictionary for the protection of personal data.

4.1 Automatic translation

We analyse the situation where a SignaComm member uses the application to translate in real time a sequence of syntagms (or text phrases) into a series of signagrams, and to engage in an informative conversation with a member or group of members. We present information to be quickly understood by members, to prevent some manipulation of personal data without their knowledge or permission and regardless of the computing device used [12], [25] (e.g., the portability of data).

While developing this machine translation prototype, we face a main difficulty, namely the non-figurative legal corpus. The suggested solutions are, on the one hand, to segment and analyse a thematic text and, on the other hand, to only translate the syntagms related to the case [29].

Here, for this machine translation, the expressions' exactness is necessary to be able to break down their relations with other encompassing units [29]. This would help by decreasing blunders and uncleanness in the translation process [39], [40]. Consequently, we utilise the National Commission for Informatics and Liberty portal's thematic text that presents reliable and relevant information [25].

Our model is composed of the ontology for the protection of personal data [41], [42], the construction of a dictionary of signagrams also related to the protection of personal data [43], [44] and the adaptation of the function translating text phrases into signagrams [28], [29], [30], [31].

We are particularly interested in the works of Palmirani et al. [41], [42] as their ontology is based on the application of the General Data Protection Regulation. The accuracy, flexibility and reliability of this ontology are well in line with our work objective. Therefore, it is appropriate to integrate it in the project.

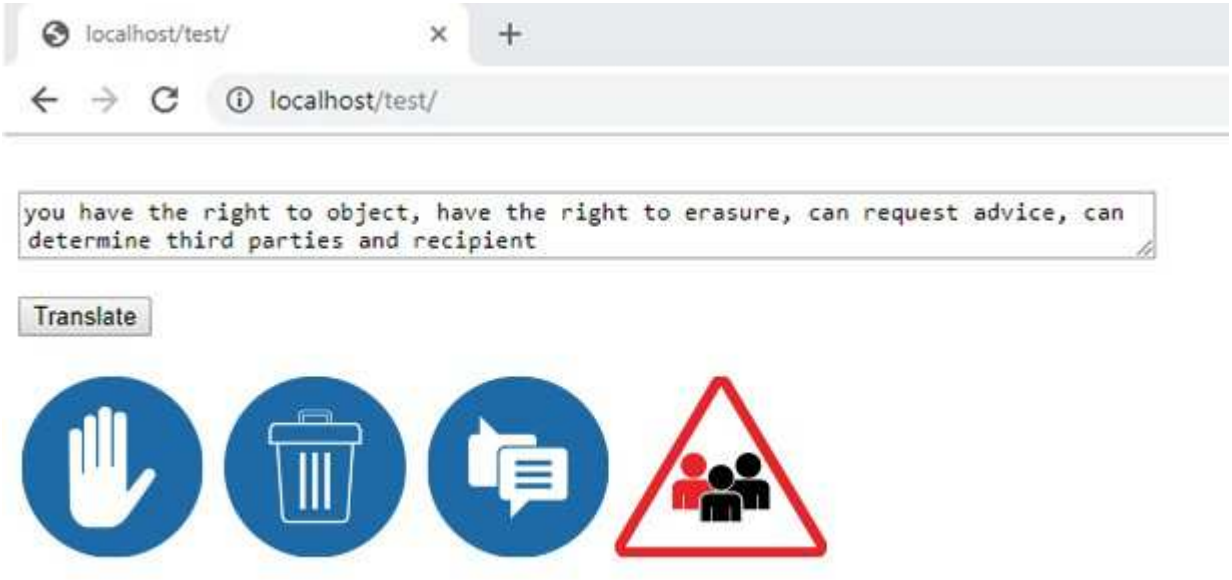


Fig. 3. Example of a machine translation result.

4.2 Dictionary of signagrams for the protection of personal data

To our knowledge, published works related to the dictionary of signagrams for the protection of personal data are limited. We rely on the works of Kordahi [45] and Takasaki [46] to design and develop this first dictionary, which is specialised. It provides information on signagrams to improve their understanding by any user.

The dictionary's design is based on the correspondence of vector signagrams to homologous semantic-based concepts [45], [46]. We program a mapping between two resources. The first semiotic graphical resource contains signagrams' external shapes (including the contours) and internal shapes [16]. The external and internal shapes, coming from that graphical resource, are stored in the dictionary. The second resource is a semantic lexical one (e.g., the WordNet [47]). The latter contains the concepts with their definitions and synonyms in English. The words, definitions and synonyms, coming from that lexical resource, are contained inside the dictionary.

We create fifty signagrams based on the works of Holtz et al. and the United Nations Economic Commission for Europe [43], [44], as well as the "Fotolia" international image bank [48]. The latter holds a large collection of images and symbols used globally. The signagrams' colours and shapes follow the international charter roads signs [44]. Fig 3 shows an example of automatic translation of text phrases into signagrams.

5. SignaComm's first technical test in the specific situation

For now, we have designed and programmed a prototype of the SNS. It is implemented in the Elgg platform and hosted on local and private servers.

The SignaComm is written with the PHP and Javascript programming languages to enable queries to be performed from a Web page. The interface is made up of a set of HTML Web pages. In this section, we choose to explain the four main patterns that are dynamically connected (section 3) [18], [19], [20], [21]. These patterns are the following: the interface (patterns 100 and 101), user profile (patterns 20 and 21), automatic translation of syntagms into signagrams (patterns 22 to 25) [28], [29], [45] and activities (pattern 30) [27].

5.1 SignaComm's interface pattern

The SignaComm's interface is used to display two sorts of information: the resulting information from an exchange between SNS members as well as interactions between the SNS system and its members. We provide the following examples, which include sending and receiving instant messages, displaying automatic translation of written texts into a sequence of signagrams, viewing a member's profile (section 3, Fig. 2).

The graphical user-interface consists of a main interface and secondary one. The main interface is used to display the Web pages' content. The secondary interface is the navigation bar. It enables the browsing between the various pages (Fig. 4).

5.2 SignaComm's user profile pattern

The user profile pattern performs three essential tasks. These are the registration of a user, invitation of a user and geolocation of members (section 3, Fig. 2). The first task allows a user to register and login to the SignaComm, which are a condition to use this SNS. The registration is done by submitting a user-name and password as well as some of the information regarding the user (e.g., choosing to share her/ his information [36], [37] and geographical position with the SNS) (Fig. 4). The login is done by submitting the member's user-name and previously saved password. The second function allows a SignaComm member to invite another user by sending an electronic invitation (e.g., instant message) while using the other patterns (e.g., pattern 21). This pattern 20 is connected to a geolocation process to allow performing the third task. The latter task automatically suggests a language of conversation [27].

This pattern comprises an application page and a PHP function. The HTML application page collects the user's registration information, including the name, physical address, email and address. The collected information is sent to the PHP function.

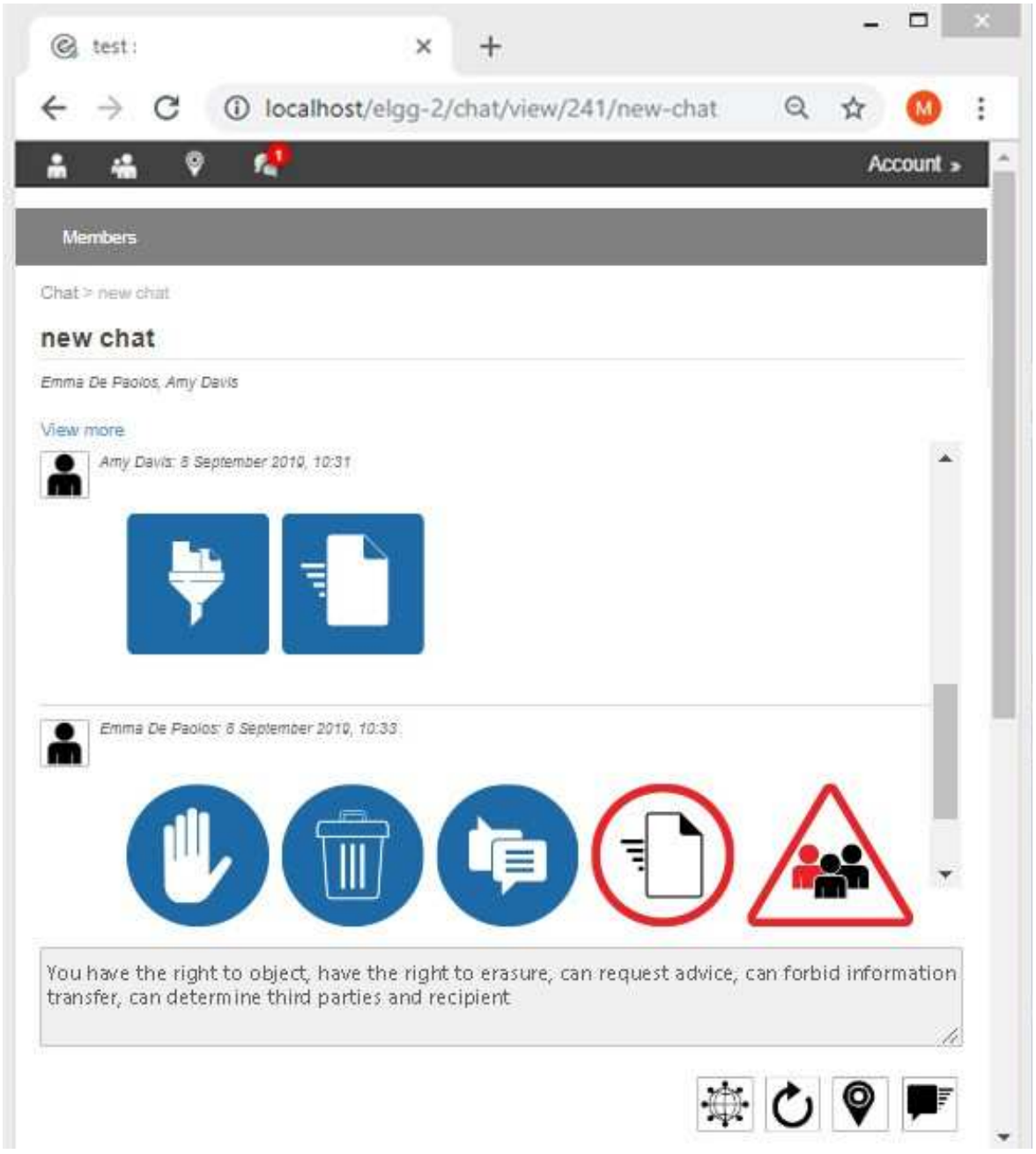


Fig. 4. Example of the chat room page.

5.3 Pattern of automatic translation

We rely on our works developed in sections 3 and 4 to implement the machine translation in the SignaComm structure.

Once implemented in the SignaComm, the translation pattern runs three consecutive tasks that are stored in this SNS database. The chat room page (written in HTML format) can receive the member's input text. A first request transmits the input text to be automatically translated into vector signagrams. A second request displays the machine translation result in the same HTML page. And a third request waits for the member's action to send the translated message to the activities pattern, or to reset the automatic translation process [28] (Fig. 4).

5.4. Pattern of activities

The pattern of activities performs two simultaneous and programmed tasks that are saved in the SignaComm. Chat histories are saved in the database's tables (section 3, Fig. 2). Through the server, the translation pattern receives requests from a member in the form of packets compliant with a common Internet protocol (e.g., the HyperText Transfer Protocol (HTTP) POST packets). These packets contain the translated information (the message is translated before delivery) [27]. The second task displays instant messaging exchange between members in the chat room page [27] (Fig. 4).

Fig. 4 shows an example of the SignaComm and the translation results. In this paper, the reported digital identities are simulated using fake profiles. Member 1 writes an input text (*we wish to collect information and transfer files*), activates its

automatic translation and then sends the resulting translation to a corresponding member 2. Member 2 replies to member 1 by writing, translating and sending a message [25]. The signagrams' reading direction is from left to right and top to bottom [15]. The result of Fig. 4 is comparable to Fig. 3.

6 Discussion and conclusion

While creating the SignaComm, with an internationally oriented communication system for the protection of personal data, we overcame at least one difficulty. To protect personal information on the Web, information accuracy, reliability, flexibility and speed of transmission are needed to assist individuals. We have formed the SignaComm of interrelated patterns. This interrelation has allowed us to synchronise the information exchange.

The obtained results demonstrate that the SignaComm is functioning correctly. In real time and instantly a sequence of text phrases is translated into a series of signagrams in order to send the results to members. Members can create their own network of contacts by inviting users of their choice. The geolocation process also identifies the member's preferred language.

Moreover, since the SignaComm for the protection of personal data is a first and a new prototype, we recommend preparing users for its utilisation with the aim to optimise its performance. This preparation should include detailed explanations regarding the SNS: the purpose, usefulness of its utilisation, interface functionalities and signage system. This preparation could be done in various ways, such as through a demonstration video, detailed guide, Questions and Answer (Q&A) forum. An online help would be interesting to design and implement in the SignaComm context. This would explain the SNS' social utility and the meaning of every signagram. Its use may be punctual, used to understand the meaning of a specific signagram or to search for a specific functionality.

In the near future, we would like to analyse and test the SignaComm with other writing systems, for instance Chinese. Furthermore, we wish to improve this first prototype. We will place the SignaComm in other areas and contexts, like the online learning one. In this context, on the one hand, we will analyse the digital identity of different SignaComm users/ members as well as the visibility models [36], [49]. On the other hand, we will conduct qualitative and quantitative studies on the user's behaviour while utilising the SNS. This study will allow us to evaluate, measure and improve the time required to understand a visual message.

Acknowledgment

We wish to thank Mohammad Haj Hussein, computer and communication engineer, for his valuable help while programming this prototype.

References

[1] A. Löwstedt, "Communication Ethics and Globalization," in *Ethics in Communication*, edited by Patrick Lee Plaisance, Volume 26 in de Gruyter Mouton Handbooks of Communication Science series, Peter Schulz and Paul Cobley, pp. 367-390, 2018.

[2] C. Alloing and J. Pierre, *Le Web Affectif: une Économie Numérique des Émotions,* INA éditions, 2017.

[3] B. Wellman, "Little Boxes, Glocalization, and Networked Individualism," In *Kyoto workshop on digital cities*, pp. 10-25. Springer, Berlin, Heidelberg, 2001.

[4] B. Wellman and M. Gulia, "Net-Surfers Don't Ride Alone: Virtual Communities as Communities," In *Networks in the global village*, Routledge, pp. 331-366. 2018.

[5] J. Fernback, "The Individual within the Collective: Virtual Ideology and the Realization of Collective Principles," *Virtual Culture: Identity and Communication in Cybersociety*, pp. 36-54, 1997.

[6] H. Rheingold, *The Virtual Community: Finding Commection in a Computerized World*, Addison-Wesley Longman Publishing Co., Inc., 1993.

[7] L. Raine and B. Wellman, *Networked: The New Social Operating System*. Massachussets: Massachusetts Institute of Technology, 2012.

[8] D. Boyd and N. Ellison, "Social Network Sites: Definition, History, and Scholarship," *Journal of computer-mediated Communication*, vol. 13, no. 1, pp. 210-230, 2007.

[9] H. Waheed, M. Anjum, M. Rehman and A. Khawaja, "Investigation of User Behavior on Social Networking Sites," PloS one, 12 (2), 2017, DOI:10.1371/journal.pone.0169693.

[10] M. Granovetter, "The Strength of Weak Ties: A Network Theory Revisited," *Sociological Theory*, vol. 1, pp. 201-233, 1983.

[11] C.M. Ma, Y. Zhuang and S. Fong, "Information Sharing over Collaborative Social Networks Using Xacml," In *8th International Conference on e-Business Engineering*, pp. 161-167, IEEE, 2011.

[12] E. Kennedy and C. Millard, "Data Security and Multi-factor Authentication: Analysis of Requirements under EU Law and in Selected EU Member States," *Computer Law & Security Review*, vol. 32, no. 1, pp. 91-110, 2016.

[13] T. Takasaki and Y. Mori, "Design and Development of a Pictogram Communication System for Children around the World," In *International Workshop on Intercultural Collaboration*, Berlin: Springer-Verlag, pp. 193-206, 2007.

[14] S. Fitrianie and L. Rothkrantz, "A Visual Communication Language for Crisis Management," I*nternational Journal of Intelligent Control and Systems* (Special Issue of Distributed Intelligent Systems), vol. 12, no. 2, pp. 208-216, 2007.

[15] M. Neurath, "Isotype," *Instructional science*, vol. 3, no. 2, pp. 127-150, 1974.

[16] M. Kordahi, "Signage as a New Artificial Communication System," *Canadian Journal of Information and Library Science*, vol. 37, no. 4, pp. 237-252, 2013.

[17] J.M. Klinkenberg, P*récis de Sémiotique Générale,* Louvain: De Boeck & Larcier, 1996.

[18] R. E. Kraut and P. Resnick, *Building Successful Online Communities: Evidence-based social design*, MIT Press, 2012.

[19] E. Gamma, R. Helm, R. Johnson and J. Vlissides, *Design Patterns : Element of Reusable ObjectOriented Software*, Addison-Wesley professional computing series, 1995.

[20] C. Alexander. *The Timeless Way of Building,* Vol. 1, New York: Oxford University Press, 1979.

[21] C. Alexander, *A Pattern Language: Towns, Buildings, Construction,* Oxford university press, 1977.

[22] T. Gruber, "A Translation Approach to Portable Ontology Specifications," *Knowledge acquisition*, vol. 5, no 2, pp.199-220, 1993

[23] N.F. Noy and D.L. McGuinness, "Ontology Development 101: A Guide to Creating your First Ontology", 2001.

[24] T. Gruber, "Collective Knowledge Systems: Where the Social Web Meets the Semantic Web," *Web semantics: science, services and agents on the World Wide Web*, vol. 6, no. 1, pp. 4-13, 2008.

[25] "National Commission on Informatics and Liberty",

hhttps://www.cnil.fr/en/home (15/01/20).

[26] C.Y. Yang and H.Y. Li, "An Instant Messaging with Automatic Language Translation," In *3rd IEEE International Conference on Ubi-Media Computing*, pp. 312-316, IEEE, 2010.

[27] Y. Seme, "Method and System for Translating Instant Messages," *U.S. Patent Application*, 10/035,085, filed July 3, 2003.

[28] M. Kordahi, "Réseau Social Numérique et Images Vectorielles: Introduction à une Communication à Vocation Internationale," *Recherches en Communication*, vol. 42, pp. 233-251, 2016.

[29] M. Kordahi and C. Baltz, "Automatic Translation of Syntagms into "Signagrams" for risk prevention", *Management des Technologies Organisationnelles*, vol. 5, pp. 23-37, 2015.

[30] M.C. Emel, M. Dorna, A. Lüdeling, H. Zinsmeister and C. Rohrer, "Semantic-based transfer," In *Verbmobil: Foundations of Speech-to-Speech Translation*, Berlin: Springer-Verlag, pp. 359-376, 2000.

[31] M. Kordahi, "La signalétique comme système de communication internationale, la protection des informations personnelles sur le Web," *International Association for Media and Communication Research (IAMCR)*, Madrid 2019.

[32] D. Spiliotopoulos, E. Tzoannos, C. Cabulea and D. Frey, "Digital Archives: Semantic Search and Retrieval," In *International Workshop on Human-Computer Interaction and Knowledge Discovery in Complex, Unstructured, Big Data,* pp. 173-182, Springer, Berlin, Heidelberg, 2013.

[33] R. Jain, J. Bose and T. Arif, "Contextual Adaptive User Interface for Android Devices," In *India Conference (INDICON)*, IEEE, pp. 1-5, 2013.

[34] D.B. Morin et al. "Method And System For A Personal Network," *U.S. Patent Application* 12/945,743, filed May 17, 2012.

[35] S. Proulx, "L'Irruption des Médias Sociaux: Enjeux Éthiques et Politiques," *Médias sociaux: enjeux pour la communication*, Québec : Presses de l'Université du Québec, 2012.

[36] D. Cardon, "The Design of Visibility, " *Réseaux*, vol. 6, no. 152, pp. 93-137.

[37] S. Bouraga, I. Jureta and S. Faulkner, "Requirements Engineering Patterns for the Modeling of Online Social Networks Features," In *4th international workshop on requirements patterns*, pp. 33-38. IEEE, 2014.

[38] R. Gross and A. Acquisti, "Information Revelation and Privacy in Online Social Networks," In *Proceedings of the ACM workshop on Privacy in the electronic society*, pp. 71-80, ACM, 2005.

[39] Y. Bar-Hillel, "The Present Status of Automatic Translation of Languages," *Advances in computers*, vol. 1, no. 1, pp. 91-163, 1960.

[40] M. McShane, S. Nirenburg and S. Beale, "An NLP Lexicon as a Largely Language-independent Resource," *Machine Translation,* vol. 19, no. 2, pp. 139-173, 2005.

[41] M. Palmirani, M. Martoni, A. Rossi, C. Bartolini and L. Robaldo, "PrOnto: Privacy Ontology for Legal Reasoning"' *In International Conference on Electronic Government and the Information Systems Perspective*, Springer, Cham, pp. 139-152, 2018.

[42] M. Palmirani, M. Martoni, A. Rossi, C. Bartolini and L. Robaldo, "Legal Ontology for Modelling GDPR Concepts and Norms," In *JURIX*, pp. 91-100, 2018.

[43] L.E. Holtz, K. Nocun and M. Hansen, "Towards Displaying Privacy Information With icons," In *IFIP PrimeLife International Summer School on Privacy and Identity Management for Life*, Springer, Berlin, Heidelberg, pp. 338-348, 2010.

[44] Economic Commission for Europe – Transport Division, Road Traffic and Road Signs and Signals Agreements and Conventions, https://www.unece.org/fileadmin/DAM/trans/conventn/Conv_road_signs_2006v_EN.pdf (15/01/20).

[45] M. Kordahi, "SignaNet: First specialised electronic dictionary for signage," *Pratiques et usages numériques H2PTM'2013,* pp. 385-387, 2013.

[46] T. Takasaki, "PictNet: Semantic Infrastructure for Pictogram Communication," in *The Third International WordNet Conference (GWC-06)*, pp. 279-284, 2006.

[47] G. Miller, *WordNet: An Electronic Lexical Database*, Cambridge, Massachusetts: MIT Press, 1998.

[48] "Fotolia", https://www.fotolia.com/ (15/01/20).

[49] S. Turkle, *Life on the Screen*, Simon and Schuster, 2011

Marilou Kordahi is a senior lecturer in information systems at Saint-Joseph University (Lebanon) and member of Paragraph laboratory, University of Paris 8 (France). She holds a PhD in Information and Communication Sciences for the University of Paris 8.

BERT and fastText Embeddings for Automatic Detection of Toxic Speech

Ashwin Geet D'Sa, Irina Illina, Dominique Fohr
Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France

Abstract - With the expansion of Internet usage, catering to the dissemination of thoughts and expressions of an individual, there has been an immense increase in the spread of online *hate speech*. Social media, community forums, discussion platforms are few examples of common playground of online discussions where people are freely allowed to communicate. However, the freedom of speech may be misused by some people by arguing aggressively, offending others and spreading verbal violence. As there is no clear distinction between the terms *offensive, abusive, hate and toxic* speech, in this paper we consider the above mentioned terms as *toxic* speech In many countries, online toxic speech is punishable by the law. Thus, it is important to automatically detect and remove toxic speech from online medias. Through this work, we propose automatic classification of toxic speech using embedding representations of words and deep-learning techniques. We perform binary and multi-class classification using a Twitter corpus and study two approaches: (a) a method that consists extracting word embeddings and then using a DNN classifier. We observed that BERT fine-tuning performed much better.

Index Terms - Natural language processing, classification, deep neural network**,** hate speech.

1. Introduction

Hate speech expresses an antisocial behavior. The topics of the hate can be gender, race, religion, ethnicity, etc. [1].

There is no clear definition of the term *hate speech*. A Committee of Ministers from Council of European Union define hate speech as: "All forms of expression which spread, incite, promote or justify racial hatred, xenophobia, anti-semitism or other forms of hatred based on intolerance, including intolerance expressed by aggressive nationalism and ethnocentrism, discrimination and hostility towards minorities, migrants and people of immigrant origin."[1]. As there is no clear distinction between the terms *offensive, abusive, hate and toxic* speech, in the following of this paper, we will consider the above mentioned terms as *toxic* speech.

Table 1 gives some examples of toxic comments from social media.

She look like a tranny.
You Asian, they will deport you when they see your eyes.
I'm not going to believe any of the stupid rumors I hear about jews being friends of Christians.
We hate niggers, we hate faggots and we hate spics

Table 1: Examples of toxic comments from social media.

Toxic speech can be expressed in different forms. Explicit toxic speech contains offensive words such as '*fuck', 'asshole'*. Implicit toxic speech can be realized by a sarcasm and an irony[2][3]. While explicit toxic speech can be identified using the lexicons that forms the toxic speech, implicit toxic speech is often hard to identify and requires semantic analysis of the sentence. Examples of implicit and explicit toxic speech are shown in Table 2.

Toxic content on Internet platform can create fear, anxiety and threat to individuals. In the case of company or online platform, the company or platform may lose its reputation or the reputation of its product. Failure to moderate these contents may cost the company in multiple ways: loss of users, drop in stocks [2], penalty from legal authority[3], etc.

Explicit toxic speech
You are a real fag aren't you?
Go fuck yourself asswipe!
Haha you are a dumb shit.
Implicit toxic speech
Affirmative action means we get affirmatively second rate doctors and other professionals.
I will remove all your organs in alphabetical order.
She looks like a plastic monkey doll!

Table 2: Examples of explicit and implicit toxic speech.

Most of the online platforms such as social media or the forums, generally cannot be held responsible for the propagating of toxic speech. However, their inability to prevent its use is the reason for the spread of hate. Manual analysis of such content and its moderation are impossible because of the huge amount of data circulating on the internet. An effective solution to this problem would be the automatic detection of toxic comments.

In many countries, online hate speech is an offense and it is punishable by the law. In this case, the social medias are held responsible and accountable if they did not remove hate speech content promptly.

1 https://www.article19.org/data/files/medialibrary/3548/ARTICLE-19-policy-on-prohibition-to-incitement.pdf

2 https://www.telegraph.co.uk/technology/2018/07/27/twitter-stock-sinks-reporting-decline-active-users/

3 https://www.cnet.com/news/german-hate-speech-law-goes-into-effect-on-1-jan/

Automatic detection of toxic speech is a challenging problem in the field of *Natural Language Processing* (NLP). The approaches proposed for automatic toxic speech detection are based on the representation of the text in a numerical form and on the using of some classification models. In the state-of-the-art on this field, word and character n-grams [4], *Term Frequency-Inverse Document Frequency* (TF-IDF), *Bag of Words* (BoW), polar intensity, noun patterns [5] and word embedding are largely used as input features. The notion of word embedding is based on the idea that, semantically and syntactically similar words must be close to each other in an n-dimensional space [8]. *Global Vectors for word representation* (GloVe) [6] and random embeddings as input to DNN classifiers has been compared in [7]. Recently, sentence embeddings [9] and *Embeddings from Language Models* (ELMo) [10] were used as input to classifiers for toxic comment classification. Multi-features based approach combining various lexicons and semantic-based features is presented in [11].

Deep-learning techniques have shown to be very powerful in classifying toxic speech [7]. For example, Convolutional Neural Network (CNN) are able to capture the local patterns in text [12]. Long Short Term Memory (LSTM) model [13] or Gated Recurrent Unit (GRU) model [14] capture the long range dependencies. Such properties are important for modelling toxic speech [7], [15].

In this article, we propose a new methodology to automatically detect toxic speech. We perform toxic speech classification using two powerful word representations: fastText and BERT embeddings. These representations are used as inputs to DNN classifiers, namely CNN and Bi-LSTM classifiers. We study two cases: binary classification and multi-class classification. In the last case, we want to classify toxic speech more finely in *hate speech and abusive speech.* Moreover, we explore the capabilities of BERT fine-tuning on both binary and multi-class classification tasks. We evaluate the proposed approaches on the a Twitter corpus.

The contributions of our paper is as follow:

- We use fastText embeddings and BERT embeddings as input features to CNN and Bi-LSTM classifiers.
- We perform fine-tuning of the pre-trained BERT model.
- We study the classification of comments from two perspectives:
  (a) **binary classification**, where we consider two classes: *non toxic speech* versus *toxic speech (hate speech* and *offensive speech* together);
  (b) **multi-class classification,** where we use three classes *hate speech, offensive speech* and *neither.* This three class classification allows to perform fine-grained distinction between hate and offensive speech within toxic speech.

The rest of the paper is organized as follows. Section 2 describes the word embeddings. Section 3 presents the proposed methodology. Section 4 describes data and the preprossessing description. The results are discussed in section 5.

2. Word embeddings

The main idea of word embeddings is to project words in a continuous vector space. In this space, semantically or syntactically related words should be located in the same area. An important advantage of word embedding is that their training does not require a labeled corpus.

The embeddings are generally learned from a very huge unlabelled corpus. This training is time consuming and often requires high-level technical conditions (big GPU, large memory, etc). Pre-trained word embeddings are made available via Internet and can be used by researchers from around the world for different NLP tasks. For example, Facebook provided fastText model, Google provided several BERT models for different languages. In this paper, we propose to use these pre-trained embeddings. In the following of this section, we will describe the embeddings used in this study.

fastText embedding: It is an extension of Mikolov's embedding [8]. The fastText approach is based on the skip-gram model, where each word is represented as a bag of character n-grams [16], [17]. A vector representation is associated to each character n-gram; words being represented as the sum of these representations. The word representation is learned by considering a large window of left and right context words. Unlike Mikolov's embeddings, fastText is able to provide an embedding for misspelled word, rare words or words that were not present in the training corpus, because fastText uses character n-gram word tokenization.

BERT embedding: Currently BERT (*Bidirectional Encoder Representations from Transformers*) is one of the most powerful context and word representations [18]. BERT is based on the methodology of *transformers* and uses *attention* mechanism. Attention is a way to look at the relationship between the words in a given sentence [19]. Thanks to that, BERT takes into account a very large left and right context of a given word. It is important to note that the same word can have different embeddings according to the context. For example, the word *bank* can have one embedding when it occurs in the context *the bank account* and a different embedding when it occurs in the context *the bank of the river*. Moreover, BERT model uses word-piece tokenization. For instance, the word *singing* can be represented as two word-pieces: *sing* and *##ing.* The advantage is, that when the word is not in the BERT vocabulary, it is possible to split this word into word-pieces. Thus, it is possible to have embeddings for rare words, like in fastText.

BERT model can be used in two ways:

- for generating the embeddings of the words of a given sentence. These embeddings are further used as input for DNN classifiers;
- for fine-tuning a pre-trained BERT model using a task-specific corpus and further to perform the classification.

3. Proposed methodology

Figure 1 shows the proposed methodology. The general idea is as follow: we use pre-trained embeddings to represent each comment in the continuous space. After this, we use these embeddings as input features for a DNN classifier.

We propose to use the word representations in two ways, ***feature-based*** and ***fine-tuning*** approaches:

- in feature-based approach, two steps are performed. First, each comment is represented as a sequence of words or word-pieces and for each word or word-piece, an embedding is computed using fastText or BERT. Secondly, this sequence of embeddings will form the input to the DNN classifiers, that takes the final decision. We use CNN and Bi-LSTM models as classifiers.
- in fine-tuning approach, everything is done in a single step. Each comment is classified by a fine-tuned BERT model.

We classify each comment as *non toxic* or *toxic speech* for binary classification and *offensive*, *hate speech* or *neither* for multi-class classification.

3.1 Feature-based approaches

For feature-based approaches, we used pre-trained fastText and BERT models to obtain the sequence of embeddings for a given comment. This sequence of embeddings is used as input features to DNN classifiers. The sequence should have a fixed size. For this, we extend the short tweets by zero padding.

fastText model: We use pre-trained fastText embedding model and apply this model to generate one embedding for each word of a given comment. Thanks to the bag of character n-grams model of fastText, every word in a given comment will have an embedding, even out-of-vocabulary and rare words.

BERT model: Word-piece tokenization is performed on the comment and then used as input to a pre-trained BERT model. BERT model provides contextual embedding for the word-pieces.

The obtained embeddings from either fastText or BERT models are then used as input to a DNN classifier.

3.2 Deep Neural Networks classifiers

For the purpose of toxic tweet classification, we use CNN and Bi-LSTM deep neural network classifiers:

- CNN were traditionally used in the application of image processing, and are good at capturing the patterns. Kim [12] demonstrated the efficient use of CNN for Natural language processing on various benchmark tasks.
- Bidirectional LSTM (Bi-LSTM) is a class of RNN models, which overcomes the problem of vanishing gradient problem. Bi-LSTMs are used for sequential processing of the data and are efficient at capturing long-range dependencies.

3.3 BERT fine-tuning

The BERT pre-trained model can be fine-tuned to a specific task. This consists in the adapting of the pre-trained BERT model parameters to a specific task using a small corpus of task specific data. Since BERT is contextual model and pre-trained BERT model is trained on a very huge corpora containing few *toxic speech* or *twitter data*, it will be interesting to fine-tune this model with the toxic and twitter specific data set. For the purpose of classification task, a neural network layer is used on top of fine-tuned BERT model. So, the weights of this layer and the weights of the other layers of the Bert model are trained and fine-tuned correspondingly using task specific data in order to perform the classification task.

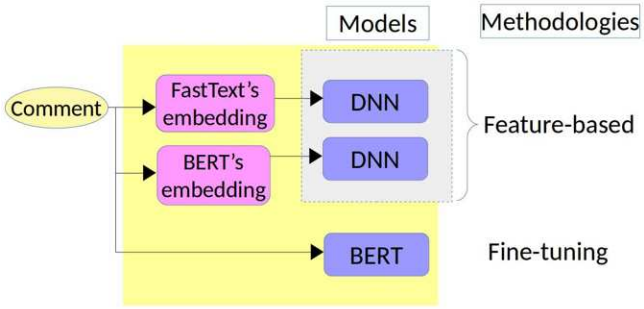


Figure 1: Proposed methodology

4. Experimental setup

4.1 Data set

For the purpose of toxic classification, we used the Twitter corpus [20]. The tweets are collected based on keywords from *hatebase.org* lexicon. The data set contains 24883 tweets and annotations performed by CrowdFlower. Each tweet is annotated by at least 3 annotators. The annotator agreement is 92%. The labels correspond to three classes: *hate speech*, *offensive language* and *neither*, representing 5.7%, 77.1% and 16.7% respectively. Thus, this data set is an unbalanced data set. Table 3 gives the statistics of the data set after pre-processing.

We followed the 5-fold cross validation procedure as in [20]. We used 70% of data as training, 20% as test set and 10% as development set. The development set is used to choose the hyper-parameters. The test set is used to evaluate the proposed approaches.

In our experiments, during the **binary classification** we merged *hate-speech* and *offensive speech* together in the single class (*toxic speech)*. Thus, we have *toxic speech class* and *non toxic speech* class. For **multi-class classification** we use the three classes and the labels provided with the data set: *hate speech*, *offensive speech* and *neither*.

4.2 Text pre-processing

Most of the text classifiers results depend on the way the input text is pre-processed. For both, fastText and BERT embeddings, we decided to remove the numbers and all the special characters except *'!', '?', ',', '.'* and apostrophe.

We also performed tweet specific pre-processing. We removed user names (words beginning with symbol '@'). and the word 'RT', indicating *re-tweet*. We split hast-tags in multiple words. For example, *#KillThemAll* is split into *Kill Them All*.

	Hate speech	*Offensive speech*	*Neither*	*Total*
Nbr of tweets	1430	19190	4163	24783
Corpus size (word count)	19.6k	259.5k	62.1k	341.2k
Nbr. of unique words	3.7k	16.2k	9.9k	21.2k
Average nbr. of words per tweet	13.7	13.5	14.9	13.8

Table 3: Statistics of Twitter data set after pre-processing. k denotes thousand.

4.3 Embedding models

- **fastText embedding:** the model is provided by Facebook[4] and pre-trained on *Wikipedia 2017*, *UMBC webbase* and *statmt.org news* data sets with total 16B tokens. The embedding dimension is 300, the vocabulary is 1M words.
- **BERT model:** In our work, we used BERT-base-uncased word-piece model (for English), provided by Google[5] and pre-trained on *BookCorpus* and *Wikipedia* corpora. The model has 12 stacked transformer encoder layers, with 24 attention heads. The embedding dimension is 768, the number of word-pieces is 30k.

4.4 DNN configurations

We perform the classification experiments with different hyper-parameters and choose the final configuration based on the best performance obtained on the development set. The best model configurations are detailed below.

For Bi-LSTM, we used one or two bidirectional LSTM layers with varying LSTM units (between 50 and 128) followed by one or two dense layers with 64 and 256 dense units in the first dense layer and 16 and 64 dense units in the second layer. For CNN we have used either one or two layers (filter size between 3 and 5), and used between 16 and 64 units, followed by two dense layers having 64 and 256 dense units in the first dense layer and 16 and 64 dense units in the second layer. The dense units use *Rectified Linear Unit* activation (ReLU), while the final output neuron uses *sigmoid* activation. We use a varying dropout upto 0.2. We use l2 regularization. The models are trained using Adam optimizer with learning rate of 0.001. For BERT fine-tuning we used maximum sequence length 256, batch size 16, learning rate $2 \cdot 10^{-5}$ and 3 epochs.

4 https://fasttext.cc/docs/en/english-vectors.html

5 https://github.com/google-research/bert

We evaluate the performance of our approaches in terms of macro-average F1-measure. **F1-measure** is a statistical measure to analyze classification performance. This value ranges between 0 and 1, where 1 indicates the best performance. F1-measure is calculated as follow:

$$F1 = \frac{2 * (precision * recall)}{(precision + recall)}$$

where ***precision*** is the ratio between number of samples correctly predicted as class A and total number of samples predicted as class A by the classifier; ***recall*** is the ratio between number of samples correctly predicted as class A and total number of samples that should be predicted as class A.

Macro-average F1-measure provides the arithmetic mean of F1-measures of all classes:

$$macro\,F1 = \frac{1}{C} \sum_{i=1}^{C} F1_i$$

where, C is the total number of classes.

For each experiment, we compute an average macro-average F1-measure obtained from the 5-folds test sets.

5. Results and discussion

Table 4 gives the macro average F1 results for binary classification task using Bi-LSTM and CNN classifiers with fastText and BERT embeddings as input features. Table 5 presents the macro-averaged F1-measure results for multi-class classification task.

From table 4 and table 5, it can be observed that both fastText and BERT embeddings provide nearly the same results. Among the classifiers, Bi-LSTM performs slightly better than CNN. The performance of binary classification, presented in table 4, is better than multi-class classification performance, given in table 5. This can be explained by the fact that it is difficult to distinguish between hate speech and offensive speech.

Finally, BERT fine-tuning performs better than feature-based approaches in both binary as well as in multi-class classification: compared to feature-based approaches, we obtained 63% and 42% of classification error reduction for binary and multi-class classification correspondingly. One reason may be that in the feature-based approach the

A. Feature-based approaches		
	CNN	Bi-LSTM
fastText embedding	91.5	91.9
BERT embedding	90.9	91.9
B. BERT fine-tuning		
BERT fine-tuning	**97.0**	

Table 4: Macro-average F1-measure for different classifiers and different embeddings. Binary classification.

A. Feature-based approaches		
	CNN	Bi-LSTM
fastText embedding	70.9	72.3
BERT embedding	71.9	72.4
B. BERT Fine-tuning		
BERT fine-tuning	**84.0**	

Table 5: Macro average F1-measures for different classifiers and different embeddings. Multi-class classification.

embeddings vectors have not been trained on hate speech or offensive data. On the contrary, the BERT fine-tuning approach is fine-tuned on twitter data to distinguish hate, offensive and non toxic speech and this allow to create more accurate model for toxic speech.

Table 6 and 7 present the confusion matrices between the 3 classes for multi-class classification: table 6 for feature-based Bi-LSTM with BERT embeddings and table 7 forBERT fine-tuning. We can notice that the main confusions occur between hate speech and offensive speech. This suggest that the model is biased towards classifying tweets as less hateful or offensive than the human annotators. This result is close to the results obtained in [20]. The feature-based approach is able to detect only 31% of the hate speech tweets, while Bert fine-tuning achieved 53%. Many fewer tweets are classified as more offensive or hateful than their true category.

True label	hate	31	60	9
	offensive	2	95	3
	neither	3	10	87
		hate	offensive	neither
		Predicted labels		

Table 6: Confusion matrix for feature-based Bi-LSTM with BERT embeddings (in %). Multi-class classification.

True label	hate	53	43	4
	offensive	1	98	1
	neither	1	4	95
		hate	offensive	neither
		Predicted labels		

Table 7: Confusion matrix for BERT fine-tuning (in %). Multi-class classification.

6. Conclusion

In this article, we proposed new approaches for automatic toxic speech detection in the social media. These approaches are based on deep learning classifiers and word embeddings. We have explored the classification from two perspectives-binary classification and multi-class classification. For binary classification we consider toxic speech (hate speech and offensive speech together) and non toxic speech. For multi-class, we considered *hate speech, offensive speech* and *neither*.

We proposed feature-based approaches and fine-tuning of pre-trained BERT model. In feature-based approaches, fastText and BERT embeddings are used as input features to CNN and Bi-LSTM classifiers. Further, we have compared these configurations with fine-tuning of pre-trained BERT model.

We observed that BERT fine-tuning performed much better than feature-based approaches on a Twitter corpus. The main confusions occur between offensive speech and hate speech. In the future work, we want to investigate this problem. Proposed methodology can be used for any other type of social media comments.

7. Acknowledgment

This work was funded by the M-PHASIS project supported by the French National Research Agency (ANR) and German National Research Agency (DFG) under contract ANR-18-FRAL-0005.

8. References

[1] R. Delgado and J. Stefancic, “Hate Speech in Cyberspace”, *Social Science Research Network,* 2014.

[2] Z. Waseem, T. Davidson, D. Warmsley, and I. Weber, “Understanding Abuse: A Typology of Abusive Language Detection Subtasks,” in *Proceedings of the First Workshop on Abusive Language Online*, 2017, pp. 78–84.

[3] L. Gao, A. Kuppersmith, and R. Huang, “Recognizing Explicit and Implicit Hate Speech Using a Weakly Supervised Two-path Bootstrapping Approach,” in *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 2017, pp. 774–782.

[4] C. Nobata, J. Tetreault, A. Thomas, Y. Mehdad, and Y. Chang, “Abusive Language Detection in Online User Content”, in *Proceedings of the 25th International Conference on World Wide Web*, 2016, pp. 145–153.

[5] M. Wiegand, J. Ruppenhofer, A. Schmidt, and C. Greenberg, “Inducing a Lexicon of Abusive Words–a Feature-Based Approach,” in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies,* Volume 1, 2018, pp. 1046–1056.

[6] J. Pennington, R. Socher, and C. Manning, “Glove: Global Vectors for Word Representation”, in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014, pp. 1532–1543.

[7] P. Badjatiya, S. Gupta, M. Gupta, and V. Varma, “Deep Learning for Hate Speech Detection in Tweets”, in *Proc. 26th Int. Conf. World Wide Web Companion - WWW 17 Companion*, pp. 759–760, 2017.

[8] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient Estimation of Word Representations in Vector Space”, *ArXiv13013781 Cs*, 2013.

[9] V. Indurthi, B. Syed, M. Shrivastava, N. Chakravartula, M. Gupta, and V. Varma, “FERMI at SemEval-2019 Task 5: Using Sentence embeddings to Identify Hate Speech Against Immigrants and Women in Twitter”, in *Proceedings of the 13th International Workshop on Semantic Evaluation*, 2019, pp. 70–74.

[10] M. Bojkovský and M. Pikuliak, “STUFIIT at SemEval-2019 Task 5: Multilingual Hate Speech Detection on Twitter with MUSE and ELMo Embeddings”, in *Proceedings of the 13th International Workshop on Semantic Evaluation*, 2019, pp. 464–468.

[11] S. Almatarneh, P. Gamallo, and F. J. R. Pena, “CiTIUS-COLE at SemEval-2019 Task 5: Combining Linguistic Features to Identify Hate Speech Against Immigrants and Women on Multilingual Tweets”, in *Proceedings of the 13th International Workshop on Semantic Evaluation*, 2019, pp. 387–390.

[12] Y. Kim, “Convolutional Neural Networks for Sentence Classification”, in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014, pp. 1746–1751.

[13] A. Baruah, F. Barbhuiya, and K. Dey, “ABARUAH at SemEval-2019 Task 5 : Bi-directional LSTM for Hate Speech Detection”, in *Proceedings of the 13th International Workshop on Semantic Evaluation*, 2019, pp. 371–376.

[14] K. Cho, B. V. Merrienboer, C. Gulcehre , D. Bahdanau, F. Bougares , H. Schwenk, Y. Bengio. “Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation,” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014, pp. 1724–1734.

[15] S. Bodapati, S. Gella, K. Bhattacharjee, Y. Al-Onaizan “Neural Word Decomposition Models for Abusive Language Detection”. In *Proceedings of the Third Workshop on Abusive Language Online,* pp. 135-145, 2019.

[16] A. Joulin, E. Grave, P. Bojanowski, M. Douze, H. Jégou, and T. Mikolov, “FastText.zip: Compressing Text Classification Models”, *ArXiv Prepr. ArXiv161203651*, 2016.

[17] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, “Enriching Word Vectors with Subword Information,” *Transactions of the Association for Computational Linguistics*, vol. 5, no. 1, pp. 135–146, 2017.

[18] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies,* Volume 1, 2019, pp. 4171–4186.

[19] A. Vaswani *et al.*, “Attention is All You Need,” in *Advances in Neural Information Processing Systems*, 2017, pp. 5998–6008.

[20] T. Davidson, D. Warmsley, M. Macy, and I. Weber, “Automated Hate Speech Detection and the Problem of Offensive Language,” in *Eleventh International Conference on Web and Social Media*, 2017.

Scientific watch based on information quality

N. Ibrahim, A. Habacha Chaibi, H. Ben Ghézala

Abstract—In this paper, we examine how diffusion of scientific information contribute to knowledge creation. Recent studies have shown that scientists can benefit from diffusion tools within collaborative networks by being able to receive more and recent information. Here we focus on the quality of diffused information and discuss the implications of this tendency for scientific performance. Base on qualitative study of the diffused scientific information, we determine our need to qualitative diffusion tool base on scientific watch. We propose a scientific watch tool based on scientometrics to select qualitative new publications to diffuse to the users. Such a tool helps researchers on selecting good resources which corresponds to their needs.

Index Terms—scientometrics, scientific quality, scientific evaluation, scientific watch, academic social networks

1 Introduction

ACCORDING to Rieger [1], scientific communication is a process by which scientific information is produced, certified, diffused, preserved and used. Scientific communication is part of the context of scientific research. The scientific community generally shares the idea in order to improve knowledge, innovation and creativity [2]. The way we access, use and analyze scientific knowledge has radically changed in the last few years due to the availability of a large amount of research databases on the Web. These databases provide us with accurate and complete information about the content of scientific papers. As more information on scientific production continues to grow, new tools are needed in order to extract and organize knowledge [3].

A scientific paper is considered to be the output of scientific research. The purpose of scientific publications is sharing research results to benefit from research work. The means of sharing and diffusing scientific information play an important role in the development of science and research activity. There are several means of sharing and diffusing used by researchers such as academic social networks such as: research gate, academia.edu, science work, scoop.it and watching tools such as: Google Scholar alerts, Google alerts, mention, talkwalker alerts and GigaAlert. These means aim to collect and diffuse scientific articles and promote research work to all readers.

In this context, we are interested in the quality of the information diffused by these tools. These tools follow a scientific watch process which recommends to researchers a set of scientific articles corresponding to their thematic needs. This information may not be of an expected quality. Why not integrate the quality of information into the scientific watch process to diffuse qualitative information?

The diversity of scientific diffusion tools makes the control of the information quality more difficult than before. Hence, we need a qualitative diffusion tool to inspect the quality of scientific information used by researchers. Our question concerns the quality of diffused information between researchers. The information communication tools (academic social networks, blogs and watching tools) do they play a role in improving scientific production quality? Do these tools' users refer to researchers' based on their quality?

We will first present several tools used for scientific information diffusion in Section 2. In Section 3, we study the quality of diffused information. In section 4, we propose a qualitative scientific watching system. We discuss our findings in Section 5, before concluding.

2 Tools of Scientific Information Diffusion

Research output is diffused as written informations published in different forms. Publications forms differ according to the discipline. Scientific information can be defined as all the information produced by research and necessary for scientific activity [4]. The researcher-author is at the heart of the knowledge production process, he is at the origin of the research that has been conducted and he is responsible for diffusing the results of his research. This diffusion is done through scientific publications. Price [5] considers that the final product of the scientific research is the publication of a written text (scientific articles, contributions to symposium, reports or any other kind of literature). Science is which is published in journals, articles, papers and scientific works [6]. Scientific journals are the main vehicles for diffusing new knowledge after validation by the review committee [7].

2.1 Scientific Watch

Scientific information must circulate as widely, quickly and efficiently as possible, while being of the highest possible quality, as this diffusion fully contributes to the functioning of research. Putting a watch in place is to monitor the environment in which we evolve.

- *N. Ibrahim is with RIADI Laboratory (ENSI), University of Manouba, Manouba 2011, Tunisia. E-mail: nedra.ibrahim@ensi-uma.tn.*
- *A. Habacha Chaibi is with RIADI Laboratory (ENSI), University of Manouba, Manouba 2011, Tunisia. E-mail: anja.habacha@ensi.rnu.tn.*
- *H. Ben Ghézala is with RIADI Laboratory (ENSI), University of Manouba, Manouba 2011, Tunisia. E-mail: henda.benghezala@ensi.rnu.tn.*

Establishing scientific monitoring of open access resources is therefore to monitor the latest publications of scientific articles and scientific literature in general in a given discipline.

The watch is a permanent activity whose main attribute is observation. It consists of keeping the user informed in a very up-to-date manner in an evolving information environment. It contributes to giving information a strategic dimension, particularly in terms of decision-making. Applied in scientific research, intelligence is a major documentary skill in an environment where information is highly mobile.

The watch methodology includes the selection of sources of information to exploit. The reading of the national but also international scientific publication is an essential tool for any activity of scientific watch. It also consists in being able to work on several levels of depth of the documents.

2.2 Diffusion of Scientific Information Via Academic Social Networks

Academic social networks are the new intersection between social media and scholarly publishing which are interesting online spaces that merit their own discussion [8]. These spaces are added to the classical scientific resources such as journals, book chapters, books and conference proceedings. Social networking sites can seem frivolous and pointless to academics. We observe that specialized academic social networking sites are gaining popularity in certain disciplines and with certain faculty [9]. Academic social networks, also known as scientific social networks, are used to accelerate the world's research. Researchers can access millions of academic papers for free, share their research and track its impact.

Following the classic definition of a digital social network, these tools orchestrate the networking of users from their profile. A network of user is formed using semi-automated connections established from published information (CVs, lines of research, articles, etc.).

With more than 3 million registered researchers, ResearchGate [10] is in the process of building the largest graph of researchers ever made on the model of Facebook and reproduces main feature of the news feed. Its strength lies in the exploitation systematic citations and reservoirs of articles freely available or in archives. ResearchGate has its own baptized index RG Score assigned to each member, based on the contribution intake of the profile and the interactions of members with it (article downloads, questions, answers to questions).

Other academic social networks use watching systems for scientific information diffusion such as: Academia[1], Scoop.it[2] and ScholarWorks[3].

Academia's mission [11] is to make every scholarly and scientific paper available for free on the internet and to enhance academic discussion and collaboration. Based on a watching system, academia allows the creation of an unlimited number of alerts to receive updates on specific search preferences.

Scoop.it [12] is a useful platform for the creation and sharing of information resources with other researchers and the wider community. Scoop.it facilitates the creation of networks for information sharing and knowledge building. Scoop.it has additional functionality that allows the user to diffuse their content via the social media platforms that are embedded within the tool and suggest content to other users.

ScholarWorks makes the intellectual output of the Walden University community publicly available to the wider world. ScholarWorks' rich repository encourages new ideas, preserves past knowledge, and fosters new connections to improve human and social conditions. ScholarWorks provide alerts tool which keep track of newly published content, tailored to the user interests. Based on keywords the user is notified via email of content fitting his/her desired criteria.

2.3 Diffusion of Scientific Information Via Bibliographic Databases

The recent development of online bibliographic databases allows us to have a quantitative description of large amount of scientific papers as well as the relations between these papers (citations). This development offers exciting new perspectives for understanding how the process of scientific production evolves over time [13].

The study by Norris et al. [14] shows that several scientific watching systems exist and they are used by researchers to be up-todate in their researches. The widely used scientific watching systems are: Google Scholar alert system[4] and Science Watch[5]

Google and Google scholar are effective search tools for finding scientific publications. These two search engines have the advantage to index the articles posted on the websites of the authors and the websites of the research laboratories. It is possible to create email alerts on Google scholar. Google Scholar will send a notification each time it will meet a new reference with the terms used during searches. It is therefore important to focus search and therefore choose relevant and targeted keywords so as not to receive too many notifications with useless references, or on the contrary, not to receive enough.

ScienceWatch.com [15] provides a look at the researchers, journals, institutions, nations, and papers selected by Essential Science IndicatorsSM from Clarivate Analytics and other products of the Research Services Group. ScinceWatch is updated weekly, new papers are added with every update, and ScienceWatch.com tracks these new additions. ScienceWatch.com highlights the most-cited of these new entries.

3 Qualitative Study of Diffused Information

The existing of free content, available in diffusing and sharing scientific information tools, allows anyone to publish anything and make it accessible by a large number of readers who may not have the ability to select

[1] http://www.academia.edu
[2] http://www.scoop.it
[3] http://scholarworks.waldenu.edu
[4] https://scholar.google.com
[5] http://archive.sciencewatch.com

qualitative information. Niazov et al. [16] find that a paper in a median impact factor journal uploaded to scientific social networks receives 16% more citations after one year than similar article not available online, 51% more citations after three years, and 69% after five years. This practice can orient researchers to rely on publications which are not judged on their qualities but on their number of shares.

This uncontrolled content can influence the quality of scientific production. This requires studying the quality of diffused information by scientific watch systems available on academic social networks and bibliographic databases. We focus on post-publication quality which reflects the impat of informations after peer evaluation and publication.

One way to study the quality of diffused scientific information is to observe the alerts of the different watch systems in the domain of computer science. We study the quality of diffused information provided by the following watch systems: Google Scholar alerts, Scoop.it alerts, ScholarWorks alerts and Academia alerts. We observed diffused information provided by these systems during 2 months for 10 queries. Each time, we analyze the quality of the alerts by extracting scientometric informations corresponding to each diffused paper. We choose to analyse the citations number, the h-index of the first author, the class of the conference or scientific journal and the impact factor of scientific journals. For scientific books we can consider their citations number, author h-index and SJR of the book. These indicators have become widely accepted measures of the scientific production quality [17]. Also, we categorized the diffused scientific papers according to their types: journal paper, conference paper, book or other. The results of the qualitative analysis carried out are shown in Fig. 1,2,3,4 and 5.

We analysed the citations number of the diffused scientific paper for each system. We calculated the percentage of papers wich did not receive citations, which received a citations number between 0 and 10 and those which received a citations number greater than 10. We observe, in Fig. 1, that more than 43% of diffused papers have received 0 citations. These results are logical considering that those papers was been recently published. On the other hand, we observe a low percentage varying between 7% and 21% corresponding to the papers having citations number greater than 10. This result is justified by the fact that scientific papers diffused by academic social networks are available for free on the internet.

Another factor that may reflect the quality of scientific paper is the h-index of the first author. We analysed this factor for all diffused papers and we present the results in Fig. 2. Less than 26% of diffused papers have the h-index of the first author qual to 0. We note that 38% to 82% of the diffused papers have the h-index of the first author between 0 and 10. Then, 9% to 36% of the diffused papers have the h-index of the first author greater than 10. These results reflect a good impact of the first authors of the diffused papers which may influence the quality of scientific papers.

In Fig. 3 we present the gatogirization of diffused papers according to their types. We observe that 64% to 80% of the diffused scientific papers are published in scientific journals. Less than 15% are conference papers. We notice that 10% to 26% are of different types such as: books, book chapters, patents, presentations and forums.

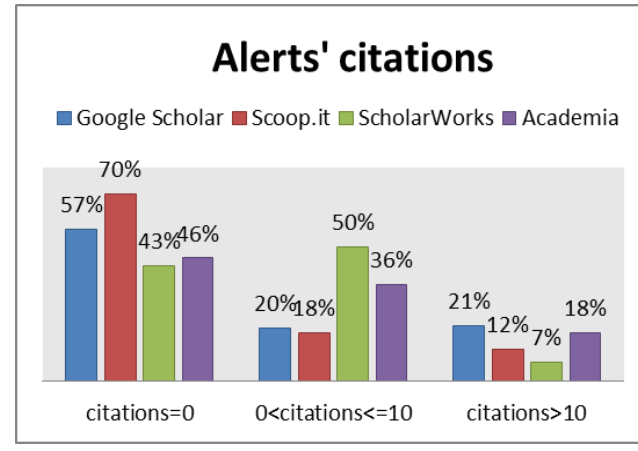


Fig. 1. Alerts' citation number

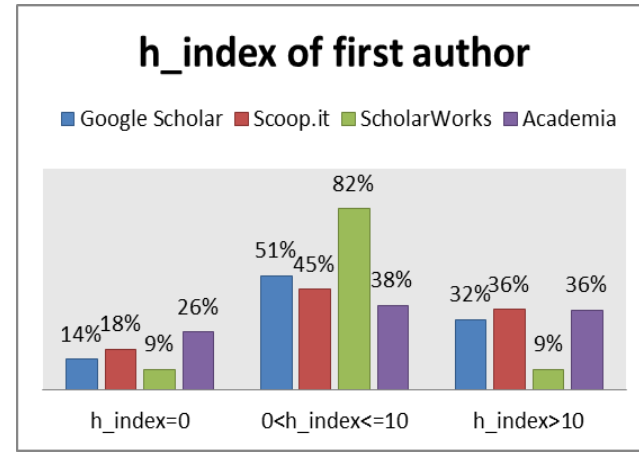


Fig. 2. H-index of the first author

We analysed the quality of scientific journals and conferences in which diffused papers are published. We used Core ranking[6] to provide these analysis. In Fig. 4, we present the percentages of papers published in conferences or journals of different classes: A*or A, B or C and no class. We observe that 5% to 26% of journals of conferences are of class A* or A. Less than 10% of conferences or journals are of class B or C. Then, 37% to 85% of journals or conferences have no class.

For diffused journal papers, we analyse the journal impact factors[7] provided by Clarivate Analytics. We present results in Fig. 5. We notice that 31% to 64% of the

[6] http://portal.core.edu.au/conf-ranks/
[7] https://jcr.clarivate.com/JCRJournalHomeAction.action

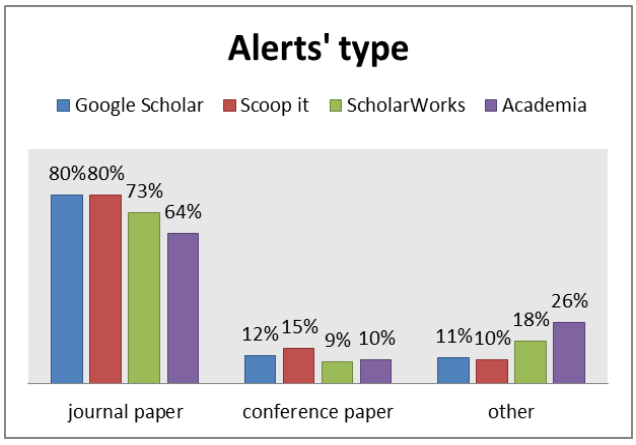


Fig. 3. Alerts' type

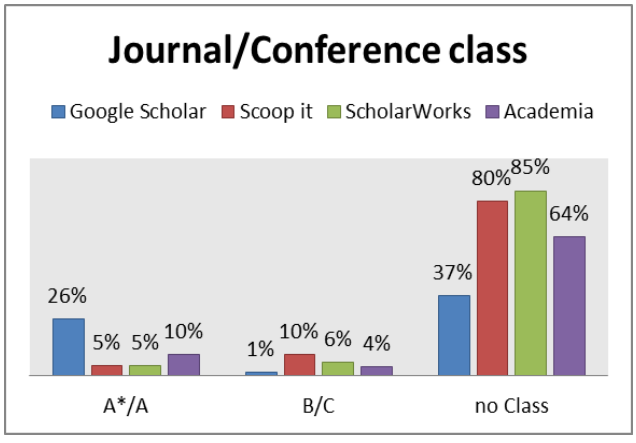


Fig. 4. Journal or conference class

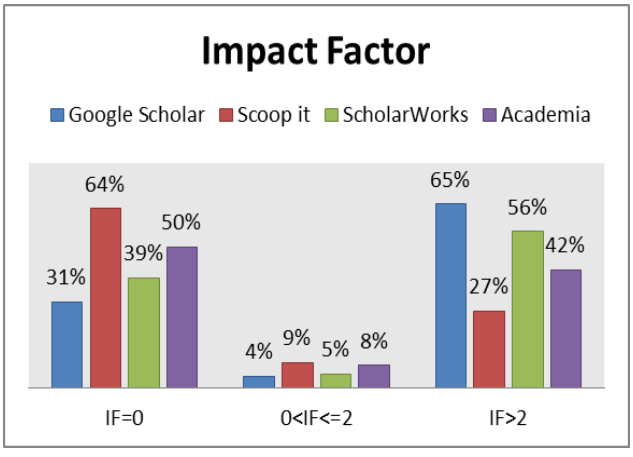


Fig. 5. Journal impact factors

journals have an impact factor equal to 0. Less than 9% of the journals have an impact factor between 0 and 2. Then, 27% to 65% have an impact factor greater than 2. We observe that the impact factors of majority publishing journals are either null or greater than 2. These results show a variation in the quality of publishing journals which, in turn, influences the quality of diffused papers.

4 Qualitative Scientific Watch

In our approach we consider scientific disciplines to extract informations published in scientific papers. Between all publications diffused by different scientific watch systems, it is not always easy for the researcher to choose qualitative ones. Our objective is therefore to present some tools that facilitate the documentary research and the establishment of a scientific watch of available scientific resources. Based on the qualitative study carried out, we justify our need for a qualitative tool to diffuse scientific information. This tool must analyze the quality of the information before alerting the users. Scientific quality is now determined by a set of metrics measuring the quality of scientific documents [18].

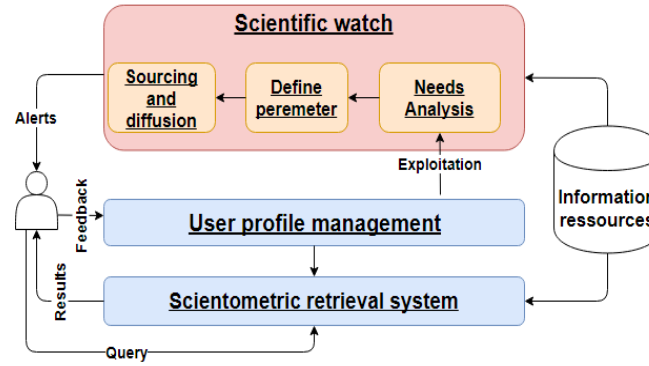


Fig. 6. Scientific watch approach

To integrate quality in scientific watch we propose to use scientometric indicators (document citations, author h-index [19], conference or journal class and journal impact factor) to analyse the quality of diffused papers. This analysis is based on the scientometric indicators of each paper detected.

We can adapt the diffusion criteria according to the preferences of researchers based on their scientometric preferences available in their profile [20]. In the context of the establishment of a complete platform for personalized retrieval and watch of qualitive scientific documents, we propose a qualitative scientific watch system based on scientometrics as shown in Fig. 6.

The process of scientific watch includes defining the themes of the scientific documents to be monitored, the identification and selection of their sources, analysis, synthesis and diffusion of these documents. This is in order to update documentary bases that will help the researchers in the detection of new scientific documents which are relevant to their research and correspond to their qualitative needs. This process requires document access tools, processing tools and communication tools.

In order to fulfil this mission, we recover the researcher qualitative preferences from his/her profile. Face with this permanent evolution, the validated scientific

document is indispensable in the documentary watch process.

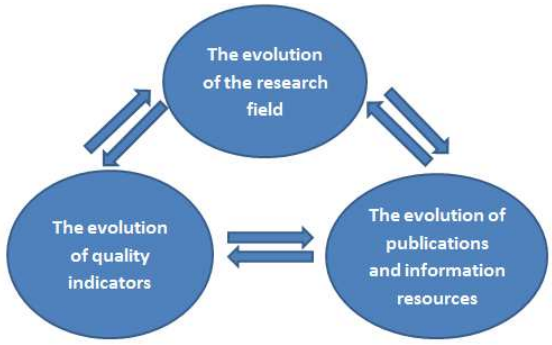


Fig. 7. Domains of researchers'needs

4.1 Needs Analysis

First of all, the environment in which the watch is going to be performed must be taken into consideration, precisely defining the needs to be covered. The context in which the researcher is placed is that of a so-called scientific information watch. In this perspective, it is necessary to know clearly what researcher needs to monitor in different identified domains as shown in Fig. 7.

The watcher must also take into account the objectives he sets and the orientation of the structure in which the watch is set up. The objectives of the watch can be of two different types: they can aim at a state of the art and / or the regular detection of novelties. The objectives are defined:

- Depending on the field: keep abreast of the production of knowledge in a discipline, discover the fields of research related to the field, identify the new most qualitative research and identify experts in the field.
- Depending on the problems, domain-specific constraints. Given the dynamic aspect of scientometrics domain, synchronization must be done to synchronize the different scientometric indicators considered from the different bibliographic databases.
- Depending on the products targeted: synthesis, state of the art, criticism, etc. ;
- Depending on the targeted audiences: watch for the researchers, watch for the research institution.
- Depending on the research institution: Each research institution is interested in improving the quality of its scientific production. Scientific quality requirements vary from one institution to another. Thus, the researcher's needs evolve according to these requirements.

4.2 Define a Perimeter for the Scientific Watch

Defining a perimeter is based on the establishment of a methodology deducted from the needs analysis and goal targeting.

- Thematic perimeter: It will be a question of defining the main themes on which to collect data, by identifying, if necessary, sub-themes and delimiting their borders and relations, to specify the subjects, to express them in questions of research and to translate them by keywords. The detection of thematic is based on the researcher preferences available on his/her profile.
- Linguistic perimeter: In any watching approach, we must think of defining a linguistic and geographical framework of reference: we may wish to limit ourselves to English resources.
- Scientometric perimeter: Based on the research institution requirements, we define the scientometric perimeter. We consider only publications on ranked scientific journals or conferences.

4.3 Sourcing and diffusion

Sourcing is identifying and selecting sources that can meet the needs of researchers. Sourcing is a fundamental step in any watch process. Source evaluation is particularly crucial in the current context of information overabundance and it is necessary to constantly assess the authority, relevance, quality, completeness, reliability and freshness of the information provided. Documents are automaticly annotated by our scientometric annotation system [21]. Our system consists on annotating scientific documents by scientometric indicators in addition to the thematic and semantic annotation [22]. In our context, we consider the quality of the document as a core feature of the watch system.

In our approach, we opted for radar watch mode. This mode covers a broad spectrum. In our case, the watch is done on many sources through queries, alerts, to identify new qualitative papers. Our proposed solution for sourcing and diffusion is:

i. Retrieving continuous alerts from existing scientific watch systems (Google Scholar, scoop.it, Academia, Scholarworks and sciencewatch)
ii. Analyzing the quality of the information received based on the scientometric annotation of each paper in addition to the user qualitative preferences.
iii. Rediffusion of qualitative documents selected by our watch system which corresponds to the user's preferences.

5 Discussion

We synthesize the qualitative analysis carried out on the various watch systems in Table 1. For each watch system, we calculate: (i) the mean of citations number received by diffused scientific papers, (ii) mean of h-index corresponding to the first author of each diffused paper, (iii) mean of the publisher score and (iv) mean of journals impact factors. The mean of the publisher score is calculated based on the type of publisher (journal,

conference or other) and its ranking (A*, A, B, C or no rank). We attribute different weights for each publisher type or rank according to Table 2.

TABLE 1
SYNTHESIS OF THE QUALITATIVE ANALYSIS OF THE DIFFUSED INFORMATION

	Mean of citations number	h-index mean	Mean of the Publisher Score	Mean of Journals Impact Factors
Google Scholar	↑ 4,48	↑ 11,27	↑ 1,56	↑ 1,81
Scoop.it	↓ 0,18	↑ 10,55	↓ 1,31	↓ 1,07
ScholarWorks	→ 2,18	↓ 5,73	↓ 1,19	→ 1,38
Academia	↑ 3,78	→ 8,29	↓ 1,23	↓ 1,27

TABLE 2
WEIGHTS CORRESPONDING TO THE PUBLISHER TYPE AND RANK

Class	Weight	Type	Weight
A*	1	Journal	1
A	0,8	Conference	0,8
B	0,6	Other	0,6
C	0,4		
No Class	0,2		

Based on the results in Table 1, we can remark that the information diffused by Google Scholar is more qualitative than other watch systems. This can be justified by the fact that Google Scholar is more comprehensive source than the others. Furthermore, the free uncontrolled information available in academic social networks can be of an unsupervised quality. However, the quality presented by these watch systems does not necessarily corresponds to the preferences of the user who is the researcher.

6 CONCLUSION

Based on our qualitative study of available scientific watch systems, we determined our need for a qualitative scientific watch. We proposed a scientific watch approach based on scientometric indicators. The user can receive alerts of new qualitative publisher papers which corresponds to his/her needs and preferences. We collect new published papers from the collection of papers diffused by scientific watch systems previously studied (Google Scholar, Scoop.it, ScholarWorks, ScienceWatch and Academia).

Establishing an effective scientific watch is a laborious exercise. It should be noted that the preparatory stages for documentary research and the setting up of a scientific watch are essential. The choice of keywords and concepts will have a direct impact on the results of the research at first, then on the effectiveness of the scientific watch in a second time from the moment when we export the results of the research carried out in the form alerts. The multiplicity of our resources makes our watch system more efficient and comprehensive.

The proposed watch system is not intended, however, to replace bibliographic databases that a researcher can access to retrieve informations. Diffused scientific informations are complementary to traditional scientific resources in order to be more effective in the process of collecting scientific and technical information.

REFERENCES

[1] O. Y. Rieger, "Understanding interdisciplinary ecosystems: social construction of scholarly communication", Corne ll University Library, Apr. 2008.

[2] S. Belli, R. Cardenas, M. Velez, A. Rivera and V. Santoro, "Open Science and Open Access, a Scientific Practice for Sharing Knowledge", 2019.

[3] Y. Luan, "Information Extraction from Scientific Literature for Method Recommendation", *arXiv preprint arXiv:1901.00401*.

[4] R. György, "*Scientific information and society*", Walter de Gruyter GmbH & Co KG, 2019.

[5] D. D. S. Price, "Quantitative measures of the development of science", Archives Internationals d'Histoire des Sciences, vol. 14, pp. 85-93, 1951.

[6] D. D. S.Price, Science since Babylon, New Haven: Yale University Press, 1961.

[7] S. Tveden-Nyborg, M. Misfeldt and B. Boelt, "Diffusing scientific knowledge to innovative experts", Journal of Science Communication, vol. 12, no. 1, 2013.

[8] S. Ovadia, "ResearchGate and Academia. edu: Academic social networks", *Behavioral & social sciences librarian*, vol. *33, no.* 3, pp. 165-169, 2014.

[9] S. Ovadia, "When social media meets scholarly publishing", Behavioral & Social Sciences Librarian, vol. 32, no. 3, pp. 194–98, 2013.

[10] J. Lee, S. Oh, H. Dong, F. Wang and G. Burnett, "Motivations for self-archiving on an academic social networking site: A study on researchgate", *Journal of the Association for Information Science and Technology*, vol. *70, no.* 6, pp. 563-574, 2019.

[11] M. Thelwall and K. Kousha, "A cademia. edu: Social network or A cademic Network?", *Journal of the Association for Information Science and Technology*, vol. *65, no.* 4, pp. 721-731, 2014.

[12] A. Antonio and D. Tuffley, "Creating educational networking opportunities with Scoop. It", *Journal of Creative Communications*, vol. *9, no.* 2, pp. 185-197, 2014.

[13] R. Lambiotte and P. Panzarasa, « Communities, knowledge creation, and information diffusion", *Journal of Informetrics*, vol. *3, no.* 3, pp. 180-190, 2009.

[14] M. Norris, C. Oppenheim and F. Rowland, "The citation advantage of open-access articles", *Journal of the American Society for Information Science and Technology*, vol. *59, no.* 12, pp. 1963-1972, 2008.

[15] A. Noruzi, "Hot Papers in Library and Information Science from the Point of View of Research Methods", *Webology*, vol. *14, no.* 2, 2017.

[16] Y. Niyazov, C. Vogel, R. Price, B. Lund, D. Judd, A. Akil and M. Shron, "Open access meets discoverability: Citations to articles posted to Academia.edu", *PloS one*, vol. *11, no.* 2, 2016.

[17] T. Zainab, Z. A. Wani, "Encyclopedia of Information Science and Technology, chapter Advancement and Application of Scientometric Indicators for Evaluation of Research Content", pp. 6739-6747, IGI Global, 2018.

[18] B. Hammarfelt and A. D. Rushforth, "Indicators as judgment devices: An empirical study of citizen bibliometrics in research evaluation", Research Evaluation, vol. 26, no. 3, pp. 169-180, 2017.

[19] Y. Huggins-Hoyt K, "African American Faculty in Social Work Schools: A Citation Analysis of Scholarship", Research on Social Work Practice, vol. 28, no. 3, pp. 300-308, 2018.

[20] N. Ibrahim, A. Habacha Chaibi and H. Ben Ghézala, "A new Scientometric Dimension for User Profile", In The 9th International Conference on Advances in Computer-Human Interactions (ACHI'16), Venice, Italy, 24-28 April 2016. p.261-267.
[21] N. Ibrahim, A. Habacha Chaibi and H. Ben Ghézela, "A Scientometric Approach for Personalizing Research Paper Retrieval", In the 20th International Conference on Enterprise Information Systems (ICEIS'18), Maderia, Portugal, March 2018. p.419-428.
[22] F. Kboubi, A. Habacha Chaibi and M. Ben Ahmed. Semantic visualization and navigation in textual corpus. International Journal of Information Sciences and Techniques (IJIST), vol. 2, pp. 53–63, 2012.

Semantic Mining Approach Based On Learning of An Enhanced Semantic Model For Textual Business Intelligenge

Hammou Fadili

Abstract— This article presents the results obtained in applying an innovative approach to textual semantic analysis in the service of decision-making. Significant improvements have been made in the existing procedures of sentiment and recommendation analysis, and in opinion mining, to enable better-motivated decisions and benefit from big data. The results obtained show the interesting contribution of the approach to the specific field of business intelligence (BI) relative to user behaviors analysis.

Index Terms—Machine learning, Modeling and prediction, Natural Language Processing, Neural models, Semantics, Sentiments analysis, Text Mining

1 Introduction

This article presents a research on applying sentiment and recommendation analysis, opinion mining and user behavior understanding to Business Intelligence (BI). The work and its results are part of a general research strategy on economic intelligence and their exploitation in the context of Big Data.

In this context, data comes mainly from the reviews and opinions of users/customers, reported via social medias, forums, blogs, sales sites, etc., on all kinds of topics, such as events, products, attitudes, etc., to express their sentiments and experiences. The formalization of sentiment and recommendation analysis, and the exploration of users' opinions and experiences, have become a major issue in the BI domain. Indeed, several studies and statistics confirm that more than 80% of users shopping on the Internet consult the comments and opinions of former users before making their own purchases [49]. Recommendations influence our opinions about products, services, etc., and therefore influence our purchases.

Moreover, despite advances in this area, several recent studies show that only 29% of companies and organizations use the data in their decisions [48]. The main reason for this is that existing solutions are not yet mature enough while the cost of implementing them is still too high. These technologies can also be used to enable companies to measure and better protect their reputation.: Individuals who use the Internet freely to express sentiments and opinions on everything can damage the reputation and activities of organizations. The latter must take these phenomena into account to protect themselves better. E-reputation has also become a new emerging field of research. Businesses and institutions therefore, have a great interest in understanding feedback from their customers. They must invest heavily in automatic systems analysis, to be helped to adapt to users' requirements and improve their profits; in the new open world of business intelligence that is the Web.

The phenomena of business data analysis for intelligence purposes in the context of Big Data challenges the traditional methods and approaches of modeling, design, implementation and evaluation of information and decision systems. Thus, the integration of new technologies, such as semantic textual analysis and artificial intelligence, is an unavoidable necessity. Whereas most existing solutions deal mainly with structured data, they must evolve to survive in the context of big and unstructured data. They encounter several problems when it comes to integrating and/or federating unstructured heterogeneous data sources as natural language texts: the exact meaning of words and thus texts in their contexts is very difficult to detect and to exploit. These problems are even more accentuated, in the case of Web texts often written with short texts (SMS), abbreviations, acronyms, styles & very particular elements of speech such as irony, humor, and so on.

To answer these questions, we have engaged in a research work as part of a global approach whose objective is to set up innovative methods to deal with the specific problems still posed by the semantic analysis of textual data in the service of decision-making. The goal is to make

• *Hammou Fadili*
Laboratoire CEDRIC du CNAM Paris
(Pôle Recherche & Prospective, Programme Maghreb) de la FMSH Paris
hammou.fadili(at)(cnam.fr , msh-paris.fr).

improvements to existing procedures used in sentiment and recommendation analysis and in opinion mining to enable better motivated decisions and benefit from big data. This will allow organizations to know how their products and services are perceived and to adapt them to the demands of their users/customers.

The issues raised during this work may relate to several areas of BI research but, with a view to simplification, we concentrated our research and development efforts on sentiment analysis and opinion mining. We were interested in the semantic analysis of unstructured and noisy data for the sentiment analysis.

This article is structured as follows. The first part will be devoted to a review of the state of the art, followed by a second part on the issues raised. In the third part, we present our contribution, before concluding the article.

2 STATE OF THE ART

This part is devoted to a review of the state of the art in sentiment analysis in the context of Big Data, according to the following major points: semantic sentiment analysis, domains of use of sentiment analysis and opinion mining, the different levels of analysis and, finally, the different approaches to calculating polarities.

A. Sentiment analysis and semantic text mining

A sentiment is a tendency to feel an emotion about an object or about a person [1] [2]. It also includes people's opinions, evaluations and feelings about entities, events, etc. [6]. In the context of Big Data, sentiments are often expressed in writing through product reviews, websites, blogs, discussion forums, and so on. In the context of the Web, these writings often convey messages and opinions difficult to detect and extract, because they are frequently expressed in languages and styles specific to the Web: SMS, acronyms, abbreviations, irony, humor, etc.

Sentiment analysis consists in attributing a relative polarity to subjective elements (words and sentences that express opinions, sentiments, emotions, ...) to decide on the orientation of a document (Turney, 2002). The interest in sentiment analysis and opinion mining, in a context where taking charge of several important elements such as speech acts, metaphor, frozen idiomatic expressions ..., naturally situates sentiments analysis in the general context of semantic text analysis of Big Data in the service of business intelligence. Big Data here is defined as consisting mainly of massive, unstructured & heterogeneous textual data

Given the nature of the data, mainly textual, the techniques used to detect the messages conveyed and the opinions, must exploit Natural Language Processing (NLP) and linguistics to identify the polarity (positive, negative, or neutral) of a word, a sentence, a text or even a corpus.

B. Domains of use

Sentiments analysis is based, for the most part, on text mining. It can target the study, the detection and the extraction of sentiments, opinions, emotions and subjectivities in the text (Pang et al., 2004) in specific areas, such as advertising, marketing, production, business, politics, psychology, etc.

C. Classification of sentiment analysis technologies

Sentiment analysis can be used in several areas and for different needs: studies of the polarities and subjectivities of words (SentiWordNet and equivalents), expressions (phrases), documents (the most common case), corpora, etc. This involves different approaches and strategies characterized by different levels of analysis.

1) Word level analyses

Word level analyses deal with the polarities of words. They represent the bases of the analyses of the other levels: sentence, expression and document. Word level analyses are of two types: those based on lexicons of sentiments and those based on corpora.

2) Lexicon-based approaches for sentiment analysis

These approaches are considered unsupervised as they exploit dictionaries and lexicons where the words are characterized by polarity scores (cf. [44]). For example, the positive score of the word “excellent” is 0.8, its neutral score is 0.3 and its negative score is 0 (see [45]). Several methods are then used to calculate the polarity of several words: opinions, sentences, paragraphs and documents.

3) Corpora-based approaches for sentiment analysis

This kind of methods is considered supervised as it exploits a pre-processed corpus, i.e., a corpus that has been already annotated in order to train and learn a polarity classifier. In general, in this case, we cannot be limited to the "Positive", "Neutral" and "Negative" polarities, but we can enrich and refine the classes of polarities with others according to needs.

4) Sentence-level analyses

The phrase refers to a small group of words, forming a conceptual unit, usually a component of a clause. Sentiment analysis of the level of the sentence is as important as sentiment analysis of the level of the word. It allows the analysis of the higher levels: expression, document, etc. It is an important element in the process of analyzing sentiments. Several studies have been conducted in this area. They exploit systems based either on machine learning [46], on statistical models [47] or on dictionaries of polarity [45].

5) Expression level analyses

The term also called "complete sentence" refers to a set of words that is complete, usually containing a subject and a predicate, conveying a statement, question, exclamation…; and consists of a main clause and sometimes one or more subordinate clauses. An expression can express a thought, a sentiment, an opinion, etc.

Several studies have also been conducted in this area (see [34]) and technologies for the classification of complete sentences were developed based on machine learning. The comments are first pretreated and annotated with grammatical annotations, among others. Their polarity is then calculated by combining partial results (words and sentences), using algorithms such as [29] based on statistical methods combined with a log-likelihood computation to calculate the scores.

6) Document level analyses

This kind of analysis looks at the sentiment of the whole document, for example, of specific piece of information, an opinion, a comment, a forum, a blog, etc. Several works have been done for the sentiment analysis of the document. They generally exploit machine learning-based approaches for classifying and inferring the polarity at the document level. In [48], the authors exploited unsupervised methods to classify documents using several steps:

- They extracted adjectives and adverbs using grammatical annotation systems,
- Then they extracted the sentences and their polarities ([31])
- Finally, they calculated the average to deduce the polarity of each document.

The results obtained were of the order of 70%.

D. Polarity measurements

From a technological point of view, the Big Data context imposes the use of machine learning for calculating polarity. Several studies have shown that: training a system on large amounts of data greatly increases its performance. Big Data is therefore a considerable asset for machine learning-based systems.

In this section, we will present some methods for classification and calculating scores and polarities; which fall into two categories:

- Unsupervised calculation methods, from lexicons of sentiments
- Supervised calculations methods, from annotated corpora

1) Unsupervised calculations from lexicons of sentiments

Lexical based methods of sentiments calculation are unsupervised because they do not require annotated corpora. They are based on the automatic extraction of discriminant characteristics (features) from text and the exploitation of sentiments lexicons from resources such as SentiWordNet [66] and SentiStrength [71] for classification.

In [49] for example, the authors calculate the sum of a synSet scores for different Tags of words, to which they associate a polarity (negative, neutral, positive). Note that for Tags that are not in SentiWordNet, they associate the score 0. They also refined the analysis by choosing the right score according to the grammatical category (noun, adjective and verb). They explain this by the following example: the word "short" admits 11 adjectives, 3 nouns, 7 adverbs and 2 verbs for the different meanings. For the adjective Tag, the term "short" is -1, for the nounTag the term is 0, and so on.

The methods in this category are based on additional features that can refine the analysis. We can target a point of view of analysis, for example a product, an organization, a service, a theme, etc. In [61], the authors explain their process as follows: they targeted a product and then extracted some features from user comments. They then identified the comments of each characteristic (by grouping or clustering), then designated all the comments of a positive, neutral or negative characteristic. Finally, they refined their experiences by combining the results obtained for each characteristic of a product.

Our approach belongs to and improves this last category. Instead of limiting the calculations directly to combinations of the scores of certain characteristics, we extend and enrich all the characteristics that we extract automatically and confront with a sentiment lexicon to submit to a recurrent neural network (RNN) with long-term memory (LSTM), on which improvements have been made, for the calculation of the global polarity (see our approach below).

2) Supervised calculations methods, from annotated corpora

The classification methods based on annotated corpora are supervised, they need an annotated corpus. They are used to train and learn automatic polarity measurement systems. The research in this area is slim, because of the difficulties involved in manual annotation processes, especially in the case of Big Data. Among research works conducted in this category, we find, for example, [47], where several news articles were classified into 7 categories ranging from 1 (very negative) to 7 (very positive) (cf. [56]).

3) « Emoticons » based calculations methods

Another way to calculate polarity scores is to use emoticons, which are icons that express an emotion, a state of

mind, a sentiment, feeling, and so on. They are often used in social networks such as Facebook, Twitter. Several works have exploited emoticons for classification. Contents containing positive emoticons have been classified as positive, those containing negative Emoticons are classified negative and those containing no emoticons are classified as neutral (see [59]). The use of emoticons is limited and does not represent a solution in itself, because, on the one hand, not all documents include emoticons and, on the other hand, several studies have shown that there is often a mismatch between the emoticons and the associated sentiments.

3 PROBLEMS OF MOTIVATION

From the point of view of content, sentiment analysis has become an important discipline in Big Data mining. Although there are several proposals and approaches in the literature, there is no solution that can support all aspects of sentiments analysis. Their numerous limitations are due to several considerations, as reflected the following findings, among others:

- The quality of sentiment analysis depends on the degree of support for the semantics of the texts that convey them;
- The presence of sentiments and opinions is conditioned by subjectivity;
- Sentiments depend on the domain, on the analytical point of view, etc.
- Sentiments analysis requires substantial data and technological resources, in addition to sophisticated processing systems.

We consider that these aspects, poorly or not at all modeled and supported in the current systems, are important in terms of optimization and efficiency. They constitute the basis of the problems to be solved in our work, which will be presented through the two prisms of subjectivity and semantics for sentiments analysis.

From the technological point of view, we believe that these aspects remain challenges for sentiment analysis, but also that they will be addressed by possible improvements in the exploitation of advanced artificial intelligence technologies to extract the knowledge inherent in the text.

Indeed, overall, there are several technologies used in the field of sentiment analysis. According to the literature, deep learning technology (Deep Learning) is now, by far, the most popular and exploited technology. It provides the best results compared to other technologies, especially in the context of Big Data.

Most of the existing technologies exploit the words in text directly, regardless of their meanings or their order. The most commonly used technology in this field "Word Embeddings" or "Word2Vec" and its various implementations (SKIP-GRAM & CBOW) [33], which does not 'learn' the meaning of words nor their order in their vector's representations. It is obvious that the loss of order implies 'automatically' the loss of semantics: moreover, we know several sentences, where the meaning changes when we change the order of words. Implementing technological solutions that take into account the meaning and order of words can make significant improvements. This is a first major technological problem we considered in current work.

Another problem encountered by the deep learning architectures exploited in the field of sentiment analysis is the fixed size of networks. A configured and trained classical neural network keeps its parameters throughout its entire life cycle. This poses a problem in the case of sentences, paragraphs, texts, etc. composed of variable numbers of words and has had a decisive impact in the choice of our solution. We adopted approaches based on an improved variant of recurrent neural networks (RNN) that are independent of the number of words to be processed and therefore the length of sentences and documents. We used N-GRAMS and sliding windows of fixed sizes to work around this problem, despite the loss of information occasioned by the fact that the meaning of a word is related to the sentence containing it and not necessarily to a window of words that surround it.

Additionally, deep learning means multilayer architectures, which in some cases can be composed of a large number of layers. It can be argued that since:

- The training of the network is done by the optimization of the parameters minimizing the prediction error, by the backpropagation of the error;
- The optimization is done by the gradient descent method, which is used to modify the coefficients (parameters of the system);
- In the case of a recurrent network, the learning is calculated by multiplying the coefficients of a layer by a specific value, λ, calculated from the loss, as many times as the rank of the layer in the network;

Then:

- If λ is less than 1, the learning of the first layers is multiplied by a number close to zero, so they do not learn anymore. We speak in this case of the Vanishing of the Gradient;
- If λ is greater than 1, the learning of the first layers is multiplied by a high number, provoking the explosion of the coefficients (no learning either). In this case, we speak of Explosion of the Gradient.

Addressing the problem of the vanishing or the explosion of the gradient is another important element we consider.

4 OUR CONTRIBUTION ANDEXPERIMENTATIONS

A. Context of the work

This section briefly recalls the general context of the work; in terms of issues and activities, and their evolution.
Initially, this work was done in the context of several projects on the analysis and exploitation of structured data and more specifically on business intelligence (BI) from databases. The evolutions of information systems and uses, generating significant amounts of heterogeneous data and nonexistent or complex data models, forced us to extend our processes to all kinds of data (not just structured) by integrating the semantic analysis of unstructured and more specifically textual data.

B. Our approach and experiments

The scientific and experimental approaches we adopted are of two types: those related to contents and their semantics, and those considering the technological and processing problems.

From the technological point of view, we have made the choice of using deep architectures because they have been tested and successfully used in the case of natural language processing, and also because they are well adapted to sentiments analysis. During our work, several configurations were tested, some of which generated a lot of difficulties and/or were abandoned because of certain layouts of symbolic and numerical/statistical layers in the learning models. The location of symbolic models in the general process was a problem; because they are difficult to set up, except for very limited cases. They require a lot of resources and knowledge, and therefore a lot of work.

To circumvent these difficulties, we integrated unsupervised models upstream of the process. These models are based on mathematical calculations in optimized, well-adapted semantic spaces. The advantage is that one does not need previously processed information; the data to be analyzed and the mathematical models are used directly to deduce the learning models and their representations.

This approach has the advantage of combining the two kind of learning (supervised and unsupervised), at different levels, and with specific layout. Its originality lies in the fact that, on the one hand, it incorporates important notions into the model of the sentiments analysis: polarity, subjectivity, etc. while, on the other hand, it uses these notions to generate a rich semantic data model on which we applied the learning algorithms.

As previously mentioned, deep learning has been successfully exploited in many areas including natural language processing. One of the best uses is the generation of dense semantic vector spaces (Mikolov 2013). More recently, deep learning networks were then enhanced by creating Recurrent Neural Networks (RNNs) to support sequential data, where each state is calculated from its previous state and the new entry. These recurrent networks can propagate information in both directions: to the input layers and to the output layers. They are an implementation of neural networks, close to the functioning of the human brain, where information can spread in all directions while exploiting memory via recurrent connections propagating the information of a subsequent learning (information stored).

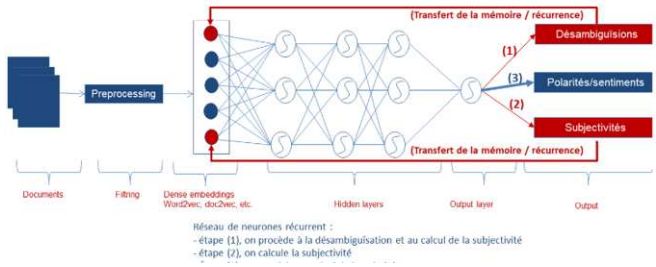


Fig. 1. Adaptation of a deep architectures for sentiments analysis

The depth of RNNs can be high because of their sequential nature, generally depending on the number of words to be processed. This can provoke:

- Vanishing of the Gradient in the first layers and stopping learning from a certain depth
- Explosion of the Gradient in the first layers and stopping learning from a certain depth

The Long Short-Term Memory (LSTM) architecture of RNNs was designed to address these issues, by optimizing control gates for the propagation of information in the network.

For the needs of Big Data, we adopted and improved the LSTM architecture initially based on three gates. We have additionally integrated the concepts of perspective, followed by the notion of attention. The new LSTM model now allows to control:

- What to use from the input
- What to use from the hidden state
- What to send to the output

and this depending on the chosen point of view or perspective and by paying attention to relevant elements.

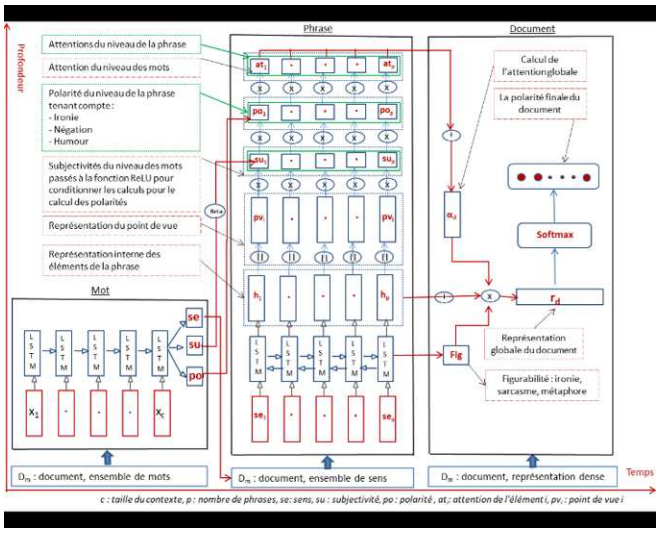


Fig. 2. Graphic version of our algorithm

These gates allow to cancel certain information that is useless for prediction and to reinforce other elements. It can be demonstrated mathematically that this architecture, in addition to optimizing the calculations in Big Data, makes it possible to solve the problems related to the vanishing and explosion of the gradient described above.

From a technological perspective, we have adapted the global semantic context model for the specific treatments of subjectivity, polarity, irony, metaphor, expressions (locutions). The notion of semantic context we are talking about here concerns all the parameters that can influence the meaning of words in the context of BI. This is a difficult notion to define, because there are no studies today defining in an exhaustive manner this notion in general and in the field of BI. It has been deduced that the context in the domain of BI can also be viewed as an aggregation of parameters of the context of the natural language (H. FADILI 2017), increased by new parameters specific to BI.
It should be noted that, in this section, only the new parameters that are to be combined with those of the natural language will be presented to form the overall context for BI. We classified them as following:

- Subjectivity: parameters to represent the subjectivity or not, of a word, a sentence, a document, etc.
- Polarity: parameters to represent the polarity, of a word, a sentence, a document, etc.
- Discourse analysis context: parameters to represent aspects of discourse analysis: irony, metaphor, expression, etc.

The detection of subjectivity consists in determining if a unit of language (word, phrase, document) expresses a personal attitude, an opinion, etc. and, if so, what is its polarity? Polarity has been conditioned by subjectivity to lighten the processing. At the level of the approach, the innovation consists of associating subjectivity to the sense and not to the words, to circumvent the issue of ambiguity negatively impacting the subjectivity and polarity calculations. We have distinguished two phases of analysis, after the disambiguation phase:

- Calculation of subjectivity, followed by
- Polarity calculation.

Indeed, a subjective word can have different polarities depending on the context. Its meaning can be modified by irony, humor, etc. Disambiguation alone is not enough. Dictionaries, mapping tables and ontologies can be used to model and represent certain elements of the context. We can, for example, use mappings to recognize the use of words in a domain, to find the complete forms of acronyms and initials or even to extract expressions for specific treatments.

As with natural language, sentiment analysis also needs the most complete representation of a word, a text (comment, recommendation, etc.), or a corpus, to capture all the discriminating information needed to deduce its exact polarity. Figure 1 models a representation of these elements, integrating several important notions for the analysis of unstructured data in the service of BI.

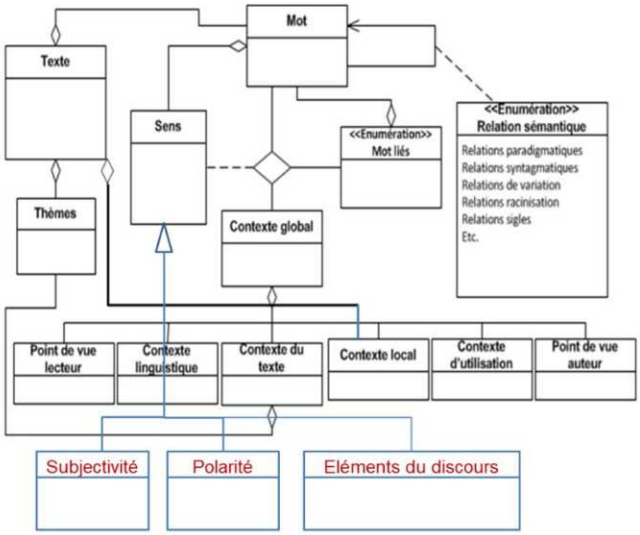


Fig. 3. Semantic model augmented by subjectivity, polarity and elements of discourse

C. Test scenarios

Several modules and programs have been developed that implement the elements of the overall process, mainly:

- Automatic extraction of the various semantic features of words and texts for sentiments analysis.
- Implementation of a deep learning system based on RNN and LSTM model with the attention and domain mechanisms described previously.

- We have also designed a development environment that centralizes access to all modules:
- Integration of all the elements in a single workflow implementing all the modules of the process.

As required in 'honest' machine learning systems, we divided the generated learning data into three parts:
A first part, representing 20% of the dataset, was used for validation, to optimize the hyper-parameters of the system: the learning step, the type of the activation function and the number of layers.

The rest of the dataset was divided into two parts:

- 60% for training, to estimate the best coefficients (wi) of the neural network function, minimizing the error between the real outputs and the desired outputs.
- 20% for tests, to evaluate the performance of the system.

During the learning phases, the system is autonomous. It generates the characteristics (features) of the text for the training, so that the trained model will be able to deduce the correct sense of each word of the text, as well as the subjectivity and the polarity of the concerned elements (words, sentences, documents, corpora, etc.), when applied to unseen data.

D. An overview of the results

The results obtained during training and testing of the system are shown in Figures 2 and 3.

1) Evolution of accuracy during training & tests

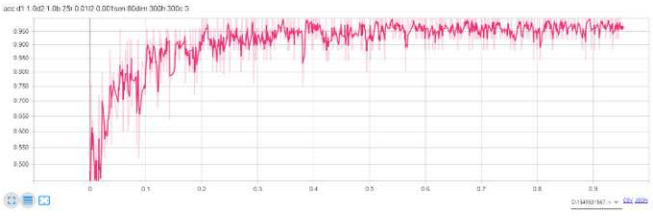


Fig. 4. An overview of the results: accuracy

2) Evolution of the loss during training & tests

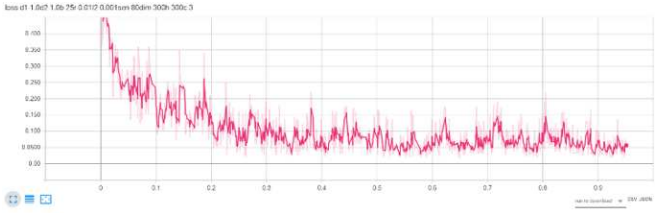


Fig. 5. An overview of the results: loss

The results of the tests show the good performance of the approach, from the evolutions of the accuracies and the errors. The system succeeds thanks to the learning on a part of the instances to deduce in the end the polarity of the concerned elements (corpus, document, paragraph or sentence) with better rates of precision and error.

7 CONCLUSION

Sentiment analysis has become a very important discipline in the exploitation of Big Data. Although there are several proposals and approaches in the literature, there is at present no solution that can handle all the requirements of the sentiment analysis process. It is in this context that we have proposed a solution that targets the following aspects:

- Improved disambiguation at the beginning of the process;
- Better use of global context to deduce semantics and then subjectivity and polarity;
- Potential for mapping ontologies and other knowledge bases to solve problems such as acronyms, SMS, phrases (expressions), etc.
- Implementation of an optimized LSTM version of a deep neural network.

At the current state of progress of the project, only a part of the problems described above has been addressed and solved; the remaining parts constitute our planned future work.

REFERENCES

[1] A Neelakantan, J Shankar, A Passos, A McCallum. Efficient non-parametric estimation of multiple embeddings per word in vector space. Conference on Empirical Methods in Natural Language Processing, 2014

[2] Cui Tao, Dezhao Song, Deepak Sharma, Christopher G. Chute, Semantator: Semantic annotator for converting biomedical text to linked data. Journal of Biomedical Informatics, Volume 46, Issue 5, Pages 882-893 (October 2013). DOI: 10.1016/j.jbi.2013.07.003

[3] Das, T. K., & Kumar, P. M. (2013). Big data analytics: A framework for unstructured data analysis. International Journal of Engineering and Technology, 5(1), 153-156.

[4] Boury-Brisset, A.-C. (2013), Managing Semantic Big Data for Intelligence., in Kathryn Blackmond Laskey; Ian Emmons & Paulo Cesar G. da Costa, ed., 'STIDS' , CEUR-WS.org, , pp. 41-47 .

[5] Delia Rusu , Blaž Fortuna , Dunja Mladenić. Automatically Annotating Text with Linked Open Data (2011).Venue: In 4th Linked Data on the Web Workshop (LDOW 2011), 20th World Wide Web Conference.

[6] Archit Gupta, Krishnamurthy Viswanathan, Anupam Joshi, Tim Finin, and Ponnurangam Kumaraguru. Integrating Linked Open Data with Unstructured Text for Intelligence Gathering Tasks. Proceedings of the Eighth International Workshop on Information Integration on the Web, March 28, 2011.

[7] Isabelle Augenstein. Lodifier: Generating Linked Data from Unstructured Tex". ESWC 2012

[8] Marin Dimitrov. From Big Data to Smart Data. Semantic Days May 2013

[9] Khalili, A.; Auer, S. & Ngonga Ngomo, A.-C. (2014), conTEXT -- Lightweight Text Analytics using Linked Data, in 'Extended Semantic Web Conference (ESWC 2014)'.

[10] E. Khan, "Addressing Big Data Problems using Semantics and Natural Language Understanding," 12th Wseas International Conference on Telecommunications and Informatics (Tele-Info '13), Baltimore, September 17-19, 2013.

[11] E. Khan, "Processing Big Data with Natural Semantics and Natural Language Understanding using Brain-Like Approach", submitted to Journal– acceptance expected by Dec. 2013 Jan 2014.

[12] James R. Curran, Stephen Clark, and Johan Bos (2007): Linguistically Motivated Large-Scale NLP with C&C and Boxer. Proceedings of the ACL 2007 Demonstrations Session (ACL-07 demo), pp.33-36.

[13] Hans Kamp (1981). A Theory of Truth and Semantic Representation. In P. Portner & B. H. Partee (eds.), Formal Semantics - the Essential Readings. Blackwell. 189-222.

[14] Minelli, Michael & Chambers, Michele & Dhiraj, Ambiga 2013. Big Data, Big Analytics: Emerging Business Intelligence and Analytics Trends for Today's Businesses.

[15] Chan, Joseph O. "An Architecture for Big Data Analytics." Communications of the IIMA 13.2 (2013): 1-13. ProQuest Central. Web. 6 May 2014.

[16] H. Fadili. Towards a new approach of an automatic and contextual detection of meaning in text, Based on lexico-semantic relations and the concept of the context., IEEE-AICCSA , May 2013.

[17] George A. Miller (1995). WordNet: A Lexical Database for English. Communications of the ACM Vol. 38, No. 11: 39-41.

[18] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten (2009); The WEKA Data Mining Software: An Update; SIGKDD Explorations, Volume 11, Issue 1.

[19] Hiemstra, P.H., Pebesma, E.J., Twenhofel, C.J.W. and G.B.M. Heuvelink, 2008. Real-time automatic interpolation of ambient gamma dose rates from the Dutch Radioactivity Monitoring Network. Computers & Geosciences, accepted for publication.

[20] Christian Bizer, Tom Heath, Kingsley Idehen, Tim B. Lee. Linked data on the web (LDOW2008), In Proceedings of the 17th international conference on World Wide Web (2008), pp. 1265-1266.

[21] Jianqing Fan, Fang Han, Han Liu. Challenges of Big Data analysis National Science Review, Vol. 1, No. 2. (1 June 2014), pp. 293-314.

[22] http://wiki.dbpedia.org/

[23] Publication MEDES 2016 : Towards an Automatic Analyze and Standardization of Unstructured Data in the context of Big and Linked Data. H. FADILI.

[24] Publication TICAM'2016 : Le Machine Learning : numérique non supervisé et symbolique peu supervisé, une chance pour l'analyse sémantique automatique des langues peu dotées. H. FADILI.

[25] Frijda, N. H., Mesquita, B., Sonnemans, J., & Van Goozen, S. (1991). The duration of affective phenomena or emotions, sentiments and passions.

[26] Shand, A. F. (1920). The foundations of character: Being a study of the tendencies of the emotions and sentiments. Macmillan and Company, limited.

[27] Zhou, X., Tao, X., Yong, J., & Yang, Z. (2013, June). Sentiment analysis on tweets for social events. In Computer Supported Cooperative Work in Design (CSCWD), 2013 IEEE 17th International Conference on (pp. 557-562). IEEE.

[28] Park, C., & Lee, T. M. (2009). Information direction, website reputation and eWOM effect: A moderating role of product type. Journal of Business research, 62(1), 61-67.

[29] Nassirtoussi, A. K., Aghabozorgi, S., Wah, T. Y., & Ngo, D. C. L. (2014). Text mining for market prediction: A systematic review. Expert Systems with Applications, 41(16), 7653-7670.

[30] Duan, W., Cao, Q., Yu, Y., & Levy, S. (2013, January). Mining online user-generated content: using sentiment analysis technique to study hotel service quality. In System Sciences (HICSS), 2013 46th Hawaii International Conference on (pp. 3119-3128). IEEE.

[31] Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment. Icwsm, 10(1), 178-185.

[32] Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012, July). A system for real-time twitter sentiment analysis of 2012 us presidential election cycle. In Proceedings of the ACL 2012 System Demonstrations (pp. 115-120). Association for Computational Linguistics.

[33] Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. Computational linguistics, 37(2), 267-307.

[34] Denecke, K. (2008, April). Using sentiwordnet for multilingual sentiment analysis. In Data Engineering Workshop, 2008. ICDEW 2008. IEEE 24th International Conference on (pp. 507-512). IEEE.

[35] Le, Q., & Mikolov, T. (2014). Distributed representations of sentences and documents. In Proceedings of the 31st International Conference on Machine Learning (ICML-14) (pp. 1188-1196).

[36] Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013). New avenues in opinion mining and sentiment analysis. IEEE Intelligent Systems, 28(2), 15-21.

[37] Agarwal, A., Biadsy, F., & Mckeown, K. R. (2009, March). Contextual phrase-level polarity analysis using lexical affect scoring and syntactic n-grams. In Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics (pp. 24-32). Association for Computational Linguistics.

[38] Besag, J. (1986). On the statistical analysis of dirty pictures. Journal of the Royal Statistical Society. Series B (Methodological), 259-302.

[39] Paltoglou, G., & Thelwall, M. (2012). Twitter, MySpace, Digg: Unsupervised sentiment analysis in social media. ACM Transactions on Intelligent Systems and Technology (TIST), 3(4), 66.

[40] Singh, V. K., Piryani, R., Uddin, A., & Waila, P. (2013, January). Sentiment analysis of textual reviews; Evaluating machine learning, unsupervised and SentiWordNet approaches. In Knowledge and Smart Technology (KST), 2013 5th International Conference on (pp. 122-127). IEEE.

[41] Rao, D., & Ravichandran, D. (2009, March). Semi-supervised polarity lexicon induction. In Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics (pp. 675-682). Association for Computational Linguistics.

[42] Esuli, A., & Sebastiani, F. (2007). SentiWordNet: a high-coverage lexical resource for opinion mining. Evaluation, 1-26.

[43] Hung, C., & Lin, H. K. (2013). Using objective words in SentiWordNet to improve sentiment classification for word of mouth. IEEE Intelligent Systems, 1.

[44] Boudia, M. A., Hamou, R. M., & Amine, A. (2016). A New Approach Based on the Detection of Opinion by SentiWordNet for Automatic Text Summaries by Extraction. International Journal of Information Retrieval Research (IJIRR), 6(3), 19-36.

[45] Forrester. (2016). Think You Want To Be "Data-Driven"? Insight Is The New Data. [online] Available at: https://go.forrester.com/blogs/16-03-09-think_you_want_to_be_data_driven_insight_is_the_new_data/.

[46] Amiri, H., & Chua, T. S. (2012, July). Sentiment Classification Using the Meaning of Words. In Workshops at the Twenty-Sixth AAAI Conference on Artificial Intelligence.

[47] Hu, M., & Liu, B. (2004, August). Mining and summarizing customer reviews. In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 168-177). ACM.

[48] Zhang, W., & Skiena, S. (2009, September). Improving movie gross prediction through news analysis. In Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology-Volume 01 (pp. 301-304). IEEE Computer Society.

[49] Zhao, J., Dong, L., Wu, J., & Xu, K. (2012, August). Moodlens: an emoticon-based sentiment analysis system for chinese tweets. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 1528-1531). ACM.

Hammou Fadili is a HDR (Habilité à Diriger des Recherches) researcher at Conservatoire National des Arts et Métiers (CNAM) in computer science. He is a member of the information systems team, named Ingénierie des Systèmes d'Information et de Décision (ISID). Also, he is a manager of the Digital Humanities projects at the research department of the FMSH and associate researcher at the Digital Humanities department of the Paris 8 university. His research work concerns especially these areas: Semantic Web, WEB 3.0, Machine learning (supervised, unsupervised, etc.), Digital Humanities, Automatic Natural Language Processing (NLP), Language & Context Modeling, Detection and anonymisation of sensitive data, Semantic mining of structured and unstructured data, (Linked, Big, Smart) data. He is a member of the program committee and member of the organization committee for various conferences. He is also reviewer for some journals. He is (or was) involved in many ANR (Agence Nationale de la Recherche) and European research projects.

An ontological Rule-Based Approach for Software Product Lines Evolution

Olfa Ferchichi, Raoudha Beltaifa, and Lamia Labed

Abstract— A software product line usually constitutes a long-term investment and, therefore, has to undergo continuous evolution to correct, improve, or extend assets or products. Software Product Line (SPL) evolution needs to gather product line knowledge to be able to successfully conduct evolution solutions. Despite various attempts, applying SPL evolution proposals remains limited and no promising approach has been proposed to evolve product lines under a common knowledge management framework. Ontologies emerge as one of the most appropriate knowledge management tools for suppoting knowledge representation, processing, storage and retrieval. Given great importance to knowledge for product line evolution, and the potential benefits of managing SPL knowledge, we propose an evolution-oriented knowledge management approach. This approach provides a continuous evolution of SPLs by means of an ontological rule-based knowledge management framework. The framework delivers formal semantics and evolution rules to help evolving SPLs by using a core ontology. This ontology, kernel of the framework, represensts a common conceptualization of SPLs knowledge. Mainly, it considers knowledge of SPLs requirements, architecture and traceability. In this paper, we present our approach, the associated framework. But for the reason that the knowkedge related to architecture and traceability is not yet implemented inside the ontology, we are confined to present the ontology of SPLs requirements knowledge. This ontology manages knowledge associated to Feature Modeling, as a key activity involved in requirements modeling of SPLs. Based on the example of the Electric Parking Brake system, the paper presents an instanciation of the proposed ontology and proposes some improvements and evolution rules of the system Feature Model.

Index Terms—knowledge management, ontology, software product line evolution, evolution evolution

1 Introduction

SOFTWARE Product Line is a set of software intensive systems that share a common, managed set of features to satisfy 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].

Commonality is a property shared by all systems of the product line; e.g., all Braking systems of a vehicule have a parking brake. Variability defines how the different systems of the product line can vary [2]; e.g., some vehicules may include automatic parking brake (also known as Electric Parking Brake), but others include manual Parking Brake.

Software Product Line Engineering (SPLE) involves domain engineering, in which core assets are developed, and application engineering, in which systems are developed and configured from the core assets. Core assets are reusable software-related assets used in the production of more than one system in a software product line. A core asset may be a requirement model, a reference architecture, a software component, a process model, a test plan, a document, or any other useful result of building a system [1].

The market and technology related to a product line change with time. Therefore, a product line development organization should maintain and optimize a product line by evolving core assets. Core asset evolution, is not only adding to a product line but also removing, extending or restructuring core assets depending on a situation.

Evolution is any change in the quality, functionality, etc. of SPL assets. Change is not carried out in an instant; it is continuous. It is necessary to evolve SPL assets continuously. Besides, an asset evolution may require changes in other related or dependent assets.

To assist engineers in dealing with the increased complexity of SPL assets when evolving them, specific approaches have been proposed. Several researchers focused on the study of the evolution of feature models (FMs), as a key requirement asset of SPLs [23][24][25]. Others focused on the evolution of technical, architectural and implementation assets by providing some evolution scenarios [19]

Some of the works propose using ontologies to represent FMs [27][11] or to identify defects in FMs [14][9][13]. Despite these various attempts, SPL assets evolution remains limited and no promising approach has been proposed to evolve product lines under a common knowledge man-

- *O.F. Laboratoire de Recherche en Génie Logiciel, Applications distribuées, Systèmes décisionnels et Imagerie intelligentes, Université de Manouba . olfaferchi@ yahoo.fr.*
- *R.B. Laboratoire de Recherche en Génie Logiciel, Applications distribuées, Systèmes décisionnels et Imagerie intelligentes, Université de Manouba. E-mail: raoudha.beltaifa2011@ensi.rnu.tn*
- *L.A. Laboratoire de Recherche en Génie Logiciel, Applications distribuées, Systèmes décisionnels et Imagerie intelligentes, Université de Manouba et Université du Tunis E-mail: Lamia.Labed@isg.rnu.tn*

agement approach. A common knowledge management gathers knowledge from various product lines and use it incrementally and continusly to evolve assets.

Our general goal is to develop an evolution-oriented approach for knowledge management of SPLs. This approach offers continuously semantic improvements and evolution rules (scenarios) of SPLs assets.
In this paper, we propose a first step to achieve this goal. In particular, we propose an ontological rule-based framework for a unified conceptualization of FMs including common caracterization of features, a formal semantic and a set First-Order Logic rules to evolve of FMs.

The remainder of the paper is structured as follows. Section 2 gives a brief overview of the necessary background for understanding the approach presented in Section 3. Section 4 presents the ontological rule-based framework. Section 5 presents implementation details. Section 6 presents the evaluation of our framework. Section 7 presents related research. Finally, Section 8 presents the conclusions and suggests future research directions.

2 BACKGROUND

2.1 Product Line Engineering

The software product line engineering (SPLE) paradigm separates two processes: domain engineering and application engineering. "Domain engineering is the process of software product line engineering in which the commonality and the variability of the product line are defined and realized "[7].
"Application engineering is the process of software product line engineering in which the applications of the product line are built by reusing domain artefacts and exploiting the product line variability "[7].

Domain artefacts (assets) constitute the platform of the software product line and are stored in a common repository. The artefacts are interrelated by traceability links to ensure the consistent definition of the commonality and the variability of the software product line throughout all artefacts. Domain artefacts are the following: domain requirements, domain architecture, domain realization artefacts and domain test artefacts [7].
Application artefacts (or application assets) encompass all development artefacts of a specific application. They are interrelated by traceability links. Traceability links between application artefacts and the underlying domain artefacts are captured to support the various activities of the application engineering. These links also support the consistent evolution of the product line. For example, if a domain artefact changes, the application artefacts affected by this change can be easily determined. Application artefact are: application requirements, application architecture, application realization artefacts and application test artefacts [7].

One of the most important assets of a product line is the domain requirements. The domain requirements define the common and variable features, functions and qualities of the product line [1]

2.2 Feature Model

Feature modeling is a key activity to capturing and managing the common and variable features of the domain requirements. Feature Model (FM) which firstly introduced by Kang with Feature Oriented Domain Analysis (FODA) became a core part of SPLE research and development [4]. The original FM has been extended by several approaches [5], one of them is Cardinality-based Feature Model (CBFM) [6].
A feature is a system property that is relevent to some stakeholders and is used to capture commonalities and distinguish between systems. Features are organized in feature diagrams. A feature diagram is a tree with the root representing a concept (e.g., Electric Parking Brake system) and descendant nodes are features. In the FODA feature diagram notation (see Table 1), features can be mandatory, optional or alternative. Feature models are feature diagrams plus additional information such as feature descriptions, binding times, priorities, etc. A feature model describes the configuration space of a product line. An application engineer may specify a product of the PL by selecting the desired features from the feature model within the variability constraints defined by the model (e.g., a choice of one feature from a group of alternative features).

Feature models are the main technique to represent domain requirements variability in SPLs. However, FMs have several limitations related to the lack of means to represent explicitly the semantics of features and their relationships/dependencies. Following some limitations of feature models [21][26]:

- Quality features are not considered in feature models: FMs represent variability of SPL functional requirements. But variability of quality attributes, as SPL requirements, are neglected.
- Types of SPL features are not expressed explicitly. A not expert can not distinguish between hardware and software features.
- Relationships between a parent feature (variation point) and its children features (variants) is ambiguous. Sometimes variants have an inheritance (is-a) relationship with their parent feature; e.g., a network, as a parent feature, has three 'is-a' children features: 3G, 4G and 5G. Other variants can have composition relationship with a parent feature, as compound feature; e.g., a car is composed of an engine, doors, etc.
- Lack of distinction between behavioral and structural features.
- Evolutions of features in time and in space are not expressed.

3 Evolution-oriented KM Approach for SPLs

The KM approach associated to SPLs we propose is shown in Figure 1. This approach is called evolution-oriented KM in the sense that evolution becomes a key activity in SPL engineering. The process is iterative and continuous. It encompasses three steps: SPL knowledge acquisition, SPL knowledge capitalization and building of a rule-based ontological framework. The first step, Knowledge acquisition, accumulates all types of information related to domain knowledge and application knowledge; e.g., Feature Models, Feature Meta-models, reference architecture models, configurations, experiences. In the second step, knowledge capitalization, the accumulated knowledge is analysed to distinguish knowledge of high quality from low. Then knowledge useful for SPLs evolution is extracted. For exemple, if a feature model lacks of precision and it is difficult to understand the relationship between its variants then an improvement of features modeling based on existing experience documents or well-known documents is extracted. The extracted knowledge is formalized by means of meta-models representing common and better concepts we need to model SPLs. The meta-models we established are Feature Meta-Model, Architecture Meta-model and Traceability Meta-Model. Each one of these meta-models is an extended meta-model including all necessary concepts to evolve existing SPL modeling concepts. In section 4, we present only the extended Feature Meta-Model because the two others are not yet validated.

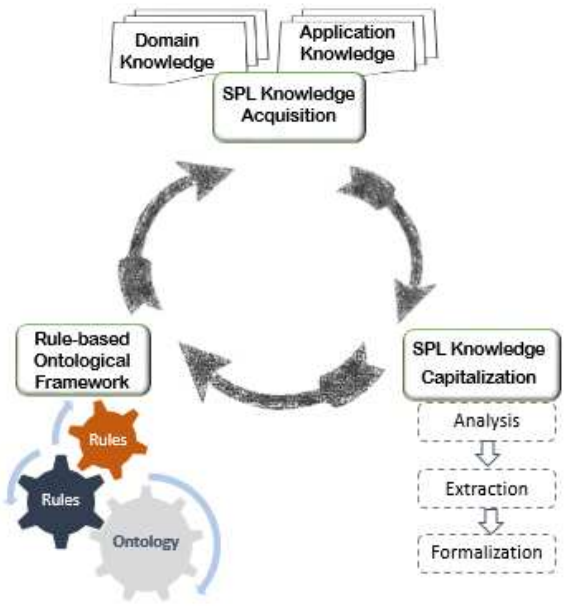
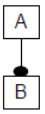
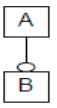
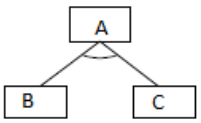
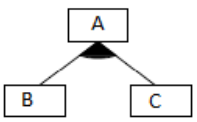
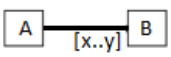
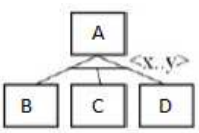


Fig. 1. Evolution-oriented KM of SPLs

The last step is very important because it shows how to validate the meta-models and how much they are useful for management knowledge by an ontology. Building a rule-based ontological framework is, in fact, the goal of the third step. The ontology is defined by three sub-ontologies. Each ontology is related to one of the extended meta-models. Rules provided by the framework are written in the first-order in order to define various evolution scenarios and contraints. For example, if a feature F1 is related to Feature F2 by versioning (which means that F1 evolves in time) then we need to represent explicitly that F1 exists before F2.

The Rule-based ontological framework and the sub-ontology (we call it 'ontology') associated to the extended Feature Meta-model are presented in the next section.

TABLE 1
TYPES OF DEPENDENCIES IN FEATURE MODELS

Notation	Type of Dependency
	**Mandatory** [4] Child feature B must be included in all valid products containing the parent feature A and vice versa. A feature is a full mandatory feature, if the feature and all its ancestors are also mandatory.
	**Optional** [4] Child feature B may or may not be included in the valid products containing parent feature A.
	**Alternative** [4] Only one child feature (B or C) can be included in the valid products containing parent feature A
	**OR** [4] At least one child feature can be included in the valid products containing parent feature A
	**Excludes** [4] Both features A and B cannot be in the valid products at the same time.
	**Requires** [4] Feature B should be included in the valid products with feature A.
	**Feature cardinality** [6] Represents the minimum (x) and maximum (y) number of clones of the child B that the product can have when its parent feature A is included in the product.
	**Group Cardinality** [6] Represents the minimum (x) and the maximum (y) number of child features (B..D) grouped in a cardinality (<x..y>) that a product can have when its parent feature A is included in the product.

4 Rule-based Ontological Framework

We present our framework through two Sub-sections. The first one presents the ontology model, which represents the concepts of the extended Feature-Meta-Model created at the SPL knowledge capitalization step of our approach. The second one presents the ontological framework that uses the Feature Meta-Model ontology for generating axioms and rules. Axioms specify formal semantics of FMs and rules represent evolution scenarios of FMs.

4.1 Feature Meta-Model Ontology: FMMO

The conceptual model associated to our Feature Meta-Model ontology, presented in Figure 2, is written in OntoUML. The FMMO ontology is shown in Figure 3.

In FMMO ontology the meta-model classes (stereotyped <<kind>> and <<subkind>> correspond to classes of the ontology; the dependencies between meta-model classes are represented as ontology object properties; and the attributes of the Cardinality, GroupCardinality and Behavioural_Feature meta-model classes are represented as ontology datatype properties. The inheritance dependencies among the meta-model classes are represented as is-A dependencies.

As mentioned earlier, FMMO ontology is built with the main goal of making the best possible specification of the FMs for establishing a common conceptualization about FM-oriented requirement modeling of SPLs. In particular, it is designed to be used as a core conceptual model to be used for several purposes, such as:

- As a basis for structuring and representing knowledge related to FMs
- As a basis for specifying FMs in a formal semantic, by means of axioms and rules,
- As a reference model for integrating software tools supporting FMs evolution.

The new FM concepts we defined to improve FMs semantic are:

- Concepts to classify requirement features according to their types: structural features, behavioral features and quality features.
- Concepts to distinguish FM nodes : Variant point feature or variant feature,
- Concepts to express precisely relationships between a Variant point and its variants: compose, aggregate or is-a. A compose relationship indicates that we have a compound component which is composed of one or many components and that removing the compound one implies the removing of its components. Agregate is a type of a composition relationship but the existence of the components does not depend to their compound one. This means, for aggregate relationship removing the compound component does not remove its components. But, for both of the composition relationships the components exist inside the compound one. The is-a relationship indicates the relationship between a supertype and its sub-types.
- Concepts to differentiate hardware features from software features,
- Concepts, defined as new constraints, to express features evolving in time (Before and During constraints). For instance F1 and F2 are two features. Before(F1,F2) expresses three possible cases: 1) If F1 and F2 are structural features then the contraint means that F1 exists before F2. 2) If F1 and F2 are behavioral features then it means that F1 happens before F2. 3) If F1 (ou F2) is a structural feature and F2 (ou F1) is a behavioral feature then F1 exists before F2 happens (F1 happens before F2 exists). During(F1,F2) expresses that F2 happens when F1 is happening.

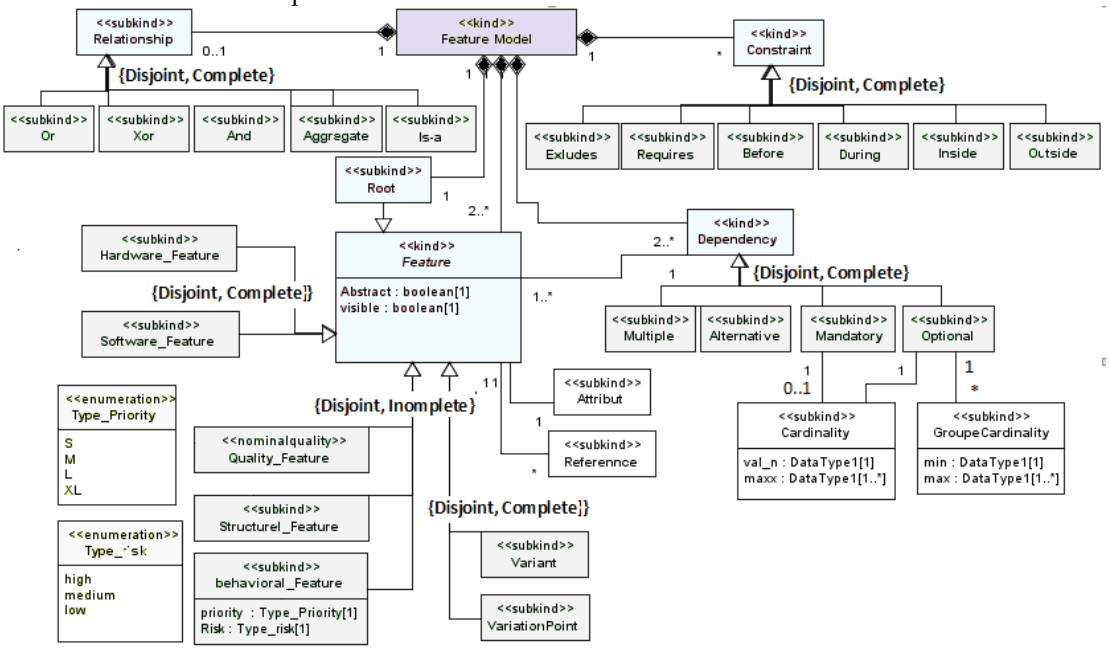


Fig. 2. Ontology Meta-Model for FMs

- Concepts related to features evolving is space. They are expressed by Inside and Outside constraints. Inside(F1,F2) indicates that F1 exists inside F2. Outside(F1,F2) means that F1 exists outside F2.

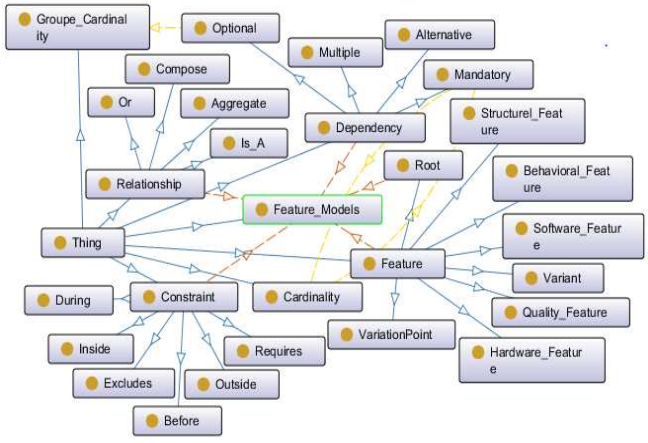


Fig. 3. Feature Meta-Model Ontology

4.2 Ontological Framework

The Framework we propose, see Figure 4, includes: 1) Two inputs: A feature model to evolve and the Feature Meta-Model ontology, 2) Three modules: Features characterization, Transformer and Rules Engine and 3) Two outputs: Feature model enriched by new constraints (formal semantic) and evolution rules.

The first module, receives the FM for which we want to add a formal semantic by means of constraints and to recommend evolution rules. In this module, the domain expert of SPL adds charateristics to each feature of the FM. These characteristics are associated to the semantic concepts defined in the Feature Meta-model ontology and which can not be represent in a FM. For instance, a feature can be structural, behavioral or a quality attribure. Adding semantic characterics to features is manual and is guided by the *Features characterization* module (see Subsection 4.2.1).

Features characteristics and the two inputs of the framework are entries of the second module. The *Transformer module* generates an ontology for the FM to evolve considering the added characteristics and the FMMO ontology (see sub-section 4.2.2).

Finally, the third module, we called *Rules Engine*, executes the rules already defined inside the ontology FMMO and generates (as outputs of the framework) a set of constraints and rules. The constraints are axioms, which specify a formal semantic as enrichment (a semantic improvement) of the initial FM. For exemple, if the domain expert adds the composition relationship between a feature F and the two features F1 and F2 then the axioms Compose (F,F1) and Compose (F,F2) are defined inside the ontology of the FM. Based on the FM semantic enrichment and the generic evolution rules defined in FMMO ontology, specific evolution rules (scenarios) are generated for the initial FM. For instance, the evolution rule which consists in removing a compound or a component feature from a FM has to be defined to ensure a consistent FM. Detailed explanations of the axioms and evolution rules are presented in Sub-section 4.2.3.

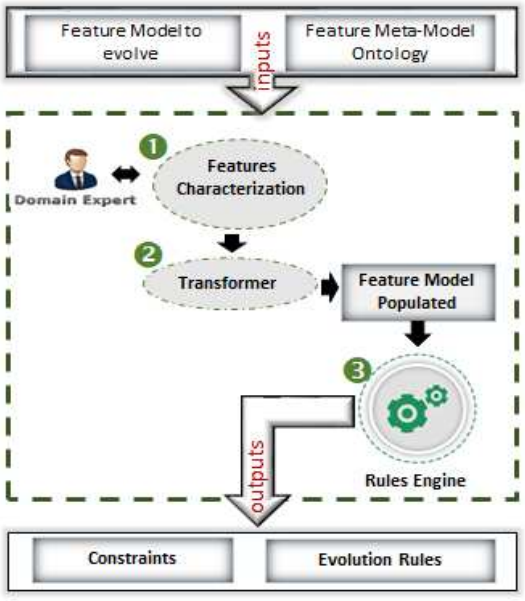


Fig. 4. Rule-Based Ontology Framework

4.2.1. Features Characterization

Features characterization is based on the SPL domain expert knowledge. For each feature of the initial FM, this module presents the characteristics of a feature that are present in the FMMO ontology; e.g., structural, behavioral, quality, component, compound, hardware and software. Then the expert selects the appropriate characteristics. In fact, for each feature he has to select one of the types: hardware or software feature. For a feature that is a Variation Point (VP) (a node that have child nodes), its relationships with its children (variants) have to be expressed clearly by is-a, compose or aggregate relationship. As explained in section 2, a feature model neglects semantic relationships between a variation point and its variants. So, the domain expert has to remove this limit by expressing explicitly the convient relationship.

To give better explanation to our framework, a running example of an Electric Parking Brake (EPB) system FM is presented in Figure 5. This FM is well described in [20].

4.2.2. Transformer

Transformer module is responsible of populating the FMMO Ontology with the elements of the FM to evolve and the characterizations defined by the domain expert. Populate an ontology consists in creating individuals for the classes of the ontology.

First, the Transformer reads each element of the input FM and the characterizations of features, and then it creates an individual to the corresponding ontology class and fills the properties of each individual.

The class in which the Transformer creates each individual depends of the type of dependency in the FM. For example, dependency between features Design_EPB and Software_architecture shown in the running example (see figure 5) is an individual of the Mandatory ontology class. Besides, for the new semantic dependencies defined by the domain expert (cf., section 4.2.2) indivuduals are created. For example, the feature Automatic is a type of the feature Parking_Brake_Service (see figure 5), then an individual of the Is-a class of the ontology is created. All features of the input FM and their characteristics are individuals by inheritance of the Feature class. Moreover, the FM root is an individual of the Root class. We consider that each feature (node) having children features have an individual of the Variation-point class and each leaf feature has an individual of the class Variant. For exemple, individuals for the class Variant are created for the features automatic, Assisted and Manual and an individual for the class Variation-Point is created for the feature Parking_Brake_Service.

Transformer fills the properties of each individual that it created using information obtained from the input FM and the features characterizations.

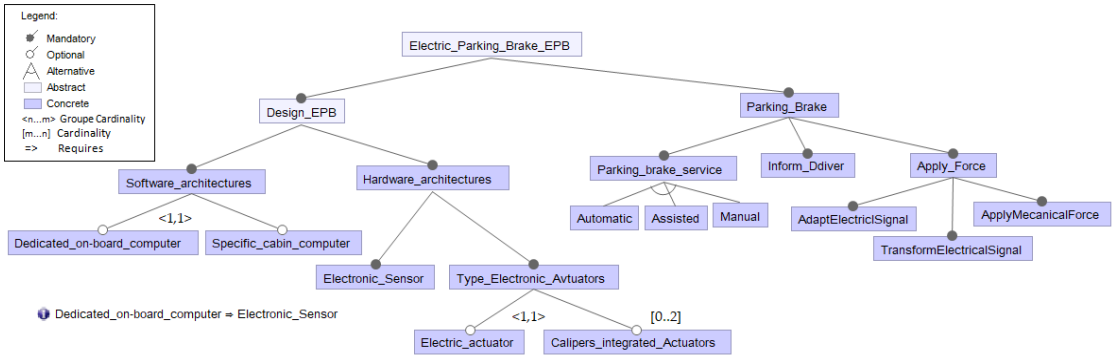


Fig.5. Elerctric Parking Brake System FM

4.2.2 Rules Engine

The Rules Engine is the module that generates axioms (constraints) to improve a FM by a formal semantic and suggests evolution rules for the FM.

We use the following first-order logic predicates, functions and sets to formalize the rules as Horn Clauses. For most of these symbols we present an example extracted from the populated ontology associated to the EPB FM (see figure 5).

- Requires (x,y): This predicate indicates that feature x requires feature y; e.g., requires (Dedicated_on_board_computer, Electronic_sensor).
- excludes(x,y): This predicate indicates that feature x and feature y are mutually exclusive.
- Inside(x,,y) : this predicate indicates that feature x exists inside feature y; e.g., inside(Software_architecture,Deign_EPF).
- Outside(x,y) : this predicate indicates that feature x exists outside feature y.
- Before(x,y): this predicate indicates that feature x happens before feature y. x and y are behavioural features; e.g., before (adapt_electric_signal,transform_electricalsignal).
- During(x,y): this predicate indicates that feature x happens during feature y. x and y are behavioral features; e.g., during(apply_mechanical_force,inform_driver).
- Structural(x): indicates that x is structural feature; e.g., structural(electronic_sensor).
- Behavioral(x): indicates that x is behavioral feature; e.g., behavioral(adapt_electric_signal).
- Quality(x): the function quality indicates that is a quality requirement.
- Compose(x,y): this predicate indicates that y composes x; e.g., compose(Design_EPB,Parking_Brake).
- Aggregate(x,y): indicates that x contains y. However, this relationship is weak (not like composition). This means that existence of y does not depend on x.
- Is-a(x,y): indicates that y is a subtype of y; e.g., is-a(automatic, Parking_Brake_Service).
- variant(x): indicates that x is a variant; e.g., variant(software_architecture).
- variation_point(x): indicates that x is a variation-point; e.g., variation_point(Parking_Brake).
- optional(x,y): indicates that features y has an optional dependence with the feature y; e.g.,
- mandatory(x,y): indicates that feature y is mandatory for y (mandatory relationship); e.g., mandatory(adaptElectricsignal,Apply_force).
- childrenSet(x): set of children features related to a feature x; e.g., childernSet(Parking_Brake_Service)={Automatic, Assisted, Manual}.
- FMFeaturesSet: collection of all features of a FM.

- OpFMSet: This set represents the collection of all optional features of a FM.

Some of the generic constraints (formal semantic) of a FM defined inside the framework are the following: inside(x,y), outside(x,y),before(x,y), During(x,y).

The following generic rules generate more constraints for any FM:

R1: ForAll x,y; before(x,y) and (before(y,z) or during(y,z)) implies before(x,z). R1 means that if x happens before y and y happens before z then x happens before z.

R2: ForAll x,y; during(x,y) and during(y,z) implies during(x,z).

R3: ForAll x,y; during(x,y) and before(y,z) implies before(x,z).

R4: ForAll x,y; if inside(x,y) and inside(y,z) implies inside(x,z).

Some of the generic evolution Rules are defined as follows: We consider FM1 as the initial FM and FM2 the evolved one.

ER1: forall f1 such that compose(F,f1); remove(f) implies remove(f1). This rule means that removing a parent compound feature implies removing its (components) children too.

ER2: FM evolves by moving a behavior from one component feature of the SPL system to another.

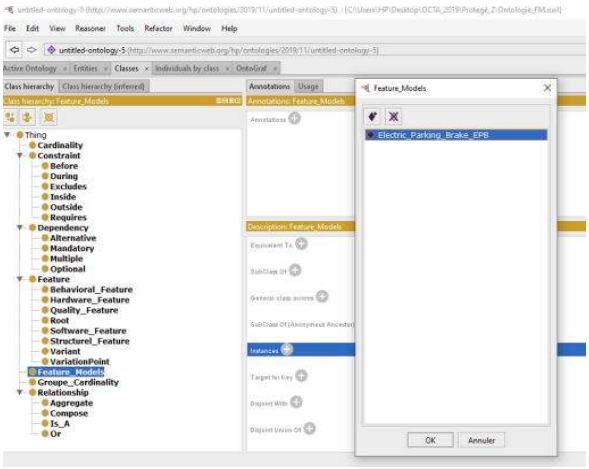
Move(FM1,f)=FM2 implies exists F1 isin FMFeaturesSet(FM1) such that inside(f,F1) and exists F2 isin FMFeaturesSet(FM2) such that inside(f,F2) and notEqual(F1,F2) and outside(f,F1).

4 Implementation Details

We implemented our Feature Model Ontology using Protégé OWL [22] and SWRL [18] to write rules. In this section we present some interface associated to the implemented modules defined inside our framework. In figure 6, we present the interface provided to the domain expert to characterize the features of the FM already stored as OWL and XML files. This module displays this interface for each feature of the FM and the domain expert selects the corresponding characteristic according to his knowledge in the domain of the SPL. In figure 7, the ontology implemented with Protégé is shown.

Fig.6. Features Characterization Interface

Fig.7. FMMO ontology with Protégé OWL tool



5 Related Work

In this section, we present works, which proposed approachs for evolving FMs based on ontologies.

João Bosco Ferreira Filho et al [8] proposed an approach to enrich FMs by relating its features to corresponding UML concepts. For example, if a feature represents a function then manually an associated use case is defined. This approach is limited because it did not provide a formal semantic to FMs or evolution rules (scenarios).

Wang, H. et al [9] presented a technique to design, verify and debug ontology-based feature models. The OWL DL language is used to represent FMs and configurations in an unambiguous way. FM and configuration verification are realized by a OWL reasoning engine to check for inconsistencies of feature configurations automatically.

Nieke and al [10] proposed an ontology to define and check FM evolution consistency rules. Temporal Feature Models (TFMs) are an extension to feature models that allow for engineers to model past feature-model evolution and plan future evolution.

Lee, S. -B., Kim et al [11] proposed a problem-oriented and value-based analysis method for software product line variability evolution analysis. The method takes into account both kinds of changes requirements and contexts during the life of an evolving SPL. The proposed method is based on the use of ontology to represent variability-intensive problem decomposition and evolution. The analysis method identifies candidate changes, detects influenced features, and evaluates their ontributions to the value of the SPL.

Zaid, L. A [12] provided an ontology framework for variability modeling wirh feature models. This framwork consists of an ontology that formally provides a specification for feature models. In addition, it provided means to integrate segmented feature models and provide a rule based model consistency check and conflict detection.

Rincón, Giraldo et al [13] proposed an ontological rule-based approach to analyze dead and false optional features in feature models as well as identifying certain causes of these defects, and explaining these causes in

natural language.

Abo, Kleinermann and De Troyer [14] proposed to use ontologies to represent FMs and facilitate their integration when they represent different views of a SPL. To validate model consistency the authors defined rules based on the Semantic Web Rule Language SWRL [18].

In our approach, we used the knowledge management 'ontology' as a tool to 1)improve FMs by a formal semantic defined by constraints (or axioms) and 2)propose evolution rules (scenarios).

7 Conclusion

In this paper, we proposed an evolution-oriented approach. This approach provides a continuous and iterative process for SPLs evolution. Any asset of a SPLE can evolve. In this paper our presentation is limited FM evolution. We proposed: 1) improvements to FMs by providing formal semantics as constraint (known as axioms) and 2) evolution rules to guide their evolution in a consistent way.

In the future, we will implement all the modules of our approach taking into account the evolution of various assets of SPLs. Evaluation of our approach with many FMS will be our future work too.

References

[1] P.Clements, L. Northrop, and L. M. Northrop. *Software Product Lines: Practices and Patterns*. Addison-Wesley Professional, August 2001.

[2] A. Metzger, P. Heymans, K. Pohl, P.-Y. Schobbens and G. Saval, "Disambiguating the Documentation of Variability in Software Product Lines: A Separation of Concerns, Formalization and Automated Analysis," Proc. 15th Int'l Requirements Engineering Conference (RE'07), NewDelhi, India, 2007.

[3] D. Carney, D. Fisher and P. Place, "Topics in Interoperability: System-of-Systems Evolution", Technical Report CMU/SEI-2005-TN-002, Carnegie Mellon, Software Engineering Institute, 2005.

[4] K. C. Kang, S. Cohen, J. Hess, W. Nowak and S. Peter- son, "Feature-Oriented Domain Analysis (FODA) Feasi- bility Study," Technical Report CMU/SEI-90-TR-021, Carnegie Mellon University Software Engineering, 1990.

[5] K. C. Kang, "FODA: Twenty Years of Perspective on Feature Modeling," Proceeding of VaMoS'10, Vol. 37 of ICB-Research Report, Universität Duisburg-Essen, 2010, p. 9.

[6] K. Czarnecki and P. Kim, "Cardinality-based Feature Modeling and Constraints: A Progress Report," Proceed-ing of International Workshop on Software Factories, 2005.

[7] K. Pohl, G. Bockle and F. Linden, "Software Product Line Engineering: Foundations, Principles and Techniques," Springer, 2005.

[8] J. A.B.F. Filho, O. Barais, B. Baudry., W. Viana, &R.M.C. Andrade. (2012). An approach for semantic enrichment of software product lines. In Proceedings of the 16th international software product line conference. SPLC '12 (Vol. 2, pp. 188–195). New York, NY, USA: ACM

[9] H.H.Wang., Li, Y. Fang., J. Sunc , H. Zhang, & J .Pan. (2007). Verifying feature models using owl. Journal of Web Semantics: Science, Services and Agents on the World Wide Web, 5(2).

[10] M. Nieke, C. Seidl, T. Thum. ,Back to the future: avoiding paradoxes in feature-model evolution, SPLC (2), 2018.

[11] S.B.Lee, J.W.Kim, , C.Y.Song,, & D.K.Baik. (2007). An approach to analyzing commonality and variability of features using ontology in a software product line engineering. In 5th ACIS international conference on software engineering research, management applications, 2007. SERA 2007 (pp. 727–734).

[12] L. A. Zaid, F.Kleinermann, & O.Troyer,. (2009). Applying semantic web technology to feature modeling. In Proceedings of the 2009 ACM symposium onapplied computing (pp. 1252–1256). ACM.

[13] L. Rincón., G. Giraldo, R. .Mazo, & C. Salinesi. (2014). An ontological rule-based approach for analyzing dead and false optional features in feature models. Electronic notes in theoretical computer science, 302(0): 111 – 132. In Proceedings of the {XXXIX} latin american computing conference (CLEI 2013).

[14] L. Abo, G. Houben, O. Troyer, and F. Kleinermann, An OWL-Based Approach for Integration in Collaborative Feature Modelling, in: Proceedings of the 4th Workshop on Semantic Web Enabled Software Engineering, Germany, 2008.

[15] J. Guo, Y. Wang, , P.Trinidad, & D.Benavides. (2012), Consistency maintenance for evolving feature models. Expert Systems with Applications,2012, 39(5), 4987-4998

[16] Object Management Group (OMG), "About the Unified Modeling Language Specification Version 2.5.1," Object Management Group (OMG), 2017.

[17] C. Barrett, & C.Tinelli. (2018). Satisfiability modulo theories. In Handbook of Model Checking (pp. 305-343). Springer, Cham.

[18] I. Horrocks, P. Patel-Schneider, H. Boley, S. Tabet, B. Grosof, M. Dean.: SWRL: A Semantic Web Rule Language Combining OWL and RuleML, http://www.w3.org/Submission/SWRL

[19] D. Weyns, B. Michalik, A. Helleboogh, and N. Boucké, "An Architectural Approach to Support Online Updates of Software Product Lines."

[20] C. Dumitrescu, "CO-OVM : A Practical Approach to Systems Engineering Variability Modeling Cosmin Dumitrescu CO-OVM : A Practical Approach to Systems Engineering," 2014.

[21] R. Mazo, "Avantages et limites des modèles de caractéristiques dans la modélisation des exigences de variabilité," Rev. Génie-Logiciel, vol. 111, no. 111, pp. 42–48, 2014..

[22] Stanford Protégé OWL, http://protege.stanford.edu/overview/protege-owl.html.

[23] A. Pleuss, G. Botterweck, D. Dhungana, A. Polzer, and S. Kowalewski, "Model-driven support for product line evolution on feature level," J. Syst. Softw., vol. 85, no. 10, pp. 2261–2274, 2012 ok

[24] G. Botterweck, A. Pleuss, D. Dhungana, A. Polzer, and S. Kowalewski, "EvoFM : Feature-driven Planning of Product-line Evolution Categories and Subject Descriptors," Proc. 2010 ICSE Work. Prod. Line Approaches Softw. Eng., no. Please, pp. 24–31, 2010

[25] F. Benbassat and P. Borba, "Safe Evolution of Software Product Lines : Feature Extraction Scenarios," 2016 X Brazilian Symp. Softw. Components, Archit. Reuse, pp. 11–20, 2016

[26] H. Herman , T. Trew,T. Matsinger, AJ. Aart AJ, Supplier independent feature modelling, SPLC,pp .191--200,2009..

[27] K. Czarnecki., C.H.P. Kim, and K.T. Kalleberg, Feature Models are Views on Ontologies, in:Proceedings of the 10th International on Software Product Line Conference, IEEE Computer Society, Washington, DC, USA, 2006: pp. 41–51.

An Adaptive Genetic Algorithm for Dynamic Vehicle Routing Problem with Backhaul and Two-dimensional loading constraints

1st Ines Sbai
Université de Tunis,
Institut Supérieur de Gestion de Tunis,
LARODEC Laboratory,
Tunis, Tunisia
sbai_ines@yahoo.fr
ORCID: 0000-0002-4163-9869

2nd Saoussen Krichen
Université de Tunis,
Institut Supérieur de Gestion de Tunis,
LARODEC Laboratory,
Tunis, Tunisia
Krichen_s@yahoo.fr

***Abstract*—In this paper, we consider a Dynamic Vehicle Routing Problem with Backhauls (DVRPB) followed by a two-dimensional loading problem (2L-DVRPB). For the Vehicle Routing Problem with Backhauls (VRPB), a vehicle can deliver (Linehaul), then collect goods from customers (backhaul) and brings back items to the depot. Once customer demands are designed as a set of two- dimensional rectangular weighted items, the problem is considered as a 2L-VRPB. To the best of our knowledge, this problem has not been analysed so far in the literature.**

The 2L-DVRPB is an NP-Hard problem aims to design the route set of minimum cost. So, in order to ensure a better solution quality for a large number of instances, we propose to use a Genetic Algorithm and five packing heuristics. Since, the problem has not been addressed in the literature; we have used the classical 2L-CVRP benchmarks to test the effectiveness of our approach. Results show that our proposed approach is competitive in terms of the quality of the solutions found.

***Index Terms*—Dynamic Vehicle Routing Problem, Backhaul, Packing, Genetic algorithm, 2L-DVRP**

I. Introduction

Vehicle Routing Problem (VRP) is a key component of distribution and logistics management. It consists of finding an optimal set of trips for a fleet of vehicles which must serve a predefined set of customers. Several works have been developed to address numerous variants of the CVRP while considering additional features such as dynamic aspect, backhaul and loading (packing) aspects. The dynamic vehicle routing problem is well-known as an *NP-H*ard combinatorial optimization problem that attract significant attention over the past few years, Abdallah et al. [1] and [3]. To solve the DVRPB, Ninikas and Minis [15] employed the heuristic Branch-and-Price algorithm. Also, Ninikas et Minis [16] solved a variant of DVRP with Multi Backhauls that allows orders to be transferred between vehicles during plan implementation using Branch-and-Price algorithm.

In recent years, new research studies combining the capacitated vehicle routing and two-dimensional loading have appeared. The 2L-CVRP was studied first by Iori et al [9] using an exact algorithm which is the branch-and-cut technique. Also, metaheuristics methods has also been proposed to solve this problem such as Tabu Search [7], Guided Tabu Searc [25], Ant Colony Optimization (ACO) [6] Extended Guided Tabu Search algorithm (EGTS+LBFH) [11], Greedy Randomized Adaptive Search Procedure and Evolutionary Local Search algorithm (GRASP∗ELS) [5], Simulated Annealing (SA) with heuristic local search [12], Promise Routing-Memory Packing (PRMP) [26], Variable Neighborhood Search approach [23], adaptive GA ([20], [21]), Simulated Annealing algorithm [24], hybrid simheuristic algorithm [8].

The combination of VRP whith backhauls and loading constraints is a recent studied problem, Bortfeldt et al. [2] proposed a Large Neighborhood Search and a Variable Neighborhood Search (LNS-VNS) for solving the three dimension VRP with backhaul in both routing and a packing procedure and in both algorithms a Tree Search heuristic (TSH) is proposed for packing boxes. Reil et al. [19] extended the last approach proposed by [2] for the VRPBTW with 3D loading constraints by considering various types of backhauls. Pinto et al. [17] studied the VRP with mixed Backhaul using an insert heuristic and a Bottom-Left heuristic (BLH) for packing aspect. Also, Dominguez et al. [4] proposed a hybrid algorithm that integrates biased-randomised versions of vehicle routing and packing heuristics within a Large Neighbourhood Search metaheuristic framework. In the same case, zachariadis [27] described a local search (LS) approach for solving the 2L-VRPSDP and the 2L-VRPCB. Pinto et al. [18] proposed a VNS algorithm for solving the 2L-VRPB.

The remainder of this paper is structured as follows. The related literature review is provided in Section 2. Section 3 present a brief description and a Mathematical formulation of the static 2L-VRPB and 2L-DVRPB problem. Section 4 describes the resolution framework and presents the proposed Genetic Algorithm for solving the 2L-DVRPB. In Section 5, a set of heuristics for the loading subproblem are given. In section 6, the efficiency of the proposed approach is investigated with experimental results. In Section 7, we end with some concluding remarks and future works.

II. Problem definition

In this paper, a general description of the static 2L-VRPB is first given. In the static 2L-VRPB all the routing information is known in advance before the optimization has begun. On the other hand, in the dynamic 2L-DVRPB, some information may exist to the planner before the optimization begins but, generally, information can change or new information can be added during optimization.

A. Static 2L-VRPB

The 2L-VRPB is defined on a complete graph $G = (V, E)$, where $V = D \bigcup L \bigcup B$ is the vertex set of $1 + l + b$ vertices composed of the disjoint subsets $D = \{0\}$, $L = \{1, \ldots, l\}$ and B=$\{$l+1,...,l+b$\}$, that represent the depot, the linehaul customers and the backhaul customers, respectively. $E = \{(i, j)|i, j \in V, i \neq j\}$ is the set of edges that connect the customers with each other and with the depot. A nonnegative cost, c_{ij}, is associated with each edge $\{i,j\} \in E$ and represents the travel cost spent to go from customer i to customer j.

In the depot, there are m identical vehicles, each vehicle has the same weight capacity Q and a rectangular loading surface that is accessible from a single side for loading and unloading operations, whose width and length are W and H, respectively. We also denote by $A = W \times H$ the total area of the loading surface. It is also assumed that $m \geq max(m_L, m_B)$, where m_L and m_B are the minimum numbers of vehicles needed to separately serve all linehaul and backhaul customers, respectively.

Each vehicle starts its tour from the depot, deliver (Linehaul) its designated customers, then collect goods from customers (backhaul) and turns back to the depot.

An amount of demands q_i with a total weight equal to d_i is associated with every customer $i \in L \bigcup B$ that represents the amount requested from or delivered to the depot, depending on whether the customer is linehaul ($i \in L$) or backhaul ($i \in B$). Each demand has a specific width w_{it} and length h_{it}, t= $\{ 1, \ldots, |q_i|\}$.

Each item t will be denoted by a pair of indices (i, t). We denote by $a_i = \sum_{t=1}^{|q_i|} w_{it} \times h_{it}$ the total area of the items of each customer i, depending on whether the customer is linehaul ($i \in L$) or backhaul ($i \in B$).

The decision variables of the problem are, x_{ij}^k defined as:

$$x_{ij}^k = \begin{cases} 1 \text{ if the vehicle } k \text{ travels from customers } i \text{ to } j \\ 0 \text{ otherwise} \end{cases}$$

In order to effectively manage the placement of the item into the vehicle k, y_t^k is defined as follows:

$$y_t^k = \begin{cases} 1 \text{ if the item } t \text{ is inside the vehicle } k \\ 0 \text{ otherwise} \end{cases}$$

As shown, Figure 1(a) presents the static case (at *t*=0) where customer orders are known in advance and two initial routes schedules (linked with solid line) are generated to service these static customers (presented with white nodes). We present the mathematical formulation of the 2L-DVRPB as follow:

$$MinZ(x) = \sum_{k=1}^{k} \sum_{i=0}^{n} \sum_{j=0, j \neq i}^{n} c_{ij} x_{ij}^k \tag{1}$$

S.t.

$$\sum_{j=1}^{n} x_{0j}^k = \sum_{i=1}^{n} x_{i0}^k = 1, \quad k \in \{1, \ldots, K\} \tag{2}$$

$$\sum_{j=0, j \neq i}^{n} \sum_{k=1}^{k} x_{ij}^k = 1, \quad i \in \{0, \ldots, n\} \tag{3}$$

$$\sum_{k \in K} \sum_{j \in V} x_{ij}^k = 1, \quad i \in N \tag{4}$$

$$\sum_{i \in b, j \in l} x_{ij}^k = 0, \quad k \in K \tag{5}$$

$$\sum_{i \in l, j \in b} x_{ij}^k \leq 1, \quad k \in K \tag{6}$$

$$\sum_{i=1}^{n} \sum_{j=0, j \neq i}^{n} x_{ij}^k q_i \leq Q, k \in \{1, \ldots, K\} \tag{7}$$

- The objective function (1) consists of minimizing the total cost of a fleet of vehicles.
- Constraints (2) express that each travel should begin and end at the depot and routes are only allowed to start with a linehaul customer.
- Constraints (3) provide that a single vehicle leaves each client i. describe the remaining available vehicle capacities, which are updated over specific time periods.
- Constraint (4) ensures that all new customer will be served.
- Constraints (5) and (6) enforce that no linehaul customers are visited with vehicle k after servicing any backhaul customer.
- Constraints (7) guarantee that the vehicle weight is not exceeded. Constraints (14) and (15) ensure that the item demanded by customer i is placed in the loading space.

B. Dynamic 2L-VRPB

In contrast to a static 2L-VRPB, the performance of the dynamic counterpart is assumed to be dependent not only on the number of customers and their spatial distribution, but also the number of dynamic events and the time when these events actually take place with respecting the packing and the backhaul constraints.

Therefore, to measure the dynamism problem [13], the degree of dynamism is defined as follows:

δ ($\delta = \frac{n_d}{n_{tot}}$). While, δ the ratio between the number of dynamic requests n_d and the total number of requests n_{tot}.

if (δ=1) all requests are known in advance (the problem is completely static)

if (δ=0) No requests are known in advance (the problem is completely dynamic)

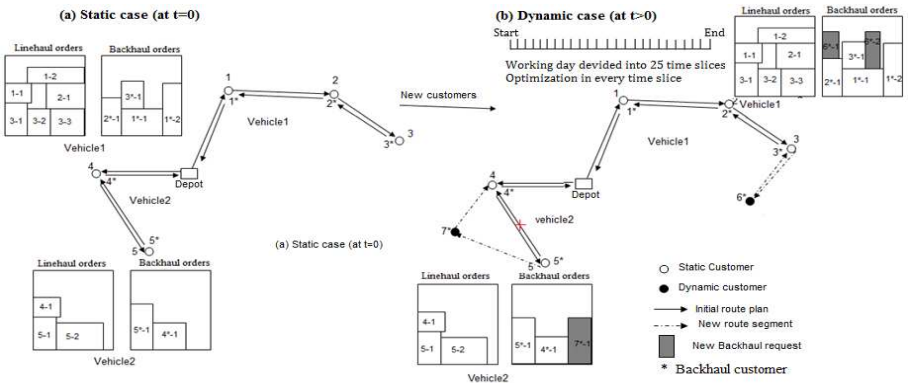


Figure 1: (a) An example of a static 2L-VRPB solution (at t=0), (b) An example of a dynamic 2L-VRPB solution (at $t > 0$)

As proposed by Montemanni et al. [14] , we decompose the 2L-DVRPB into a sequence of static 2L-VRPBs.
Therefore, to solve the 2L-DVRPB, the working day, T, is divided into n_{ts} time slices of $\frac{T}{n_{ts}}$ length. If any new customers order arrives during a time slice, it is postponed to the end of it and optimized in the next algorithm's run (in the next time slice). During each time slice, the problem would be similar to a static VRP.
In the DVRP, the solution(s) obtained from the previous time slice can be reused as an initial population for the next time slice.
Figure 1(b) presents the dynamic case, where new backhaul customers (presented with black node) are considered. Therefore, new route segments (presented with dashed line) are created.

C. An Adaptive Genetic Algorithm for the 2L-DVRPB

Genetic Algorithms (GAs) proved to be able to solve many NP-hard problems reaching near optimum solutions. In addition, GAs are good at solving dynamic problems AbdAllah et al. [1]. An adaptive GA is developed to solve the proposed 2L-DVRPB. In our case, our AGA starts by generated an initial population using the insertion heuristic, to determine whether a route-sequence of customers-is feasible in terms of the loading constraints of the examined problem, we designed a bundle of six packing heuristics $Heur_i$ (i = 1, 2, ..., 6). Then, a fitness function using the objective function is used to evaluate each individual. Two solutions are selected randomly from the population using the tournament selection. After that, a two point crossover is used in order to maintain the feasibility of the random moves. Then, an inversion mutation operation is defined as a perturbation of the structure with a random element that may effect the next generation. At each generation, the fitness is then selected to enhance an improved population for the subsequent steps. Finally, the old population is replaced by the new population of offspring solutions. This process is repeated until a number of generations are reached (1000 generations). Algorithm 1 describes the main steps of our AGA. In the following subsections, we describe our proposed AGA in more details.

D. Initial population

First, a randomly initial population is generated.
An insertion heuristic is used to dispatch requests to vehicle routes. First we start with opening R empty routes. A starting region is chosen randomly from the depot. After that, iteratively insert one customer at a time. Since the definition of the VRPB requires that every vehicle visits at least one linehaul customer the heuristic insert a randomly selected linehaul customer at the beginning of each route as an initialization step. At each iteration, one customer is randomly selected and inserted in a randomly chosen route. If the selected customer is a linehaul customer, it is inserted at the beginning of the route. If it is a backhaul customer, it is inserted at the end of the route. The customers are inserted in the solution with considering the capacity and the packing constraints of the problem . Once the initial static routes are generated, the simulation of the operations day can start. A new customer request is inserted at minimum additional cost into one of the planned routes. The solution is updated. The above process is repeated until a feasible 2L-DVRPB solution is obtained and will be updated automatically with the arrival of a new customers.

E. Solution encoding

In our AGA representation, the solution is represented as an integer string. In our approach, a chromosome representation has the form of a vector of length (L+B+m+1), where L is the number of Linehaul customers, B is the number of backhaul customers and m is a set of identical vehicles. There are also the depot 0 in the vector representing the start and the end of each vehicle route. The sequence between two 0 is the sequence of nodes to be visited by a vehicle.
Figure 3 presents an example of a chromosome of 2 routes with 12 customers (5 linehaul customers and 6 backhaul customers) where the node 0 indicates the center Depot. The positive nodes represent the static customers and the negative ones represent the dynamic customers (when a new customer is newly added).
The two routes are presented as follows:

Algorithm 1 The Adaptive GA approach for the 2L-DVRPB

Input: :
N: set of customer ; Q: vehicle capacity; i_l, i_b: linhaul and backhaul requests,
population $P(t)$, t: number of generation { GA parameters}
1: Begin
2: t=0 {generation counter}
3: Create an Initial Population ← Insertionheuristic
4: Evaluate each chromosome in $P(t)$ {Evaluate $P(t)$: Population of chromosomes using fitness function}
5: **while** Stopping criterion is not satisfied {Stopping criterion: no more feasible moves can be made} **do**
6: Initialize P' {P': a temporary population}
7: **for** i=1 **to** $|P(t)|$ **do**
8: Select two parents from $P(t)$
9: Apply Crossover (offspring) {Using the Two point Crossover operator}
10: **if** the offspring and parents are identical **then**
11: Improve each offspring by the Mutation operator {Using the inversion operator}
12: **end if**
13: **end for**
14: $t = t + 1$; {Update generation counter}
15: Replace the old Population $P(t)$ by the new P';
16: **end while**
Output: Best found solution {routes with shortest total travelled distance for serving linhaul and backhaul customer nodes}
17: End

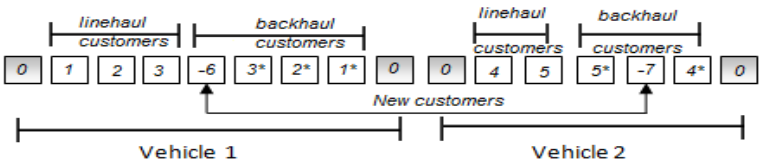


Figure 3: A Chromosome representation

Route 1: [0 1 3 -6 3* 2* 1* 0]
Route 2: [0 4 5 5* -7 4* 0]

F. Fitness function

Each individual is evaluated using the fitness value $F(x)$. The fitness function of our 2L-DVRPB problem is to find the shortest routes. So, the fitness value is calculated as follows that is the total distance travelled:

$$F(x)= \sum_{k=1}^{k} \sum_{i=0}^{n} \sum_{j=0, j\neq i}^{n} c_{ij} x_{ij}^{k}$$

G. Selection operator

In this paper, we choose to use Tournament method as a selection operator that requires the following steps:
Step 1: Select randomly two individuals from the population.
Step 2: Compares their fitness values.
Step 3: Select parents with the better fitness value as P_1 and P_2.

H. Crossover and mutation operators

In this paper, we use the two point crossover operator. The two point Crossover operator selects two parents for crossover and then randomly selects two crossover points. Two offspring are created by combining the parents at crossover point. it requires three steps:
Step 1: Select two parents used for crossover.
Step 2: Randomly select crossover linehaul point p_i (i = 0 to l-1) and crossover backhaul point p_i (i = 0 to b-1) called mapping section.
Step 3: Two offspring are created by combining the parents at crossover linehaul and backhaul point. In Figure 4 an example is given to illustrate the two point crossover operator. For the mutation operator, we use the inversion which generates two cut-points from the linehaul customers or from the backhaul customers of the chromosome, and then reverses a part of customers between these two cut-points. Figure 5 describes an example of the inversion mutation operator.

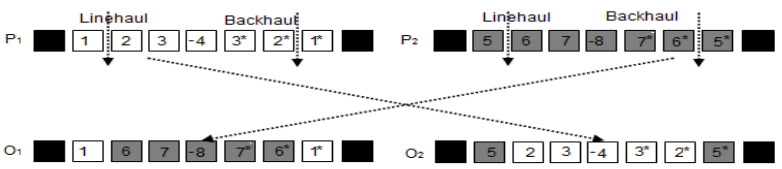


Figure 4: An example of the two point crossover operator

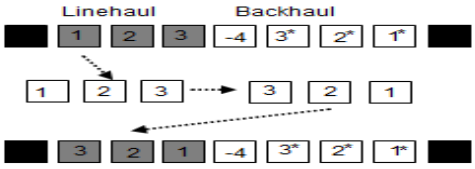


Figure 5: Example of the inversion mutation operator

I. Stopping criterion

This above process are repeated until a number of 1000 generation is reached.

III. Heuristics for the loading subproblem

In this section, we described whether a route-sequence of customers-is feasible in terms of the loading constraints of the examined problem.
In our case of sequential version, first, the rectangles are sorted by the reverse visiting order of customers. Then, the rectangles demanded by the same customer are sorted in decreasing order of the area of the rectangle.
Furthermore, the feasibility of loading an item into the vehicle loading space is checked using five packing heuristics. The five heuristics $Heur_i$ (i = 1, 2, ..., 5) are based on the work by Zachariadis et al. [25].
Let $Load\_pos$ denote a list of available loading positions for the items. So, the first available loading position lies in the front left corner (0,0) of the vehicle and the $Load\_pos$=$0, 0$. When an item is successfully inserted, four new positions are added onto the list and the $Load\_pos$ is updated. This bundle of heuristics are employed for the linehaul and the backhaul cases.

IV. Computational experiments

This section presents the computational results based on a set of benchmark instances introduced by [9], [7], but with a different types of 2L-VRP constraints (2L-VRP with backhaul and dynamic 2L-VRP with backhaul).
In order to demonstrate the performance of our proposed AGA, we design and solved it with the following steps:

1) We start our experiments by generating a new set of instances for the static 2L-VRP whith backhaul from the classical 2L-VRP using the method described by Toth and Vigo [22] to generate VRPB instances from classic Euclidean VRP ones.
   Performance comparison between proposed algorithm and other algorithms in the literature that deal with the 2L-VRPB [4] is given as well.

2) Once the efficiency of our AGA approach has been proved for the static 2L-VRPB case, we generate 180 new set of instances for the dynamic 2L-VRPB. To the best of our knowledge, no test instances are available in the literature for the version of the dynamic 2L-VRPB studied in this article. Therefore, we cannot compare them against other state-of-the-art approaches for the problem.

The proposed AGA approach is implemented using Java Language version 7. All experiments were performed on a PC equipped Intel (R) Core (TM) i3-4005U CPU with 4 GB(Gigabytes) of RAM under Microsoft Windows 7.

A. Parameters setting

All the tests were performed with the same configuration of the AGA. The algorithm was run 10 times on each instance. Table 2 reports the parameters of our AGA algorithm.
We take one problem version, depending on the loading constraint configuration under consideration: $2|SO|$L version (sequence constraint, fixed orientation).

B. Computational Results for the static 2L-VRP with backhaul

Notice that the 2L-VRPB can be seen as an extension of the 2L-VRP, i.e. every 2L-VRP instance may be considered as a especial case of 2L-VRPB.
In class 1, each customer is associated a single item of width and length equal to nil. For this reason, our algorithm does not need to be modified to solve 2L-VRP instances. We tested our GA approach contemplating sequential oriented loading ($2|SO|$L).
We generate a new set of instances for the static 2L-VRP whith backhaul from the classical 2L-CVRP using the method described by Toth and Vigo [22] to generate VRPB instances from classic Euclidean VRP ones. Thus, we have generated three new 2L-VRPB instances for each 2L-VRP one. These new instances contain 50%, 60%, and 80% linehaul customers. To obtain such linehaul-backhaul distributions, we select a customer every two, three, or five customers, respectively, to be a backhaul location. These linehaul/backhaul configurations are represented in Table 2. Accordingly, we produced a total of 180 2L-VRPB instances derived from the 180 2L-VRP instances introduced by [9] and [7]. Then, in Table 3, our results are compared with BR-LNS of [4], to the best of our knowledge, only this work have investigated the 2L-VRPB for the sequential oriented loading ($2|SO|$L).

Therefore, on 60 instances with 50% Linehaul and %50 backhaul of the Sequential 2L-VRPB, our AGA could find better value for 42 (in bold) instances (70%) and matches the same value than BR-LNS for 11 (18%) instances from class 1 to 5. We reveals that our AGA is able to provide best solutions for 29 and 22 optimal solutions out of 60 total instances. Finally, for 80% linehaul and 20% backhaul customers, our AGA provides 35 best results compared to the results obtained by the BR-LNS algorithm.

Based on the generated results, we can remark that HGA-VNS is very efficient for solving the 2L-DVRPB. For 60% linehaul and 40% backhaul customers,

C. Computational Results for the dynamic 2L-VRP with backhaul

The degree of dynamism dod is fixed to 0.5; this means that a half of the customers is considered as static, while the other half is dynamic. The 2L-DVRPB configurations are presented in table 3.

V. Conclusion and future work

A Dynamic Vehicle Routing Problem with Backhauls and two-dimensional loading problem (2L-DVRPB) has been studied in this article. We present an adaptive genetic algorithm for

TABLE I
A META-TUNING OF THE GA

Parameters	Values
Population Size (N)	100
Selection	The tournament selection
Crossover rate	0,65
Mutation rate	0,03
Replacement strategy	The elitism operator
Maximum number of generation	1000

TABLE II
RESULTS FOR THE STATIC 2L-VRPB FROM CLASS 1 TO 5 WITH (50 %, 60%, 80% LINEHAUL CUSTOMERS RESPECTIVELY) AND (50%,40%, 20% BACKHAUL CUSTOMERS RESPECTIVELY)

Ist	Class 1			Class 2			Class 3			Class 4			Class 5		
	BR-LNS	Our AGA	%gap	BR-LNS	Our AGA	%gap	BR-LNS	Our AGA	%gap	BR-LNS	Our AGA	%gap	BR-LNS	Our AGA	%gap
50 % linehaul customers and 50% backhaul customers															
1	301.99	**300.74**	-0.41	308.76	**308.75**	-0.003	308.76	**301.33**	-2.40	312.12	**310.20**	-0.61	307.63	307.63	0.00
2	308.76	308.76	0	308.76	308.76	0	308.76	**305.66**	-1.004	308.76	308.76	0	308.76	308.76	0.00
3	335.54	**335.47**	-0.02	336.40	**334.37**	-0.6	345.66	**344.33**	-0.39	335.54	335.54	0.00	335.54	336.31	0.22
4	375.12	**373.10**	-0.54	375.12	**373.10**	-0.53	375.12	**373.44**	-0.45	375.12	**374.25**	-0.23	375.12	375.12	0.00
5	372.12	**368.98**	-0.84	376.84	**375.72**	-0.3	373.71	**373.27**	-0.12	372.12	**370.63**	-0.4	372.12	373.78	0.47
6	432.30	**430.61**	-0.40	428.88	429.97	0.25	432.30	**429.89**	-0.55	432.30	**427.31**	-1.15	432.30	429.30	0.69
7	689.32	**686.44**	-0.42	692.26	**690.25**	-0.69	691.85	695.85	0.57	699.27	**698.99**	-0.04	689.32	**687.83**	-1.52
8	689.32	**688.75**	-0.08	698.87	**697.89**	-0.14	718.89	**717.79**	-0.15	692.26	**691.96**	-0.04	677.52	**676.91**	-0.21
9	494.03	**494.11**	-0.01	501.48	**499.98**	-0.29	494.03	**493.69**	-0.07	500.57	**499.97**	-0.12	494.03	**494.01**	-0.004
10	502.77	**501.86**	-0.18	610.45	**609.54**	-0.15	536.29	**535.92**	-0.069	589.43	**588.34**	-0.18	571.68	**568.98**	-0.472
11	502.77	502.77	0.00	603.37	**602.85**	-0.08	581.42	**581.37**	-0.008	644.27	**644.21**	-0.01	573.31	573.57	0.04
12	471.46	471.46	0.00	482.63	482.63	0.00	471.46	471.46	0.00	475.76	**475.4**6	-0.06	471.46	471.46	0.00
60 % linehaul customers and 40% backhaul customers															
13	2276.57	2276.57	0.00	2399.98	2399.98	0.00	2384.40	2384.40	0.00	2354.57	2354.57	0.00	2326.80	2356.80	1.28
14	751.69	751.69	0	870.04	870.04	0.00	878.23	878.23	0.00	777.60	777.60	0.00	771.31	771.31	0.00
15	751.69	**750.78**	-0.12	850.73	**848.99**	-0.20	853.62	853.62	0.00	909.02	909.23	0.21	907.13	**906.55**	-0.63
16	543.09	**542.98**	-0.20	549.86	549.86	0	544.24	**544.21**	-0.005	543.39	**543.21**	-0.03	542.60	**542.46**	-0.025
17	638.14	**637.95**	-0.03	635.94	**634.89**	-0.16	635.94	**634.96**	-0.15	638.14	638.14	0	635.94	635.98	0.006
18	834.86	**834.82**	-0.005	937.03	**936.92**	-0.11	919.65	**919.56**	-0.01	918.57	918.57	0	845.35	845.38	0.003
19	562.83	562.83	0.00	655.44	655.44	0.00	655.97	655.97	0.00	637.33	637.33	0.00	617.50	617.50	0.00
20	319.72	**319.27**	-0.14	419.92	**419.29**	-0.15	397.05	**396.83**	-0.05	398.26	398.26	0.00	375.20	375.23	0.008
21	721.78	**720.87**	-0.13	876.37	**876.33**	-0.004	892.75	**892.68**	-0.08	844.89	**844.81**	-0.01	783.33	**783.15**	-0.03
22	721.68	**721.35**	-0.045	872.10	**872.06**	-0.004	862.65	862.67	0.002	899.02	899.08	0.006	805.17	805.17	0.00
23	746.90	**746.55**	-0.05	880.09	**879.99**	-0.01	860.55	**860.05**	-0.06	862.22	862.25	0.003	802.86	802.86	0.00
24	838.96	**838.69**	-0.03	920.51	**919.98**	-0.06	890.40	**890.28**	-0.013	896.58	**896.24**	-0.037	844.15	844.15	0
80 % linehaul customers and 20% backhaul customers															
25	889.59	889.59	0	1144.05	1144.07	0.002	1102.54	1102.54	0	1091.96	1091.96	0	984.61	984.61	0.00
26	779.21	**778.99**	-0.09	1031.22	**1031.08**	-0.01	1039.09	1039.09	0.00	1096.63	1096.63	0.00	903.86	903.88	0.002
27	964.88	**962.89**	-0.20	1073.48	1073.48	0	1089.58	**1085.85**	-0.34	1058.67	**1057.76**	-0.09	1012.70	**1011.70**	-0.09
28	1022.91	**1022.90**	-0.09	1780.33	**1779.33**	-0.05	1801.48	**1800.84**	-0.03	1813.13	1813.17	0.002	1616.89	1616.89	0.00
29	1217.36	1217.36	0.00	1727.00	1727.00	0.00	1638.68	1638.68	0.00	1667.36	1667.36	0.00	1625.58	1625.58	0.00
30	1050.11	**1050.09**	-0.002	1415.14	**1415.08**	-0.004	1396.25	**1395.52**	-0.09	1385.71	1385.71	0	1236.57	1236.59	0.002
31	1216.24	**1215.85**	-0.03	1686.66	**1684.95**	-0.10	1698.68	**1697.89**	-0.04	1730.54	**1728.86**	-0.09	1545.89	**1542.98**	-0.18
32	1202.83	**1201.38**	-0.12	1700.82	**1700.52**	-0.017	1679.53	**1679.35**	-0.01	1687.62	**1686.86**	-0.04	1521.70	**1515.61**	-0.40
33	1213.71	1213.71	0.00	1716.05	**1716.00**	-0.003	1715.24	**1715.14**	-0.006	1732.86	**1730.68**	-0.12	1505.30	1505.30	0.00
34	702.84	**701.48**	-0.2	890.10	**890.07**	-0.003	908.90	**906.85**	-0.22	877.18	**875.96**	-0.14	808.02	808.32	0.37
35	747.01	747.01	0	1006.72	**1005.27**	-0.14	1020.11	**1018.96**	-0.11	1027.38	**1026.83**	-0.05	893.69	**892.71**	-0.11
36	488.96	**488.66**	-0.06	1090.58	1090.58	0.00	1126.35	1126.35	0	1052.64	**1052.60**	-0.003	946.13	947.31	0.12
Avg	741.085	**740.48**	-0.17	912.61	**912.08**	-0.13	906.39	**905.68**	-0.13	884.13	**859.23**	-0.24	846.30	**839.21**	-0.019

solving the 2L-DVRPB. In addition, five heuristics are used to load items into the vehicle. To the best of our knowledge, the problem has not been analysed so far in the literature. Therefore, our approach is tested on an extensive set of instances which have been adapted from existing benchmarks for the 2L-VRP.

Since there has been no other algorithm in the literature for solving such problems, we could not compare the performances of different algorithms. In addition, our algorithm was applied to a series of newly constructed 2L-DVRPB benchmark instances. The obtained results indicate that our method is stable and capable of achieving very high utilization of the vehicle loading spaces. For future work, we can apply our method of AGA for the Two dimensional Dynamic Vehicle Routing Problem with Backhaul and Time windows constraints.

REFERENCES

[1] AbdAllah, A. M. F., Essam, D. L., and Sarker, R. A. On solving periodic re-optimization dynamic vehicle routing problems. *Applied Soft Computing*, *2017* **55**, 1-12.

[2] Bortfeldt, A. , Hahn, T. , Mannel, D. and Monch, L. Hybrid Algorithms for the Vehicle Routing Problem with Clustered Backhauls and 3D Loading Constraints. *European Journal of Operational Research*,*2015*, **243**, 82-96.

[3] Chen, S., Chen, R., Wang, G. G., Gao, J., and Sangaiah, A. K. (). An adaptive large neighborhood search heuristic for dynamic vehicle routing problems. *Computers & Electrical Engineering*,2018, **67**, 596-607.

TABLE III

RESULTS FOR THE DYNAMIC 2L-DVRPB FROM CLASS 1 TO 5 WITH 50 % LINEHAUL CUSTOMERS AND 50% BACKHAUL CUSTOMERS FOR A DOD= 10% AND 25% RESPECTIVELY

DoD= 10%					DoD=25%				
Cls 1	**Cls 2**	**Cls 3**	**Cls 4**	**Cls 5**	**Cls 1**	**Cls 2**	**Cls 3**	**Cls 4**	**Cls 5**
375.83	393.25	385.63	389.34	398.36	550.74	549.75	559.25	572.26	572.32
384.76	394.52	392.53	392.35	393.68	508.76	508.79	507.99	508.77	508.94
435.47	427.69	428.61	439.85	441.36	585.47	585.02	601.35	587.52	588.92
448.1	448.65	448.32	448.35	457.29	623.73	639.86	629.75	620.55	625.69
442.98	442.32	442.89	441.15	444.65	619.06	635.36	635.27	623.85	637.28
508.01	510.02	509.16	509.63	510.96	700.94	700.89	700.89	700.26	700.97
776.44	792.36	786.35	777.63	794.39	1187.07	1194.35	1199.75	1187.09	1189.45
838.75	858.97	963.75	851.36	967.98	1209.00	1229.92	1337.98	1298.36	1255.96
594.11	594.36	597.13	596.38	599.18	794.37	846.98	843.96	846.97	443.96
601.86	710.98	636.02	698.48	698.51	852.22	966.86	910.36	896.32	886.95
622.77	732.39	689.35	697.52	698.96	941.81	602.85	581.37	644.21	573.57
571.69	582.67	571.26	577.64	578.62	797.15	797.22	803.96	797.15	799.36
2976.57	3018.94	3020.96	3019.35	3014.36	3633.26	3789.32	3779.36	3780.96	3789.36
911.69	1024.39	1098.36	982.36	998.36	1307.05	1458.32	1459.22	1359.67	1359.98
911.18	1001.99	1024.64	1128.39	1121.16	1291.76	1398.35	1398.75	1485.39	1488.35
663.21	678.86	684.36	689.31	697.85	893.63	893.77	894.52	896.84	897.85
788.28	798.25	782.69	798.36	798.88	1040.31	1042.36	1047.66	1052.96	1055.63
1015.18	1245.36	1203.97	1198.36	1101.39	1347.18	1452.36	1457.63	1458.36	1347.22
686.61	789.36	786.39	789.64	791.38	913.19	1028.33	1024.36	1032.69	915.26
399.52	501.39	501.87	499.86	501.97	569.51	672.85	673.95	678.25	688.65
871.10	998.37	997.85	998.73	999.18	1177.22	1278.55	1298.55	1278.96	1299.36
872.21	984.37	992.37	992.72	994.87	1172.04	1284.36	1298.55	1299.52	1299.89
906.77	1008.96	1008.95	1007.85	1009.18	1186.00	1296.32	1294.56	1298.63	1295.86
1019.44	1192.25	1193.68	1124.39	1194.85	1299.67	1399.36	1399.25	1399.54	1399.75
1079.95	1279.73	1307.98	1279.35	1298.95	1412.95	1725.36	1739.35	1798.23	1798.66
969.91	1398.36	1397.84	1397.84	1399.87	1345.44	1665.36	1668.96	1678.25	1687.56
1073.13	1173.48	1187.97	1188.94	1197.98	1595.54	1695.86	1695.84	1696.85	1698.55
1243.33	1992.38	1994.87	1918.37	1997.86	1693.70	2396.52	2396.36	2396.85	2398.77
1567.94	2057.98	2078.94	2126.38	2132.98	1967.61	2417.52	2418.25	2427.25	2455.36
1350.09	1415.08	1395.52	1385.71	1236.59	1700.50	2102.36	2109.26	2121.3	2123.63
1536.16	1978.84	2009.38	2197.38	2009.87	1968.21	2365.32	2356.36	2365.32	2365.78
1521.86	1978.58	2009.87	2158.96	2007.89	1960.92	2469.32	2478.23	2485.22	2485.86
1534.66	2001.32	2278.96	2287.94	1978.89	1974.56	2470.23	2484	2484.32	2488.63
851.84	990.28	1101.39	998.87	998.91	1158.30	1286.45	1352.63	1286.85	1289.63
907.67	1209.96	1291.69	1297.82	1198.34	1227.33	1526.27	1532.16	1592.36	1584.96
587.34	1590.58	1663.53	1552.67	1449.87	857.51	1486.36	1487.25	1426.77	1488.56

[4] Dominguez, O., Guimarans, D., Juan, A. A., and de la Nuez, I. A biased-randomised large neighbourhood search for the two-dimensional vehicle routing problem with backhauls. *European Journal of Operational Research*, *2016*, **255(2)**, 442-462.

[5] Duhamel, C. and Lacomme, P. and Quilliot, A. and Toussaint, H. A multi-start evolutionary local search for the two-dimensional loading capacitated vehicle routing problem.*J. Computers & Operations Research. 2011* **38**, 617-640.

[6] Fuellerer, G. and Doerner, K. and Hartl, R. and Iori, M. Ant colony optimization for the two-dimensional loading vehicle routing problem. *J. Computers & Operations Research 2009* **36**, 655-673.

[7] Gendreau, M. and Iori, M. and Laporte, G. and Martello, S. A Tabu search heuristic for the vehicle routing problem with two-dimensional loading constraints. *Networks*, *2008* **51** 4-18.

[8] Guimarans, D., Dominguez, O., Panadero, J., Juan, A. A. A simheuristic approach for the two-dimensional vehicle routing problem with stochastic travel times. *Simulation Modelling Practice and Theory*, 2018, **89**, 1-14.

[9] Iori, M. and Salazar, J.J. and Vigo, D. An exact approach for the vehicle routing problem with two-dimensional loading constraints. *J. Transportation Science*, *2007* **41**, 253-264.

[10] Kilby, P., Prosser, P., and Shaw, P. Dynamic VRPs: A study of scenarios. *University of Strathclyde Technical Report*, *1998*, 1-11.

[11] Leung, S. and Zhou, X. and Zhang, D. and Zheng, J. Extended guided tabu search and a new packing algorithm for the two-dimensional loading vehicle routing problem: *Computers & Operations Research. 2011* 205-215.

[12] Leung, S. and Zhang, Z. and Zhang, D. and Hua, X. and Lim, M. A meta-heuristic algorithm for heterogeneous fleet vehicle routing problems with two-dimensional loading constraints.*J. Computers & Operations Research,2013* **225**, 199-210.

[13] Lund, K., Madsen, O.B.G., and Rygaard. Vehicle routing problems with varying degrees of dynamism. Technical report, Institute of Mathematical Modelling, Technical University of Denmark, 1996.

[14] Montemanni, R., Gambardella, L. M., Rizzoli, A. E., and Donati, A. V. . A new algorithm for a dynamic vehicle routing problem based on ant colony system.*In Second international workshop on freight transportation and logistics*, *2003, April* (Vol. 1, **No. 1**, pp. 27-30).

[15] Ninikas, G., Minis, I. Reoptimization strategies for a dynamic vehicle routing problem with mixed backhauls. *Networks*, *2014* **64(3)** , 214231

[16] Ninikas, G., and Minis, I. Load transfer operations for a dynamic vehicle routing problem with mixed backhauls. *Journal on Vehicle Routing Algorithms*, *2018* **1**(1), 47-68.

[17] Pinto, T., Alves, C., de Carvalho, J. V. and Moura, A. An Insertion Heuristic for the Capacitated Vehicle Routing Problem with Loading Constraints and Mixed Linehauls and Backhauls.*FME Transactions*, *2015* **43(4)**, 311-318.

[18] Pinto, T., Alves, C., and de Carvalho, J. V. Variable neighborhood search algorithms for pickup and delivery problems with loading constraints. *Electronic Notes in Discrete Mathematics*, *2017* **58**, 111-118.

[19] Reil, S., Bortfeldt, A., and Monch, L. Heuristics for vehicle routing problems with backhauls, time windows, and 3D loading constraints. *European Journal of Operational Research*, *2018* **266(3)**, 877-894.

[20] Sbai, I., Limem, O., and Krichen, S. An Adaptive Genetic Algorithm for the Capacitated Vehicle Routing Problem with Time Windows and Two-Dimensional Loading Constraints. *In Computer Systems and Applications (AICCSA), 2017 IEEE/ACS 14th International Conference* on *2017, October*(pp. 88-95). IEEE.

[21] Sbai, I., Limem, O., and Krichen, S.(in press). An effective Genetic Algorithm for solving the Capacitated Vehicle Routing Problem with Two-dimensional Loading Constraint. *International Journal of Computational Intelligence Studies*, *2018*

[22] Toth, P., and Vigo, D. An exact algorithm for the vehicle routing problem with backhauls. *Transportation science*, *1997* **31(4)**, 372-385.

[23] Wei, L., Zhang, Z., Zhang, D., and Lim, A. A variable neighborhood search for the capacitated vehicle routing problem with two-dimensional loading constraints. *European Journal of Operational Research*, *2015* **243(3)**, 798-814.

[24] Wei, L., Zhang, Z., Zhang, D., Leung, S. C. A simulated annealing algorithm for the capacitated vehicle routing problem with two-dimensional loading constraints. *European Journal of Operational Research*, 2018, **265(3)**, 843-859.

[25] Zachariadis, E. and Tarantilis, C. and Kiranoudis, C. A guided tabu search for the vehicle routing problem with two-dimensional loading constraints. *J. European Journal of Operational Research*, *2009* **195**, 729-743.

[26] Zachariadis, E. and Tarantilis, C. and Kiranoudis, C. Integrated distribution and loading planning via a compact metaheuristic algorithm.*J. European Journal of Operational Research*, *2013* **228**, 56-71.

[27] Zachariadis, E. E., Tarantilis, C. D., Kiranoudis, C. T. Vehicle Routing Strategies for Pick-up and Delivery Service Under Two Dimensional Loading Constraints. *Operational Research*, *2017*, **17(1)**, 115143.

HIGHLIGHTING THE USER'S PRIMARY CRITERION IN TEXT MINING-BASED RECOMMENDATION SYSTEMS

Mariem BRIKI
LARODEC
Tunis, Tunisia
briki.mariem@hotmail.com

Sabrine BEN ABDRABBAH
LARODEC
Tunis, Tunisia
abidrabbah.sabrine@gmail.com

Nahla BEN AMOR
LARODEC
Tunis, Tunisia
nahla.benamor@gmx.fr

***Abstract*—Recommendation systems (RS) were proposed to filter data in order to provide users with items that match their needs and interests. Most of the existing recommendation algorithms are designed to deal with the numerical notes (i.e. ratings) that may not be consistent (i.e. users may not interpret the rating scale in the same way). The semantic web is introduced in the context of RS in order to consider a new type of data so-called textual data. The textual data is exploited to improve the personalization task in many application fields. In fact, analyzing textual data can reveal hidden information that helps us to better capture users' preferences and interests and consequently to determine correctly the users' profiles. However, the existing works are unable to consider the multi-criteria feature of items which is quite important to represent more complex users' preferences when generating recommendations. This paper proposes a text mining-based RS that consists in exploiting the users' reviews in order to find out the primary criterion of the active user. Finally, we recommend the items that match that criterion by computing the percentage of its existence into the items reviews content. We went through improvement versions that we tested on a real database extracted from the TripAdvisor website. The proposed solution is able to greatly improve accuracy compared to the traditional approaches.**

***Index Terms*—Data analysis, Multi-Criteria Decision Making, Primary criterion, Recommendation systems, Text mining, Textual data, Unstructured data.**

I. INTRODUCTION

Due to the frequent accumulation of Web resources, finding useful and relevant information is becoming more and more a difficult task to manage. RS have been integrated into the online information systems (e.g. e-commerce platforms, social media sites, etc.) to solve the increasing information overload problem by providing users with new items that may interest them. Most of recommendation algorithms handle numerical data since it can be easily expressed. However, the numerical data may not present the evaluation of the user's notes correctly.

Rizkallah in [1] estimated that 80% of all generated data is unstructured. This data can be in the form of textual data derived from several sources (e.g. comments from shared publications on social networks, replies to registration questionnaires, etc.) and can be represented in users' reviews. These reviews may be a powerful source of information that can be exploited in RS in order to model the users' preferences. At this point, text mining techniques (e.g. categorization, entity extraction, sentiment analysis and natural language processing) are used to find out the hidden knowledge in text content and revealing useful patterns, trends and insights. However, the existing text mining-based RS focuses on computing the users' preferences without considering the multi-criteria aspect of the objects. In fact, the multi-criteria based systems provide an opportunity to compute more accurate recommendations by maintaining the details of users' preferences in multiple dimensions instead of considering the object in its globality.

Our objective is to improve the recommendation task by combining the text mining techniques, the multi-criteria aspect and the creation of new information corpus. The main idea behind the proposed solution is to model correctly the user's profile in order to create a RS. This system computes the percentage of the existence of user's criteria in user's reviews using the created information corpus to find the most interesting items.

Therefore, we propose a text mining-based RS based on 4 steps: the first step consists in the use of the information extraction techniques to obtain users' comments from the data source. The second step consists in the use of text mining techniques to clean and to analyse the users reviews. The third step consists in computing the percentage of the existence of the corpus of information of each criterion in the user's comments. At this level, we use a Multi-Criteria Decision Making (MCDM) method to detect the importance of each criterion according to the users. The fourth step consists in determining the user's primary criterion based on the computed percentage of the existence and generating recommendations accordingly.

Throughout this work, we implemented algorithms, including *Primary criterion's percentage-based recommendation system (PCPRS)* that uses text mining techniques to highlight the primary criterion of users, finds the similarity with the information corpus that we created, in order to provide the most suitable items and *Multi-criteria text mining-based recommendation system (MCTMRS)* that uses text mining techniques to model the user's profile and a multi-criteria decision making method to make recommendations. The performed

experimentation has shown a considerable improvement for the accuracy compared to the traditional methods.

The rest of the paper is organized as follows: Section 2 describes the basics on RS and related works. Section 3 details the proposed solution and Section 4 presents the endings emerged from the experiments.

II. Basics on recommendation systems and related work

RS have become an indispensable tool for users. They offer to online information systems an intelligent, responsive and useful service by providing suggestions for items or services that match the user's expectations. These systems have been widely used in various domains and diverse applications and have drawn increasing intention from different research communities such as machine learning, electronic commerce and information retrieval. RS needs to know as much as possible information from the user in order to provide reasonable recommendations right from the onset. RS rely on different types of input data :

- *Explicit feedback* which is the information we obtain by directly questioning the user about the proposed items.
- *Implicit feedback* which is the information we obtain by analyzing the behaviour of the user such as clicks/ queries/ watches...

The output of RS can be twofold [2]:

- A *prediction* expressed as a numerical value, $r_{a,j}$, which represents the anticipated opinion of active user u_a for item i_j. This predicted value should necessarily be within the same numerical scale (e.g. 1-bad for 5-excellent) as the input referring to the opinions provided initially by active user u_a. This form of recommendation system's output is also known as *individual scoring*.
- A *recommendation* expressed as a list of N items, where $N \leq n$, which the active user is expected to like the most. The usual approach in that case requires this list to include only items that the active user has not already purchased, seen or rated. This form of recommendation system's output is known as *top-N recommendation* or *ranked scoring*.

Fig. 1 highlights the recommendation phases.

The use of efficient and accurate recommendation methods is very important for a system that will provide good and useful recommendations to its users. This process explains the importance of understanding the features and potentials of different recommendation approaches. The recommendation methods are classified into three main classes, including *content-based*, *collaborative filtering* and *hybrid*. First, the *collaborative filtering approach* [3] relies on a matrix of user-item ratings to predict unknown matrix entries, and thus to decide which items to recommend. There are two main CF approaches: *item-based* [4] where the recommendations are generated based on computing similarity between items that the active user liked in the past and *user-based* [5] where the recommendations are computed based on what similar users liked in the past. Second, the *content-based approach* [6] uses the contents of the items that the active user liked in the past and suggests items having similar features to them. And third, the *hybrid approach* [7] that combine between content-based filtering and collaborative filtering in order to take advantage from the representation of the content and the rating information of users to produce recommendations.

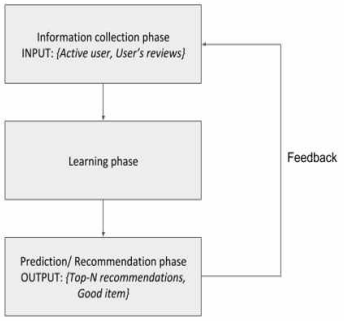


Fig. 1. Recommendation phases

Most of RS algorithms consider the quantitative preferences of users which are generally expressed on a numerical scale. However, the quantitative notes may be not consistent. In fact, they can be affected by many key factors including user's mood when one user may react differently with the same item according to his situation, the limited scale when the user may give the same rating to two items that he appreciates differently since the scale of possible values is generally reduced, etc. Thus, numerical data of users are not reliable and cannot represent the precise degree of users' liking. For these reasons, many works highlight the use of textual data as it may be a strong support to consider in the recommendation process.

A. Text mining-based recommendation systems

Text mining is a process to extract interesting and significant patterns to explore knowledge from textual data sources. It is a multi-disciplinary field based on information retrieval, data mining, machine learning, statistics, and computational linguistics [8].

Several works have used text mining techniques in RS to analyze unstructured data and capture the interest of users. This field is constantly evolving and researchers are constantly proposing new solutions in order to improve the recommendation and best meet the needs of customers.

Current RS are mainly based on customers' personal information and online behavior. Those systems have a lack in efficiency and accuracy. At the same time, we observe the large amount of review data with exponential growth. Based on this observation, Li et al. proposed in [12] a RS based on opinion mining. With text mining methods, they extract the opinion related information from the massive reviews to find

the most suitable products for customers. Ziani et al. presented in [13] a basic tool which can be used to analyze Algerian reviews and comments and detect their polarity in order to generate meaningful recommendations for users. Ganu et al. enhanced the RS in [11] by manipulating topic and sentiment information at sentences level. They estimated ratings from text comments written by users about restaurants in multi-point rating scales, instead of the two bipolar classes: positive or negative. They used the regression model to estimate scaled sentiment points from written reviews and they were the first who integrated the sentiment information from reviews into RS. The proposed solution in [9] was a personalized hotel recommendation approach based on both textual and contextual data. The authors identified the user's preferred aspects through tracking the browsing behavior on the mobile devices. The proposed solution in [32] was a system that implements efficient recommendation system by using proposed key extraction algorithm, Content-based Filtering (CBF) method and Jaccard Coefficient that will help the users who want to buy the car by providing relevant car information.

However, these works do not consider the multi-criteria aspect of the objects in the context of RS.

B. Multi-criteria-based recommendation systems

Multi-criteria decision making (MCDM) [10] is a branch of operational research dealing with finding optimal results in complex scenarios including various indicators, conflicting objectives and criteria. Applications of MCDM include areas such as integrated manufacturing systems, evaluations of technology investment, water and agriculture management, etc. Several methods have been introduced in the multi-criteria decision making field to enable users make decisions based on different parameters. The most know are:

1) *Weighted sum method (WSM)* [33] is used for evaluating a number of alternatives in accordance to the different criteria which are expressed in the same unit. It is the most commonly used approach. For instance, let M be the alternatives and N be the criteria. Then the best choice is the one that respond the following equation:

$$A^*_{WSM} = Max \sum_{i}^{j} a_{ij} w_j \quad for \quad i = 1, 2, 3, ..M \quad (1)$$

where A^*_{WSM} is the WSM score of the best alternative, N is the number of decision criteria, a_{ij} is the actual value of the i^{th} alternative in terms of the j^{th} criterion, and w_j is the gravity of products' attention of the j^{th} criterion.

The amount value of each preference is equal to the sum of items.

2) *Weighted product method (WPM)* [33] is related to the previous method, WSM. The main contrast is that there is multiplication instead of addition in the model. Every criteria is compared with the others by multiplying a number of ratios, one for each alternative. Each ratio is raised to the power equivalent to the relative weight of the current criterion. Principally, the following product is obtained in order to compare the alternatives A_K and A_L :

$$R(\frac{A_K}{A_L}) = \sum_{j=1}^{N} (\frac{a_{kj}}{a_{Lj}})^w_j \quad (2)$$

where N is the number of criteria, a_{kj} is the actual value of the k^{th} alternative in terms of the j^{th} criterion, and w_j is the weight of importance of the j^{th} criterion. If $R(\frac{A_K}{A_L})$ is better than one, then alternative A_K is more desirable than alternative A_L. The top alternative is introduced as the one that is greater than or at least correspond to all the other alternatives.

3) *Analytical hierarchy process (AHP)* is created by [34] [35]. The main idea of the process is the division of a complex challenge into a hierarchy with objective at the top of the hierarchy, criteria and sub-criteria at levels and sub-levels of the hierarchy, and decision alternatives at the bottom of the hierarchy. Objects in a specific hierarchy level are compared in dual to assess their relative preference with respect to every element at the next higher level.

The multi-criteria aspect is introduced in the context of RS to enable them to detect the users' interests by defining their primary criterion. Defining a correct user profile requires an extra effort to properly detect its expectations. In [14], Ko et al. enable the multi-aspect collaborative filtering with the datasets that include only ratings on a single-criterion. In [15], Nilashi et al. proposed a new recommendation method that uses the regression and clustering techniques to improve the predictive accuracy of the multi-criteria collaborative filtering. They used expectation maximisation in the recommendation process. A new multi-criteria recommendation system for tourism domain using a set of machine learning techniques was proposed in [16]. Nilashi et al. used the Adaptive Neuro-Fuzzy Inference Systems and Support Vector Regression as prediction techniques, Principal Component Analysis as a dimensional reduction technique. Self-Organizing Map and Expectation Maximization as two well known clustering techniques. The main contribution of this work was the use of clustering ensemble method in the recommendation model. Hdioud et al. extended in [17] the concept of mono-criterion ratings to multi-criteria ones to meet the requirements of more practical RS by using the Correlation Coefficient and Standard Deviation Approach for the Weights (CCSDW) of attributes selection method. In [19], Chamoso et al. extracted relevant information from job offers and from the text content and links shared by users. Carrying out the categorization process allows the determination of affinity among users. Based on the content and profiles of users, new relationships are recommended to those with similar interests.

The first hierarchical representations for describing user profiles and documents have been proposed in [18]. In order to customize the recommended products, we should identify the user's preferences which can be described by three elements:

generation, *maintenance* and *exploitation of user profile*, while using behind a recommendation algorithm.

1) The *generation* is among the essential aspects of a user profile. There are several methods of generation:
   - Empty profile: Once a user makes a simple registration in a system, an empty profile will be created automatically which will subsequently be updated after any user interaction with the system.
   - Manually: The users have a personal space that will allow him to register their interests (e.g. by checkboxes, by keywords, by tags, etc), their demographic information also (age, marital status, sex, etc), as well as their geographic data (e.g. city, country, street, etc) and psychographic data (e.g. lifestyle, nature of music, etc).
   - Import: The users can be asked by the system to fill in and add information available via external sources (e.g. files to import in pdf, links to other pages on the web, etc)
   - Learning Set: After consulting or trying a specific article, the user is asked to provide feedback on these articles (e.g. feedback on experience), indicating whether they are relevant or irrelevant to his interests.
   - Stereotype: If a new user ever joins the system, he will be assigned to a group of users who have the same interests as him. Also represented by their stereotype, taking into consideration certain demographic, geographic or psychographic information.
2) The *maintenance* highlight the update of user preferences. Once the user profile is initialized, the system will automatically consider the actions that the user is doing and the preferences they are mentioning and will update the user profile.
3) The *exploitation the user's profile* and preferences takes into consideration the user's actions that the system detect. We recommend finally the items that meet the user's needs.

C. Text-mining recommendation systems introducing the multi-criteria aspect

In [20], authors presented a multi-criteria recommendation system for hotel recommendations to choose the best suited hotel in a city according to a user's preference and other user's reviews. In order to determine the rating of a hotel from previous users with respect to different parameters the paper uses various NLP approaches on a hotel review corpus and builds a user-item-feature database.

In [21], authors used a text mining approach to mine product features, opinions and their semantic similarity from Web opinion sources. The consumer can clearly see the strengths and Weaknesses of each product in the minds of existing consumer's opinion. The system assists online shoppers or goal oriented shopper by suggesting the most effective navigation products for their specified criteria and preferences.

In [24], authors proposed a recommendation system for public facilities by utilizing both structured and unstructured data gathered from multi-channel data sources. The system uses single-criteria rating, multi-criteria-rating, and text data as inputs. The challenge is how to handle data variety such that any kind of data from any channel can be integrated. The second challenge is how to extract location-related data from the raw data. There are four data channels used in the system. Three of them are social media channels (i.e. Twitter, Instagram, and Foursquare) while the other is internal data channel built as a part of the system itself. The system deals with three categories of public facility (i.e. park, hospital, and mosque). The whole system consists of two subsystems (i.e. the extractor system, including the rating, input module and the recommendation system).

In [22], authors proposed a highly accurate hotel recommender system, implemented in various layers. Using multi aspect rating system and benefit from large-scale data of different types, the recommender system suggests hotels that are personalized and tailored for the given user. The system employs NLP and topic modelling techniques to assess the sentiment of the users' reviews and extract implicit features. The entire recommender engine contains multiple sub-systems, namely users clustering, matrix factorization module, and hybrid recommender system. Each sub-system contributes to the final composite set of recommendations through covering a specific aspect of the problem.

In [23], authors proposed a framework that provides users with a ranked list of alternatives, while it also permits them to submit their evaluations on existing objects of the database. Much attention is given to the extent in which the user evaluation may affect the values of the stored objects. The applicability of our approach is demonstrated through a web-based tool that provides recommendations about visiting different cities of a country.

III. Proposed solution

Nowadays, users tend to express better their opinion and preferences by text. However, the existing works does not treat deeply the multi-criteria aspect and user's preferences which is a crucial feature to consider when building user profiles and generating recommendations. Exploiting the textual data has become a necessity especially with the rapid growth of the unstructured data coming from different sources such as: comments on items on websites, emails, etc. The availability of huge volumes of textual data is an opportunity to extract valuable information. In this context, text mining techniques are used to analyze and identify relevant information effectively.

To this end, we propose a text mining-based recommendation system that consists in exploiting the users' reviews using text mining techniques to analyse and find out the primary criterion of the active user to provide recommendations.

Our solution is based on the use of text mining techniques to find out the user's interest on different items' criteria. We

assume that by identifying the criteria in which users are interested, a better representation of users' profiles and preferences is generated. Informally, we define the primary criteria as the most mentioned vocabulary in users' reviews. For example, a user can talk about the criterion 'service' expressed by different words that describe the service more than the other criteria. The main idea is to use the unstructured data in order to consider the deep aspect of users' preferences and to detect their primary criteria. This provides them corresponding recommendations.

To this end, we propose recommendation algorithms. The *Primary criterion's percentage-based recommendation system* is based on the use of text mining techniques to highlight the primary criterion of users in order to provide the most suitable items and the *Multi-criteria text mining-based recommendation system* is based on the use of text mining techniques to model the user's profile and a multi-criteria decision making method to make recommendations.

A. Primary criterion's percentage-based recommendation system algorithm

Five corpus describing each object criterion have been created to better describe each concept. For example, for the criteria "service", we created a list of words (i.e verbs, adjectives, etc) that can be related or that can describe the relation between clients and persons who works into a restaurant. The information corpus is a way to express the real meaning of each criteria. It is denoted as CC_i where i is the current criteria.

Given the users' reviews, we start by putting all the active user's comments in one sentence denoted as AUR. The AUR is defined as the union of all the active user's reviews. Formally, AUR is:

$$AUR = \cup_{j \in t} AUR_j \tag{3}$$

where t is the number of the active user's reviews and j is the current review.

Then, to find out the primary criterion of the $activeuser$, we compute the percentage of the existing of each criterion i in the active user's reviews denoted by $UPEC_i$. The criterion with the highest $UPEC_i$ value is selected as the primary criterion of the active user. Formally, $UPEC_i$ is:

$$UPEC_i = \frac{AUR \cap CC_i}{\sum_{j \in n} AUR \cap CC_j} \tag{4}$$

where n is the number of criteria, j is the current criterion and i is the information corpus of the current criterion.

Let m be the number of all the items not yet seen by the active user. Given the users' reviews of input data, we put all the comments of each unseen item in one sentence denoted as IR_k. The IR_k is defined as the union of all the reviews given to item k. Formally, IR_k is :

$$IR_k = \cup_{j \in s} R_{j,k} \tag{5}$$

where s is the number of all the reviews to item k and $R_{j,k}$ is the j^{th} review to item k.

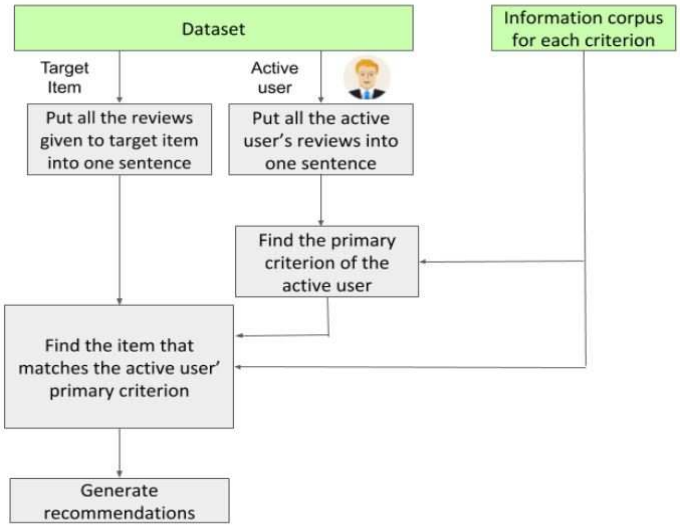


Fig. 2. Process of *Primary criterion percentage-based recommendation system*

Finally, to find out the items matching the primary criterion of the active user, we compute the percentage of the existing of each criterion in the reviews content of the unseen item k denoted by $IPEC_{i,k}$. Formally, $IPEC_{i,k}$ is:

$$IPEC_{i,k} = \frac{IR_k \cap CC_i}{\sum_{j \in n} IR_k \cap CC_i} \tag{6}$$

where i is the current criterion and k is the current unseen item.

The items having the highest $IPEC$ value with the active user's primary criterion are then selected for the recommendation list. To better explain, we have summarized the process of this solution in Figure 2. Algorithm 1 presents the pseudo-code of the *Primary criterion's percentage-based recommendation system* algorithm.

```
Initialization:
CriteriaPercentageUser ← getCriteriaPercentage(CleanListUser,
BibOfEachCriteria)
// getCriteriaPercentage(CleanListUser,
BibOfEachCriteria) is a function that returns
a list of existence's percentages of each
criterion in user/ item reviews.
for item in notvisitedItems do
   CriteriaPercentageItem ←
   getCriteriaPercentage(CleanListItem, BibOfEachCriteria)
   TopU, TopI ← 0
   maxU ← max(CriteriaPercentageUser)
   maxI ← max(CriteriaPercentageItem)
   if Index(maxU) is Index(maxI) then
      RecommendedList ← Item
   end if
end for
sort(RecommendedList)
return RecommendedList
```

Algorithm 1: Primary criterion's percentage-based recommendation system

B. Multi-criteria text mining-based recommendation system algorithm

The added value of the *Multi-criteria text mining-based recommendation system* algorithm is to use the WSM method of the Multi-Criteria Decision Making aspect to measure the weight of importance of each criterion in the object in its globality.

We followed the same process of the *Primary criterion's percentage-based recommendation system* algorithm by integrating WSM method to compute the weight of each criterion in order to obtain a better recommendation.

Algorithm 2 presents the pseudo-code of the *Multi-criteria text mining-based recommendation system* algorithm.

```
Initialization:
CriteriaPercentageUser ← getCriteriaPercentage(CleanListUser,
BibOfEachCriteria)
// getCriteriaPercentage(CleanListUser,
BibOfEachCriteria) is a function that
computes the WSM of each criterion in
user/item reviews.
for item in notvisitedItems do
   CriteriaPercentageItem ←
   getCriteriaPercentage(CleanListItem, BibOfEachCriteria)
   TopU, TopI ← 0
   maxU ← max(CriteriaPercentageUser)
   maxI ← max(CriteriaPercentageItem)
   if Index(maxU) is Index(maxI) then
      RecommendedList ← Item
   end if
end for
sort(RecommendedList)
return RecommendedList
```

Algorithm 2: Multi-criteria text mining-based recommendation system

IV. Experimental study

We perform experimentation to evaluate the performance and effectiveness of the proposed solution.
This section describes at first our experimental data and the evaluation metrics. Then, we present the experimental protocol. Finally, we will discuss the experimental results relatives to different algorithms and tunings.

A. Dataset and metrics

The present work focuses on the domain of restaurants. In this domain, the overall preferences often come in the form of a rating and textual reviews. We are interested in TripAdvisor[1] website as it is one of the most famous online websites.

TripAdvisor is an American website that offers tourists advises from consumers on hotels, restaurants, cities and regions, places of leisure, etc., internationally. It also provides accommodation and airline ticket booking tools that compare hundreds of websites to find the best deals. Each week hundreds of users visit TripAdvisor to rate and receive recommendations for new restaurants.

We start by using the WebCrawler as a technique to extract the data used to test the performance of the proposed solution.

[1]https://www.tripadvisor.fr/

Number of distinct users	10000
Number of distinct restaurants	1000
Average of reviews' number per user	5
Average of reviews' number per restaurant	5

TABLE I
DESCRIPTION OF TRIPADVISOR DATASET

In fact, this technique is generally designed to explore the Web. It is a robot which is generally designed to collect resources (e.g. Web pages, images, videos, Word documents, PDF , etc.) and allows a search engine to index them.
The extracted data are available in excel format. It can be described by two main description formats, including *the textual description* and *the statistical description*.

1) **Textual description of the database**: The database extracted is a corpus of comments on restaurants in the Lille region. It contains 47168 comments and ratings collected from 10000 users of the Website on 1000 restaurants, from 18/11/2007 to 01/11/2016. Five criteria are considered, including *food*, *service*, *ambience*, *quality/ price* and *restaurant* in its globality. The score of preferences' rating is ranged from 1 to 5. 20 attributes describe the TripAdvisor's dataset that provides us with all the information needed for efficient operation.
2) **Statistical description of the database**: A statistical description allows to better analyze the database by providing statistics that can be used to better align the proposed solution. Table I and II presents a statistical description of TripAdvisor dataset. To begin the treatment, many conditions have been established:
- Divide the dataset into two sets: 80% for the training set and 20% for the test set.
   - A training set is a dataset of examples used for learning.
   - A test set is a dataset used to provide an unbiased evaluation of a final model fit on the training dataset.

   Given an algorithm that has as input the initial database and generating as output two files in .csv format, the conditions that have been taken into account are as follows:
   - If a user has only one comment, then it will be deleted.
   - If a user has two comments, then:
     - One line will be inserted in the training set.
     - The second line will be inserted in the test set.
   - If a user has more than 3 comments, then special treatment will be done:
     - If all the comments belong to completely different restaurants, then divide the instances in 80% for the training set and the rest for the test set.
     - If not, insert all comments made for the same restaurants in the training set and add more lines until you reach 80% of the user base and insert what remains in the test set.

Number of users having 1 review	Number of users having 3 reviews	Number of users having 4 reviews	Number of users having more than 4 reviews
4241	1410	2115	39362

TABLE II
DESCRIPTION OF TRIPADVISOR DATASET ACCORDING TO USER'S REVIEWS' NUMBER

	Suggested	**Non-Suggested**	**Total**
Relevant	$N_{r,s}$	$N_{r,n}$	N_r
Irrelevant	$N_{i,s}$	$N_{i,n}$	N_i
Total	N_s	N_n	N

TABLE III
ITEMS' CATEGORIZATION

After applying the conditions mentioned above, the database is now divided as follows: 37898 lines from the training set and 5029 lines of the test set. The data selection should take into account the temporal order of instances, so that, instances of the testing set should be chosen over those of the training set. The idea behind these conditions is to be sure that our models are learning correctly from the training dataset to recommend the right restaurants.

B. Evaluation metrics

Recommendation systems research has used several types of measures to evaluate the quality of the provided recommendation. The Evaluation metrics can be categorized into three classes, including *predictive accuracy metrics*, *classification accuracy metrics* and *rank accuracy metrics* [30]. In this work, we are interested in classifying metrics because they are most commonly used in the natural language processing and easier to interpret directly [30]. In fact, these metrics measures how many times a recommendation system makes correct or incorrect decisions about whether an item is good. Table III shows the possible categorization of items, where N is the number of items in the database and how the item's are categorized. We can conclude that recommended items can be either successful recommendations (i.e. relevant) or unsuccessful recommendation (i.e. non relevant) and the relevant items can be either suggested in the recommendation list or not. Precision, recall and F-Measure [31] are the most popular metrics for evaluating information retrieval systems.

- *Precision* is used to evaluate the validity of a given recommendation list and it is defined as the ratio of relevant items selected by the active user relative to the number of items recommended to him.

$$P = \frac{N_{r,s}}{N_s} \tag{7}$$

- *Recall* computes the portion of favored items that were suggested for the active user relative to the total number of the objects actually collected by him.

$$R = \frac{N_{r,s}}{N_r} \tag{8}$$

Algorithm/ Evaluation measure	**F-Measure**
Sentence similarity using cosine similarity measure	0.033
Sentence similarity using Jaccard similarity measure	**0.039**

TABLE IV
EVALUATION METRICS RESULTS FOR 1000 USERS USING COSINE & JACCARD SIMILARITY MEASURES

- *F-Measure* is a measure of a statistic test's accuracy. It considers both precision and recall measures of the test to compute the score. We could interpret it as a weighted average of the precision and recall, where the best F-Measure has its value at 1 and worst score at the value 0.

$$F = 2 * \frac{P * R}{P + R} \tag{9}$$

C. Experimental protocol

Table IV shows the performance of the proposed algorithm for the different similarity measures. This confirms that the measure of Jaccard performs better than that of cosine. That's why we choose to compare with the Jaccard similarity measure for the existing work. Several types of entries can be included in the recommendation systems such as digital data and textual data that will rely on our algorithms *Primary criterion's percentage-based recommendation system* and *Multi-criteria text mining-based recommendation system*. We used text mining techniques to clean and transform user's reviews as follows:

1) Eliminate special characters from user comments.
2) Create a list of stop words and add an extension. We use this latter to eliminate the stop words of the comments of the users.
3) Concatenate all the comments of a single user in one sentence.
4) Use the tokenization library to have a list of words that builds a sentence.
5) Use the stemmatization library to extract the origin of the words from the obtained sentence.

Once the comments are analysed and then cleaned, we move to the exploitation step. We perform experiments to study the existing algorithm that uses text mining techniques using Jaccard and the cosine measure. We aim to compute similarity between texts using Jaccard and cosine similarity metrics in order to detect which users are sharing the same opinion. We choose to compare our proposed solution with the traditional algorithms, including *Criterion's ratings average-based recommendation system* [36] that uses users' ratings in the multi-criteria aspect to provide a recommendation and *Reviews similarity-based recommendation system* [32] that uses text mining techniques and a similarity text measure in order to find similarity between reviews to make the recommendation. Though this choice, we study the impact of each feature in the performance of the recommendation. We test later the effectiveness of the proposed algorithms 1000 users, which are selected randomly. The results of the different algorithms are listed in Tables V and VI.

D. Experimental results

We starts by analyzing the results of the *Criterion's ratings average-based recommendation system* that uses digital data taking into consideration the multi-criteria aspect. It can be noticed, from Table V that the results of the performance of the recommendations of the *Criterion's ratings average-based recommendation system* are the worsts. In fact, the weight of each criterion may change and differs from user to user. Therefore, it is not objective to adopt such a ranking mechanism in recommendation systems. This confirms what is stated in [29]. Researchers note that the algorithm which is based on computing the average of the criteria is defined as a traditional way to rank multiple items.

When it comes to textual data, analyzing the algorithm so called *Reviews similarity-based recommendation system* that uses user's reviews to make recommendation based on the score obtained from computing similarities. Comparing the latter with the *Criterion's ratings average-based recommendation system*, we record a negligible difference. But this comparison results that, the more we exploit textual data, the more performed recommendations we'll have.

To go further, we compare the traditional algorithms with the *Primary criterion's percentage-based recommendation system* that uses text mining techniques to highlight the primary criterion of users in order to provide the most suitable items. We noticed the evolution of the results by increasing the number of users on which we test algorithms. In fact, the calculation of the percentages of the existence of each corpus of information on the 5 criteria created enabled us to detect the primary criterion of the active user. We use this criterion to finally model a more accurate user profile. This representation is the key element on which the recommendations were based. Finally, the results obtained are improving accuracy.

To go more further, we compare the previous algorithm with the proposed *Multi-criteria text mining-based recommendation system* that uses text mining techniques to model the user's profile and a multi-criteria decision making method, WSM, to make recommendations.

Table V shows that the textitMulti-criteria text mining-based recommendation system has the best results regarding the F-Measure evaluation metric than all other algorithms. This improvement is due to the use of the WSM method and to the creation of new information corpus that enabled our system to better detect the user's profile. Good results are highlighted by gold text in Table V.

This can only be a good sign towards the use of text mining techniques and multi-criteria decision making aspect. The results obtained could push researchers to exploit textual data rather than digital data and apply multi-criteria methods that consider the customers' preferences and primary criterion in order to make good recommendations and improve accuracy.

To go further, we tested the proposed solution and the traditional algorithms taking into consideration the number of comments that each user has made. We tried it on users who have 3, 4 and more than 4 comments in the TripAdvisor dataset. F-Measure metric is used to consider both precision and recall results. Table VI shows that the more reviews the user have, better is the quality of the recommendation. We could mention that the algorithm *Multi-criteria text mining-based recommendation system* is always the most efficient. This highlights the power of the proposed system in learning the users' preferences. Finally, we recommend products that better match the user's profile. Curves in Figure 3 are presenting F-Measure results according to the number of user's reviews of each algorithm. Horizontal axis presents the $number of user's reviews$ and the vertical axis presents the F-Measure results. The growth of curves is highlighting that the more the number of user comments increases, the better the recommendation is in each algorithm.

Algorithm	**Precision**	**Recall**	**F-Measure**	**Response time**
Ratings average [36]	0.034	0.035	0.021	0.265564 sec
Sentence similarity [32]	0.048	0.113	0.040	0.291961 sec
Preferences percentage	0.223	0.753	0.218	158.294949 sec
Multi-criteria	0.232	0.772	**0.235**	183.132370 sec

TABLE V
EVALUATION METRICS RESULTS FOR 1000 USERS

Algorithm	**Users having 3 reviews**	**Users having 4 reviews**	**Users having more than 4 reviews**
Ratings average	0.026	0.054	**0.133**
Sentence similarity	0.065	0.098	**0.171**
Preferences percentage	0.211	0.226	**0.270**
Multi-criteria	0.223	0.221	**0.290**

TABLE VI
F-MEASURE RESULTS DEPENDING ON THE NUMBER OF USERS' REVIEWS

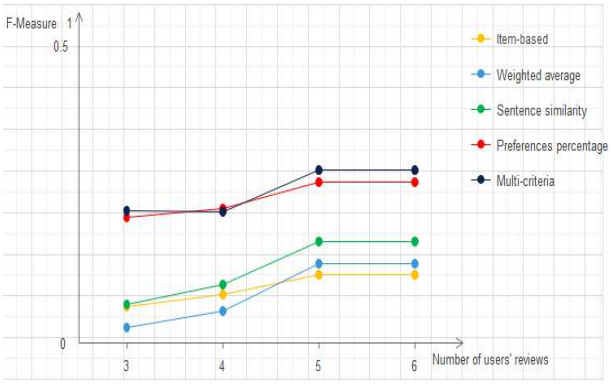


Fig. 3. F-Measure results depending on the number of users' reviews

V. Conclusion

In this work, we went through 2 improvement versions that we tested on a real database extracted from the TripAdvisor website. We presented as well the evaluation metrics used to evaluate the accuracy of the proposed solution. The experimental study presented shows that our proposed solution provides motivating results comparing with the existing works. Moreover, it deals with user's reviews where user's preferences are highlighted. Finally, the results obtained are improving accuracy. This can only be a good sign towards the use of text mining techniques and pushing researchers to exploit textual data rather than digital data in order to consider the customers' preferences and primary criterion. In parallel, information on the web is largely unstructured, with web pages authored by many people with a diverse range of topics. This makes simple browsing too time consuming to be practical. Web page filtering has thus become necessary for most web users in order to find the items they need. The Semantic Web offers some potential to help, allowing more intelligent search of web pages by utilizing semantic metadata. In this context, ontologies are introduced to analyze items comparing them with others following their semantic links. It is a way to increase the performance of recommendation systems.

References

[1] Rizkallah, "The big (unstructured) data problem.," Forbes, 2017.

[2] Vozalis, Emmanouil and Margaritis, Konstantinos G, "Analysis of recommender systems algorithms," in The 6th Hellenic European Conference on Computer Mathematics & its Applications., 2003, pp.732–745.

[3] J.S. Breese, D. Heckerman and K. Kadie *Empirical Analysis of Predictive Algorithms for Collaborative Filtering*, 1998.

[4] B. Sarwar, G. Karypis, J. Konstan and J. Riedl, *Item-Based Collaborative Filtering Recommendation Algorithms*, 2001.

[5] C. Sneha and G.P.S. Varma *USER-BASED COLLABORATIVE-FILTERING RECOMMENDATION*, 2015.

[6] F.O. Isinkaye, Y.O. Folajimi and B.A. Ojokoh *Recommendation systems: Principles, methods and evaluation*, 2015.

[7] W.T. Chu, and Y.L. Tsai *A hybrid recommendation system considering visual information for predicting favorite restaurants*, 2017.

[8] R. Talib, M. Hanif, S. Ayesha and F. Fatima, *Text Mining: Techniques, Applications and Issues*, 2016.

[9] K.P. Lin, C.Y. Lai, P.C. Chen and S.Y. Hwang *Personalized Hotel Recommendation Using Text Mining and Mobile Browsing Tracking*, 2015.

[10] H.A. Simon *The new science of management decision*, 1960.

[11] G Ganu, N. Elhadad and A. Marian *Beyond the Stars: Improving Rating Predictions using Review Text Contents*, 2009.

[12] X. Li, H. Wang and X. Yan, *Accurate Recommendation Based on Opinion Mining*, 2015.

[13] A. ZIANI, N. AZIZI, D. SCHWAB, M. ALDWAIRI, N. CHEKKAI, D. ZENAKHRA and S. CHERIGUENE *Recommender System Through Sentiment Analysis*, 2018.

[14] H. Ko, J. Son and I. Konstan, *Multi-Aspect Collaborative Filtering based on Linked Data for Personalized Recommendation*, 2015.

[15] M. Nilashi, M. Dalvi-Esfahani, M.Z. Roudbaraki, T. Ramayah and O. Ibrahim *A Multi-Criteria Collaborative Filtering Recommender System Using Clustering and Regression Techniques*, 2016.

[16] M. Nilashi, K.B. Fard, M. Rahmani and V. Rafe *A Recommender System for Tourism Industry Using Cluster Ensemble and Prediction Machine Learning Techniques*, 2017.

[17] F. Hdioud, B. Frikh, B. Ouhbi and I. Khalil *Multi-Criteria Recommender Systems: A Survey and a Method to Learn New User's Profile*, 2017.

[18] E. Savia, T. Kurki and S. Jokela, *Metadata Based Matching of Documents and User Profiles*, 1998.

[19] P Chamoso, A. Rivas, S. Rodríguez and J. Bajo *Relationship recommender system in a business and employment-oriented social network*, 2017.

[20] Sharma, Yashvardhan and Bhatt, Jigar and Magon, Rachit *A multi-criteria review-based hotel recommendation system*, 2015.

[21] Daoud, Mohammad and Naqvi, SK and Ahmad, Asad *Opinion Observer: Recommendation System on ECommerce Website*, 2014.

[22] Ebadi, Ashkan and Krzyzak, Adam *A hybrid multi-criteria hotel recommender system using explicit and implicit feedbacks*, 2016.

[23] Karacapilidis, Nikos and Hatzieleftheriou, Lefteris *Exploiting similarity measures in multi-criteria based recommendations*, 2003.

[24] Putri, Alifa Nurani and Akbar, Saiful and Sunindyo, Wikan Danar *Public facilities recommendation system based on structured and unstructured data extraction from multi-channel data sources*, 2015.

[25] Jain, Abhishek and Jain, Aman and Chauhan, Nihal and Singh, Vikrant and Thakur, Narina *Information retrieval using cosine and Jaccard similarity measures in vector space model*, 2017.

[26] Sailaja, D and Kishore, M and Jyothi, B and Prasad, NRGK *An overview of pre-processing text clustering methods*, 2015.

[27] J. HERLOCKER, J. KONSTAN, L. TERVEEN, and T. RIEDL *Evaluating Collaborative Filtering Recommender Systems*, 2004.

[28] C. Basu, H. Hirsh and W. Cohen *Recommendation as Classification: Using Social and Content-Based Information in Recommendation*, 1998.

[29] Lee, Hsin-Hsien and Teng, Wei-Guang *Incorporating multi-criteria ratings in recommendation systems*, 2007.

[30] Herlocker, Jonathan L and Konstan, Joseph A and Terveen, Loren G and Riedl, John T *Evaluating collaborative filtering recommender systems*, 2004.

[31] Basu, Chumki and Hirsh, Haym and Cohen, William and others *Recommendation as classification: Using social and content-based information in recommendation*, 1998.

[32] Naw Naw and Ei Ei Hlaing, *Relevant Words Extraction Method for Recommendation System*, 2013.

[33] Pohekar, SD and Ramachandran, M, *Application of multi-criteria decision making to sustainable energy planning—A review*, 2004.

[34] Wind, Yoram and Saaty, Thomas L, *Marketing applications of the analytic hierarchy process*, 1980.

[35] Thomas, L Saaty, *Decision Making for Leaders The Analitycal Hirarchy Proses for Decisions*, 1992.

[36] Hill, Will and Stead, Larry and Rosenstein, Mark and Furnas, George, *Recommending and evaluating choices in a virtual community of use*, 1995.

On solving the Hazardous Health-Care Waste Transportation Problem: a Real Case Study

Nasreddine Ouertani
Université de Tunis
ISG de Tunis, LARODEC
Université d'Artois, France
EA 3926 LGI2A, F-62400
nasreddine.ouertani@gmail.com

Hajer Ben-Romdhane
Université de Tunis
ISG de Tunis, LARODEC
Tunis, Tunisia
hajer.bn.romdhan@gmail.com

Issam Nouaouri
Université d'Artois
EA 3926 LGI2A, F-62400
France
issam.nouaouri@univ-artois.fr

Hamid Allaoui
Université d'Artois
EA 3926 LGI2A, F-62400
France
hamid.allaoui@univ-artois.fr

Saoussen Krichen
Université de Tunis
ISG de Tunis, LARODEC
Tunis, Tunisia
saoussen.krichen@isg.rnu.tn

***Abstract*—Health-care wastes are special category of wastes generated by medical facilities and containing infectious materials. According to the Word Health Organization, more than 15% of the medical wastes are very dangerous and represent a serious threat for humans and for the environment. In this paper, we investigate a major issue related to the collection and transportation of these wastes from a set of dispersed hospitals to the disposal center. The problem is modeled as a Capacitated Vehicle Routing Problem (CVRP) with the aim of minimizing the total travel distance. To solve this problem, three solutions approaches are considered: a hybrid evolutionary algorithm, a Genetic Algorithm, and a Variable Neighborhood Search. These approaches are tested over real data sets. The experimental results show that the hybrid approach is very competitive in comparison with the other methods.**

***Index Terms*—Health-care waste, Capacitated VRP, Evolutionary Algorithm, Genetic Algorithm, Variable Neighborhood Search**

1. Introduction

Medical waste -also known as Health-Care Waste (HCW) or Hospital Solid Waste- is all kind of garbage generated while providing health-care (e.g., treatment, diagnosis, surgical operation) to patients in hospitals, clinics, pharmaceutical manufacturing plants, and laboratories. The HCW include : human blood, wastes from patients carrying highly contagious diseases, contaminated sharps, wastes of cytotoxic chemicals and cytostatic drugs and other infectious materials [1]. The generated quantity of HCW depends essentially on the hospital applied practices, their capacity (number of beds) and the medical staff number [2].

During the last two decades, the HCW became one of the important category of waste due to its negative impact and potential risk to both human (public health, patients, medical staff, etc) and environment [3] [4]. The World Health Organization (WHO), stated that around 15% (or more) of the generated waste is very dangerous, while the remaining quantity is non-hazardous. Managing this kind of wastes is crucial and involves complexes task as: segregation, collection, transportation, storage, and final disposal [5]. Many studies analyzed the HCW management and in particular the costs and the risk of the HazHCW as [6].

The collection and transportation of these wastes constitute around $80 - 95\%$ of the total cost of solid waste management in general [7]. In the literature, the transportation procedure of the HCW management has been mainly modeled as a Vehicle Routing Problem (VRP) [8]. The VRP is a highly popular combinatorial optimization problem with the objective of minimizing the transportation cost (distance) while provide services to a set of geographical dispersed customers using a fleet of vehicles. Various VRP variants were investigated depending on the studied problem (e.g., VRP with time windows, the Green VRP, the dynamic VRP, the inventory routing problem), each of which have a special characteristics whether in the objective function or in the system constraint [9].

A real case study have discussed in [10]. The authors divided the problem in two stages. The first one is about how to select the best waste treatment and disposal, and in the second stage they modeled the problem as CVRP at the aim to design the best appropriate routes with the lowest distance while collecting HCW from a set of private and public hospitals in Sfax, Tunisia. The authors in [1] modeled the problem of transporting HCW as CVRP [11]. In their work, they used real-life case study and solved the

problem with exact methods. A bi-objective VRP with the aim of minimizing the cost of transportation and the vehicle emissions when collecting and transporting the HCW from a set of Health-care Facilities (HCF) is investigated in [12]. In [13], the authors investigated the problem of HCW stored at pharmacies as a bi-objective stochastic inventory routing problem. The first objective of their model is the minimization of the routing cost and the second objective is to reduce the inventory costs. Three heuristic approaches were proposed to solve the problem and tested over a real-world case study. In [14], two solutions approaches were proposed for the stochastic VRP in order to design the adequate route for the collection of Haz-HCW in the region of Provence Alpes-Côte d'Azur, in France. In [15], the authors studied the problem of collecting and transporting HCW from group of HCF as a periodic VRP. A HCW management review is presented in [16].

In this article, we model the collection and the transportation of the hazardous HCW (Haz-HCW) as CVRP. In this problem, we seek to collect and transport Haz-HCW from a set of dispersed HCFs then to carry them to the disposal center at the minimum transportation distance. To deal with the problem, we adopt an efficient hybrid evolutionary algorithm, previously proposed in [17], and evaluate it against state-of-the-art algorithms using real data sets.

The remainder of the paper is structured as follows. In section 2, the problem description and the mathematical formulation are presented. In Section 3, the proposed approach is explained. The computational study is provided in Section 4. Finally, the concluding remarks are illustrated in Section 5.

2. Problem Description

In this paper, we are interested to design the minimum distance routes to collect and transport Haz-HCW from a set of dispersed HCF with a predefined Haz-HCW quantity in the following way: the collection and the transportation are provided using a fleet of homogeneous vehicles starting and ending at the disposal center. Each HCF is visited once in such a way that each quantity of Haz-HCW is affected to a single vehicle. Moreover, the vehicle capacity should be respected.

This problem is referred to as: Hazardous Health-Care Waste Vehicle Routing Problem (HazHCW-VRP). Formally, it is defined as follows: let N and q_i indicate respectively the number of HCF and the quantity of the Haz-HCW generated in the i^{th} HCF, and the vehicle capacity is denoted by Q, also the starting and ending point of each vehicle is the disposal center (depot). Each HCF has a spatial coordinates x_i and y_i, and d_{ij} represents the travel distance between two HCF $i, j \in N$ where $i \neq j$.

In what follows, we define the required notations : decision variables and parameters. Then, we present the mathematical formulation of the problem [18].

Decision variables

$$x_{ijk} = \begin{cases} 1 & \text{if vehicle } k \text{ travels from the } i^{th} \text{ to the } j^{th} HCF. \\ 0 & \text{otherwise.} \end{cases}$$

$$y_{ik} = \begin{cases} 1 & \text{if vehicle } k \text{ visits the } i^{th} HCF. \\ 0 & \text{otherwise.} \end{cases}$$

Parameters

N	Set of HCF.
V	Number of vehicles where $k \in \{1, .., V\}$.
Q	Capacity of vehicles.
q_i	The HazHCW of the i^{th} HCF, where $i \in \{1, .., N\}$.
d_{ij}	Transportation distance between the i^{th} and the j^{th} HCF, where $i \neq j$.

$$Min \quad Z(x) = \sum_{k \in V} \sum_{i \in N} \sum_{j \in N} d_{ij} x_{ijk} \tag{1}$$

Subject to

$$\sum_{k \in V} y_{ik} = 1 \quad \forall i \in N \tag{2}$$

$$\sum_{i \in N} q_i y_{ik} \leq Q \quad \forall k \in V \tag{3}$$

$$\sum_{k \in V} y_{0k} = V \tag{4}$$

$$\sum_{i \in N} x_{i0k} = 1 \quad \forall k \in V \tag{5}$$

$$\sum_{i \in N} x_{ihk} - \sum_{j \in N} x_{hjk} = 0 \quad \forall h \in N, \quad \forall k \in V, \quad i \neq j \tag{6}$$

$$\sum_{i,j \in S} x_{ijk} \leq \mid S \mid -1 \quad S \subset N, \quad 2 \leq \mid S \mid \leq n+2, \quad \forall k \in V, \quad i \neq j \tag{7}$$

$$\sum_{i \in N} x_{ijk} = y_{jk} \quad \forall j \in N, \quad \forall k \in V, \quad i \neq j \tag{8}$$

$$\sum_{j \in N} x_{ijk} = y_{ik} \quad \forall i \in N, \quad \forall k \in V, \quad i \neq j \tag{9}$$

$$x_{ijk} \in \{0, 1\} \quad \forall i, j \in N, \quad \forall k \in V, \quad i \neq j \tag{10}$$

$$y_{ik} \in \{0, 1\} \quad \forall i \in N, \quad \forall k \in V \tag{11}$$

Formula 1 represents the problem objective function: the minimization of the total travel distance. The set of constraints 2-11 have a conflicting natures and are detailed as follows:
Constraint 2 ensures that each HCF is visited only once by

only one vehicle. Constraint 3 guarantees that the capacity of the vehicle is not exceeded. Constraint 4 states that each vehicle starts from the disposal center. Constraints 5 and 6 implies that all vehicles end at the collection center and that a vehicle k departs from a HCF that has already been visited. Constraint 7 ensures the sub-tour elimination. Constraints 8 and 9 define the relationship between both decision variables whereas constraints 10 and 11 assume the binary decision variables.

3. Solution Approaches

To solve the HazHCW-VRP, three solutions approaches are considered. We present first an evolutionary method combining together genetic operators and Local Search (LS) operators. Then, a generational Genetic Algorithm (GA), an a population based Variable Neighborhood Search (VNS).

3.1. EGALS: An Evolutionary Approach based on GA and LS

We adopt a competitive evolutionary approach proposed in [1] to solve the HazHCW-VRP. EGALS combines the genetic operators (selection and crossover) with LS for a better balance between the exploration and the exploitation, and to prevent convergence to local optimum. While the original version of EGALS was proposed for a dynamic optimization problem, we adapt it here for the static case by looking over the re-initialization step made after every environmental change.

Subsequently, we consider an adapted version of the algorithm that operates as follows: we start by creating the initial population that contains N solutions. Next, we select the fittest chromosomes using the Tournament Selection (TS) technique [19]. The reproduction phase is performed using (1) the Best Cost Route Crossover (BCRC) [20] and (2) Local Search (which is applied here instead of the mutation operator. The obtained offsprings are then evaluated and the best ones take the place of the worst solutions in the current population. EGALS pseudo-code is presented in Algorithm 1.

Algorithm 1 Pseudo-code of EGALS

Randomly generate an initial population P of size N
Evaluation : Evaluate the fitness for each chromosome in P
repeat
  Selection: Select chromosomes from the current population
  Crossover: Apply the BCRC
  Local_Search: Apply the Local_Search
  Evaluation : Evaluate the obtained solutions
  Replacement: Replace the current population with the new chromosomes
until (Stopping criterion)

3.2. Genetic Algorithm

The GA [21] is a meta-heuristic search algorithm inspired from the natural evolution process, and largely used to solve NP-hard problems. The GA uses the process of natural selection where the best solutions are selected for reproduction phases (using the crossover and the mutation operators), in order to produce new offsprings for the next generation. Algorithm 2 contains the main steps of the used GA.

Algorithm 2 Pseudo-code of GA

Randomly generate an initial population P of size N
Evaluation : Evaluate the fitness for each chromosome in P
repeat
  Selection: Select chromosomes from the current population using the Tournament Selection
  Crossover: Apply the BCRC
  Mutation: Apply the Swap mutation
  Evaluation : Evaluate the obtained solutions
  Replacement: Replace the current population with the new population
until (Stopping criterion)

3.3. A Population-based Variable Neighborhood Search

The Variable Neighborhood Search (VNS) is a well known meta-heuristic proposed in [22]. The VNS has demonstrated its effectiveness in solving hard problems, and is widely used to deal with VRPs [23] [18].

In this paper, we consider a population-based VNS which operates as follows. Firstly, a set of solutions is generated randomly. Afterwards, we select the fittest solution for the next step which is the Local Search phases. The goal of this step is to generate a set of new solutions from the selected one. The shaking phase (or perturbation phase) of the used algorithm is performed by the CROSS exchange operator [24]. Next, the swap operator is applied on the new generated chromosomes. Finally, the algorithm replaces the worst solution with the new obtained one. These steps are repeated until the stopping condition is reached. The pseudo code of the used algorithm is presented in the Algorithm 3.

4. Experimental Study

In this section, we first describe the experimental settings, then we present the computational results and analyze them.

4.1. Health-Care Waste Data Sets

The above described algorithms were tested on several realistic instances from the region north of France (pas-de-Calais). This region has 14 Hospital Group Territory (HGT) with a total of 79 hospitals.

Algorithm 3 Pseudo-code of population-based VNS

Randomly generate an initial population P
Evaluation : Evaluate the fitness for each chromosome in P
repeat
  Selection: Select the best chromosome solution
  Insert: Apply the insert method
  Selection: Select the best out of the resulting solutions
  Shaking phase: Apply the shake operator to that solution
  Swap: Apply the swap process for the obtained chromosome
  Replacement: Replace the worst solution with the new obtained one
until (Stopping criterion)

A vehicle can transport a maximum of 1 680 Kg of HCW, spread across red containers (as recommended by WHO), each one of them can hold about 60 Kg. Furthermore, each vehicle can carry 28 red containers [1]. The quantity of generated HazHCW is about 0.78 Kg per bed per day [1].

4.2. Design of the experimentation

All the algorithms were implemented using C++ language, and experiments were run on Windows 10 operating system with $8GB$ memory and Intel Core (TM) $i7$ CPU. The used parameters are presented in the following table.

TABLE 1. PARAMETERS SETTINGS

	Parameters	**Values**
Experimental settings	Number of runs	30-40
	Number of evaluations	1000
GA	Population size	100
	Selection	TS
	Crossover type	BCRC
VNS	Shaking	Cross Exchange
	Local Search	- Insert - Swap

4.3. Comparative Study

The obtained results and the comparison study between the GA, the VNS and the EGALS are summarized in Tables 2-3. In these tables, columns HGT, #HCF , Qte , V represent respectively Hospital Group Territory, the number of HCFs, the monthly quantity of HazHCW generated in each HCF, the number of used vehicles. In addition, bold values indicate the best obtained results.

To better summarize the obtained results of each algorithm, we use the $NormalizedScore$ (NS) [25] performance measure. It is calculated as follows :

$$NS(A,i) = \sum_{i=1}^{N} \frac{\mid (best(i) - Alg(A,i)) \mid}{\mid (best(i) - worst(i)) \mid} \qquad (12)$$

The NS computes the performance of the i^{th} algorithm by normalizing the results obtained by this algorithm over the total number of instances in the range(0,1). The higher the value of NS, the better is the algorithm.

Based on the experimental results reported in Tables 2-3, we can state the following remarks :

- In Table 2, the EGALS produced improved results compared with the GA and the population-based VNS in 8 instances among 14 HGT. In the remaining instances, the population-based VNS outperforms the considered algorithms in 4 instances while the GA came the latest algorithm and provide 2 better results.
- According to the NS performance measure, it is clear that the EGALS outperforms the other algorithms and give a competitive performance results (0.647).
- In Table 3, we summarize the total routing distance providing by each algorithm for all HCFs. According to the obtained results listed in this table, we can state that the EGALS was capable to find the minimum travel distance despite the higher number of HCFs.

Through the computational experiment, the EGALS has very good performance compared with the competing algorithms.

5. Conclusion and Perspectives

In this research, we modeled the problem of collecting and transporting HazHCW from a set of HCFs as CVRP with the objective of finding the optimal routes with the minimum travel distance while respecting the CVRP constraints. To deal with the problem, a hybrid evolutionary algorithm is adopted and tested against state of the art algorithms using real data sets. The experimentation study show that the presented method outperform others algorithms in most cases. As a perspective, we seek to take into account several constraints specific to the health-care waste management.

References

[1] W. Hachicha, M. Mellouli, M. Khemakhem, and H. Chabchoub, "Routing system for infectious healthcare-waste transportation in tunisia: A case study," *Environmental engineering and management journal*, vol. Vol.13,, pp. pp. 21–29, 2014.

[2] M. Tsakona, E. Anagnostopoulou, and E. Gidarakos, "Hospital waste management and toxicity evaluation: A case study," *Waste Management*, vol. 27, no. 7, pp. 912 – 920, 2007.

[3] Y.-C. Jang, C. Lee, O.-S. Yoon, and H. Kim, "Medical waste management in korea," *Journal of Environmental Management*, vol. 80, no. 2, pp. 107 – 115, 2006.

[4] D. Makajic-Nikolic, N. Petrovic, A. Belic, M. Rokvic, J. A. Radakovic, and V. Tubic, "The fault tree analysis of infectious medical waste management," *Journal of Cleaner Production*, vol. 113, pp. 365 – 373, 2016.

[5] T. S. Aung, S. Luan, and Q. Xu, "Application of multi-criteria-decision approach for the analysis of medical waste management systems in myanmar," *Journal of Cleaner Production*, vol. 222, pp. 733 – 745, 2019.

TABLE 2. THE TOTAL TRANSPORTATION DISTANCE FOR EACH HCF

HGT	#HCF	Qte	V	GA	VNS	EGALS
HGT-1	2	648.18	1	**1.45248e+06**	**1.45248e+06**	**1.45248e+06**
HGT-2	2	714.48	1	**1.33411e+06**	**1.33411e+06**	**1.33411e+06**
HGT-3	3	1237.86	2	2.97605e+06	2.41463e+06	**1.49942e+06**
HGT-4	3	508.56	2	2.64513e+06	2.62327e+06	**1.33518e+06**
HGT-5	3	1127.10	2	1.48867e+06	2.86464e+06	**1.46235e+06**
HGT-6	4	1475.76	3	2.91906e+06	2.91905e+06	**2.86712e+06**
HGT-7	4	757.38	3	4.42336e+06	4.41882e+06	**3.51322e+06**
HGT-8	4	1039.74	3	1.52024e+06	2.99441e+06	**1.32065e+06**
HGT-9	5	1965.60	3	3.34957e+06	3.66178e+06	**2.68563e+06**
HGT-10	6	762.84	3	4.05423e+06	**4.04859e+06**	4.05370e+06
HGT-11	10	2445.05	5	**7.39625e+06**	7.41324e+06	7.43187e+06
HGT-12	10	3132.48	5	7.13989e+06	**7.10001e+06**	7.18362e+06
HGT-13	11	1999.92	5	7.08010e+06	**7.04973e+06**	7.09704e+06
HGT-14	12	3967.94	5	7.30312e+06	**7.30083e+06**	7.35187e+06
Normalized Score				0.501	0.494	**0.649**

TABLE 3. THE TOTAL TRANSPORTATION DISTANCE FOR ALL HCFS

#HCF	Qte	V	GA	VNS	EGALS
79	7 409.61	40	1.18278e+11	1.18277e+11	**1.06145e+11**

[6] V. Ferreira and M. Teixeira, "Healthcare waste management practices and risk perceptions: Findings from hospitals in the algarve region, portugal," *Waste management*, vol. 30, 12 2010.

[7] A. Z. Alagöz and G. Kocasoy, "Improvement and modification of the routing system for the health-care waste collection and transportation in İstanbul," *Waste Management*, vol. 28, no. 8, pp. 1461 – 1471, 2008.

[8] G. Ghiani, D. Laganà, E. Manni, R. Musmanno, and D. Vigo, "Operations research in solid waste management: A survey of strategic and tactical issues," *Computers & Operations Research*, vol. 44, pp. 22 – 32, 2014.

[9] P. Toth and D. Vigo, *Vehicle routing: problems, methods, and applications*. SIAM, 2014.

[10] D. Baati, M. Mellouli, and W. Hachicha, "Designing a new infectious healthcare-waste management system in sfax governorate, tunisia," 2014, pp. 350–355.

[11] A. N. Letchford and J.-J. Salazar-González, "The capacitated vehicle routing problem: Stronger bounds in pseudo-polynomial time," *European Journal of Operational Research*, vol. 272, no. 1, pp. 24 – 31, 2019.

[12] H. Alshraideh and H. Abu Qdais, "Stochastic modeling and optimization of medical waste collection in northern jordan," *Journal of Material Cycles and Waste Management*, vol. 19, no. 2, pp. 743–753, 2017.

[13] P. C. Nolz, N. Absi, and D. Feillet, "A bi-objective inventory routing problem for sustainable waste management under uncertainty," *Journal of Multi-Criteria Decision Analysis*, vol. 21, no. 5-6, pp. 299–314, 2014.

[14] P. Nolz, N. Absi, and D. Feillet, "An inventory routing problem for infectious medical waste collection," *Networks*, vol. 63, pp. 82–95, 2014.

[15] L.-H. Shih and H.-C. Chang, "A routing and scheduling system for infectious waste collection," *Environmental Modeling & Assessment*, vol. 6, pp. 261–269, 2001.

[16] V. Thakur and R. Anbanandam, "Healthcare waste management research: A structured analysis and review (2005-2014)," *The journal of the International Solid Wastes and Public Cleansing Association, ISWA*, vol. 33, 2015.

[17] N. Ouertani, H. B. Ramdhan, S. Krichen, I. Nouaouri, and H. Allaoui, "A new evolutionary method to deal with the dynamic vehicle routing problem," in *2018 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD)*. IEEE, 2018, pp. 1–5.

[18] M. Amous, S. Toumi, B. Jarboui, and M. Eddaly, "A variable neighborhood search algorithm for the capacitated vehicle routing problem," *Electronic Notes in Discrete Mathematics*, vol. 58, pp. 231 – 238, 2017, 4th International Conference on Variable Neighborhood Search.

[19] P. Chand and J. Mohanty, "A multi-objective vehicle routing problem using dominant rank method," *International Journal of Computer Application*, pp. 29–34, 2013.

[20] B. Ombuki, B. J. Ross, and F. Hanshar, "Multi-objective genetic algorithms for vehicle routing problem with time windows," *Applied Intelligence*, vol. 24, no. 1, pp. 17–30, 2006.

[21] M. Mitchell, *An introduction to genetic algorithms*. MIT press, 1998.

[22] N. Mladenović and P. Hansen, "Variable neighborhood search," *Computers & operations research*, vol. 24, no. 11, pp. 1097–1100, 1997.

[23] G. A. Bula, C. Prodhon, F. A. Gonzalez, H. M. Afsar, and N. Velasco, "Variable neighborhood search to solve the vehicle routing problem for hazardous materials transportation," *Journal of Hazardous Materials*, vol. 324, pp. 472 – 480, 2017.

[24] É. Taillard, P. Badeau, M. Gendreau, F. Guertin, and J.-Y. Potvin, "A tabu search heuristic for the vehicle routing problem with soft time windows," *Transportation science*, vol. 31, no. 2, pp. 170–186, 1997.

[25] T. T. Nguyen, S. Yang, and J. Branke, "Evolutionary dynamic optimization: A survey of the state of the art," *Swarm and Evolutionary Computation*, vol. 6, pp. 1 – 24, 2012.

A Multi-agent architecture for modeling organizational planning against terrorist attacks in urban areas

Oussama Kebir*, Issam Nouaouri†, Mouna Belhaj‡, Lamjed Bensaid* and Kamel Akrout§
*Institut Supérieur de Gestion de Tunis
Strategies for Modelling and ARtificial inTelligence Laboratory
University of Tunis
Tunisia
†Laboratoire de Génie Informatique et d'Automatique de l'Artois
University of Artois
France
‡Institut supérieur des études technologiques de Bèja
Strategies for Modelling and ARtificial inTelligence Laboratory
Tunisia
§Tunisian Presidency
Tunisia

***Abstract*—Nowadays the world is suffering from the emergence of a new concept of war, it is the asymmetric warfare created by the terrorists' new combat doctrine. As the plans to face classic enemies have become unusual against terrorism, this calls for innovative concepts and technologies to support the units and to improve the capability of leaders and structure their choices. In this paper, we propose a multi agent architecture for action planning against terrorist attacks. It is characterized by rapid decisive responses and methodical steps to handle the situation, and by the flexibility to adapt a contingency scenario. We aim to create a multi-agent model that describes the relation between actors during the terrorist attack in order to find the best possible units distribution to neutralize the enemy.**

1. Introduction

Counter terrorism planning is based on real time events, in such way that actual conditions and facts replace assumptions and projections following the attack's chronological progress. To refine the fast selection of a course of action, peacetime planning facilitates anticipating potential crises. The decision-making process due to the selection of the organizational plan may be greatly accelerated if the terrorist attack conditions overlay with the assumptions in a previously developed plan. The plan hypothesis should submit to some rules such as the types and forms of operations, the forms of maneuver as well as the tactical mission tasks and the organization of available forces. Then, the activities must be arranged in time, space and purpose. The Risks that the Commander is willing to take in the allocation of resources should be taken in charge [1].

There are many forms of plans with difference in scope, complexity, and length of planning perspective. Strategic plans cover the general conduct of a war. Operational plans take in charge the organization of military operations aimed to accomplish a strategic or operational objective with a predefined time and space. The employment of units in operations-including the ordered arrangement and maneuver of units in relation to each other and to the enemy- are the role of the tactical plans. The latter try to push forces to use their full potential. All these plans are merged to handle a coherent warfare-conducting process. Not only bad planning could lead to a military failure but also it could induce a bad execution or a major blunder, general lack of skill or ability and the lack of military technology. Modern war is becoming more and more complex and thus requires high technologies.

Therefore, many nations are racing to develop new warfare technology to create more sophisticated weapons, train units by simulating a virtual environment of real situations and even help the Commander to make a better decision during operations.

In fact, with the development costs of major weapon systems, simulation that studies combats is the solution for a reduced army budget. It is considered to be one of the key factors in achieving military goals with the lowest costs by saving material losses and reducing the need for costly prototypes and hazardous testing due to the paucity of training ranges. Simulation helps to resolve time-induced tensions between political and military imperatives in the conduct of warfare actions. Therefore, simulation helps to achieve political-military congruence by saving time in employing forces and military power with an appreciation of their abilities to contribute soldiers.

Congruence will be solidified by helping those in authority positions compile enough information to make an informed decision about combat changing situations. In

fact, all the previous reasons added to the huge number of stakeholders and the necessity of taking decisions in a short period of time with an evolving situation push us to resort to the multi agent simulation.

Modern war simulation is very different from traditional war simulation. This difference is a problem of a concept of war. On the one hand, modern war is generally based on the asymmetric warfare described as strength disproportion between the opponents at the outset, and the difference in essence between their assets and liabilities [2]. On the other hand, traditional war is a Blitzkrieg, interstate war characterized by its proportional clarity.

The approaches of asymmetric warfare generally seek a major modeling impact, it often employs innovative, non-traditional tactics, weapons, or technology, and it can affect all levels of warfare (strategy, operations and tactics) across the military operations spectrum.

There are three kinds of asymmetric warfare: guerrilla warfare, independence warfare and terrorism. Guerrilla is composed of a certain number of militant groups, acting as a military unit, and thus takes the name of paramilitary forces or combatants. Guerrilla tactics include ambush, sabotage as well as hit and run technique. They often avoid open battle and aim to harass and weaken the enemy's strength, herewith forcing the enemy to withdraw from the contest. Combatants are formed like soldiers, directed by a commander; they wear a distinctive uniform to show that they are guerrillas. The guerrillas' goal is usually to exert pressure on the government in order to control or dominate territory and population, and they target only military units.

Independence war is a civil war that struggles against colonial domination or foreign occupation. It is defined as "any armed conflict that involves military action internal to the metropole, the active participation of the national government, and effective resistance by both sides" [3]. With reference to the CIA definition, the term "Terrorism" means premeditated politically motivated violence perpetrated against noncombatant targets by subnational groups or clandestine agents. While terrorism across the globe has markedly increased in recent decades, it continues to be relatively rarely simulated, and even the organization's counter terrorism plans are under testing. The efforts of the Tunisian military in some counter terrorism missions were marked by a failure to bring significant results.

The failure to recognize and to understand the enemy, to provide an adequate plan and resources lead to a mismatch between forces, capabilities, missions, and goals. Therefore, we aim to create a coherent organizational intervention plan that abides by the strategic, operational and tactical laws. To deal with the huge number of components and information about the multidimensional needs present in this model, we suggest a multi agent model that outlines the role of every actor. This model will support decision makers with an integrated tool.

The paper is structured as follows: The next section addresses various simulations associated with counter insurgency situations and provides a brief description of simulations context, tools, and the procedures used to model and schedule operations in a different point of view. In the following two sections, we present our Model of Organization Plan Against Terrorist Attack in Urban Area called MOPATAUA. Subsequently, we show the difference between MOPATAUA and the model that inspired us in this work. We finish with a discussion and a conclusion.

2. Background

In the past, Armed Conflicts (AC) modeling was wide ranging and involved heavily armed state actors and strategic weapons. This motivated the development of large-scale force on force models. Although these types of conflicts are always feasible and require known modeling, current armed conflicts like terrorism are smaller in scale and involve non state actors. Studying the effect of such asymmetric warfare like terrorism attacks and analyzing response steps are of great importance to the government and first responders.

Probabilistic estimations could be used to achieve acceptable expectations and outcomes, while examining patterns of violence in terrorist events, classifying attacks and their levels of urgency, and identifying a common pattern regarding the size distribution and timing of such events. To deal with the large number of parameters, many models based on analytical models capture only some key aspects of armed conflict; a set of ordinary differential equations is proposed by Lanchester to describe the dynamics of force-on-force engagements [4]. These equations have been developed and widely used over the past 60 years by the U.S. military. Even Lanchester's models provide a convincing physical description of military actions; they have been criticized in the literature mainly because of the discrepancy between the decision made in armed conflict and the decision made by simulation and the difficulty of validating them [5], [6].

Some models dealing with terrorist attacks are based on the notion of coalescence and fragmentation of insurgents or terrorist organizations. To model counter-terrorism operations, we have two major challenges to focus on: detection and protection. To detect and intercept terror plots, it is critical to create an effective intelligence collection and improvement analysis of terrorists' behavior. Advances in data mining and the application of specially tailored techniques will result in better resources allocation and improve the processes during the analysis phase since, generally, state-controlled forces are significantly larger and better equipped and trained than non-state actors. The latter hide within the civilian population and use simple, yet lethal weapons to avoid their detection and targeting by state forces. So, the civilians become a key component in irregular warfare modeling, which is not the case in legacy (AC) models as the civilians are a source of information to the state forces and sometimes to terrorists, and a target of terrorist attacks. Protection models should be used to support needs in homeland security on topics such as evolving national security infrastructures, border screening, and controlling monetary circuits [7].

An important model of civil violence [8] simulates two scenarios. The first is about rebellion against central authority while the second treats ethnic violence between two groups. This model is widely cited, it uses simple reactive agents and relevant variables to model the agents' behaviour. To defeat insurgency, their recruitment circuits should be broken by limiting the increment of population support to insurgents, which will exhaust their recruitment sources. These hypotheses are the results of Iruba agent - based guerrilla warfare model which also shows that the high mobility and uses of guerrillas' tactic against government forces could make a huge difference [9].

The model developed by Johnson and Madin [10] is a simple population-growth model that studies the dynamics of an insurgent population. In order to focus on the failure and success conditions of an insurgency, they valorize insurgent population size and their recruitment rates and the population carrying capacity [11].

Recently, inspired from the Deitchman's classic guerrilla model [12] and the Lanchester model [4], the work of Kress and Szechtman [13] focuses on the dynamics of insurgencies.

Some models were developed like an attrition reinforcement models demonstrating that some insurgencies cannot totally be eradicated by force but, at best, can be contained at a stable level, which could only be solved by political circumstances [6].

Cil proposes a Multi Agent Based Simulation Model of Military Unit Combat (MABSIM), based on the hierarchical structure popularized by Wooldridge and Jennings [14], supporting commanders of small units to make decisions in stressful situations with complex adaptive systems. Due to the difficulty of validating these systems [15], Cil uses a multi-agent architecture. According to his point of view, hybrid multi-agent architecture will be a better solution to bridge the breach between cognitive and reactive agent systems: even if cognitive systems are capable of reasoning about actions for a large number of agents, they become quickly unanalysable, where reactive systems subsist better with scale but it is difficult to understand and validate their unreasonable behavior [17] [18].

The Multi-agent based asymmetric combat simulation architecture ACOMSIM [16] is an extension of the MABSIM model that seeks to answer the needs of future multidimensional warfare. It is a simulation of small units combats in asymmetric warfare.

ACOMSIM is divided into two layers, the first of which focuses on decision-making, while the second layer examines the links between different agents and the environment.

The first layer of ACOMSIM includes seven cognitive agents and six associated databases. Those agents and databases are presented in table 1.

The second layer uses MANA (Map Aware Non-Uniform Automata), which evaluates warfare as a complex adaptive system to assess high-level plans and provide a response to the commander. MANA was created as a model for exploring scenarios to solve a variety of problems.

Cognitive agent	**Database**
Mission analysis agent	Intelligence database
Mission time scheduling agent	Environment database
Enemy situation analyzing agent	Terrain database
Own situation analyzing agent	Enemy tactics, techniques and procedure database
Terrain analysis agent	Tactics, techniques and procedures database
Logistic agent	Logistic Database
Action Generating Agent	

TABLE 1. AGENTS AND DATABASES IN ACOMSIM MODEL

Another example illustrates the economics of counter insurgency armed revolts of Berman [19] represents another way of explaining asymmetric warfare where civilian demonstrations and social unrest turn into an armed conflict. Unlike classical blitzkrieg war models, the outcome of the asymmetric warfare is independent of the initial force sizes; it only depends on the number of the population supporting each side and on the combat effectiveness of the government forces and the rebels. This model highlights conditions for a stalemate and underscores the critical effect of foreign intervention as a factor of rocking the power.

The big challenge is to combine attrition models, political, social and behavioural science, and economic theory into a unified model. The dissemination of ideas that show how people change their attitudes and understand the effect of private and public preferences on their governments' decisions is of great interest to social modelling. Social networks are the road map to highlight connectivity within the population, its dynamics and its impact on the actions of state and non-state actors. For those reasons, models that describe social and behavioural components should be integrated into social networks. Makowsky and Rubin [20] put forward an agent-based model that tests the impact of government and social network technology on the falsification of preferences and their racial effect in changing the conduction way of revolutions because the topology of these self-organizing networks' changes over time.

Finally, Pechenkina and Bennett's [21] agent-based model offers a hybrid combination of minimizing collateral damage and ensuring the involvement of military counter-insurgency strategy. They assume that we have a better chance of winning an insurgency if we use a balance of military action that targets insurgents without injuring civilians. Figure 1 summarises the various works carried out in the field of asymmetric warfare simulation.

Our model is based on the ACOMSIM model to handle the three warfare levels (strategic, operational and tactic). We choose this model [16] as a starting point for the reason that it allows a full range of military strategies from a tactic point of view.

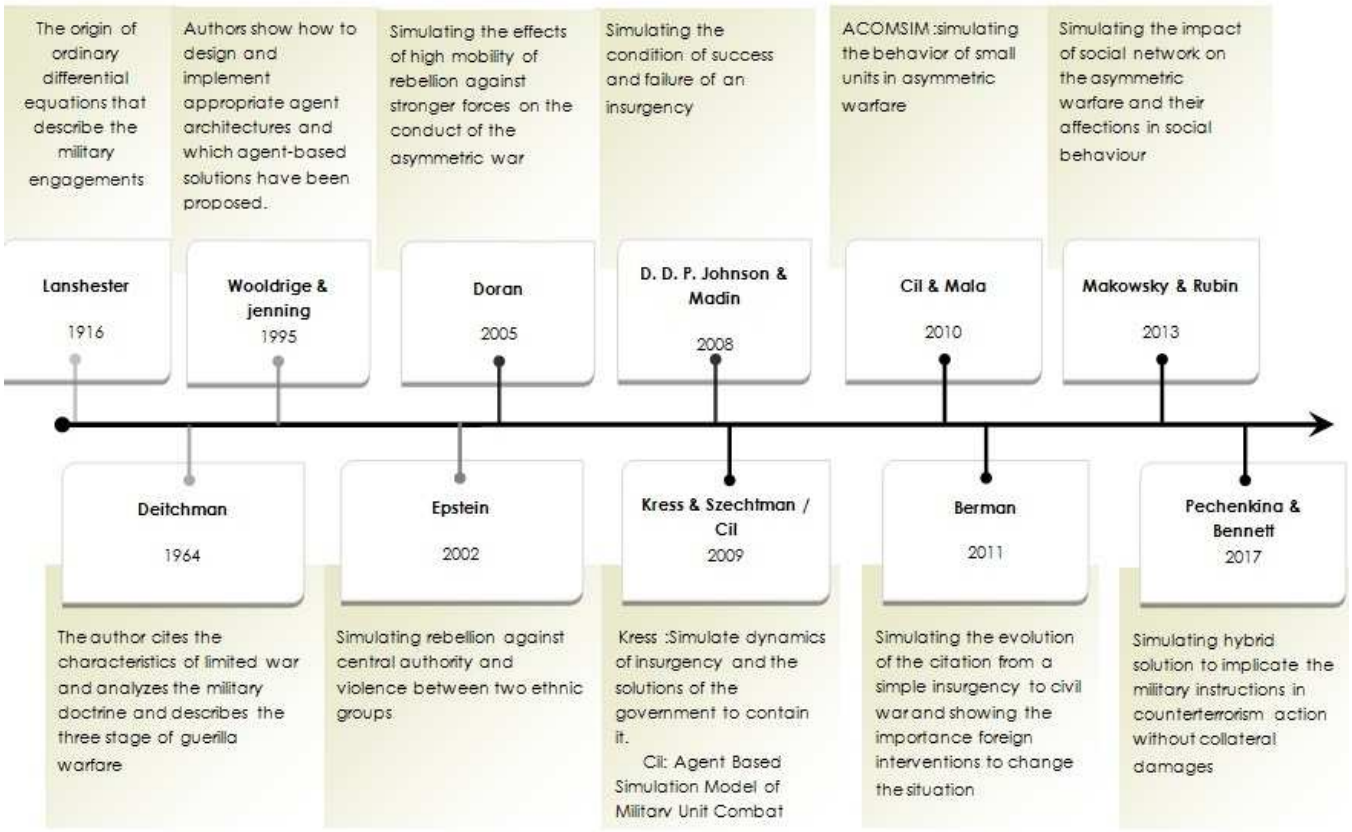


Figure 1. Related works on asymmetric warfare simulation

3. Description of the MOPATAUA model

3.1. Model architecture

Counter terrorism missions designed to defeat an armed group should include military, paramilitary, political, psychological, and civic actions. Mostly, they are financeored by a non-state opponent aimed at the overthrow or territorial change of an incumbent state. In this paper, we consider that our model studies two core aspects types.

The first one analyzes the dynamic of behavioral decisions as presented in figure 2, it takes in charge the decisional and organizational sides of the organization plan and the second layer interested in the management of human resources and the duties of the different actors in the military terms.As time has a major impact during terrorist attacks, it has become essential to make the right choice that one may solve problems and reduce losses. These choices must not be arbitrary or subject to the whims of the commander, but must follow a pre-established organizational plan. The relation between the different actors is hierarchical with a certain degree of freedom in the execution. Indeed, leaders need to have more authority and flexibility in managing their human resources in such a manner that permits the creation of a better work environment without missing the predefined goals and overflowing the limits of the task to be done accounting for the use of that authority.

The multi agent system is one of the most appropriate systems that manage the ordering procedures according to the situation. It is appropriate with the interactions between different agents with conflicting goals and proprietary information. In the first layer, we model the relationships between the different actors without focusing on human resource management. Therefore, a well-designed small-scale program would be more beneficial to the government's effectiveness in managing terrorist attacks than a poorly designed program.

The second layer represents the operation tactical phase which deals with how units are employed during combat. Given that there are no set of tactics checklists for asymmetric warfare since each application is unique and has never used a particular options combination before every situation, we might have to incorporate new and novel options to counter asymmetry [22].

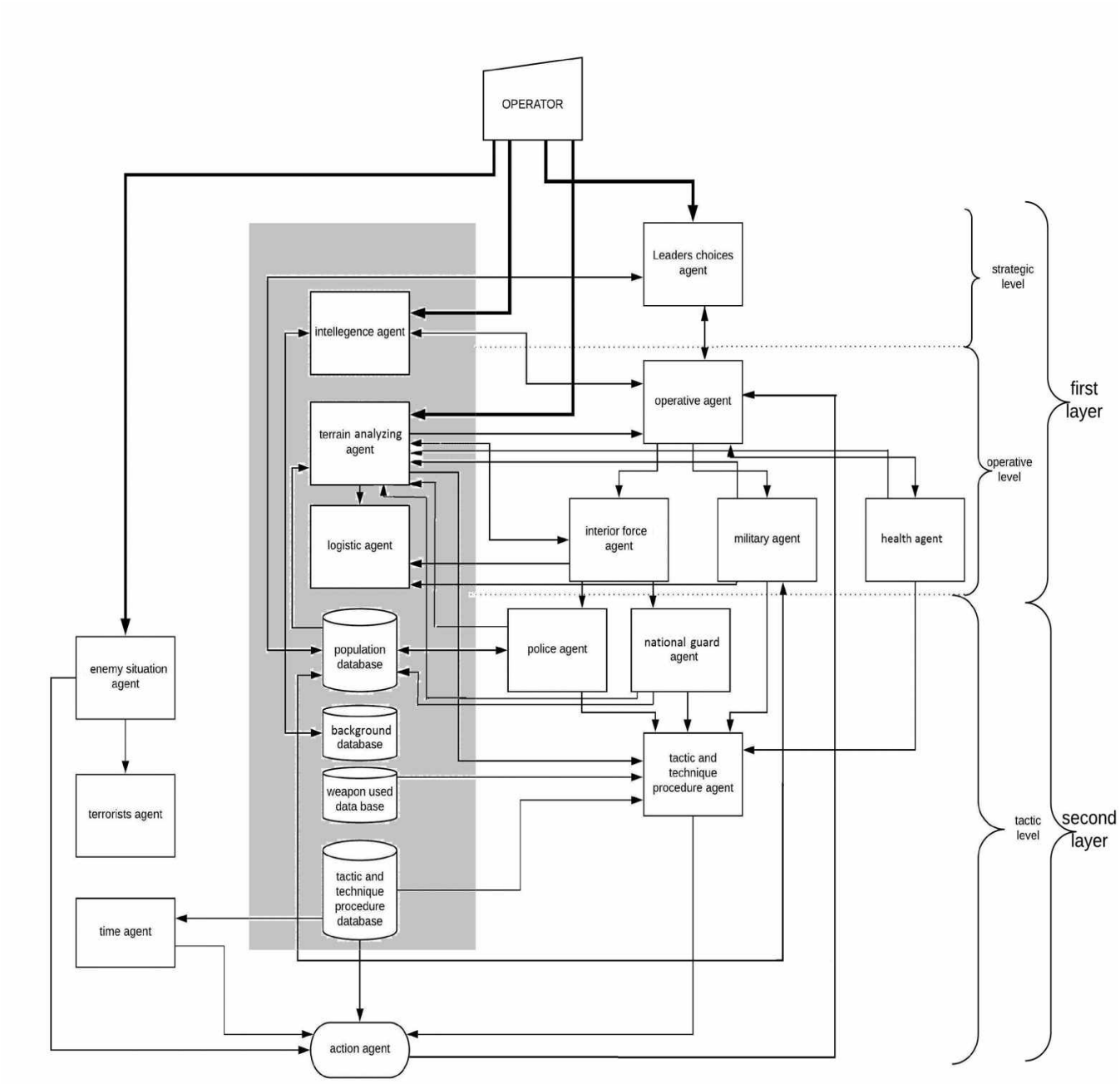


Figure 2. MOPATAUA Model

3.2. Model mechanism

Indeed, we model the organization plan to counter a terrorist attack. The leader choice agent will analyze, reffering to the political situation, the gravity of the attack depending on the time of the attack, its location and the population' size. At the strategic level, those parameters have more effect than other ones. The resulting decisions can support confidence in the functioning and authority of the state.

In our case, the leader choice agent provides the time proposed for this mission to the operative agent. Concurrently, a predefined plan will be executed at the tactical level; the attack location is the only input parameter that runs this tactical process. This pre-organized anti-terrorist reaction could be useful anywhere without strategic or operational directives.

The operative agent classifies the level of emergency of

the terrorist attack in a predefined scale. Results obtained from the intelligence agent and the terrain analyzing agent are matched by the operative agent to give birth to a new classification of the level of threat on a conventional scale that preserves the confidentiality of information.

The military, health and interior force agents receive the level of emergency. Then, each one of them choose the optimal number of interveners for the mission. The military agent support tactic and technique procedure agent with soldiers which reinforces the number sended in predefined plan. Moreover, the same process will be carried out by the health agent to contribute to the exact number of ambulances needed.

In urban areas, anti-terrorist actions are governed by a territorial distribution of authorities between the different governmental organisations. In order to choose the principal agent to neutralize terrorists, the interior force agent is based on territorial distribution from the terrain analyzing agent. Besides, it points out the effectif needed from each base concerned by the intervention. Those data will be sent to the police agent and the National Guard agent, to create a tempo-dispatching of intervenient based on the number of soldiers available in the bases.

The tactic and technique procedure agent receives the results from the police, National Guard, health and military agents. Then, it handles the actors' dispatching in the field. To obtain the optimal results and limit the terrorists' ability, this agent makes the tactic choices.

The action agent takes in charge the tactic engagements based on tactic and techniques database; that one may neutralize the terrorists with minimum losses, the action agent combines the terrorists deployment plan provided by the enemy situation agent and the soldiers' geographical positioning from the technique and tactic procedure agent.

4. Agents and Databases description

We mentioned that MOPATAUA is composed by fifteen agents and four databases. Throughout this section, we discuss the roles of each agent and database.

4.1. Agents description

4.1.1. Intelligence agent. Intelligence agent estimates the number of victims based on:

- Number of victims in the real time
- Number of terrorists
- Population in the attack' location
- Proposed governmental forces' support

This agent classifies terrorists by referring to the following information:

- Weapons used in the attack (operator)
- Backgrounds of terrorists (background database)

4.1.2. Enemy situation agent. The enemy situation agent takes in charge the tactical phase of the enemy; these tactic reactions depend on the terrorists' type and the way of access to the battle field (terrain analyzing agent).

4.1.3. Leaders choices agent. The strategic level is presented by Leaders choices agent: decisions of the political leaders are based on:

- Period: The period of the attack has an important effect on the number of civilians in the targeted place
- Location: Strategic decisions are related with the location of the attack

This agent imposes the time dedicated for the mission based on the previous factors.

4.1.4. Terrain analyzing agent. The Terrain analysis agent receives location from the operator, then it generates information about:

- The access ways to the attack location.
- Key points in the field, Avenues of approach, obstacles.

The system of location UTM WGS84 used by this agent indicates the position following a coordinate system.

4.1.5. Logistic agent. The aim of this agent is to prove the units' logistic level and support them with the necessary logistic items at the right time and place. Therefore, he has to make decisions about the necessary logistic level based on:

- Duration of the mission
- Number of soldiers needed for the mission

After that, it has to choose the right course of actions for the execution using information about location from the terrain analyzing agent. Outputs are sent to the military and interior forces agents. The logistic agent is not always used because generally missions against terrorist attacks are limited in time, so they do not need a lot of logistic sources.

4.1.6. Operative agent. The aim of this agent is to identify a categorization of the level of emergency creating a common language to produce a system that allows a quick reaction between all intervening actors. The classification is based on:

- Classification of terrorists
- Estimation of the evolution of the number of victims
- Imposed time for the mission
- Location

This classification offers the military, health, and interior forces agent an estimate of the threat level without disregarding the confidentiality of the mission.

4.1.7. Interior Force agent. The interior force agent plays the role of the Ministry of the Interior according to the geographic location of the attack. Based on the attack classification, this agent will specify the

- Necessary number of police and national guard men
- Bases concerned with the intervention

4.1.8. Military agent. In urban areas, military intervention acts are well determined; soldiers usually isolate the zone by encircling the area in such a way that interior forces focus only on neutralizing the enemies. The military agent obtains the classification of the attack from the operative agent and then identifies the number of soldiers required from each base. This agent gives the best distribution and planning of soldiers during the mission in coordination with the terrain analyzing agent and the population database.

4.1.9. Health agent. In case of a terrorist attack, hospitals must be able to treat the massive influx of victims. Thus, optimization of different medical resources is fundamental to save human lives. This agent works on dispatching of available ambulances over time and the hospitals' capacities. The health agent is a broad subject of research, so we are limited to the management of ambulances and their distribution in the area of the terrorist act.

4.1.10. Police agent. The objective of this agent is to ensure that the sufficient police forces requested by the Interior force agent is available in the bases involved in the intervention in accordance with the terrain analyzing agent. The number of available police officers will be extracted from the population database.

4.1.11. National Guard agent. This agent has the reel number of National Guard forces available in bases; his work consists in providing the number needed by the interior force agent.

4.1.12. Time agent. The time agent is responsible for the time organization; it assigns the time needed for every step at the model' tactic level. This agent replicates the Mission time scheduling agent that was used in the ACOMSIM model.

4.1.13. Action agent. As the last agent and summarizing agent in the first layer, the Action agent completes the plan and formats in the input form of the simulation. The final output should cover optimal decisions about the

- Planned execution (tactic plan)
- Concept of operation
- Loss probability

4.1.14. Tactic and technique procedure agent. This agent is concerned with employing forces in the operations theater to obtain an advantage over the enemy that will bend him to your will in the entire operation. This is achieved by attaining strategic goals through the design, organization, and conduct of different actors. This agent controls the dispatching of forces on the field based on inputs from the terrain analyzing, military, National Guard and health agents.

4.1.15. Terrorists agent. This agent works on the notion of action and reaction in terrorist behaviour and models the way terrorists do things in combat by instructions of the the enemy situation agent.

4.2. Databases description

4.2.1. Population database. This database provides:

- Estimated number of population presented in the place of attack.
- Key places that usually have massive crowd of people (e.g. stations, big commercial centers, museums, etc.)

The population database answers the questions of how many people are in that place at that time. Also, it gives the police and national guard agents the number of soldiers available in the bases and police offices.

4.2.2. Weapon used database. This database gives information about the useful characteristics of the enemy's weapons including:

- Type
- Rang
- The energy of the explosives

That information has major impact on the tactical phase because it allows us to differentiate suspescious areas and the range targeted by terrorists.

4.2.3. Background database. The purpose of this database revolves around the perception of the enemy's reaction, the entries will be the type of terrorist group to which they belong, their religious attributes from a database and the proposed purpose of the attack given by the intelligence agent. This database is the starting point for the classification of terrorists.

4.2.4. Tactic and technique procedure database. The aim of this database is to generate military rules of combat based on the field commander's tactics.

5. Differences between the ACOMSIM and MOPATAUA

The frame of this part highlights the innovation of MOPATAUA model and focuses on some different concepts of the two models. The ACOMSIM is a model and simulation about the asymmetric warfare; by analyzing the model proposed by Cil and Mala, as we mentioned before they focus on guerrilla warfare. In MOPATAUA, our target is to model terrorism. The figure 3 shows the different warfare types.

5.1. Civilian and collateral damage

Civilians are generally present in the terrorist attacks; they could be injured victims or hostages and even terrorists' Human shields. Citizens are helpful when they give armed forces information and collaborate with them. However, they become obstructions if they betray their countries by supporting terrorists. In our model, we try to give meaning of civilians' presence in the battle field creating an estimation of victims present in the attack.

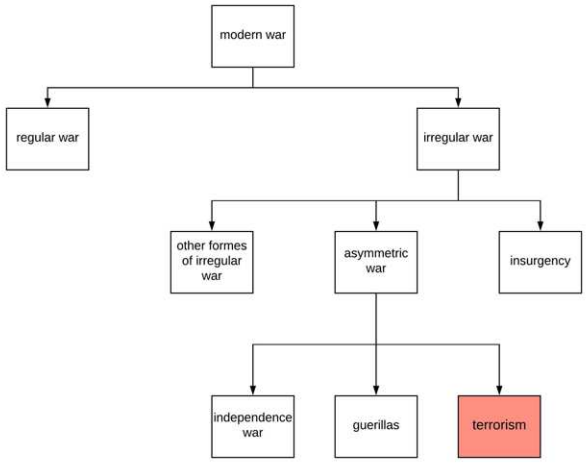


Figure 3. types of warfare

		ACOMSIM	MOPATAUA
Type of the asymmetric warfare		Guerilla warfare	Terrorism
Civilian and collateral damage		Not modeled	modeled
Classification of terrorists and attacks levels		no	yes
Level of modern war modeled	strategic	no	yes
	operative	no	yes
	tactic	yes	yes

TABLE 2. DIFFERENCE BETWEEN THE TWO MODELS

5.2. Classification of terrorists and attacks levels

The ACOMSIM is about planned combat between two armed groups, it does not take in charge the specificity of terrorists and their backgrounds, and if they were trained somewhere or belonged to another terrorist group in the past. This information is very important because terrorists do not behave like a traditional army. Therefore, we may figure out a terrorists' classification to reveal their threat level.

5.3. Level of warfare modeled

Warfare consists of three levels: the strategic, operational and tactic levels. The strategic level focuses on defining and supporting national policy and relates directly to the war outcome or other conflict as a whole. The war's operational level lies between the strategic and tactical levels. The lowest tactical level is that where individual battles and engagements are fought. ACOMSIM has organized the mission only from one point of view, treating the tactical level. It simulates the management of the combat from the local level, assigned to tactic units, where activities and engagements are planned and executed to achieve military goals whereas MOPATAUA model treats the three warfare levels.The difference between the two models is showing in table 2.

6. Conclusion and future work

The MOPATAUA model makes leaders at all levels capable of understanding higher mission intent while tailoring and utilizing the guidelines in order to begin their own planning and orders development process. The goal of designing MOPATAUA model lies in being able to direct the planning of focused actions and tasks to produce the right combination of effects in terms of time, field, and purpose.

We enumerate the contributions of other researches made in the same field. Even if we could find theoretical solutions in the literature, it is still hard to approve the effectiveness of their counter terrorism process. This data may not be the most representative model and simulation of counter terrorism worldwide, but it is an unclassified document.

The actors' relations are described by the proposed bi-layers MOPATAUA model of interactions analyses and tasks affectation. The architecture of the interveners' connexion as modeled organization or person arises from the plan's hierarchical vertical interactions structure. MOPATAUA architecture provides a simplified reasoning architecture that conforms to the basic concepts of military planning. The architecture facilitates synchronization between commanders and their subordinate units to link the strategic level concept of operations to their operational design structure. We have used the ACOMSIM architecture as a baseline to study the relationship between actors at the tactical level and we have shown that MOPATAUA provides a complementary tool for the different levels of courses-of-action.

In the future, we will work on other specificities of irregular conflicts. furthermore, we will simulate MOPATAUA model based on building models, optimizations, data fusion and machine-learning models. We will carry out qualitative and quantitative analyses for a better understanding of asymmetric conflicts.

References

[1] D. of the Army, *Army Planning and Orders Production*, H. D. O. T. ARMY, Ed. U.S ARMY, 2005, vol. FM 101-5, no. January.

[2] D. Galula, *Counterinsurgency Warfare: Theory and Practice*, 1964.

[3] M. Small and J. D. Singer, *Resort to arms: International and civil wars, 1816-1980*. Sage Publications, Inc, 1982.

[4] F. W. Lanchester, *Aircraft in warfare: The dawn of the fourth arm*. Constable limited, 1916.

[5] J. W. R. Lepingwell, "The laws of combat? Lanchester reexamined," *International Security*, pp. 89–134, 1987.

[6] J. Scheffran, M. Brzoska, J. Kominek, P. Link, and J. Schilling, "Climate change and violent conflict," *Science(Washington)*, vol. 336, no. 6083, pp. 869–871, 2012.

[7] M. Kress, "Modeling armed conflicts," *Science*, vol. 336, no. 6083, pp. 865–869, 2012.

[8] J. M. Epstein, "Modeling civil violence: An agent-based computational approach," *Proceedings of the National Academy of Sciences*, vol. 99, no. Supplement 3, pp. 7243–7250, 2002. [Online]. Available: http://www.pnas.org/cgi/doi/10.1073/pnas.092080199

[9] J. Doran, "Iruba : An Agent-Based Model of the Guerrilla War Process," pp. 198–205, 2005.

[10] D. D. P. Johnson and J. S. Madin, "Population models and counterinsurgency strategies," *Darwinian Security: Perspectives from Ecology and Evolution*, pp. 159–185, 2008.

[11] D. Johnson, "Darwinian selection in asymmetric warfare: the natural advantage of insurgents and terrorists," *Journal of the Washington Academy of Sciences*, pp. 89–112, 2009.

[12] S. J. Deitchman, *Limited war and American defense policy*. The MIT Press, 1964.

[13] M. Kress and R. Szechtman, "Why defeating insurgencies is hard: The effect of intelligence in counterinsurgency operations—a best-case scenario," *Operations Research*, vol. 57, no. 3, pp. 578–585, 2009.

[14] M. Wooldridge, C. Street, M. Manchester, and N. R. Jennings, "Intelligent Agents : Theory and Practice," vol. 10, no. January, pp. 1–62, 1995.

[15] A. Yang, H. A. Abbass, and R. Sarker, "Evolving agents for network centric warfare," in *Proceedings of the 7th annual workshop on Genetic and evolutionary computation*. ACM, 2005, pp. 193–195.

[16] I. Cil and M. Mala, "A multi-agent architecture for modelling and simulation of small military unit combat in asymmetric warfare," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1331–1343, 2010.

[17] P. Gowlett, "Moving forward with computational red teaming."

[18] I. Cil, "Mabsim : A multi agent based simulation model of military unit combat turkish general staff department of force development and resource management," pp. 731–736, 2009.

[19] E. Berman, *Radical, religious, and violent: The new economics of terrorism*. MIT press, 2011.

[20] M. D. Makowsky and J. Rubin, "An agent-based model of centralized institutions, social network technology, and revolution," *PloS one*, vol. 8, no. 11, p. e80380, 2013.

[21] A. O. Pechenkina and D. S. Bennett, "Violent and non-violent strategies of counterinsurgency," *Jasss*, vol. 20, no. 4, 2017.

[22] C. J. Ancker and M. D. Burke, "Doctrine for Asymmetric Warfare," *Military Review*, no. August, 2003.

Genesis of the Diophantine equation: $x^2 + n = {y_1}^2$ and $x^2 - n = {y_2}^2$ in Arabic mathematics

Saif-Eddin Toumi, Foued Nafti

Résumé :
Nous comptons dans cet article décrire une partie de la genèse de la théorie des nombres relatifs à la résolution de l'équation $x^2 + n = {y_1}^2$ et $x^2 - n = {y_2}^2$ en mettant l'accent sur les travaux des arithméticiens arabes, à partir du X^e siècle, comme, par exemple, al-Khāzin, Ibn al-Laith, al-Khallāṭ.

Nous présenterons par la suite les apports des mathématiciens de l'école occidentale comme par exemple Fibonacci, Fermat, Frenecle, Euler traitant de la même équation.

Mots-clefs : mathématiques arabes, théorie des nombres, Diophante, al-Khāzin, Ibn al-Laith, Al-Khallāṭ, Fibonacci, Fermat, Euler, Frenecle.

Abstract:
We intend in this paper to outline one part of the number theory genesis relatively to the equation $x^2 + n = {y_1}^2$ and $x^2 - n = {y_2}^2$ focusing on the Arabic mathematicians' works, like, for example, those of al-Khāzin, Ibn al-Laith, and al-Khallaṭ. We will then describe the occidental tradition relative to his subject threw the works of Fibonacci, Fermat, Frenecle, Euler, and others.

Keywords**:** Arabic mathematics, the number theory, Diophantus, al-Khāzin, Ibn al-Laith, al-Khallāṭ, Fibonacci, Fermat, Euler, Frenecle

Introduction

Translated into Arabic language by Ibn Lūqā, Diophantus's *Arithmetica*[1] opened a loophole for the mathematicians of Bagdad. Abū Kāmil is, according to our knowledge, the first to profit from this legacy. He devoted a whole book to Diophantine equations: *The Funny book on calculus* (*Al-Ṭarīf f-il ḥisāb*), in which he studied several problems of undetermined analysis[2] known by the Diophantine problems. He cited three types of these problems: those who have a unique solution, those who have more than one solution and those with no solution:

> Many <persons> among specialists and non-specialists asked me about problems of this type. I answered them giving for the same problem one solution, if it has only one, two or three or may be more for others. The solution may however not exist. Until I found a problem that I resolved and I found many solutions for it.

After him, this subject will be tackled by al-Khāzin, al-Khuğandī[3] and Ibn al-Laith who was able to regenerate new results related to Diophantine equations, like for example the theorem of al-Khāzin al-Khuğandī: «there is no cube which can be written the sum of two cubes» which we can translate in modern notations as follows:

For every three positive integers x, y and z, we can't have the equality[4]:

$$x^3 + y^3 = z^3$$

One should not here that this problem was the subject of many researches in Arabic mathematics and was studied by occidental mathematicians to be solved in its general form by Wiles (Fermat-Wiles theorem, 1993-95).

During the thirteenth century Ibn al-Khawwām al-Baġdadī carried on the work on the Diophantine problems. In his book The Bahāiyya benefits in calculation (*al-*

[1] This text was translated into Arabic language by Qustā ibn Lūqā al-Baʿlabakkī in the Xth century.

[2] Roshdi Rashed asserts that Arabic mathematicians reserved the expression *al-masā'il sayyāla* (undetermined analysis) to the Diophantine equations, in order to distinguish between it and the algebraic equations. See: R. Roshdi, Diophantine analysis and the number theory. Encyclopedia of History of Arabic sciences, Vol II, p. 514.

[3] Abu Muhamad al-Khuğandī, Persian astronomer and mathematician. He was active during the end of the tenth century. He participated in the construction of the astronomical observatory in Rayy.

[4] Al-Khuğandī tried to prove the following theorem: «the sum of two cubes cannot be a cube». According to al-Khāzin, the proof of al-Khuğandī is not complete. Al-Khāzin tried however to prove that: «the sum of two squares cannot be a square and a cube cannot be divided in the sum of two cubes. However a square can be divided into a sum of two squares»

Fawāid al-Bahāiyya fi-l ḥisāb) he mainly focused on the impossible ones. He claimed that he couldn't afford a demonstration about its impossibility. [5]

On the other hand, Roshdi Rashed (Rashed, 1982) asserts that the book *The base of rule in the origin of calculation* (*Asās al-qa'ida fī aṣl al-ḥisāb*) of Kamāl al-Dīn al-Fārisī, which is actually a commentary of Ibn al-Khawwām's one, is the greatest one of al-Fārisī works. In this text Al-Fārisī deals with the impossible Diophantine problems reaching almost the same results discovered by Ibn al-Khawwām. [6]

Ibn Mālik al-Dimashqī is another Arabic mathematician who worked on Diophantine problems. He retaken in his book «*The complete succor by pen's calculation*» (*al-Is'āf al-'atamm bi ḥisāb al-qalam*) the works of Ibn al-Khawwām. Finally, we note that in the fifteenth century Bahā' al-Dīn al-ʿĀmilī, tackled one more time the same subject, with tiny differences, in his *Summary of calculation* (*Khulāṣat al-ḥisāb*).

The equation $x^2 \pm n = y_i^2, i = \{1, 2\}$

One of the equations studied by Arabic mathematicians like al-Khāzin (Xth century), al-Karağī (XI^{e} century) is the famous:

$$x^2 + n = y_1^2 \text{ and } x^2 - n = y_2^2$$

Fibonacci of Pisa seems to be the first to settle this tradition in European area. After him, Fermat, Frenecle, Euler and Gennocci carried on researches about the same subject to find other results.

The contribution of al-Khāzin

In the tenth century, Al-Khāzin wrote two epistles on the equation mentioned above. In the one devoted to the construction of right-angled triangles with natural sides and their importance.
He declares that several issues related to this subject have been erased. Like how to now odd numbers which can be decomposed into two squares? And also the method to establish right-angled triangles using two or three successive numbers or using odd numbers.

[5] The manuscript: *The complete succor with pen's calculation* of Ibn Mālik al-Dimashqī. Library of King Sa'ūd university, nº 511, p. 225. In which al-Dimashqī quotes the following: «Ibn al-Khawwām god blesses him, quotes that it is not possible to prove the existence of a solution for it and asserts that he did not establish a proof for it regarding its impossibility».

[6] Roshdi Rashed, «History of amicable numbers», History of Arabic mathematics, Vol. 7, 1982, p. 7.

About the purpose behind these methods related to the construction of right-angled triangles al-Khāzin says in a fragment of his text:

> <The purpose> is to find a squared number if we add to it or we subtract from it a number the result is a squared number.[7]

Then, he proved the existence of a solution to this problem using Euclidian geometry. He precisely used in his demonstration the seventh proposition of the *Elements* Book II

The proof of al-Khāzin:

Assume that we have a right-angled triangle with sides x, y and z verifying: $x^2 + y^2 = z^2$, then we have:

$$z^2 + 2xy = (x + y)^2 \text{ and } z^2 - 2xy = (x - y)^2$$

So the number added in the first equation and subtracted in the second one is

$$n = 2xy.$$

To argue how he found the added-subtracted number *n*, Al-Khāzin says:

> If we take any two successive numbers and we multiply them one by the other, we then multiply the result by their sum and divide what we have by their difference so what comes is the added-subtracted number. Then we multiply every one of them by itself, we take the half of the sum and we take the divide it by the difference of the two numbers what is found is the root of the number that when we add to it the added-subtracted number then the result has a root and if we subtract from it that number the rest has a root.[8]
>
> Using our modern notation, this fragment can be translated as follows:
>
> The added-subtracted number is:

$$n = \frac{ab[a + b]}{a - b}$$

And we have:

$$\left(\frac{a^2 + b^2}{2(a - b)}\right)^2 \pm \frac{ab[a + b]}{a - b} = \left(\frac{a + b}{2} \pm \frac{ab}{a - b}\right)^2$$

After that, al-Khāzin explains how to solve the equation

[7] Abū Jaʿfar al-Khāzin, *Epistle on building up right angled triangles*, Ms. nº 2457, National Bibliotheca of France, Paris, 91ʳ-92ᵛ.
[8] *Ibid*.,fol. 92ᵛ

$$x^2 \pm n = y_i^2, i = \{1,2\}$$

He says that starting from two different numbers, we can construct a right-angled triangle with sides equal to integers. This means:

If a and b are two integers and $a > b$. Suppose that:

$$x = (a + b)(a - b)$$

$$y = 2ab$$

$$z_1 = (a + b)(a - b) + 2ab$$

$$z_2 = (a + b)(a - b) - 2ab$$

Then the triangle with sides x, y and z_1 and the one with sides x, y and z_2 are both right-angled. The hypotenuse is $z_1$or z_2.

And the added-subtracted number is:

$$.n = 4ab(a + b)(a - b)$$

However, in his second epistle, and after presenting several results al-Khāzin explains the main purpose behind writing it:

> After what we have introduced, we arrive to our purpose which is to prove that if we have a number how we can get a squared number that if we add to it this number and we subtract it from it the result and the difference are both squared numbers.[9]

Proposition 1:

> How we ask for a squared number that if we add to it a given number and we subtract it from it the sum and the different we get are both squared numbers[10].

Before he answers this questions al-Khāzin introduces first the following proposition:

> **Lemma:**
>
> Every even number which can be divided into two squared numbers then its half can be divided into two squared numbers and the half of its half and so on.
>
> **The proof:**

Suppose that $2a = b^2 + c^2$, then

[9] Adel Anbouba, *The epistle of Abū Ja'far al-Khāzin, on building up right angled triangles*, Revue of History of Arabic sciences, p. 167.
[10] *Ibid*., p. 167..

$$a = \frac{b^2 + c^2}{2} = \left(\frac{b + c}{2}\right)^2 + \left(\frac{b - c}{2}\right)^2$$

Proof of proposition 1:

Al-Khāzin says that after analyses he comes to prove the existence of three numbers verifying the system of equations

$$x^2 + n = y^2 \text{ and } x^2 - n = z^2$$

He first remarks that $z > x > y$ and then he concludes that x^2 is a sum of two squares since its double is a sum of two squares (referring to the lemma above):

$$2x^2 = y^2 + z^2$$

$$x^2 = (\frac{z + y}{2})^2 + (\frac{z - y}{2})^2$$

$$n = 2\left(\frac{z - y}{2}\right)\left(\frac{z + y}{2}\right)$$

$$2n = z^2 - y^2$$

So *n* is an even number and we have:

$$\frac{n}{2} = \left(\frac{z + y}{2}\right)\left(\frac{z - y}{2}\right)$$

Proposition 2:

The imposed number should be in the form $4m(2n + 1)$. Otherwise the problem is impossible.

Proof of proposition 2:

First, al-Khāzin proved that if $a^2 = b^2 + c^2$ then *b* and *c* cannot be odd numbers nor even-even numbers. So *b* and *c* are both even or one is even and the other is odd and in this case the number *a* is on the form:

$$a = 4m(2n + 1)$$

Otherwise the problem is impossible.

The proof of the opposite sense:

If it's we have a double even and odd (*zawj al-zawj w'al-fard*) number and we search a number when we add to it that number the result is a square and when we subtract from it that <same> number the rest is <also> a square.[11]

He supposes the divisors t and s of the number $\frac{a}{2}$.

So we have:

$\frac{a}{2} = st$ and $s^2 + t^2$ is a square. It means: $x^2 = s^2 + t^2$

Al-Khāzin calls s and t the two peers. He then concluded that the first number of this category is 24 because its half is 12 and it is the inferior one of double-even and odd numbers. He takes all the numbers coming after 12 verifying the property: the sum of their squares equals a square and their product equals 12. So we find the two numbers 3 and 4. He says:

The test/sign for the two sides of the supposed number <if you want to know> if the sum of their two squares is a square or not, is two divide the square of the smallest by the double of the greatest, so if the result has a root what we search is easy otherwise it is impossible.

Interpretation in modern symbols:

If we have number *a* and : $\frac{a}{2} = s.t$ with $t > s$ and we want to know if

$$s^2 + t^2 = square \longrightarrow s^2 + t^2 = u^2$$

Then, we divide s^2 by *2t* and the rest of the division is a square:

$$s^2 = 2tq + q^2$$

$$t^2 + s^2 = t^2 + 2tq + q^2 = (t + q)^2$$

Otherwise, i.e when we don't have the condition:

$$s^2 = 2tq + q^2$$

It is difficult to resolve the equation $x^2 \pm a = y_i^2, i = \{1,2\}$ or even impossible to do it.

[11] *Op., cit*, p. 165.

Optimization of scientific workflow scheduling in an energy-aware cloud environment

Hatem Aziza
LARODEC, Institut Supérieur de Gestion
Université de Tunis
Tunis, Tunisia
hatem.aziza@isg.rnu.tn

Saoussen Krichen
LARODEC, Institut Supérieur de Gestion
Université de Tunis
Tunis, Tunisia
saoussen.krichen@isg.rnu.tn

***Abstract*—Today, we live in an increasingly dematerialized world. The cloud is the generator of the next wave of technologies, the detonator for all cutting-edge developments like artificial intelligence and blockchain. The cloud is mainly composed of datacenters which consume almost 20% of the total energy footprint of the digital world. The host machines that make up these datacenters are becoming more and more powerful and perform calculations more quickly. Indeed, this increase in computing power is accompanied by increased consumption of electrical energy even in the event of inactivity. In this context, we are talking about a very energy-consuming development in terms of electricity consumption and even in terms of CO_2 emissions, the internet of which pollutes 1.5 times more than air transport. This exponential consumption of digital makes it very hard to feed with renewable energies. This problem leads us to focus on solutions that optimize the use of resources in the cloud environment in order to schedule scientific workflows especially since there are still today few techniques to limit energy consumption. This article is part of the resolution of this problem of minimizing energy consumption by datacenters in the cloud environment. For that, we try to propose a solution based on the technique Dynamic Voltage and Frequency Scaling (DVFS) which allows to moderate the frequency of the CPU thus influencing the energy consumption while having a positive impact on the execution time. The proposed solution is compared with other techniques in order to show its power to achieve the key objective of this article which is the optimization of energy consumption.**

***Index Terms*—cloud, energy consumption, dvfs, scientific workflows, scheduling**

REFERENCES

[1] X. Xu, J. Wu, G. Yang, and R. Wang, "Low-power task scheduling algorithm for large-scale cloud data centers," *Journal of Systems Engineering and Electronics*, vol. 24, no. 5, pp. 870–878, 2013.

[2] T. Gurout, S. Medjiah, G. Da Costa, and T. Monteil, "Quality of service modeling for green scheduling in clouds," *Sustainable Computing: Informatics and Systems*, vol. 4, no. 4, pp. 225–240, 2014.

[3] C.-M. Wu, R.-S. Chang, and H.-Y. Chan, "A green energy-efficient scheduling algorithm using the dvfs technique for cloud datacenters," *Future Generation Computer Systems*, vol. 37, pp. 141–147, 2014.

[4] pp. 494–495, 2015.

[5] Z. Tang, L. Qi, Z. Cheng, K. Li, S. U. Khan, and K. Li, "An energy-efficient task scheduling algorithm in dvfs-enabled cloud environment," *Journal of Grid Computing*, vol. 14, no. 1, pp. 55–74, 2015.

[6] W. Wu, W. Lin, and Z. Peng, "An intelligent power consumption model for virtual machines under cpu-intensive workload in cloud environment," *Soft Computing*, vol. 21, no. 19, pp. 5755–5764, 2016.

[7] M. Zotkiewicz, M. Guzek, D. Kliazovich, and P. Bouvry, "Minimum dependencies energy-efficient scheduling in data centers," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 12, pp. 3561–3574, 2016.

[8] P. Arroba, J. M. Moya, J. L. Ayala, and R. Buyya, "Dynamic voltage and frequency scaling-aware dynamic consolidation of virtual machines for energy efficient cloud data centers," *Concurrency and Computation: Practice and Experience*, vol. 29, no. 10, 2017.

[9] F. Farhadian, M. M. R. Kashani, J. Rezazadeh, R. Farahbakhsh, and K. Sandrasegaran, "Withdrawn: An efficient iot cloud energy consumption based on genetic algorithm," *Digital Communications and Networks*, 2019.

[10] N. Kumar and D. P. Vidyarthi, "A green sla constrained scheduling algorithm for parallel/scientific applications in heterogeneous cluster systems," *Sustainable Computing: Informatics and Systems*, vol. 22, pp. 107–119, 2019.

A Decision Support System for Steel Bending Machines Maintenance

Djamila Bouhalouan
Department of Computer Science
University of Oran 1 Ahmed Ben Bella
Oran, Algeria
bouhalouan.djamila@cubbat.fr

Abdelkader Adla
Department of Computer Science
University of Oran 1 Ahmed Ben Bella
Oran, Algeria
adla.abdelkader@univ-oran1.dz

Abstract*—In industrial plants, the profitability of the plant is significantly affected by the quality of machines maintenance. To ensure continuous production, the high valued machines should be kept in good working conditions. This brings plants to search for means to control and reduce equipment failures. When faults emerge in plants, appropriate actions for fault diagnosis and troubleshooting must be executed promptly and effectively to prevent large costs due to breakdowns. To provide reliable and effective maintenance support, the aid of advanced decision support technology utilizing previous repair experience is of crucial importance for the expert operators as it provides them valuable troubleshooting clues for new faults. Artificial intelligence (AI) technology, particularly, knowledge-based approach is promising for this domain. It captures efficiency of problem solving expertise from the domain experts; guides the expert operators in rapid fault detection and troubleshooting. This paper focuses on the design and development of a Knowledge-Intensive Decision Support System (KI-DSS) for Maintenance, Repair and Service in industrial plants to support better maintenance decision and improve maintenance efficiency. With integration of case-based Reasoning and ontology, the Ki-DSS not only carries out data matching retrieval, but also performs semantic associated data access which is important for intelligent knowledge retrieval in decision support system. A case is executed to illustrate the use of the proposed KI-DSS to show the feasibility of our approach and the benefit of the ontology support.

Keywords: -Decision Support Systems, Knowledge-Intensive DSS, Case-Based Reasoning, Ontology, Owl, Maintenance, Diagnosis

I. INTRODUCTION

Knowledge capitalization is of a considerable contribution during organizational problem solving activities which are often critical and recurring in nature. The development of a shared memory that stores the knowledge of expert members and their experience invoked in prior solutions can be useful for expert decision makers engaged in similar problem solving activities and will clearly assist them [1].

For an expert decision maker, it would be easier to reuse solutions and resolution schema corresponding to similar problems that have worked in the past than to analyze and solve the problems in scratch. Therefore, mechanisms to capture the experiential knowledge of experts can be of significant value to the organization in general, and the decision makers in particular.

A knowledge based approach shows significant promise for improving the effectiveness of maintenance activities support. It allows accumulating, organizing, storing, and sharing knowledge coming from past experience [2].

Among existing AI technology, Case-based reasoning (CBR) as an alternative reasoning paradigm and computational problem solving method has increasingly attracted more and more attention and grown in importance for businesses and academics over the past few years. The main principle of CBR is: *similar problems have similar solutions.* But, existing CBR systems lack semantic understanding, which is important for intelligent knowledge retrieval in knowledge-based systems [3]. To overcome this drawback, ontology technology is an ideal selection for realizing knowledge-based systems because ontology has not only powerful ability of knowledge representation, but also good semantic understanding. However, the explicit use of ontology based reasoning to support repetitive problem solving activities has received less attention.

To develop a such effective systems, two issues are critical: the first is how to find an effective method for case representation, which ensures domain knowledge can be acquired in an accurate easy manner, thus laying a good foundation for case retrieval; the next is how to find an appropriate method for case retrieval, which assures the right knowledge can be retrieved to solve a specific problem when a new task takes place. The system can not only carry out data matching retrieval, but also perform semantic associated data access, and improve the traditional keyword-based search. Through the semantic search capability, the hidden, but previously defined relations among data and concepts could be shown and represented if needed. These relations allow the user to understand the knowledge behind the stored data. Semantic technology could lead to the improvement in knowledge extraction, dissemination and management. Moreover, the use of common and unified domain ontology can improve the problem solving process where most of the decisions are dependent on individual experiences and domain knowledge of relevant managerial personnel.

The objective of this paper is to construct a maintenance intelligent knowledge-based system that can leverage the support of semantics. We suggest that the integration of ontology and CBR within a knowledge-based system is likely to provide additional information processing support. Ontology is used as a means to acquire domain knowledge and construct a case-base and use ontological semantic retrieval method as the case retrieval. Besides the case base, the system uses ontology (domain and task ontology related to the combustion machines to be maintained. This system will allow a more efficient searching in the case base by exploiting the semantic relations which exist between the cases. The system uses the semantic relations existing between the concepts within each of the ontology.

We experiment our system in maintenance domain, a semi-structured problem solving environment involving multiple attributes.

The remaining part of the paper is organized as follows. First, we present the case study related to maintenance in industrial plants. Then, we outline the related work on knowledge systems using ontology and case-based reasoning. Next, we present our Knowledge-Intensive Decision Support System. Finally, we present an example relating to maintenance of industrial machine before concluding

II. RELATED WORK

Knowledge management encompasses various practices of managing knowledge such as knowledge generation, capture, sharing, and application. Within these practices, effective sharing and use of knowledge depends – to a large extent – on the organization's ability to create and manage its knowledge. This knowledge can be described as the way organizations store it from the past to support present activities [4].

Knowledge management and Case-Based Reasoning are two intertwined topics. Case-based reasoning is a problem solving paradigm that in many respects is fundamentally different from other major AI approaches. Instead of relying solely on general knowledge of a problem domain, or making associations along generalized relationships between problem descriptors and conclusions, the case-based reasoning formalism was proposed as a way of storing human experiences and retrieving stored cases similar to the current item through a process of analogical search. It draws its knowledge from a reasonably large set of cases contained in the case library of past problems and by adapting their solutions solves new problems rather than only from a set of rules. Furthermore, case-based reasoning systems are claimed to "learn" through addition of further significant cases to the case-base and by forms of abstraction which may then be applied to this collection of cases [5].

Reasoning by re-using past cases is a powerful and frequently applied way to solve problems for humans. However, one of the drawbacks of CBR is the lack of flexibility of the knowledge representation. Indeed, the structure of the case is considered as constraining and strict which does not allow dealing with a carried out experiment in its semantic context, really limiting the performances of the system. As a way to deal with needs, ontology technology is an ideal selection for realizing knowledge-based decision support systems because ontology has not only powerful ability of knowledge representation, but also good semantic understanding. Ontologies provide a semantic based approach to explicitly represent information in a computable manner so that information can be automatically processed and integrated. Ontology also provides shared understanding of a domain to overcome differences in terminology from various sources [6]. The integration of an ontology-based model and CBR within a knowledge-based system has its advantages in: (1) Facilitating knowledge sharing by providing a formal specification of the semantics for context information; (2) Supporting for logic reasoning, referring to the capability of inferring new context information based on the defined classes and properties; (3) Enabling knowledge reuse by use of existing and mature ontology libraries without starting from scratch; (4) Having the stronger ability for expressing complex context information.

Several studies have given empirical evidence for the dominating role of ontologies integrated with specific, previously experienced situations (what we call cases) in human problem solving. Park and his colleagues [7] propose an ontology-based fuzzy CBR support system for ship's collision avoidance to prevent the cumbersome tasks of creating a new solution each time a new situation is encountered. A case-based decision support system applied to loan evaluation is developed in [8]. The approach uses AHP method to select important features and fuzzy sets technique to measure similarity between cases. A proposal presented in [3] aimed at knowledge reuse, during the decision activities by means of interwoven concepts from the knowledge management research. In [9] the constructed decision support CBR prototype system of marketing strategy contains more than 600 cases. The evaluation shows that with the support of semantics, they can not only carry out data matching retrieval, but also perform semantic associated data access. Kobti and Chen [10] construct domain ontology of mold design and propose an ontology-based search model to improve the traditional keyword-based search for the mold design domain. In [11], the authors proposed an approach based on the integration of three techniques: a CBR-personalized retrieval mechanism designed to provide a user with an optimum itinerary that meets his personal needs and preferences; a semantic web rule language considered to provide the system with enhanced semantic capabilities and support personalized case representation; and a user-oriented ontology used as source of knowledge to extract pertinent information about stakeholder's preferences and needs. To facilitate decision making within collaborative design, a Decision Support Ontology (DSO) is developed in [12]. The structure of the information model developed reflects a priori knowledge of decision making and supports the communication of information independent of any specific decision method. A case-based reasoning (CBR) system for the Semantic Web is presented in [13]. It implements a generic case-based inference mechanism in which adaptation consists in retrieving similar cases and in replacing some features of these cases in order to obtain one or more solutions for a given query. In [14] the authors propose a knowledge base for the Process Equipment Failures (PEFs) through semantic feature, embedded in the ontological approach and to construct a base frame for its further applications in the PEFs and process equipment related incident investigations and other knowledge extraction processes. A knowledge-based approach to support decision making in human resource management is proposed in [15]. The appropriate support of decision making is implemented using case-based reasoning and ontology. The problems of knowledge and case representation are considered, as well as the algorithm of case retrieval. Among other systems we cite the platform PROTEUS [16], and the Risks Analysis Support Tool in industrial domain [17].

Many research efforts for decision modelling and support have been systematically applied to the field of ontologies. However, there is no complete method that would define how to model decisions in ontologies, and a few isolated cases in which an established decision making method was used in ontology for a specific

domain, and often the reasoning procedure is based only on domain ontology.

In our approach, we consider particularly the case where the reasoning process is enriched by exploring ontology. Thereby the purpose is to retrieve and provide a set of possible solutions relating to source case showing the semantic relations between them. Afterwards, it is the duty of the decision-maker, according to his/her expertise, to opt for the decision which will seem to him appropriate to the target problem. An important goal of our work was to structure decision model in such a way that the problem solution can be obtained by reasoning upon three ontologies (domain, task, and decision). The ontologies with reasoning support can be used in the function of a case base reasoning system.

III. THE PROPOSED DECISION SUPPORT SYSTEM

The frame of our work is to integrate a knowledge-based tool in a Group Decision Support System (GDSS) that will be exploited by the actors (decision makers) for the purpose of decision support [18][19]. . The proposed system will assist the actors involved in maintenance session by offering them a set of decisions for the new problem and it is for the actor to situate each solution in its semantic context and then choose a particular solution based on his expertise.

We are in the context where typically incidents are not entirely identical to each other (some symptoms are not observed) but the knowledge of past incidents enables decision makers to recognize a similar situation and tailor their strategies by taking a course of action that experience has shown is effective and successful. This can happen when there is failure at some sensors so that lights or alarms cannot be triggered. The search in the database of cases can then be disoriented.

The benefits of using the system is to provide a more convenient retrieving process in information retrieval system in order to reach conclusions and give recommendations based on knowledge from previous cases (experiences) and ontologies.

A knowledge-based problem-solving approach adopting CBR is used to solve a new problem (target case) by remembering a previous similar situation (source case) and by reusing information and knowledge of that situation. The effectiveness of this approach is further improved by the application of ontologies as a mechanism for reasoning about the domain concepts and dealing with the inconsistencies that can arise in the applied vocabulary when multiple decision makers are involved. Thus, our approach to knowledge-based systems is towards integrated applications that combine case specific knowledge with models of general domain knowledge. The more knowledge is embedded into the system, the more effective is expected to be. Semantic CBR processes can take advantage of this domain knowledge and obtain more accurate results.

A. *The Case Base*

Knowledge representation is essential in building a knowledge-based system since on this presentation depends the effectiveness and the fastness of the system case retrieving mechanism. It is therefore necessary to well identify information to be stored in each case and to choose the more efficient representation scheme of this information. A case is a contextualized piece of knowledge representing an experience. The information encoded about the past experiences, depends on the domain of application as well as on the goal for which the cases are used. Case indexing and storage are an important aspect in designing efficient knowledge-based decision support systems in that, it should reflect the conceptual view of what is represented in the case and take into account the indices that characterize the case.

Knowledge considered in our knowledge-based decision support system is represented by cases and ontologies. The case base is composed of all the structured cases which will be explored during retrieving step (recall stage). Every case consists of a breakdown problem already experienced and solved. A case represents a diagnosis experience, and thus consists of two main parts: a problem part describing the failure, and a solution part. Each part is represented by a set of simple or complex descriptors among which some are defined in an ontology.

Problem part: the task to be solved;

Solution part: the solution, the problem solving method used and, the object concerned by recommended solution

The case base is a finite set of source cases (S), denoted by CB = {S1, S2,... Sn} where a source case Si = (PbS, Sol[PbS]) / PbS is the source problem part and Sol [PbS] is the solution part of the source case. Sol [PbS] = {[A1/V1], [A2/V2], ... [An/Vn]} where [Ai/Vi] means [Attribute/Value]. The case base is created manually with typical failures. The target case (C) is denoted by C = (PbC, Sol [PbC]). The target case problem part (PbC) is the structural representation of the new failure to be repaired; the solution of the target case (Sol[PbC]) is the structural representation of the resolution of this failure. The attributes of the PbC are filled upon an analysis phase of the problem parameters introduced by the user. This analysis involves the task ontology and allows structurally representing the new problem to be solved. Initially, the attributes of the Sol(PbC) part remain empty. Fig. 1 shows the UM classes diagram relating to the modelling of the case base. The descriptors are entries to the ontologies (e.g. Id-Task, Id-Symp and Id-Cause are entries for the task ontology; the descriptors Id-Object is an entry for the domain ontology, and the descriptor Id-Solution-Id is an entry for the application ontology).

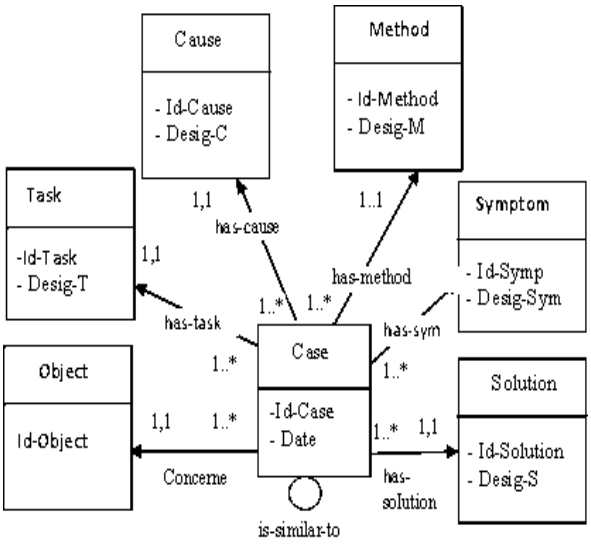


Figure 1: UML Class Diagram of the Case Base

B. Ontology Modeling

1) Conceptualizationl: The ontology development methodology is usually composed of several strategies on defining classes and class hierarchy, defining properties and naming considerations. We used the METHONTOLOGY method [20] to build the ontology. The ontology is created based on documentation resources as all the potential decisions that might be made by the decision makers are listed in an appropriate documentation. Similarly, the description of the equipment to be maintained is get from specific documentation while the specification of the task ontology is built with the support of an expert in industrial maintenance.

The ontology is composed of three parts related to equipment domain, maintenance task and application. Fig. 2 present UML classes diagram of the proposed ontology. The equipment domain part consists of a specification of the concepts relating to the equipment to maintain as well as the relations between these concepts. The latter are principally aggregation and composition relations between the equipment components. The task part described all the maintenance problems related to the combustion machine (equipment) in terms of task, symptom, cause and solution concepts, and the relations between them while the *application domain part* represents the domain of decisions.

2) Ontologization: The ontology is created using Protégé before their generation in OWL format [21]. Fig. 3 illustrates a partial view of the ontology.

3) Operationalization : An operationalized ontology is expressed in an operational language and endowed with operational semantics. In this sense the ontology operationalization consists of a computer specification of all the operations made on concepts in an operational language. The use of an operational ontology assumes its representation in an operational but also formal language, i.e. providing reasoning mechanisms appropriate to the targeted knowledge manipulations. To do this, we used the NetBeans developing environment associated to Java language [22]. Furthermore, we used the Jena framework Jena [23] to manage the ontology. Jena provides a programming environment for RDF, RDFS [24]) and OWL as well as a query engine allowing SPARQL queries execution (Simple Protocol And RDF Query Language) [25] which is a RDF query language.

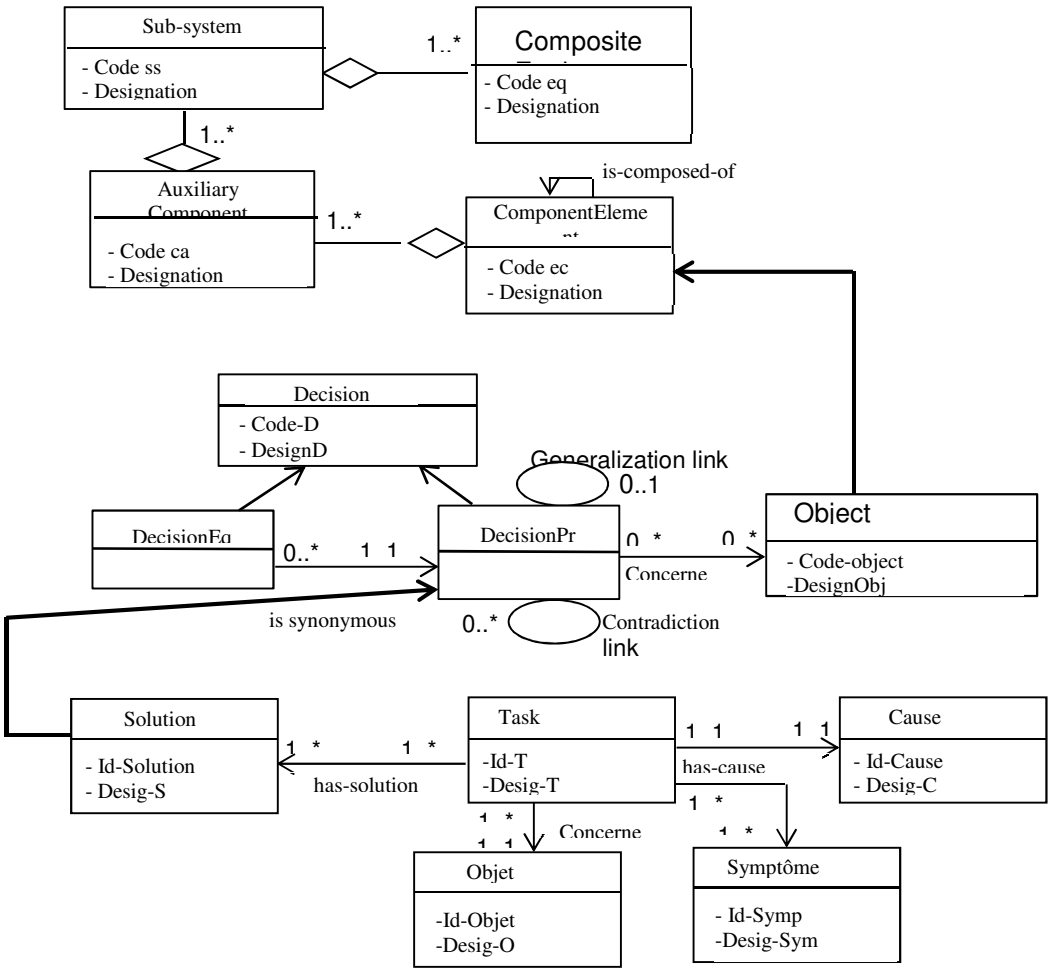


Figure 2: Conceptual model of the ontology

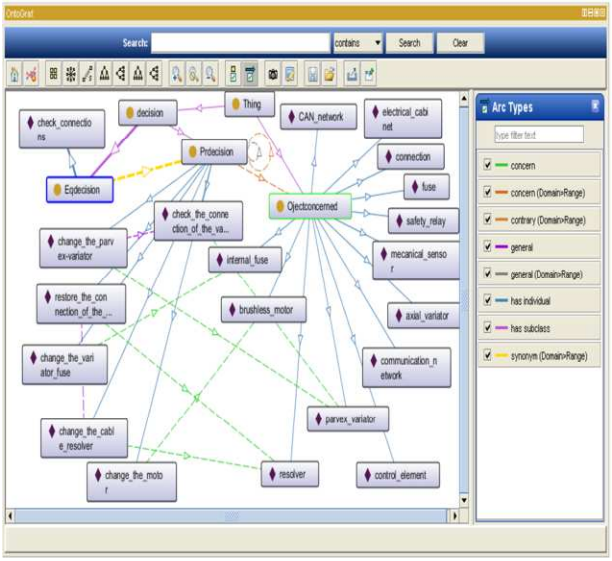


Figure 3: Partial view of the decision ontology

OWL language [21] is used to represent the case base. This would allow managing the case base as a knowledge base upon which inferences may be made. It is possible to define semantic relations between cases as for instance the transitive relation "is-similar-to" which relates the source cases already identified as being similar. Furthermore, as the remained knowledge (i.e. the ontologies) is also expressed in OWL, this would allow having to some extent compatibility between languages formalizing the different knowledge manipulated by the system, as well as, the knowledge operating tools such as SPARQL.

TABLE 1: EXAMPLE OF THE SOURCE CASE T3

<owl:NamedIndividual rdf:about="http://www.basedecas.org/ontologycases# 3">
<rdf:type rdf:resource="http:// www.basedecas.org/ontologycases#Cases"/>
<rdf:type rdf:resource="http:// www.basedecas.org/ontologycases#Cases"/>
<has-as-task rdf:resource="http:// www.basedecas.org/ontologycases#T3"/>
<has-as-cause rdf:resource="http:// www.basedecas.org/ontologycases# Failure to turn up the variator"/>
<has-as-method rdf:resource="http://www.basedecas.org/ontologycases#M1"/>
<concerns rdf:resource="http://www.basedecas.org/ontologycases# Internal fuse"/>
<has-as-solution rdf:resource="http://www.basedecas.org/ontologycases#change internal fuse"/>
<possesses rdf:resource="http://www.basedecas.org/ontologycases#display variator off"/>
<possesses rdf:resource="http://www.basedecas.org/ontologycases# Failure to turn up the variator, the machine is shut down"/>
<is-similar-tordf:resource="http://www.basedecas.org/ontologycases#15"/>
</owl:NamedIndividual>

C. The Reasoning Process

The proposed case based system should reflect human knowledge by storing data about previous significant events as "cases" within a computerized system. In this regard, the system uses the case base to retrieve similar cases to the problem to be solved. But, when the retrieving process fails or the cases retrieved are not satisfactory for the decision maker, the system uses ontologies. It makes use semantic relations between concepts within the same ontology to derive other solutions to the problem.

By making use of the ontology, the system derives more specific or more general decisions than those initially retrieved by the system. It can also set the solution relating to the equipment by visualizing the concerned component. Then, it uses the ontology to set the involved component relating to the neighboring ones or to the component in which it's comprised. Similarly other case descriptors may be used as entries to the ontology to enlarge or reduce solution space. When a solution is retained, then tested and validated, it is stored in the case base as a new case (with all its descriptors).

The reasoning process consists of the following steps (Fig. 4):

Problem description: The decision maker describes the problem to be solved. This description can be made of different ways: by providing the task to be solved, the observed symptoms, or the faulty object, etc.

Retrieving: It consists to search in the case base and retrieve similar cases to the problem to be solved. Here, we consider the usual local and global similarities measures to retrieve similar cases to the targeted problem. Once the target problem is introduced, the goal of this step is to recall the source case that is most similar to the target case, applying two measures of similarity (local, global) between the target problem and the source problem.

Local Similarity: This similarity measure is computed between the value of an attribute in target problem (PbC) and the value of the same attribute in source problem (PbS). It is evaluated in two different ways depending on whether the attribute is simple or complex.

Simple attribute: The simple attribute in our case is the symptom attribute which has a unique value. The local similarity between the attributes is equal to 1 if the values of the two attributes are equal (see equation (1)), else it is equal to 0.

$$\mathrm{Sim}_i(s_i,c_i) = \begin{cases} 1 \text{ pour } ci = si \\ 0 \text{ pour } ci \neq \text{si} \end{cases} \qquad (1)$$

C_i: Value of the attribute *i* in PbC part.
S_i: Value of the same attribute *i* PbS part.

Complex Attributes: The complex attribute in our case is the symptom attribute which has a list of values. The local similarity between the values of the symptom attribute of the PbC and the values of the same symptom attribute of the PbS is given by equation 2 below:

$$\text{sim (si, ci)} = \frac{ni * 2}{n1 + n2} \qquad (2)$$

ni: number of equal values in both target and source problem for the i attribute.
n1: number of values relating to the attribute in the source problem.
n2: number of values relating to the attribute in the target problem.

Global similarity: This similarity measure is calculated between a set of attributes in the PbC and the same set of attributes in the PbS, it corresponds to the mean of the local similarities and its value is in the range [0,1]. It is defined by the similarity function below:

$$\text{SIM (S, C)} = \frac{1}{n} + \sum_{i=1}^{n} \text{simi (si, ci)} \qquad (3)$$

C: target case.
S: source case.
n: number of attributes in which the local similarity has been calculated.

sim_i (s_i, c_i): value of local similarity for attribute *i*.

To search a PbS, the most similar to the PbC. For this, we apply the measure of local then global similarities. In the global similarity measure, a similarity threshold is determined at 0.5. If SIM (S, C) <0.5 the source case is considered negligible (not similar) else if SIM (S, C) ≥ 0.5 the source case is considered important (partially similar). In case SIM (S, C) = 1 the source case is considered perfectly similar.

Reasoning: If a perfectly similar case is not found, we move on ontology-based reasoning in order to search other source cases. For each of the initially retrieved source cases, we use semantic relations between the source cases to derive other source cases semantically close to similar cases retrieved. These if they exist are presented to the user. The latter, if the suggested solutions are satisfactory, the process of recalling is stopped and we pass to the next phase of the reasoning process. Furthermore The ontology may be used to enlarge or to reduce the solution search space. According to the object of widening, the ontology is used. For example when the object of widening is a task or a symptom, the task ontology part is used; when the object of widening is a faulty component, the equipment domain ontology part is explored, but when the object of enlargement is the problem solution then the decision ontology part is used.

Validation: Once the decision is made, executed and validated the process will skip to the next step.

Learning: The new case is added to the case base. It referees to all the similar source cases if exist.

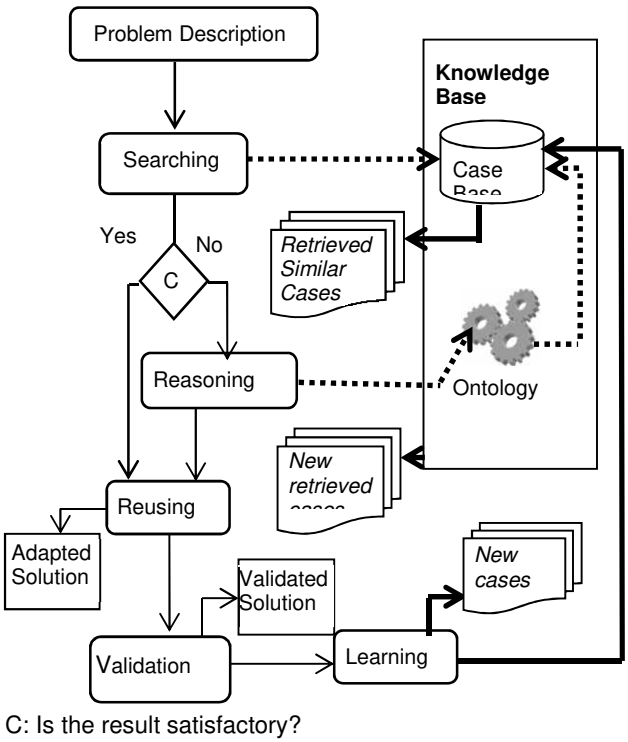


Fig. 4: Ontology-based CBR Process

The reasoning step is useful as it allows revealing semantic knowledge from ontologies between the different parameters of the problem to be solved. Given a problem to be solved, this would allow: 1) Converging to the semantically close case in the case base, or 2) Retrieving first a structurally close case from the case base then, according to the case descriptors, exploiting ontologies in order to derive other possible solutions to the problem. The decision-maker will choose among the suggested solutions that he considers being the most appropriate one to the problem.

IV. CASE STUDY

We experiment our approach in the maintenance field. The application concerns steel bending machines maintenance. The machines are built for years of operation at full capacity. The machines are manufactured for outdoor and indoor furniture, supermarkets and household electrical appliances, etc. Each machine can bend wires with diameters up to 7 mm. The FS Series machine is a combination of 13 different components. The Parvex drive in the switch cabinet is a particularly critical component.

The machines are numerically controlled tools that bend and fold sheets and tubes, precision and industrial sheet metal, steel sheet and strip. The controllers perform a variety of functions including protecting the machine from damage by performing an automated shutdown when dangerous conditions are detected and archiving sensor data.

A statistical study was carried out on the failures encountered in the company. Some of the failures are common between the different machines. In most cases, the failure is due to the electrical problem. The Parvex drive is the component the most affected by the failures. The duration of repairing can last up to a few weeks for complex problems.

Currently, the machine maintenance process was as follows. When a machine has broken down or a malfunction of a machine is detected, the on-site operator will call the company and will send a message. The breakdown will be assigned to a company technician for analysis. The technicians take into account the information provided by the customer, will access the data from the machine, review key values, draw on their previous know-how and experience, create a hypothesis about the breakdown cause, create plots specific to the breakdown type hypothesized, confirm the cause of the breakdown as best as can be done using the available data, then call the site to provide assistance and confirm the breakdown cause. Each repairing is recorded. The saved records are not well formalized and thus unexploited. These sheets should be used as a basis of experience for use in future repairs.

The goals are to improve machine and system reliability, reduced machine operating/maintenance costs, and produce the greatest possible sustained availability from the machine. For the company, it is a matter of better formalizing the experiences of the experts on maintenance, using reasoning from experience. The knowledge-based approach is to be used

to automate the data review, hypothesis generation, and hypothesis confirmation of this process whenever possible and assist the user when it does not have confidence in a single cause.

The objectives are to reduce the diagnosing and problem solving time, quickly analyze the breakdown, provide technicians with tools to help them be more effective at the diagnosis stage, and to transfer knowledge on unformed trades and train people for more efficiency and performance. The company wants to reduce the diagnostic time for more frequent breakdowns.

TABLE 2. THE CASE BASE

Id-T	Task	Cause	Method	Object	Solution
1	T1	Resolver break failure	M3	Resolver cable	Check the resolver cable
2	T2	Breaking of one or more wires of the encoder	M2	Encoder	Changing the encoder
3	T3	Failure to power up the variator	M1	Internal fuse	Change the internal fuse
4	T4	Internal drive failure	M4	Drive	Changing the drive
5	T5	High temperature	M6	Radiator	Clean the radiator
6	T6	Excessive engine speed	M5	Engine	Changing engine
7	T7	Drive power incident	M1	Internal fuse	Changing the internal fuse
…	…	…	…	…	…

TABLE 3. LIST OF SYMPTOMS

Id-T	Symptoms
1	Symp 1(Failure to power up the machine, the variator is shut down)
1	Symp 3 (Fault 2 appears on the axis drive display)
2	Symp 2 (No information at the encoder display)
3	Symp 1(Failure to power up the machine, the variator is shut down)
3	Symp 4 (At the time of automatic start the orientation head moves downwards)
3	Symp 10 (Fault 4 appears on the display of the axis drive)
3	Symp 18 (No information displayed)
3	Symp 20 (Electrical axis incident, drive power incident)
4	Symp1 (Failure to turn up the machine, the variator is shut down)
4	Symp 20 (Electrical axis incident, drive power incident)
4	Symp 21 (Absence of “READY” (Power E1)) (Symp displayed on the computer screen)
5	Symp 7 (Fault 4 appears on the display of the axis drive)
5	Symp 20 (Electrical axis incident, drive power incident)
5	Symp 21 (absence of “READY” (Power E1)) (Symp displayed on the computer screen)
6	Symp 9 (High engine temperature)
6	Symp 18 (No information displayed)
6	Symp 1 (Failure to power up the machine, the variator is shut down)
…	…

Let consider the following description of the problem: "Variator failure”. This problem is characterized by the following symptoms: "Symp1 (Symp1: impossible to power up the variator), Symp18 (No information on the drive displayed). We do search for similar cases to the problem occurred.

New Problem	**Task**	Variator failure
	Cause	
	symptoms	Impossible to power up the variator
		No information on the drive displayed
New Problem Solution	**Method**	
	Objet	
	Solution	

The task ontology is used to infer the features lacking (Cause) in the target problem structure.

New Problem	**Task**	Variator failure
	Cause	Drive power failure
	symptoms	Impossible to power on the machine
		No information on the drive displayed
New Problem Solution	**Method**	
	Object	
	Solution	

We calculate the similarity measures between the attributes: task, cause, symptom.

Case	Attributes			Global Similarity
	Task	Cause	Symptom	
T1	0	0	0.5	0.16
T2				
T3	1	1	0.5	0.83
T4	0	1	0.4	0.46

At the end of the second recall step, a source case (PbS) similar to the task T3 is retrieved. If the user would accept the proposal, he will apply the Solution to the target case (PbC).

Similar source case	**Id-T**	3
	Task	T3
	Cause	Drive power problem
	symptoms	Symp1
		Symp2
Similar source case solution	**Method**	M1
	Object	Internal Fuse
	Solution	Changing the fuse

V. CONCLUSION

In this paper, we focused on the design and development of a Knowledge-Intensive Decision Support System (KI-DSS) to support better maintenance decision and improve maintenance efficiency in industrial plants. We integrated case-based reasoning and ontology. We claim that this combination is useful for the design of KI-DSS and strengthens its reasoning process as it allows the knowledge engineer to use knowledge already acquired, conceptualized and implemented in a formal language; reducing considerably the knowledge acquisition bottleneck.

The effectiveness of CBR can be further improved by the application of ontologies as a mechanism for reasoning about the domain concepts and dealing with the inconsistencies that can arise in the applied vocabulary when multiple decision makers are involved. Moreover, the reuse of ontologies from a library also benefits from their reliability and consistency.

We believe that our approach is useful in several aspects. First, it enables to formalize the case base in OWL what allows managing it as a knowledge base. Indeed, by exploiting the semantic relations within the case base, it is possible to derive new knowledge from those stored. Also, as a result of memorizing a source case base with its descriptors, the ontologies exploration will allow deriving new knowledge which will serve for a new research cycle in case base.

As future work, we aim to collect much information to use quantitative indices for performance evaluation of our approach. Further experiments are to be organized to evaluate the performance of our system with fault coverage rate, diagnosis effectiveness ratio, and other quantitative indices.

REFERENCES

[1] M. S. Ackerman, and C.A. Halverson, Reexamining organizational memory, *Communications of the ACM*, , 2001, pp. 58-64.

[2] A. Aamodt, and E. Plaza, Case-based reasoning: foundation issues, methodological variations and system approaches, *Artificial Intelligence Communications, vol.* 7, 1994, pp. 39-59.

[3] J. L. Garrido, M. V. Hurtado, M. Noguera, and J.M. Zurita, Using a CBR approach based on ontologies for recommendation and reuse of knowledge sharing in decision making, in Proceedings of 8th international conference on hybrid intelligent systems, 2008, pp. 837–42.

[4] B. Gallupe, Knowledge management systems: surveying the landscape, *International Journal of Management Reviews*, vol. 3, no 1, 2001, pp. 61-77.

[5] J. Kolodner, *Case-Based Reasoning*, Morgan Kaufmann Publishers, Inc, 1993.

[6] G. Antoniou, and F. Harmelen, *A Semantic Web Primer*, MIT, 2004.

[7] K. Park, R. M. Benedictos, C.S. Lee, and M. H. Wan, Ontology-based fuzzy-CBR support system for ship's collision avoidance, in Proceedings of the 6th international conference on machine learning and cybernetics, 2007.

[8] N. Benmessaoud, A. Adla, Intelligent Semantic Case Based Reasoning System for Fault Diagnosis, *Journal of Digital Information Management (JDIM)*, Volume 17 Number 2, 2019.

[9] W. Dong, X. Yang, Z. Guobing, Z. Bo, Research on ontology-based case indexing in CBR in Shanghai, In: International conference on artificial intelligence and computational intelligence, ICAICI 09, IEEE Press, 2009, pp. 238–42.

[10] Z. Kobti, and D. Chen, A domain ontology model for mold design automation, *Canadian AI 2010,* vol. 6085, 2010, pp. 336–9.

[11] J. A. Rockwell, I. R. Grosse, S. Krishnamurty, and J. C. Wileden, A Decision Support Ontology for collaborative decision making in engineering design, in Proceedings of Center for e-Design, 2009.

[12] A. Bouhana, A. Zidi, A. Fekih, H. Chabchoub, and M. Abed, An ontology-based CBR approach for personalized itinerary search systems for sustainable urban freight transport, *Expert Systems with Applications*, vol. 42, 2015, pp. 3724–3741.

[13] L. Lamontagne and E. Plaza, Tuuurbine: A Generic CBR Engine over RDFS, in proceedings of ICCBR, 2014, pp. 140–154.

[14] I. Mohammadfama, O. Kalatpour, R. Golmohammadi, and H. Khotanlou, Developing a process equipment failure knowledge base using ontology approach for process equipment related incident investigations, *Journal of Loss Prevention in the Process Industries*, vol. 26, 2013, pp. 1300-1307.

[15] M. Zhukova, M. Kultsova, A. Navrotsky, A. Dvoryankin, Intelligent Support of Decision Making in Human Resource Management Using Case-Based Reasoning and Ontology In Proceedings of JCKBSE 2014, pp. 172–184.

[16] I. Rasovsca. B. ChebelMorello. N. Zerhouni, A case elaboration methodology for a diagnostic and repair help system based on CBR, in "20th International Florida Artificial Intelligence Research Society Conference, FLAIRS'07., Key West, Florida (USA), 2007.

[17] B. Debray, C. Duval, C.A. Jovanovic, O. Salvi, Integrated management of emerging risks: challenges and objectives of the iNTeg-Risk

European project, 16ème Congrès Lambda-Mu, Avignon, France, 2008.

[18] A. Adla, , J.L. Soubie, P. Zaraté, A cooperative Intelligent Decision Support System for Boilers Combustion Management based on a Distributed Architecture, *Journal of Decision Systems (JDS)*, Vol. 16 No 2, 2007, pp. 241-263.

[19] A. Adla, P. Zarate, J.L. Soubie, A Proposal of ToolKit for GDSS Facilitators, Group Decision and Negotiation (GDN), Vol. 1, 2011.

[20] M. Fernandez, A. Gomez-Perez, N. Juristo, METHONTOLOGY: From Ontological Art Towards Ontological Engineering, In proceedings of the Ontological Engineering AAAI-97 Spring Symposium Series, 1997, Stanford University (USA), 1997.

[21] OWL 2 Web Ontology Language New Features and Rationale (Second Edition)W3C Recommendation 11 December 2012 http://www.w3.org/TR/owl-new-features/

[22] Java SE http://*java*.sun.com, 2012.

[23] Apache Jena http://jena.sourceforge.net/, 2012.

[24] RDF (Resource Description Framework) Model and Syntax Specification W3C Proposed Recommendation 05 January 1999 http://www.w3.org/TR/PR-rdf-syntax

[25] SPARQL Query Language for RDF W3C Recommendation 15 January 2008, http://www.w3.org/TR/rdf-sparql-query/

A New Sum Coloring-based Integer Linear Programming For Solving The University Course Timetabling Problem

Olfa Harrabi * , Jouhaina Chaouachi Siala †

Abstract

The drawing up of university course timetables is an ongoing challenge and laborious task which must be performed by an administrative staff working on the strength of their knowledge to fill a specific institution requirements. This paper proposes a novel integer programming model as an alternative to solve the treated optimization problem. Specifically, we address providing an optimum solution satisfying at most teachers' preferences. The computational results on a real case of study well illustrate a significant performance over the manual process on both computational time and solution quality.

Index Terms

Course timetabling; teacher preference; sum coloring problem; integer linear programming

I. Introduction

For educational institutions, constructing timetables is a classical combinatorial problem that embraces the scheduling of encounters between teachers and students so as to ensure some given requirements. Generally, the optimization process focuses on assigning a number of events or lectures to a prefixed period of time throughout the working days of the week. Treated as an optimization problem, the University Course Timetabling (UCT) was proven to be NP-complete for nearly all of its variants [1]. In light of the progress achieved in modern search methodologies, much research has been invested to automate the desirable timetables. While there exists a large and growing body of literature related to this issue, little attention has been paid to consider the human factors during the automation process. Indeed, most of resolution approaches emphasize mainly on feasible timetable which requires the satisfaction of hard constraints. The latters are usually defining by no simultaneous occupation of any technical or human resources (classrooms, teachers and subjects). To fill this gap, this paper attracts more attention from researchers to this crucial yet less explored area in the field of UCT problems. Indeed, our work is aware from some soft constraints which are mostly concerned with human factors (fondness, weakness, non disponiblity, etc) and the reasonable management of the availability times. The resolution framework integrates a new Integer Linear Programming model (ILP) that mainly treats preferences of teacher as priority to ensures the smooth progress of an education task. From a theoretical point of view, our proposed ILP is based on the Sum Coloring Problem (SCP) proved to be NP-hard. The developed method is mostly recommended as it results both the proper distribution of timetables and the satisfaction of teacher preferences with moderate computational time efforts.

II. A new Sum Coloring-based formulation

Solving educational timetabling problems through graph-based approaches has a long and varied history [2]. These approaches are rationally used to schedule the events into times slots in order to obtain *feasible* timetable that satisfy all the hard constraints. In this early research work [3], a clear illustration was performed to show how a course timetabling problem could be mapped onto a graph model. The formulation followed in this work is novel and addresses two being hard combinatorial problems: the sum coloring problem and the university course timetabling problem. Indeed, this model provides the optimal timetables using the following sum coloring-based formulation. We list bellow the necessary notation to explain our proposed model:

Let

- A is a set of all the students' groups
- S is a set sections (lecture of courses attended by only a group of students)
- E_m is a set of sections in conflict (sharing common students of instructors); the mth row of the conflict matrix
- M is a set of values of m

O. Harrabi is with Higher Institute of Management of Tunis, Tunis University, 41, Liberty Street - Bouchoucha - 2000 Bardo, Tunisie (olfa-harrabi@hotmail.fr).

J. Chaouachi Siala is with Institute of Advanced Business Studies of Carthage, Carthage University, IHEC Carthage Presidency-2016 Tunis, Tunisie (siala.jouhaina@gmail.com) .

- F is a set of instructors
- T is a set of time slots
- P_f is a set of time slots preferred by the instructor f
- C_{it} is the cost of time slot t if it is assigned to the section i
- C_{it} $\begin{cases} 0 \text{ if the time slot } t \text{ belongs to the set of preferred schedules for section } i \\ (i.e.\ \text{t} \in P),\ \forall t \in T, \forall i \in S \\ 1 \text{ otherwise} \end{cases}$

- x_{it} $\begin{cases} 1 \text{ if the section } i \text{ is assigned to a time slot } t, \\ \forall t \in T, \forall i \in S \\ 0 \text{ otherwise} \end{cases}$

Using the foregoing notations, the final specification of the proposed ILP model takes the following form:

$$Minimize \quad \sum_{i \in S} \sum_{t \in T} c_{it} x_{it} \tag{1}$$

$$subject\ to: \quad \sum_{t \in T} x_{it} = 1 \quad \forall i \in S \tag{2}$$

$$x_{it} + x_{jt} \leq 1 \quad \forall (i,j) \in E, \forall t \in T \tag{3}$$

$$x_{it} \in \{0,1\} \quad \forall t \in T, \forall i \in S \tag{4}$$

The objective function (1) is a cost function that needs to be minimized. It defines the preferences of scheduling a section i in a time slot t in term of cost C_{it}. In our context, we assign each section to some time slots in $\{1, \ldots, t\}$ by affecting a cost value 0 if the scheduling is preferred; 1 otherwise. Constraints (2) ensure that each section is taught once (i.e. it must be assigned to only one time slot). Constraints (3) force a conflicting sections pair (i, j) to be assigned to different time slots. Finally, constraints (4) state that the decision variables x_{it} is binary-valued.

III. Empirical testing

The system was implemented on a i3 processor with 2.53 Ghz of clock speed and 4 Gb of available memory. The ILP model was coded using C++ and solved by Cplex 12.6. The collected data correspond to the winter semester of the Institute of *Advanced Business Studies of Carthage (IHEC)*. An optimal solution was obtained in 2 seconds of CP time. The preferences, specified by low objective function values were almost completely satisfied in most of the tests (about 80% of the instructors were satisfied).

IV. Conclusion

In this study, a new ILP model for UCT problem has been developed. One of the significance of our model is satisfying teacher preferences by using suitable objective function coefficients. We would expect that this work could be an important stage in the complete solution of the classical university course timetabling problem.

References

[1] Asratian, A. S., de Werra, D. (2002). A generalized classteacher model for some timetabling problems. European Journal of Operational Research, 143(3), 531-542.
[2] De Werra, D. (1985). Graphs, hypergraphs and timetabling. Methods Operational Research, 49(ROSE-CHAPTER-1985-001), 201-215.
[3] de Werra, D. (1985). An introduction to timetabling. European journal of operational research, 19(2), 151-162.

An overview of Machine Learning Technologies and their use in E-learning

Ramzi FARHAT[1], Yosra MOURALI[1,2,3], Mohamed JEMNI[1], Houcine EZZEDINE[2]

[1]*Research Lab. LaTICE, University of Tunis, Tunis, Tunisia*
ramzi.farhat@esstt.rnu.tn
Mohamed.jemni@fst.rnu.tn

[2]*Univ. Polytechnique Hauts-de-France, CNRS, UMR 8201 - LAMIH Valenciennes, France*
Houcine.ezzedine@uphf.fr

[3]*Faculty of Economics and Management, University of Sfax, Sfax, Tunisia*
yosra.mourali@etu-uphf.fr

Abstract—Thanks to new technologies, internet, connected objects we produce a phenomenal amount of data. Putting these data in context, organizing them to be able to perceive, understand and reflect them is very important. Traditionally, human have analysed data. However, as the volume of data surpasses, human increasingly turn to automated systems that can imitate him. Those systems able to learn from both data and changes in data in order to solve problems are called machine learning. Artificial intelligence has a major impact on e-learning research and the machine learning based methods can be implemented to improve Technology Enhanced Learning Environments (TELE). This paper is an overview of the recent findings in this research field. At first, we introduce the key concepts related to machine learning. Then, we present some recent works using machine learning in e-learning context.

Index Terms—E-learning, Technology Enhanced Learning Environments, Data, Learners' traces, Machine Learning, Deep Learning

1 INTRODUCTION

Almost everything we do today leaves a digital trace that describes our activities, specifies our location, and provides many other information about what we say, what we buy, etc. Thanks to both data storage capacity and digitalisation of society, most of devices, machines and everything we use, produce data. We can, as example, extract information from pay stations, parking, smart phone, social networks, videos, photos, etc. It is necessary to benefit and find meaning to all this collected data.

Analysing data makes it possible to understand phenomena, to model behaviours and to make predictions. Before, humans analyse data, wrote algorithms and the machine applied them to solve problems. Today, humans introduce data and allows the machine to learn on its own from these data without being explicitly programmed. We talk about the power of data. This is the principle of machine learning.

In reality, there is an awareness of the richness that data can hold and the importance of valuing it. Actually, analysing complex data through machine learning methods has emerged as an important era in several scientific research domains such as medicine [1] [2], e-commerce [3], industry [4][5], education [6][7], social networks [8][9], economics and finance [10], etc.

Figure 1 shows machine learning relationships to some other concepts of data science and artificial intelligence. In fact, data mining use statistics to extract hidden information (patterns) from raw data [11]. However, machine learning as a subfield of computer science and artificial intelligence, learns from patterns to predict. The Deep Learning is one of the main technologies of machine learning and artificial intelligence. We can say that it is the new generation of machine learning which characterized by learning by layer and on each layer the machine has to learn a little more.

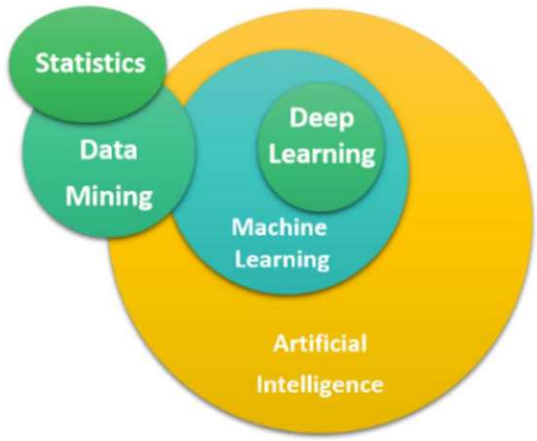


Fig. 1 Machine Learning relationships to other related fieleds

2 Machine Learning

In machine learning, a computer learns from example data how to perform tasks. We know that if we give more experiences (E) with a defined task (T) to a machine, its performance (P) improves [12]. For example, let suppose that we want an email client to classify emails as spam or not. The experience E in this case should be a set of emails already classified as spam or not. The Task T perform is to classify automatically new emails. The performance P that should increase is the accuracy rate of the classification made by the machine on a set of new emails.

2.1 Machine learning process

The generic machine learning process consists of seven steps as described next [13]. The first step is to collect data. It is a very important task because it will determine how good predictive model can be. But, data we gathered are, in most times, unstructured, contain a lot of noise or have to take other forms to be useful for our machine learning. So, data need to be cleaned and pre-processed.
After that we can begin building our machine learning model. For this, we start by the feature engineering in which we choose the most relevant features from data, then we try to select the best machine learning algorithm for the problem at hand. It is imperative in getting the best possible results.

The next task is training. In this step we use a part of our data to incrementally improve machine learning ability to predict. Once training is complete, it is time to test the model and observe how it might perform against the other part of data unseen. The performance evaluation is measured by various parameters like accuracy, precision and recall. Sometimes, it is possible to go back and improve training then test again. The last step is the result given by the machine learning. It can be a prediction or inference.

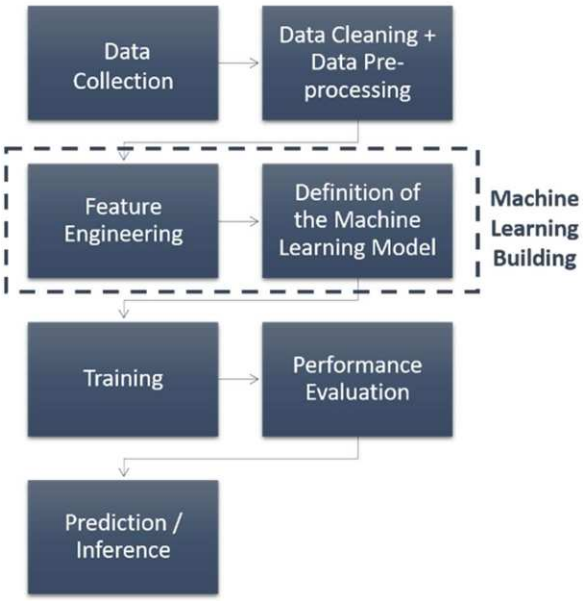


Fig. 2 Components of a Generic machine learning model

2.2 Machine learning paradigms

Machine learning can be classified based on the approach used for the learning process. Four main categories were identified: supervised, unsupervised, semi-supervised, and reinforcement learning [12].

In supervised learning, we have a set of training data or labelled data in which we know the structure and the outcome. We take this data and train a machine learning model, so it can understand patterns in the data. Once the model has been trained, we can use it to predict results of data in which outcomes are unknown [14].
Conversely, unsupervised learning methods learn structure from the data itself without the need for prior labelling [15]. That is mean we can apply unsupervised machine learning to find patterns that exist within labelled data.

However, full label information is not available at all times. Semi-supervised learning provides a powerful framework for leveraging unlabelled data when labels are limited or expensive to obtain [16].

3 Machine Learning E-leanring Applications

Nowadays, everyone wants to learn and develop his knowledge in many fields, students, employers, etc. With the spread of lifelong learning, education systems are seriously facing the modernisation and e-learning is becoming more and more popular. All this leads to a massive growth in the number of Technology Enhanced Learning Environments (TELE) offering open or private online courses and other different types of services. Ana-

lysing the large amount of data produced by TELE through machine learning methods has emerged. It is useful to study how to exploit this powerful, new technology to enhance e-learning.

3.1 Sentiment analysis

Recently, Massive Open Online Course (MOOC) success is considered as the extent of student satisfaction with the course [18]. Sentiment analysis can be used to identify complex emotions [19] aiming the prediction of learner satisfaction. In [19], researchers want to determine throw forum messages in MOOC the polarity of learners' sentiments, positive sentiments and negative sentiments. They compare five supervised machine learning algorithms which have been used more frequently in contributions related to prediction in MOOCs: Logistic Regression, Support Vector Machine, Decision Tree, Random Forest and Naïve Bayes. Results show that the most reliable technique was Random Forest.

Understanding the role of emotions in MOOC students' learning experiences is very important. In one hand, according to [20] the control of achievement emotions may serve to improve learning engagement. Based on SVM, [20] build a supervised machine learning model in order to automatically categorise achievement emotions. SVM was adopted as it gives better performance results than Naïve Bayes, Logistic Regression and Decision Tree. On the other hand, [21] track the emotional tendencies of learners in order to analyse the acceptance of the courses using big data from homework completion, comments and forums. Based on semantic analysis and machine learning, [21] investigate the relationship between emotional tendencies and learning effects.

3.2. Student behaviour prediction

An interesting literature review [22] has addressed the question of machine learning use in predicting student behaviour. Tow research goals were identified: student classification and dropout prediction.

3.2.1 Student classification:

Certainly, personalities, backgrounds knowledge, skills and preferences play a crucial part in the learning process. Recommender systems serve to give most suitable content to each learner. Profiling and classifying learners is a primordial task not only to personalize learning but also to identify abandonment factors and many other purposes. We summarise in table 1, some recent works focusing on the student classification using machine learning.

TABLE I
STUDENT CLASSIFICATION

Paper	Machine Learning Algorithm	Classification goal	Results
[23]	k-means Support Vector Machine (SVM) Naïve Bayes	Classification of engaged and disengaged faces of students with dyslexia	accuracy with 97–97.8%
[24]	Backpropagation (BP), Support Vector Machine (SVM), Gradient Boosting Classifier (GBC)	classification of student performance	Accuracy: BP = 87.78%, 83.20%= 83.20%, GBC= 82.44%
[25]	Decision Tree, Logistic regression, k-nn, SVM, random forest algorithms	Classification of successful and unsuccessful students	K-nn gives the higher accuracy = 85%
[26]	K-modes clustering algorithm Naive Bayes classifier	Classification of learner's learning style	Accuracy = 89%

3.2.2 Dropout prediction:

Various machine learning techniques have been applied to analyse interactive behaviour traces left across TELE. According to [27] who focuses on learners' clickstream data, Logistic regression (LR) has been the most frequently used technique to predict student dropout in MOOC environment achieving 89% as accuracy. SVM and Decision Tree occupy the second position, however, Natural Language Processing Technique come in the third place.

3.3 Self-Regulated Learning

With the little external teacher's monitoring in majority of TELE, learners are required to make decisions related to their own activities [28]. In that case, individuals with strong self-regulated learning (SRL) skills, characterized by the ability to plan, manage and control their learning process, can learn faster and better than those with weaker SRL skills [29]. As it is one of e-learning platforms supporting SRL strategies [30], MOOC aims learners to self-evaluate the quality or the progress of their work, to set goals and plan and give them the possibility to reread notes, logs, tests, or learning materials to prepare for testing, etc. Despite all those features, it remains important for many researchers to enhance student SRL based on machine learning approach.

Based on learners' log traces and responses to a survey, [31] contribute to enhance understanding of how students learn, and how instruction should be designed to support SRL in an asynchronous online courses at a women's university in South Kore. In this study, researchers proceed to the discovery of student profiles and the examination of student SRL process overtime. At first, they suggested three key SRL attributes; time investment in content learning, study regularity and help-seeking that apply to asynchronous online courses to serve as the basis for the analytics of SRL, and guided the selection of log variables. Second, they identified student subpopulations using K-medoids clustering algorithm by silhouette method. After discovering existing clusters and their learning patterns, [31] use random forest classification as a decision tree-based machine learning algorithm to predict cluster membership by referring to each week's log variable.

4 Conclusions

E-learning researchers have spent considerable effort on analysing learners' data through machine learning methods to enhance learning experience. This seems to be wise since the learner is considered as the main component in the e-learning sphere. However, no research work, the best of our knowledge does carry out in order to use learning data in order to measure content quality in order to improve it.

Thus, in our future work will focus on e-learning content evaluation by using machine learning. The main objective is to help course designers in the educational reengineering process based on machine learning finding and based on many factors, especially past learners' interactions.

References

[1] Rakhmetulayeva, S. B., Duisebekova, K. S., Mamyrbekov, A. M., Kozhamzharova, D. K., Astaubayeva, G. N., & Stamkulova, K. (2018). Application of classification algorithm based on SVM for determining the effectiveness of treatment of tuberculosis. Procedia computer science, 130, 231-238.

[2] Kabyshev, M. V., & Kovalchuk, S. V. (2019). Development of personalized mobile assistant for chronic disease patients: diabetes mellitus case study. Procedia Computer Science, 156, 123-133.

[3] Zhu, G., Wu, Z., Wang, Y., Cao, S., & Cao, J. (2019). Online Purchase Decisions for Tourism E-commerce. Electronic Commerce Research and Applications, 100887.

[4] Brik, B., Bettayeb, B., Sahnoun, M. H., & Duval, F. (2019). Towards Predicting System Disruption in Industry 4.0: Machine Learning-Based Approach. Procedia Computer Science, 151, 667-674.

[5] Han, Y., Zeng, Q., Geng, Z., & Zhu, Q. (2018). Energy management and optimization modeling based on a novel fuzzy extreme learning machine: Case study of complex petrochemical industries. Energy conversion and management, 165, 163-171.

[6] Hew, K. F., Hu, X., Qiao, C., & Tang, Y. (2019). What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach. Computers & Education, 103724.

[7] Hmedna, B., El Mezouary, A., & Baz, O. (2019). How Does Learners' Prefer to Process Information in MOOCs? A Data-driven Study. Procedia computer science, 148, 371-379.

[8] Birjali, M., Beni-Hssane, A., & Erritali, M. (2017). Machine learning and semantic sentiment analysis based algorithms for suicide sentiment prediction in social networks. Procedia Computer Science, 113, 65-72.

[9] Kumari, K. V., & Kavitha, C. R. (2019). Spam Detection Using Machine Learning in R. In International Conference on Computer Networks and Communication Technologies (pp. 55-64). Springer, Singapore.

[10] Ghoddusi, H., Creamer, G. G., & Rafizadeh, N. (2019). Machine learning in energy economics and finance: A review. Energy Economics, 81, 709-727.

[11] Liu, J., Kong, X., Zhou, X., Wang, L., Zhang, D., Lee, I., ... & Xia, F. (2019). Data Mining and Information Retrieval in the 21st century: A bibliographic review. Computer Science Review, 34, 100193.

[12] Portugal, I., Alencar, P., & Cowan, D. (2018). The use of machine learning algorithms in recommender systems: A systematic review. Expert Systems with Applications, 97, 205-227

[13] Alzubi, J., Nayyar, A., & Kumar, A. (2018, November). Machine learning from theory to algorithms: an overview. In Journal of Physics: Conference Series (Vol. 1142, No. 1, p. 012012). IOP Publishing.

[14] Schrider, D. R., & Kern, A. D. (2018). Supervised machine learning for population genetics: a new paradigm. Trends in Genetics, 34(4), 301-312.

[15] Rodriguez-Nieva, J. F., & Scheurer, M. S. (2019). Identifying topological order through unsupervised machine learning. Nature Physics, 1.

[16] Oliver, A., Odena, A., Raffel, C. A., Cubuk, E. D., & Goodfellow, I. (2018). Realistic evaluation of deep semi-supervised learning algorithms. In Advances in Neural Information Processing Systems (pp. 3235-3246).

[17] François-Lavet, V., Henderson, P., Islam, R., Bellemare, M. G., & Pineau, J. (2018). An introduction to deep reinforcement learning. Foundations and Trends® in Machine Learning, 11(3-4), 219-354.

[18] Hew, K. F., Hu, X., Qiao, C., & Tang, Y. (2019). What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach. Computers & Education, 103724.

[19] Moreno-Marcos, P. M., Alario-Hoyos, C., Muñoz-Merino, P. J., Estévez-Ayres, I., & Kloos, C. D. (2018, April). Sentiment Analysis in MOOCs: A case study. In 2018 IEEE Global Engineering Education Conference (EDUCON) (pp. 1489-1496). IEEE.

[20] Xing, W., Tang, H., & Pei, B. (2019). Beyond positive and negative emotions: Looking into the role of achievement emotions in discussion forums of MOOCs. The Internet and Higher Education, 100690.

[21] Wang, L., Hu, G., & Zhou, T. (2018). Semantic analysis of learners' emotional tendencies on online MOOC education. Sustainability, 10(6), 1921.

[22] de Souza, V. F., & Perry, G. (2019). Identifying student behavior in MOOCs using Machine Learning. International Journal of Innovation Education and Research, 7(3), 30-39.

[23] Hamid, S. S. A., Admodisastro, N., Manshor, N., Kamaruddin, A., & Ghani, A. A. A. (2018, February). Dyslexia adaptive learning model: student engagement prediction using machine learning approach. In International Conference on Soft Computing and Data Mining (pp. 372-384). Springer, Cham.

[24] Sekeroglu, B., Dimililer, K., & Tuncal, K. (2019, March). Student performance prediction and classification using machine learning algorithms. In Proceedings of the 2019 8th International Conference on Educational and Information Technology (pp. 7-11). ACM.

[25] Fedushko, S., & Ustyianovych, T. (2019, January). Predicting pupil's successfulness factors using machine learning algorithms and mathematical modelling methods. In International Conference on Computer Science, Engineering and Education Applications (pp. 625-636). Springer, Cham.

[26] EL AISSAOUI, O., EL MADANI, Y. E. A., OUGHDIR, L., & EL ALLIOUI, Y. (2019). Combining supervised and unsupervised machine learning algorithms to predict the learners' learning styles. Procedia computer science, 148, 87-96.

[27] Dalipi, F., Imran, A. S., & Kastrati, Z. (2018, April). MOOC dropout prediction using machine learning techniques: Review and research challenges. In 2018 IEEE Global Engineering Education Conference (EDUCON) (pp. 1007-1014). IEEE.

[28] Wong, J., Baars, M., Davis, D., Van Der Zee, T., Houben, G. J., & Paas, F. (2019). Supporting self-regulated learning in online learning environments and MOOCs: A systematic review. International Journal of Human–Computer Interaction, 35(4-5), 356-373.

[29] Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. Computers & education, 104, 18-33.

[30] Garcia, R., Falkner, K., & Vivian, R. (2018). Systematic literature review: Self-Regulated Learning strategies using e-learning tools for Computer Science. Computers & Education, 123, 150-163.

[31] Kim, D., Yoon, M., Jo, I. H., & Branch, R. M. (2018). Learning analytics to support self-regulated learning in asynchronous online courses: A case study at a women's university in South Korea. Computers & Education, 127, 233-251.

Cooperative Reinforcement Multi-Agent Learning System For Sleep Stages Classification

Université de Tunis, ISGT, SMART Lab, 2000, Le Bardo, Tunisia.

1st Rahma Ferjani
Tunis, Tunisia
ferjanirahma@yahoo.com
0000-0002-7954-585X

2nd Lilia Rejeb
Tunis, Tunisia
lilia.rejeb@isg.rnu.tn
0000-0002-5740-1556

3rd Lamjed Ben Said
Tunis, Tunisia
lamjed.bensaid@isg.rnu.tn
0000-0002-5740-1556

***Abstract*—Sleep analysis is considered as an important process in sleep disorders identification and highly dependent of sleep scoring. Sleep scoring is a complex, time consuming and exhausting task for experts. In this paper, we propose an automatic sleep scoring model based on unsupervised learning to avoid the pre-labeling task. Taking advantage of the distributed nature of Multi-agent Systems (MAS), we propose a classification model based on various physiological signals coming from heterogeneous sources. The proposed model offers a totally cooperative learning to automatically score sleep into several stages based on unlabeled data. The existing heterogeneous adaptive agents are dealing with a dynamic environment of various physiological signals. The efficiency of our approach was investigated using real data. Promising results were reached according to a comparative study carried out with the often used classification models. The generic proposed model could be used in fields where data are coming from heterogeneous sources and classification rules are not predefined.**

***Index Terms*—Sleep Scoring, Reinforcement Learning, Cooperative Agents, Learning Classifier Systems**

Sleep analysis has monopolized the interest of different areas of research in the last decades [Zhang et al., 2016], [Zhang et al., 2018], [Alickovic and Subasi, 2018], [Phan et al., 2018] due to the importance of healthy sleepiness in establishing a fit life routine. Sleep disposes of several serious abnormalities that threat the individual's life such as sleep apnea, insomnias, hyper-somnias, sleep-related breathing disorders, etc. Thus, sleep analysis is primordial in several diseases' diagnosis and treatments. A mandatory step in sleep analysis is sleep scoring that aims to identify the different stages of a human sleep according to different standards. Most researches assumed that the Rechtschaffen and Kales rules are considered as an efficient and most used standard for sleep scoring, thus we identify six sleep stages: Awake stage (W), Non Rapid Eye-Movement (NREM1, NREM2, NREM3, NREM4: from light to deep sleep) and Rapid Eye-Movement (REM) sleep stage [Quan et al., 1999]. The scoring process is based on polysomnographic recordings (PSG) extracted from a human subject during sleepiness. The most used physiological signals are Electroencephalogram (EEG), Electromyogram (EMG), Electrooculogram (EOG), Electrocardiogram (ECG), and other indicators such as blood pressure and body temperature. The recordings are traditionally divided into segments of 30-second each, called epochs [Loomis et al., 1938] that are observed and staged manually by a clinician. However, experts need to observe at least the recording 8-hours of continuous sleep and identify over 950 epochs. The process is time-consuming, complex and subjective since it is highly dependent of the expert's preferences. Therefore, many works tried to propose approaches that automatically score sleep based on different standards, criteria and machine learning techniques. Among the most used techniques, we distinguish the Artificial Neural Networks (ANNs) where some works [Sinha, 2008], [Kurt et al., 2009] scored sleep to three main stages: Drowsy, REM sleep and Awake and succeed to reach high rates of successful classification within the 90%. Other works considered five sleep stages: Awake, REM, NREM1, NREM2 and sleep waves sleep (SWS) [Hsu et al., 2013], [Becq et al., 2005], [Ebrahimi et al., 2008]. All the previous detailed works were based on a mono EEG recording channel and different ANNs. In order to ameliorate the accuracy rate, other works [Özşen, 2013], [Akin et al., 2008], [Tagluk et al., 2010], [Becq et al., 2005] switched to multiple recording channels where other signals rather than EEG signals were used. The work of Ozsen et al. [Özşen, 2013] was based an ANN architecture composed of five multilayer perceptrons ANNs to stage sleep into five stages. They considered three recording channels belonging, respectively, to EMG, EOG, ECG signals. The obtained rate of successful classification was estimated to be 90.93%. Akin et al. [Akin et al., 2008] had considered EEG and EMG signals and a regression ANN model to reach a 99% of accuracy. In [Tagluk et al., 2010], the authors used EEG, EMG and EOG signals to perform their feed-forward ANN and attained 74.7% as an accuracy rate. Artificial Neural Networks were widely used to score sleep and depicted high rated of successful classification but they remain unable to build explicit classification models that deliver explicable results. Consequently, the obtained results could not be evaluated and approved a posteriori by field's experts. Therefore, Chapotot et al. [Chapotot and Becq, 2010] combined an artificial neural network architecture and flexible decision rules to score EEG recordings. In [Yildiz et al., 2009], Yildiz et al. presented an Adaptive Neuro Fuzzy Inference

System (ANFIS) that succeed to deliver an explicit classification model obtaining an accuracy rate of 98.34%. Koley et al. [Koley and Dey, 2012] also proposed a hybrid approach for sleep stages classification by combining Support Vector Machine (SVM) and the One Against All (OAA) and achieved 85% of agreement. To deliver an explicit model, Youssef et al. [Khaoula and Lilia, 2012] combined a Multi-agent System (MAS) and two different Learning Classifier Systems (LCS): eXtended Classifier System(XCS) and sUpervised Classifier System (UCS). The model delivered a set of rules reaching 98% of successful classification.

The majority of the existing approaches are supervised methods that depend, highly, on pre-labeled data which considered as an exhausting task. Thus, few works [Rodríguez-Sotelo et al., 2014], [Grube et al., 2002] proposed unsupervised approaches to assist the work of experts. Ouanes et al. [Ouanes and Rejeb, 2016] proposed an unsupervised classification model by combining LCSs and ANN to identify sleep stages. The classification was based on unlabeled EEG recordings and the model reached 78.7% of classification agreement. Boostani et al. [Boostani et al., 2017] affirmed that even with the changes of the new sleep manual AASM the identification of some sleep stages requires extra information in addition to EEG. In fact, for the awake sleep stage labeling a high EMG tone and frequent eye movement are present. However, in REM stage lowest muscle movement and rapid eye movement are observed. Zoubek et al. [Zoubek et al., 2007] used several classification models trained first by features extracted from only EEG signals to obtain 71% of agreement, then, they added features extracted from EMG and EOG signals to attend about 80% of agreement.

Although the different proposed works for sleep stages classification, there is no consensus on a specific framework for sleep scoring that can handle multiple recording channels, deliver explicable results and generate a relative rules' background. Indeed, the number of the used PSG recordings is very limited, restricted mainly to EEG, EOG and EMG, and most of the works were based only on a single recording channel of EEG [Schulz, 2008]. However, the using of multi-signals and multi-derivation of EEG signal leaded to a better accuracy as shown in [Chapotot and Becq, 2010], [Kurt et al., 2009], [Akin et al., 2008], [Piñero et al., 2004]. Generally, the proposed approaches are unable to give results that could be interpreted a posteriori by experts [Vural and Yildiz, 2010], [Akin et al., 2008], [Yildiz et al., 2009].

Taking advantage of the distributed nature of multi-agent systems and specifically the cooperative MAS that ensures an interactive learning among the different existing agents, we propose a Cooperative Reinforcement Multi-Agent Learning (CRMAL) system for sleep stages classification based on unlabeled data. Multi-agent learning and Reinforcement Learning (RL) have shown their efficiency in solving several real world problems [Claus and Boutilier, 1998], [Abbeel et al., 2007], [Abdallah and Lesser, 2008], [Busoniu et al., 2008] where the agent learns interactively how to solve complex problems referring to the environment sensors and feedback. Among the RL techniques, we distinguish the Learning Classifier Systems (LCS) that were successfully applied in different fields thanks to their capability of building explicit models by generating, automatically, their rules set [Urbanowicz and Moore, 2009]. Therefore, we are interested in LCS to build our adaptive agents. The structure of this paper is as follows: Section 2 details the used data set and the materials of the paper. Section 3 presents the learning classifier systems, their architecture and functioning. Section 4 introduces the proposed system for sleep stages classification. In Section 5, we present the parameters settings and the discussion of the obtained results.

I. Materials

A. DataSet Description

The data set used for sleep analysis is provided by the physio-bank site which contains an important number of databases considered as references for researches working in the medical area [Ouanes and Rejeb, 2016], [Zhang et al., 2018], [Kemp et al., 2000]. The sleep data set is formed of 61 polysomnograms (PSGs), each one is accompanied with its annotations file which contains the labels assigned by experts for each epoch. The PSG files contain EEG including Fpz_{Cz} and Pz_{Oz} electrodes location, submental chin EMG and EOG signal records for a whole night for a given subject. They are used to score sleep visually based on Rechtschaffen and Kales rules. Almost 20-hours recordings were divided into 30-s epochs. In order to increase the accuracy rate, we created a balanced learning set which contains 1901 epochs for each sleep stage obtained from 20 healthy subjects (10 males and 10 females). The demographic range was between 25 and 34. The physiological signals are stored in a file with a special format which is EDF (European Data Format). This extension represents advantages for researches since it becomes an easy way to store multichannel biological and physical signals.

B. Pre-processing and feature extraction

Sleep scoring is given by three main steps: Data pre-processing, feature extraction and classification. In this section, we detail the two prior steps of sleep scoring and the used techniques. Since the physiological data may contain artifacts occurred from the body movements and eyelids during sleep, a pre-processing phase is required to obtain only useful data. For that, we used the Blind Source Separation (BSS) technique usually used on signals [Cardoso, 1998], [Kemp et al., 2000] to separate the data spatial components in addition to the SOBI algorithm to seek the pertinent spatial components [Klemm et al., 2009] as indicated in Figure 1.

The second phase is the feature extraction that allows the extraction of the pertinent classification inputs. To do so, we used the discrete wavelet transform (DWT) extraction technique since it assumes the signal to be non stationary and considers the time variations as the frequency ones. According the to literature, some measures are the most pertinent thus we computed some statical parameters for each used signal.

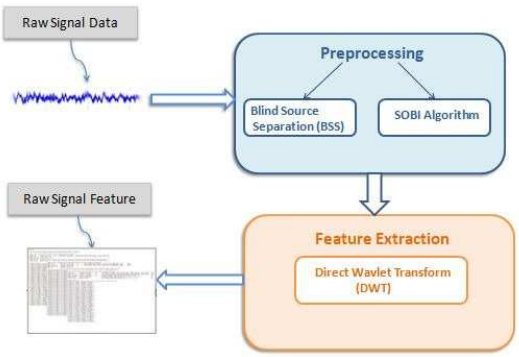


Fig. 1. Preprocessing and Feature Extraction.

II. LEARNING CLASSIFIER SYSTEMS

Learning Classifier Systems (LCS) are rule-based systems that combine evolutionary computation with artificial intelligence [Urbanowicz and Moore, 2009]. Their generalization property and the evolutionary process make them able to efficiently overcome the brittleness of classic expert systems by allowing an automatic building of a rule set. LCS are reinforcement technique that were successfully applied in classification problems, sequential decision problem resolution, knowledge discovery problem and autonomous robotics. Bacardit et al. used LCS to predict protein structure in both [Bacardit et al., 2007] and [Stout et al., 2009]. Llora and al. in [Llorà et al., 2007] were based on LCS to diagnose prostate cancer by classifying Infrared Spectroscopic Image Data.
Knowledge in LCS is represented through a population classifiers. Each classifier is given by three main parameters: condition, action, and a reward. LCS have three types namely: Strength-based systems: **Zeroth Level Classifier System** (ZCS), Anticipation-based systems: **Anticipatory Learning Classifier System** (ALCS) and Accuracy-based systems: **eXtend Classifier Systems** (XCS). We are mainly interested in the XCS since they are the most adequate LCS to real world complex problems.

A. eXtended Classifier Systems

The eXtended Classifier System XCS proposed by (Wilson,1995) is a reinforcement learning technique suitable for real world complex problems [Wilson, 1995]. XCS uses standard classifier condition-action rules where a condition is presented by a set of binary values. However, Wilson proposed an extension of XCS to surpass this limit [Wilson, 2000a]. The so-called XCSR deals with numeric inputs which is more appropriate to real world problems. Thus, Each classifier is a condition-action rule given as follows:

- The condition: Is a concatenation of intervals where each interval refers to a specific attribute of the epoch instance i. The interval is represented as follows: $[c_i - s_i, c_i + s_i]$ where c is the center value and s is the spread. For each new epoch instance i, a comparison is made to decide either the classifier's condition matches the input X_i or not. An input X_i is represented by a set of values $x_1 ... x_n$. If each value x_i verifies the condition: $((c_i - s_i) \leq x_i \leq (c_i + s_i))$ then the input matches. The input X_i consists of the current epoch features extracted using DWT. For each signal agent, an epoch is given by 79 extracted features.
- The action: Each classifier proposes an action that refers to a sleep stage (Wake, NREM1, NREM2, NREM3, NREM4, and REM) and these stages are represented respectively by integers between 0 and 5. For example: 0 refers to the Wake stage, 1 refers to the NREM1 stage, etc.

At the presence of a new input, XCSR builds a matching set $[M]$ of classifiers having the same condition of the input. In the case where the number of rules in $[M]$ is lesser than the given threshold, the covering operation is performed to generate new classifiers. A prediction value is calculated for each action in $[M]$ among all the classifiers proposing that action in the action set $[A]$. To finally choose an action, two methods are proposed: Exploration (choosing randomly an action among the possible actions) or Exploitation (nominating the action having the highest prediction value). A reward r is returned as a consequence of the made decision to update the parameters of $[A]$. In order to illustrate the functioning of XCS, we propose the scheme in Figure 2. Each classifier is given by a set of fundamental parameters: prediction reward, prediction error and fitness. The parameters are automatically updated depending on the reward. the latter is given with reference to the proposed action by two possible values: Max payoff (if the proposed action corresponds to the right action) and Min payoff (if the proposed action is wrong). After a

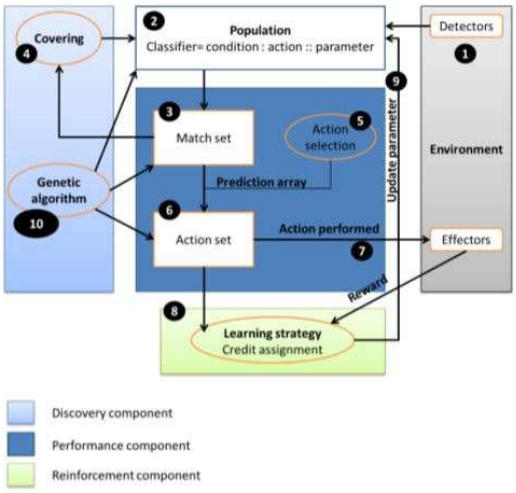


Fig. 2. eXtended Classifier System Architecture [Urbanowicz and Moore, 2009].

number of predefined iterations, the Genetic Algorithm (GA) is performed to update the action sets, to generate new rules for the population and take off useless classifiers if the population is full [Bernadó-Mansilla and Garrell-Guiu, 2003].

III. PROPOSED SOLUTION

With the objective of establishing an unsupervised sleep classification system based on different physiological signals having heterogeneous natures, our system disposes of two types of agents: Adaptive agents and a supervisor agent. Figure 3 presents the global architecture of the proposed system for a specific epoch's instance i. Each adaptive agent

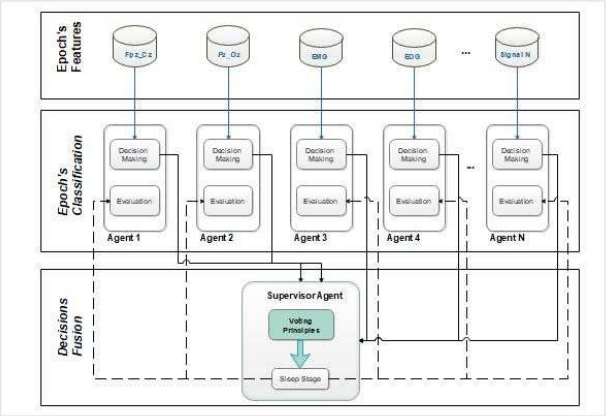


Fig. 3. Unsupervised multi-agent system for sleep scoring: Global Architecture.

is responsible of a specific physiological signal encapsulating each a XCS classifier. Hence, adaptive agents are given by two main components: Decision making component and Evaluation component. Figure 4 depicts the architecture of an adaptive agent. For each new epoch instance i, each adaptive

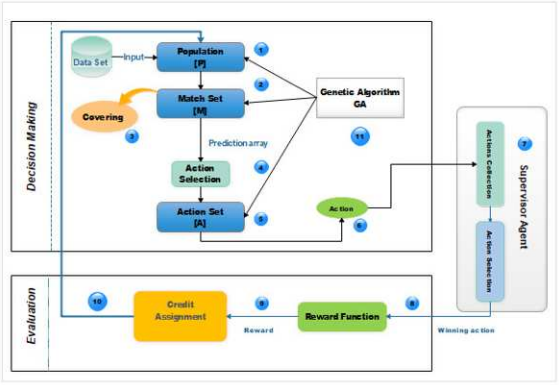


Fig. 4. Architecture of an XCS Agent.

agent performs its decision process component, given by steps from 1 to 6, to propose the adequate sleep stage. The decision making policy balanced between the total exploration and the exploitation operations. Once each adaptive agent j decide on the common entry instance i, it sends the proposed action W_{ij} to the supervisor agent as cleared in step 7. The latter is responsible of managing the interactive learning of the adaptive agents. The supervisor agent ensures the parallel learning of the existing agents for each instance and handles the reception of all their proposed actions to decide on the wining action WA_i for epoch instance i.

Each adaptive agent receives the wining action for the current epoch from the supervisor agent and then performs its evaluation component according to its reward r. Usually, in XCS functioning the reward is given with reference to the proposed action by two possible values: Max_{reward} (if the proposed action corresponds to the right action) and Min_{reward} (if the proposed action is wrong). Instead of assigning either one of two standard rewards, we propose an adaptive reward function:

$$r \leftarrow Max_{reward} - \alpha \ * \ |WA_i \ - \ A_{ij}|$$

where ($Min_{reward} \leq \alpha \leq Max\_reward$) is the learning rate. Each adaptive agent will be evaluated accordingly to how much its proposed action is close to the wining one. Once the reward is assigned, step nine is performed to adjust the learning parameters of the signal agent j at epoch i [Wilson, 2000b]: the prediction P_i, the prediction error ϵ_i, the classifier accuracy K_i and the fitness F_i.

IV. EXPERIMENTS AND RESULTS

A. Parameters Setting

In order to evaluate the obtained results, we used the Percentage of Correct Classification (PCC) and the confusion matrix. The performance of XCS depends on the setting of some parameters [Urbanowicz and Moore, 2009]. After some preliminary tests, we found that the following parameters are discriminant in our work:

- Population size = 6400,
- The exploration/exploitation rate =0.5,
- Number of iterations 100,000,
- Mutation probability μ=0.06,
- Genetic Algorithm $\theta_{Genetic}$=50.

During our experiments, we noted that the eXtended classifier system is very sensitive to some of its parameters. We are interested in the more representative ones which are the population size, number of iterations and exploration/exploitation rate. Table I shows the impact of population size variation in the PCC. The size of population may increase the performance of XCS. In our case we retain a population size of 6400 since it leads to the best result.

Table II depicts the influence of the exploration probability

TABLE I
SENSITIVITY ANALYSIS OF POPULATION SIZE PARAMETER.

Pop / PCC	1000	3000	6400
	32.02	54.63	96

on the successful classification rate. We can conclude that high exploration rates impact the obtained results and prevents the analyze of the obtained knowledge during the learning. Thus, we choose to work with an equal probability of 50%

for exploration and exploitation operations based on results given in Table II. It is obvious that the number of iterations

TABLE II
SENSITIVITY ANALYSIS OF EXPLORATION COEFFICIENT (Ex_{Cf}).

PCC \ Ex_{Cf}	0.2	0.5	0.8
	88	96	25.06

affects the performance as shown in Table III where increasing its value may lead to a high rate of accuracy. Based on the results illustrated in Table III, we retain the value of 100.000 iteration since it leads to the best PCC.

TABLE III
SENSITIVITY ANALYSIS OF ITERATION NUMBER (Nb_{It}).

PCC \ Nb_{It}	30,000	80,000	100,000
	35.1	45.2	96

B. Results and discussion

In the aim of evaluating the performance of our distributed approach, we implemented, based on the same real data, a single-agent classification model. The latter is composed of one XCS classifier having, like input, an epoch stream. Each stream is composed of N labeled epochs that refer to N physiological signal sharing the same instance number. The one-agent classification model showed a percentage of correct classification of 63%. Table IV depicts a comparison between the two approaches by presenting the PCC of each sleep stage.

The PCC provided by the proposed MAS for Awake and NREM2 stages are higher than those delivered by the single agent classification model which can be clarified by Table V and Table VI. These tables show, respectively, the confusion matrices obtained from a single agent classification and the results reached using our proposed approach. As can be seen in Table VI among 1901 epochs of NREM1, 398 epochs are classified as REM while only 159 epochs of NREM1 are classified as REM using the adaptive MAS. Thus, we notice that our approach is efficient to minimize the classification's conflict usually obtained when classifying epochs as NREM1 or REM stages. Indeed, it is always difficult to distinguish between NREM3 and NREM4 [Berry et al., 2012] as shown in Table V. However, our method succeed to, correctly, classify 1381 epochs among 1901 NREM3 epochs as NREM3 and 1479 epoch among 1901 NREM4 epoch as NREM4 which increases the scoring precision and avoid the side effects of dangerous classification on sleep analysis. One of the main causes of the classification success is the cooperation of different physiological signals especially the EOG signal that helps to avoid the confusion between the Awake and NREM2 stages. Among the works that used the Sleep-EDF database for classification, we note the work of Bajaj and al. [Bajaj and Pachori, 2013]. They used the existing PSG signals but based on labeled epochs and reached 92.93% as accuracy rate which reflects the influence and importance of the use of multiple PSG signals. Indeed, our approach, based on the same data, used unlabeled epochs and reached an accuracy rate of 96% which reflects the advantage of using less prelabeled data.
In order to evaluate the performance of using unlabeled data, we established a comparison with the work of Ouanes and Rejeb [Ouanes and Rejeb, 2016]. The comparison was established based on two major criteria. First, the so-called XCSR-based SOM used an eXtended classifier system to train its classifiers. Second, they proposed an unsupervised approach based on EEG signals extracted from the same data set. They used only the fronto-central Fpz_{Cz} channel of EEG. The data set was based on an unlabeled balanced data set composed of 9000 epochs includes 1500 epochs of each sleep stage. They reached an over all accuracy of 78.7%. As can be seen, the using of a set of different PSG signals as well as the interaction among them have enhanced the performance of the sleep stages classification.

V. CONCLUSION

We have presented a cooperative reinforcement multi-agent learning for sleep stages classification based on unlabeled data. The effectiveness of the proposed approach was tested on real physiological signals. We reached a classification accuracy of 96%. The multi-agent classification system proved its performance in delivering explicit results that could be evaluated a posteriori by field's experts. Indeed, the proposed system provided a flexible architecture that can handle a very large number of learning agents based on different kinds of learning techniques. The resulted classification system could prove its efficiency in solving other problems where standards are not pre-known and inputs are coming from different heterogeneous sources such as economy problems, epidemiology, medical diagnosis and emotion recognition.

REFERENCES

[Abbeel et al., 2007] Abbeel, P., Coates, A., Quigley, M., and Ng, A. Y. (2007). An application of reinforcement learning to aerobatic helicopter flight. In *Advances in neural information processing systems*, pages 1–8.

[Abdallah and Lesser, 2008] Abdallah, S. and Lesser, V. (2008). A multiagent reinforcement learning algorithm with non-linear dynamics. *Journal of Artificial Intelligence Research*, 33:521–549.

[Akin et al., 2008] Akin, M., Kurt, M. B., Sezgin, N., and Bayram, M. (2008). Estimating vigilance level by using eeg and emg signals. *Neural Computing and Applications*, 17(3):227–236.

[Alickovic and Subasi, 2018] Alickovic, E. and Subasi, A. (2018). Ensemble svm method for automatic sleep stage classification. *IEEE Transactions on Instrumentation and Measurement*, 67(6):1258–1265.

[Bacardit et al., 2007] Bacardit, J., Stout, M., Hirst, J. D., Sastry, K., Llorà, X., and Krasnogor, N. (2007). Automated alphabet reduction method with evolutionary algorithms for protein structure prediction. In *Proceedings of the 9th annual conference on Genetic and evolutionary computation*, pages 346–353. ACM.

[Bajaj and Pachori, 2013] Bajaj, V. and Pachori, R. B. (2013). Automatic classification of sleep stages based on the time-frequency image of eeg signals. *Computer methods and programs in biomedicine*, 112(3):320–328.

[Becq et al., 2005] Becq, G., Charbonnier, S., Chapotot, F., Buguet, A., Bourdon, L., and Baconnier, P. (2005). Comparison between five classifiers for automatic scoring of human sleep recordings. *Classification and Clustering for Knowledge Discovery*, pages 113–127.

TABLE IV
COMPARISON OF THE PCC OF EACH SLEEP STAGE BETWEEN MAS AND SINGLE-AGENT.

	MAS	Single Agent
Awake	91	78.95
NREM1	78.49	68.83
NREM2	78.96	57.68
NREM3	72.65	27.84
NREM4	77.8	32.87
REM	74.75	77.01

TABLE V
CONFUSION MATRIX FOR SINGLE XCS.

Real class \ Estimated Class	Awake	NREM1	NREM2	NREM3	NREM4	REM
Awake	1501	359	9	12	9	11
NREM1	51	1300	115	33	4	398
NREM2	7	370	1100	354	9	61
NREM3	34	16	103	523	1219	6
NREM4	19	2	5	1250	625	0
REM	8	353	64	9	3	1464

TABLE VI
CONFUSION MATRIX FOR THE PROPOSED MULTI-AGENT SYSTEM.

Real class \ Estimated Class	Awake	NREM1	NREM2	NREM3	NREM4	REM
Awake	1730	109	12	20	18	12
NREM1	45	1492	144	47	14	159
NREM2	8	154	1501	168	13	57
NREM3	35	23	113	1381	339	10
NREM4	19	5	8	389	1479	1
REM	14	322	125	10	9	1421

[Bernadó-Mansilla and Garrell-Guiu, 2003] Bernadó-Mansilla, E. and Garrell-Guiu, J. M. (2003). Accuracy-based learning classifier systems: models, analysis and applications to classification tasks. *Evolutionary computation*, 11(3):209–238.

[Berry et al., 2012] Berry, R. B., Brooks, R., Gamaldo, C. E., Harding, S. M., Marcus, C., and Vaughn, B. (2012). The aasm manual for the scoring of sleep and associated events. *Rules, Terminology and Technical Specifications, Darien, Illinois, American Academy of Sleep Medicine*.

[Boostani et al., 2017] Boostani, R., Karimzadeh, F., and Nami, M. (2017). A comparative review on sleep stage classification methods in patients and healthy individuals. *Computer methods and programs in biomedicine*, 140:77–91.

[Busoniu et al., 2008] Busoniu, L., Babuska, R., and De Schutter, B. (2008). A comprehensive survey of multiagent reinforcement learning. *IEEE Trans. Systems, Man, and Cybernetics, Part C*, 38(2):156–172.

[Cardoso, 1998] Cardoso, J.-F. (1998). Blind signal separation: statistical principles. *Proceedings of the IEEE*, 86(10):2009–2025.

[Chapotot and Becq, 2010] Chapotot, F. and Becq, G. (2010). Automated sleep–wake staging combining robust feature extraction, artificial neural network classification, and flexible decision rules. *International Journal of Adaptive Control and Signal Processing*, 24(5):409–423.

[Claus and Boutilier, 1998] Claus, C. and Boutilier, C. (1998). The dynamics of reinforcement learning in cooperative multiagent systems. *AAAI/IAAI*, 1998:746–752.

[Ebrahimi et al., 2008] Ebrahimi, F., Mikaeili, M., Estrada, E., and Nazeran, H. (2008). Automatic sleep stage classification based on eeg signals by using neural networks and wavelet packet coefficients. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 1151–1154. IEEE.

[Grube et al., 2002] Grube, G., Flexer, A., and Dorffner, G. (2002). Unsupervised continuous sleep analysis. *Methods Find Exp Clin Pharmacol*, 24(Suppl D):51–56.

[Hsu et al., 2013] Hsu, Y.-L., Yang, Y.-T., Wang, J.-S., and Hsu, C.-Y. (2013). Automatic sleep stage recurrent neural classifier using energy features of eeg signals. *Neurocomputing*, 104:105–114.

[Kemp et al., 2000] Kemp, B., Zwinderman, A. H., Tuk, B., Kamphuisen, H. A., and Oberye, J. J. (2000). Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the eeg. *IEEE Transactions on Biomedical Engineering*, 47(9):1185–1194.

[Khaoula and Lilia, 2012] Khaoula, Y. and Lilia, R. (2012). système multi-agents pour la classification des stades de sommeil basé sur les signaux physiologiques.

[Klemm et al., 2009] Klemm, M., Haueisen, J., and Ivanova, G. (2009). Independent component analysis: comparison of algorithms for the investigation of surface electrical brain activity. *Medical & biological engineering & computing*, 47(4):413–423.

[Koley and Dey, 2012] Koley, B. and Dey, D. (2012). An ensemble system for automatic sleep stage classification using single channel eeg signal. *Computers in biology and medicine*, 42(12):1186–1195.

[Kurt et al., 2009] Kurt, M. B., Sezgin, N., Akin, M., Kirbas, G., and Bayram, M. (2009). The ann-based computing of drowsy level. *Expert Systems with Applications*, 36(2):2534–2542.

[Llorà et al., 2007] Llorà, X., Reddy, R., Matesic, B., and Bhargava, R. (2007). Towards better than human capability in diagnosing prostate cancer using infrared spectroscopic imaging. In *Proceedings of the 9th annual conference on Genetic and evolutionary computation*, pages 2098–2105. ACM.

[Loomis et al., 1938] Loomis, A. L., Harvey, E. N., and Hobart III, G. A. (1938). Distribution of disturbance-patterns in the human electroencephalogram, with special reference to sleep. *Journal of Neurophysiology*, 1(5):413–430.

[Ouanes and Rejeb, 2016] Ouanes, A. and Rejeb, L. (2016). A hybrid approach for sleep stages classification. In *Proceedings of the 2016 on Genetic and Evolutionary Computation Conference*, pages 493–500. ACM.

[Özşen, 2013] Özşen, S. (2013). Classification of sleep stages using class-dependent sequential feature selection and artificial neural network. *Neural Computing and Applications*, 23(5):1239–1250.

[Phan et al., 2018] Phan, H., Andreotti, F., Cooray, N., Chén, O. Y., and De Vos, M. (2018). Joint classification and prediction cnn framework for automatic sleep stage classification. *IEEE Transactions on Biomedical*

Engineering.

[Piñero et al., 2004] Piñero, P., Garcia, P., Arco, L., Álvarez, A., Garcıa, M. M., and Bonal, R. (2004). Sleep stage classification using fuzzy sets and machine learning techniques. *Neurocomputing*, 58:1137–1143.

[Quan et al., 1999] Quan, S., Gillin, J. C., Littner, M., and Shepard, J. (1999). Sleep-related breathing disorders in adults: Recommendations for syndrome definition and measurement techniques in clinical research. editorials. *Sleep*, 22(5):662–689.

[Rodríguez-Sotelo et al., 2014] Rodríguez-Sotelo, J., Osorio-Forero, A., Jiménez-Rodríguez, A., Cuesta-Frau, D., Cirugeda-Roldán, E., and Peluffo, D. (2014). Automatic sleep stages classification using eeg entropy features and unsupervised pattern analysis techniques. *Entropy*, 16(12):6573–6589.

[Schulz, 2008] Schulz, H. (2008). Rethinking sleep analysis: Comment on the aasm manual for the scoring of sleep and associated events. *Journal of clinical sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine*, 4(2):99.

[Sinha, 2008] Sinha, R. K. (2008). Artificial neural network and wavelet based automated detection of sleep spindles, rem sleep and wake states. *Journal of medical systems*, 32(4):291–299.

[Stout et al., 2009] Stout, M., Bacardit, J., Hirst, J. D., Smith, R. E., and Krasnogor, N. (2009). Prediction of topological contacts in proteins using learning classifier systems. *Soft Computing*, 13(3):245.

[Tagluk et al., 2010] Tagluk, M. E., Sezgin, N., and Akin, M. (2010). Estimation of sleep stages by an artificial neural network employing eeg, emg and eog. *Journal of medical systems*, 34(4):717–725.

[Urbanowicz and Moore, 2009] Urbanowicz, R. J. and Moore, J. H. (2009). Learning classifier systems: a complete introduction, review, and roadmap. *Journal of Artificial Evolution and Applications*, 2009:1.

[Vural and Yildiz, 2010] Vural, C. and Yildiz, M. (2010). Determination of sleep stage separation ability of features extracted from eeg signals using principle component analysis. *Journal of medical systems*, 34(1):83–89.

[Wilson, 1995] Wilson, S. W. (1995). Classifier fitness based on accuracy. *Evolutionary computation*, 3(2):149–175.

[Wilson, 2000a] Wilson, S. W. (2000a). Get real! xcs with continuous-valued inputs. In *Learning Classifier Systems*, pages 209–219. Springer.

[Wilson, 2000b] Wilson, S. W. (2000b). Get real! xcs with continuous-valued inputs. In *Learning Classifier Systems*, pages 209–219. Springer.

[Yildiz et al., 2009] Yildiz, A., Akin, M., Poyraz, M., and Kirbas, G. (2009). Application of adaptive neuro-fuzzy inference system for vigilance level estimation by using wavelet-entropy feature extraction. *Expert Systems with Applications*, 36(4):7390–7399.

[Zhang et al., 2018] Zhang, B., Lei, T., Liu, H., and Cai, H. (2018). Eeg-based automatic sleep staging using ontology and weighting feature analysis. *Computational and mathematical methods in medicine*, 2018.

[Zhang et al., 2016] Zhang, J., Wu, Y., Bai, J., and Chen, F. (2016). Automatic sleep stage classification based on sparse deep belief net and combination of multiple classifiers. *Transactions of the Institute of Measurement and Control*, 38(4):435–451.

[Zoubek et al., 2007] Zoubek, L., Charbonnier, S., Lesecq, S., Buguet, A., and Chapotot, F. (2007). Feature selection for sleep/wake stages classification using data driven methods. *Biomedical Signal Processing and Control*, 2(3):171–179.

Modeling an agent-based cooperative dynamic behavior in a context of SME's sustainable supply chain under uncertainty

Ben Mekki Aicha
SMART Lab
Université de Tunis, Higher Institute of Management of Tunis
Bardo, Tunisia
aycha.bm@ gmail.com

Tounsi Jihene
SMART Lab
Université de Tunis, Higher Institute of Management of Tunis
Bardo, Tunisia
Tounsi.jihene@yahoo.fr

Ben said lamjed
SMART Lab
Université de Tunis, Higher Institute of Management of Tunis
Bardo, Tunisia
lamjed.bensaid@isg.rnu.tn

Abstract—**by drawing upon the concept of competitiveness and the dynamism of the environment, a knowledge base model is developed to design the behavior of autonomous actors, having different objectives, within a cluster of Small and Medium Enterprises' (SMEs'). These SMEs are designed as a group of cluster located in different geographical area, for that purpose, actors must collaborate in order to achieve the cluster's objective. This later is influenced by sustainable responsibilities and should be optimized over a feasible set of constraints. In this work, a mechanism for monitoring and evaluation is needed to assess the SME's cluster performance. For that, we model a multi-agent dynamic behavior as a dynamic multi-objective optimization (MOO) problem and resort a dynamic multi-objective evolutionary method to choose a common solution for all actors, and then, the cluster's preferences are operated to determine the most suitable one (posterior).**

Keywords—Multi-agent system, dynamic Multi-objective evolutionary algorithms (D-MOEA), Cluster of SMEs, uncertainty.

I. INTRODUCTION

Sustainable Supply Chain Management (SSCM) takes into account the environmental and social impact of supply chain operations, in addition to economic performance in the management of information, materials, and capital flow [Seuring and Müller, 2008]. [Zhou et al., 2000] highlighted the fact that these sustainable goals are often conflictual. For that, achieving a balance among a range of conflictive objectives pertaining to the supply chain is crucial for the design of SSCN.In a context of a Small and Medium Enterprises (SME), the ability to monitor the global sustainable supply chain, often across several clusters and to optimize decisions has been increasingly recognized as a crucial competitive factor, for this purpose, an SME's cluster needs to collaborate and integrate sustainability principles into their strategy and daily operations [Seuring and Müller, 2008],[Carter and Rogers, 2008] right from raw materials purchase stage to product distribution stage, in order to retain and strengthen their competitiveness in the market. In addition, the SMEs treated are located in a different geographical area, which makes the process more complex. According to our research context, developing an efficient decision making system needs a proper communication mechanism between decision makers in order to predict and evaluate the sustainable global performance.

II. RESEARCH OBJECTIVES

Our aim behind a monitoring mechanism is to find the best optimization process with which the sustainable supply chain can achieve a high level of performance. In this regard, a decentralized architecture is required in order to help achieving this process. For that, we had recourse to Multi-Agent Systems (MAS), where a set of agents endowed by a dynamic evolutionary behavior perform communication and coordination between each other in order to achieve the cluster's common goal. Thanks to its autonomous and reactive nature, intelligent agents can incorporate coordination strategies, thus enabling them to deal with the uncertainty in a dynamic environment and perform complex tasks [Bonhomme and Arnould, 2012]. In this regard, Belief-Desire-intention (BDI) agents, that have cognitive capabilities [Bratman et al., 1988], are used to highlight our dynamic behavior model. This later is modelled through our BDI agent behavior process starts from acting as an autonomous entity which is highlighted by a fuzzy BDI agent's dynamic individual behaviour [Ben Mekki et al, 2019], then, as a reactive entity that perform collaboration with other entities in order to achieve the SME cluster's common goal. Our focus for the remainder of this paper is on this dynamic collective behavior. In fact, this collective process highlights an approach of resolution that handles the issue of the simultaneous presence of several objectives (often conflicting) and constraints. Several issues lead us to treat our problem as a dynamic multi-objective optimization problem (DMOP) and this is explained by different challenges such as:

- At the level of each cluster, each BDI agent will have partial visibility of the variables to be instantiated and the objectives to be achieved, which can lead to conflicting situations. Thus, objectives cannot be attained at the same time.
- The simultaneous presence of several objectives, which are influenced by sustainable responsibilities, to be optimized over a feasible set of constraints.
- Furthermore, we are dealing with an uncertain environment in perpetual change. It is therefore difficult to manipulate the changes of situations each time to adapt them to the requirements imposed by the external environment.

For that, the main objective of our work is to model a collective dynamic behavior treating BDI agents as autonomous entities that interact together in order to achieve the cluster's common goal, with sustainable considerations, and helping them to deal with the uncertainty in a dynamic environment. In order to reach this objective, several issues mentioned above are treated thoroughly in this paper. The subsequent sections of this paper contain the following: Section 3 presents a review of the relevant literature on using dynamic multi-objective evolutionary algorithm to monitor a sustainable supply chain based on multi-agent system. The description of our knowledge base model is proposed in Section 4. The design of the collective dynamic behavior using dynamic multi-objective evolutionary algorithm is outlined in Section 5. Concluding and perspectives are given in Section 6.

III. RELATED WORKS:

A. *Decision making in SME's cluster within a SSC*

Decision making in SME's cluster within a SSC involves complex decisions and several autonomous actors, having different objectives, that collaborate together to achieve a common goal. In addition to that, it is influenced by economic, social and environmental responsibilities. As a result, extensive compromise and trade-offs due to inherent conflict among the different actors are involved which make the monitoring mechanism more complex. For that, Multi Objective Evolutionary Algorithms (MOEAs) have successfully been used in different supply chain problems, such as production and transportation planning [Zegordi et al., 2010], vendor-managed inventory problem [Diabat, 2014], lot-sizing and delivery scheduling [Torabi et al., 2006] and supply chain design and planning [Altiparmak et al., 2006]. However, most studies established single objective or summed up objectives with weights and adopted mixed linear integer programming methods. In this regard, [Saghaeeian and Ramezanian, 2018] proposed a hybrid genetic algorithm for multi-product competitive supply chain (SC) network using Karush-Kuhn-Tucker (KKT) conditions in order to transform the bi-level model into a single-level mixed integer nonlinear programming (MINLP). Other authors like [Stefansson et al. 2009; Zhou et al., 2009; Fahimnia et al., 2013c] have also followed the same strategy to solve a particular problem of a supply chain. Only a few researchers proposed the multi-objective optimization algorithms (MOEAs) to optimize the conflict objectives simultaneously. In fact, [Doolun et al., 2018] proposed a multi-objective mathematical model to solve location-allocation problem in a multi echelon supply chain network to optimize three objectives simultaneously. A data driven hybrid evolutionary analytical approach were used. [Azadeh et al, 2017] also presented a multi-objective mathematical model for integrating upstream and midstream segments of crude oil supply chain in the context of environmental indicators. A multi-objective evolutionary algorithm based on decomposition (MOEA-D) approach is employed to solve the proposed mixed integer nonlinear programming model. Our main objective is to deal with uncertainty when monitoring a Sustainable Supply Chain. Most recent works has failed to take into account conditions of uncertainty in the planning of SSC. In fact, they had recourse to mathematical approaches, which can lead to increased risk over the long term. [Tsao et al., 2017] propose an interactive method based on two-phase stochastic programming and fuzzy probabilistic multi-objective programing to design a Sustainable multi-echelon Supply Chain under Uncertain Environments. Nonetheless, to the best of our knowledge, no previous research has devolop a stochastic model to deal with the dynamism using multi objective evolutionary algorithm, which is more efficient, in a context of sustainable supply chain. Furthermore, only a reduced structure of the supply chain (particular case) was designed in most of the works. In addition to that, few have considered social issues together with environmental and economic ones. This is why in this paper, we aim to cover these gaps by several mechanisms used to monitor the global supply chain; first, integrating the dynamic mechanism within the chosen MOEA, then, considering the three pillars of sustainability simultaneously with the same degree of importance, as well as, designing a generic model that can be integrated within whatever the supply chain's structure is. Hence, we have considered the problem as a dynamic multi-objective optimization problem and have used a dynamic multi objective evolutionary method (DMOEA) within the monitoring mechanism.

B. *Proposed Method Background*

Dynamic Multi-objective Optimization is a challenging research topic. In fact, there are only few studies that have been developed to solve Dynamic Multi-objective Optimization Problems (DMOPs) in a context of SMEs with sustainable challenges. Several algorithms have been proposed for dynamic optimization problems (DOPs). Among them, some metaheuristic algorithms with population-based framework, such as EAs and Particle Swarm Optimizations (PSOs) . EAs have been considered as popular approaches to solve DOPs and prove to be useful in dynamic optimization problems because they are inspired in the natural evolution and this is a continuous process of adaptation. In fact, EA were classified into four categories [Jin and Branke, 2005]: (1) maintaining the diversity throughout the run, such as sharing or crowding mechanisms [Jiao et al. 2017], (2) or generating diversity after a change detection by taking explicit actions such as re-initializing population[Greeff and Engelbrecht,2008], a Self-adaptive Penalty function [Azzouz et al. 2015]; Also, many other approaches have been proposed such memory-based approaches [Wang and Li 2009] to recall useful information from past generations, if the environment did not change very severely; (4) multi-population approaches were also introduced into EAs for dynamic optimization, such as hierarchical clustering method [Li and Yang 2012], (5) predictive approaches [Cao et al. 2018], (6) parallel approaches [Cámara et al. 2008], and other approaches [Azzouz et al. 2014]. Exploration and exploitation are the two main mechanisms that are very influential on the performance of the dynamic EA. In fact, exploration, which is designed by the diversity maintenance strategy, can enhance the ability to search the entire solution space, making it adapt to the environmental change with a great extent. Exploitation can improve the accuracy of local search, making the algorithm to have a faster response to environmental change particularly in the solution set having relevance in the environment .This mechanism is modelled based on a Gaussian local search strategy. For that, we propose a dynamic version of the MOEA NSGA-II [Deb 2000] which is used to conduct the main part dynamically including selection, crossover, mutation and elite maintenance. To make dynamic NSGA-II more suitable for

incorporating memory scheme, an archive is used to store a well distributed non-dominated solution. For that, a memory-based dynamic NSGA-II is proposed. This algorithm aims to select, when a change occurs, the most adequate archived individuals to be included in the population to participate to the next optimization iteration. An Archive updated strategy is adopted using an improved non-dominated selection [Zhang et al, 2008] in order to remove dense solutions one by one and recalculates the crowding distances before deciding which solution should be deleted.

C. *Dynamic monitoring mechanism based on Multi agent system*

Since evolutionary algorithms are distributed by nature and since agents are able to perform many complex operations it was then natural that the idea of combining these two approaches in order to have a quick solution return in a dynamic environment. Thus, the main contribution of this work is to develop a dynamic monitoring mechanism based on evolutionary algorithms driven by agents. In fact, in a context of SMEs, BDI agents can incorporate collaboration strategies, thus enabling them to deal with the uncertainty in a dynamic environment and perform complex tasks [Bonhomme and Arnould, 2012].

IV. DESCRIPTION OF OUR KNOWLEDGE BASE MODEL

A. *Dynamic NSGAII:*

To maintain diversity, [Deb et al, 2007] has presented two dynamic optimization techniques. Their main difference is only the way of generating the initial population after a change. In the first case (DNSGA-II-A) where a ζ% of the new population is replaced with randomly created solutions. In the second version (DNSGA-II-B), ζ% of the population is replaced with mutated solutions of existing solutions (chosen randomly).

B. *Conceptual model of a macro-process:*

This section highlights the architecture of a macro-process considered as an SME. In fact, each SME's cluster is considered as a macro process and each macro-process is one of the global supply chain sub-system along with other subsystems such as manufacturing, Return and logistics macro process, and so on. Each of them can be modeled as a multi-agent system with sustainable considerations.

a) Agent-based modeling of SSC in SME's

Among characteristics which have been taken into consideration to assess the presented model based on MAS are: decentralizing decision-making, sharing the information along the SME's system by considering the global sustainable SC in the decision-making process, uncertain environment and dealing with sustainable objectives that are often contradictory. For that, an agent based framework is designed to simulate the interaction between the different monitoring agents in order to achieve the SME's goal. Thus, in the first step, multi objective evolutionary algorithm is applied in our agent-based framework to model our problem. In the next step, characteristics of dynamic aspect in multi objective evolutionary algorithm are taking into consideration on the designed agent-based framework. As shown in the figure below, to create an agent-based platform, we consider SME's management problem as a multi-agent system where each agent in this system is responsible for making decisions autonomously about an uncertain state and then collaborate and interact with other agents. The aim of this system is to achieve the common goal of the SME's cluster. Each SME's cluster is considered as a macro process and each macro-process is one of the global supply chain sub-system along with other subsystems such as manufacturing, Return and logistics macro process, and so on. Each of them can be modeled as a multi-agent system with sustainable considerations.

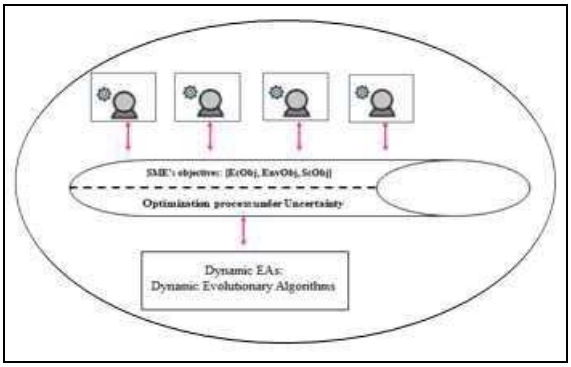


Fig.1. The agent-based framework of SME's management system.

V. THE PROPOSED COLLECTIVE DYNAMIC MODEL

In this paper, we proposed a dynamic EA based on the framework of NSGA-II [Deb, 2000]. We chose to use this algorithm as a basic framework for our proposal to conduct selection, crossover and mutation operators since there are multiple sustainable objectives to handle. Since when dealing with DMOPs, the goal is not only to converge to evolve to a near-optimal PF, but also to continually and rapidly discover the desired one before the next change occurs. The proposed dynamic algorithm must ensure a high convergence speed. To meet this challenge, another issue is showed, how to re-use past information of non-dominated solutions to effectively conduct the new population when the environment changes. For that, a memory based approach is then proposed. Thus, we propose to accelerate convergence by adding an extra storage space that store the elitist solutions found in each generation in order to use them for later generations or environment. When a change occurs, explicit actions should be taken to increase diversity and, thus, to facilitate the shift to the new optimum. The issue encountered here is what explicit actions BDI agents can take to handle this change. This is why; we propose several dynamic environment handling strategies that allow selecting individuals to participate in the next optimization iteration whenever a change is detected. These strategies may be able to effectively conduct the new population towards the new PF. The collective dynamic model is designed by a memory based DMOEA algorithm and follows an optimization process in order to propose solutions that meet the cluster goals. During this process, a set of solutions, often contradictory, must be optimized simultaneously in order to obtain a multitude of solutions that best meet the sustainable issues established by the SME's cluster. The figure 2 highlights the flowchart of the collective dynamic model that is explained in detail above.

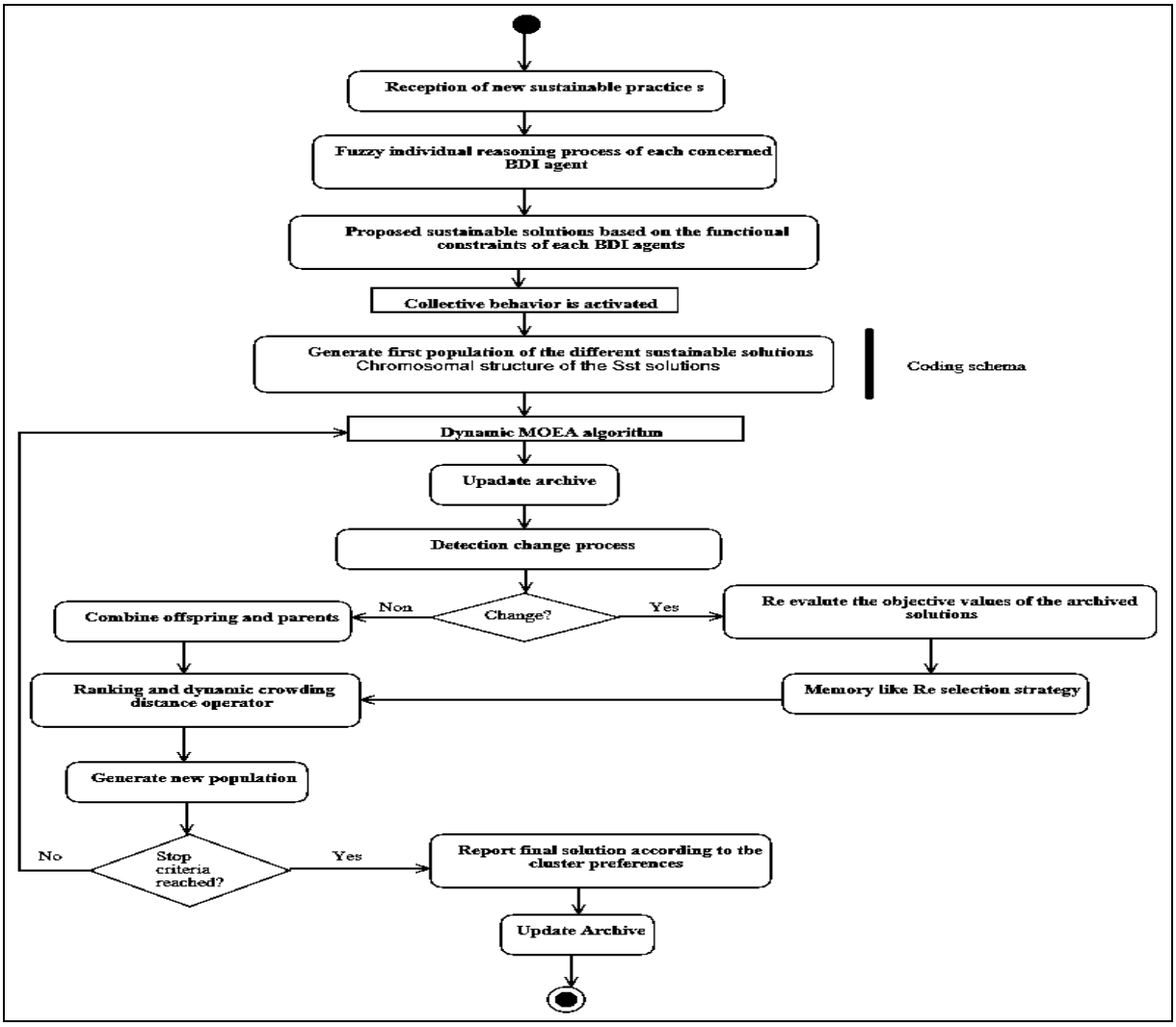


Fig. 2. The flowchart of the collective dynamic model.

In the first step of the process, each sustainable solution is transformed into a chromosome representation which represents the code schema of the model.

A. Coding schema

Each chromosome can be represented as a decision variable vector proposed by the BDI agent at each process level. The genes of the chromosome define the critical parameters of the system that are based on the direct and indirect interrelation between the proposed solution and these parameters. Thus, any change in one of the variables of the parameter causes some changes in other variables and, therefore, the representation of the chromosome will encompass all key parameters. The figure below illustrates the chromosomal structure of a proposed sustainable solution. We encode the different genes of the chromosome as follows: the objectives related to the sustainable solution: economic objective, environmental objective and a social objective, the functional constraint (s) related to each BDI agent, sustainable practice (s) S_p value. Since most of these parameters are real numbers, we use a real-valued coding technique to represent the vector of decision variables related to the solution proposed by a BDI agent. The collective dynamic model is a recursive model where the chromosome at the level of each process will be represented according to a structure that will keep the same parameters at the macro process level until reaching the global level of the sustainable supply chain. After that the D-MOEA algorithm is implemented and the optimization process is described thoroughly.

B. The Memory based D-MOEA algorithm

At the first step, many solutions presented by different BDI agents are considered as the initial parent population with size N. The population as well as the archive will be initialized. Then, the population is sorted based on the non-domination. Each solution is assigned a fitness (or rank) equal to its non-domination level (1 is the best level, 2 is the next-best level, and so on). The rank of each solution corresponds to its non-domination level. Non-dominated solutions will be generated and a copy will be stored in the archive. For each generation, we execute the following steps.

At first, the usual genetic operators (selection, then, we randomly pick a set of 'sentry' individuals and we re-evaluate their objective function values to detect environment changes. If a change of environment has occurred, we re-evaluate the objective function values of the archived individuals then a population management mechanism will be processed. Otherwise, the population is updated using the non-domination sorting tri and the crowded-comparison-operator. In fact, the parent and the offspring populations are combined. Then, solutions are sorted into different non domination levels. The new population is formed by selecting the solution with the lower rank when comparing two solutions of different non-domination levels and preferring the solution located in a lesser crowded region between two solutions having the same rank. Finally, the archive is updated using improved using improved non-domination selection and finals solutions will be reported. These steps are repeated until a maximum number of generations are reached .

C. Detection change process

In order to In order to identify if a change has taken place, as presented in [Deb et al, 2007], we introduce a test to identify whether there is a change or not. For this purpose, we randomly pick a set of solutions from the archive (10% population members) in every generation to construct 'sentry' individuals and we re-evaluate the fitness values of them. If there is at least a change in the objective functions' values of one individual comparing to the old ones, then we admit the occurrence of a change. This process allows both offspring and parent solutions to be evaluated using the changed objectives functions 'values. After that, an environment handling scheme is launched to reset the population (environmental selection) and the archive.

D. Memory like re-selection strategy

The memory based strategy is designed to lead the search towards proper region and to accelerate the convergence speed. In fact, two main processes should be treated when controlling the memory. First procedure is how to manage the re-use information. If the environment did not change severely, the stored information can provide a much better starting point for the new environment. In fact, when detecting a change and before the beginning of the next optimization process may conduct the population towards the new promising search directions; a local search scheme inspired by [Wang and Li, 2009] is used. Adding to that, the memory size is limited, for that an update of the archived information is needed, thus an improved non-domination selection is used [Zhang et al, 2008]. The main drawback of memory-based approaches is that memory is very dependent on diversity and should, thus, be used in combination with diversity-preserving techniques.

a) Gaussian Local search mechanism

Memory re-selection strategy is first used to effectively re-use the useful past information to conduct the new population when the environment changes. To re-use past information, one natural idea is to reevaluate past elite solutions and form the new population by them. However, overmuch exploitation around the memory optimal POS may cause loss of diversity and shrinking towards the local optimal POS. Therefore, the population for the new environment should not be composed of the memory solutions only but of a great mass of randomly generated solutions to effectively maintain the diversity. Gaussian Local search scheme has been testified by many researches to be a good strategy to enhance the ability of elaborate search [Yao et al, 1999]. A Gaussian local search based memory described below has good effect on the convergence ability of the algorithm. The Gaussian local search is defined as follows:

$$v_{id}(t+1) = c3 * gaussrand \quad (1)$$

Where i = 1, 2, . . ., m, d = 1, 2, . . ., D, gaussrand is a random number generated from a standard normal distribution (mean μ = 0, variance $\delta 2 = 1$), $c3$ is a positive constant. The archive (A) is considered as the input population where solutions stored into it are considered the best known solutions with different objective function values and the output P′ is used as the main population in the next generation. The proportion of local search is controlled by a probability Pl setting to be 0.2. Rand is a uniformly generated number within [0, 1]; σ is a positive constant. The neighborhood solution is generated using Gaussian mutation. Each archive solution has a chance to provide a neighborhood solution in the new population. The ideal situation is that the distribution of individuals generated by local search inherits general manifold of the old POS. Thereby, the local search step is stopped if the neighbor's objective value is better than the current solution one.

1. Gaussian local search algorithm

Input: A (Initial population), N_p: Size of the archived population.
Output: P' (New population).
Begin
For i: = 1 to N_p
  If rand > P_l
    Randomly select the i^{th} solution from the archive A.
  Else
    Randomly select a solution s from the archive A.
    Generate i^{*th} solution by performing Gaussian mutation on s
    If i^{*th} solution is better than s solution **then**
      Insert the i^{*th} solution in the new population P'
End if
End for
Update P' with i^* set solutions.

b) Archive update strategy

Since the PF contains a large number of solutions, it may be difficult to store and exploit the useful past information. For that, an archive is used to store a well distributed non-dominated solutions found so far. The archive updated strategy is an improved non-dominated selection that has been proposed, which removes dense solutions one by one and recalculates the crowding distances before deciding which solution should be deleted. It can improve the diversity property compared with the selection of NSGA-II.

E. Diversity maintenance strategy

Maintaining the diversity in the optimization process is another issue to ensure flexibility and adaptability. In fact, good diversity of the PF can provide more reasonable choices to BDI agents. For that reason, NSGA-II has been modified to ensure good diversity maintenance and a well distribution of non-dominated solutions. Dynamic crowding distance (DCD) proposed by [Luo et al, 2008] to improve distribution of non-dominated solutions are incorporated in our DMOEA algorithm.

a) Dynamic crowding distance

In a dynamic crowding distance, the individuals' DCD are varying dynamically during the process of population maintenance. In fact, individual with lowest DCD value is

removed every time and the DCD for the remaining individuals is recalculated. The individuals DCD are calculated as follows (2) and (3):

$$DCD = \frac{CD_i}{Log(\frac{1}{v_i})} \quad (2)$$

$$V_i = \frac{1}{r}\sum_{k=1}^{r}(|f_{i+1}^k - f_{i-1}^k| - CD_i)^2 \quad (3)$$

Where V_i is the variance of CDs of individuals which are neighbors of the i^{th} individual.

1. *Dynamic Crowding Distance (DCD) algorithm:*

Suppose initial population size is N, the non-dominated set at g^{th} generation is stored in the archive A(t), the size of A(t) is M; If M ⩾ N, then use DCD to wipe off (M −N) individuals from the archived non-dominated set. The process of DCD algorithm is given below:

Step 1: If, A (t) ≤ N then go to step 5, else go to step 2.

Step 2: Calculate individuals' DCD in the A (t) based on (2).

Step 3: Sort the non-dominated set, A (t) based on DCD.

Step 4: Remove the individual with the lowest DCD value in the A (t).

Step 5: If, A (t) ≤ N stop the population maintenance, otherwise go to step 2 and continue.

F. Performance metrics

Performance metrics could help in studying deeply the performance of the algorithm, such as convergence and distribution. These metrics are indicators that judge the value of an obtained Pareto approximation set from different angles. Different performance metrics are selected in this work such as Hypervolume Metric (HV) and the inverted generational distance metric (IGD) .

VI. CONCLUSION & PERSPECTIVES

The cooperative dynamic behavior, which we have proposed in this paper, highlights the optimization process based on Memory based-DMOE algorithm in order to propose solutions that meet the cluster goals. In fact, our collective dynamic model is a recursive model where the population at the level of each process will be represented according to a structure that will keep the same parameters at the macro process level until reaching the global level of the sustainable supply chain. As a perspective, the implementation of the dynamic collective behavior will be treated.

REFERENCES

[1] S. Seuring, M. Müller, "From a literature review to a conceptual framework for sustainable supply chain management". Journal of Cleaner Production,2008, 16, (15), 1699–1710.

[2] Z. Zhou, S. Cheng, and B. Hua, 'Supply chain optimization of continuous process industries with sustainability considerations',Computers and Chemical Engineering, 2000,Vol. 24, pp.1151–1158.

[3] C. R., Carter, D. S., Rogers, "A framework of sustainable supply chain management: Moving toward new theory". International Journal of Physical Distribution & Logistics Management, 2008, 38(5), 360-387.

[4] C.Bonhomme and G. Arnould, "Dynamic carpooling mobility services based on secure multi-agent platform". In Global Information Infrastructure and Networking Symposium,2012, 1–6.

[5] M. E. Bratman, D. J. Israel and M.Pollack, "Plans and resource-bounded practical reasoning". Computational Intelligence, 1988,Vol. 4, pp. 349-355.

[6]A.Ben Mekki, J.Tounsi and B.S. Lamjed, "Fuzzy Multi-Agent approach for monitoring SMEs sustainable SC under uncertainty", Procedia Computer Science, 2019.

[7] S.H. Zegordi, I.N.K. Abadi, M.A.B. Nia, "A novel genetic algorithm for solving production and transportation scheduling in a two-stage supply chain". Computers & Industrial Engineering, 2010,58, 373-381.

[8] A. Diabat, "Hybrid algorithm for a vendor managed inventory system in a two-echelon supply chain". European Journal of Operational Research,2014, 238, 114-121.

[9] S.A.Torabi, S.M.T. Fatemi Ghomi, B. Karimi. "A hybrid genetic algorithm for the finite horizon economic lot and delivery scheduling in supply chains". European Journal of Operational Research,2006, 173, 173-189.

[10] F. Altiparmak, M. Gen, L. Lin, T. Paksoy, "A genetic algorithm approach for multiobjective optimization of supply chain networks". Computers & Industrial Engineering , 2006, 51, 196-215.

[11] A. Saghaeeian, R. Ramezanian, "An efficient hybrid genetic algorithm for multi-product competitive supply chain network design with price-dependent demand". Applied Soft Computing,2018, S1568-4946.

[12] H. Stefansson, P. Jensson, N. Shah, "Procedure for reducing the risk of delayed deliveries in make-to-order production". Production Planning & Control,2009, 20 (4), 332–342.

[13] C.C. Zhou, G.F.Yin, X.B. Hu, "Multi-objective optimization of material selection for sustainable products: artificial neural networks and genetic algorithm approach". Materials and Design, 2009, 30, 1209–1215.

[14] B. Fahimnia, E. Parkinson, N.P. Rachaniotis, Z. Mohamed, a. Goh, "Supply chain planning for a multinational enterprise: a performance analysis case study". International Journal of Logistics Research and Applications,2013c, 16, 349-366.

[15] I.S. Doolun, S.G. Ponnambalam, N. Subramanian, G. Kanagaraj, "Data driven hybrid evolutionary analytical approach for multi objective location allocation decisions: Automotive green supply chain empirical evidence". Computers and Operations Research, 2018, doi: 10.1016/j.cor.2018.01.008.

[16] A. Azadeh, F. Shafiee, R. Yazdanparast, J. Heydari, A.M. Fathabad, "Evolutionary multiobjective optimization of environmental indicators of integrated crude oil supply chain under uncertainty". Journal of Cleaner Production, 2017, (152) 295-311.

[17]Y. Tsao, V. Thanh, J. Lu, V. Yu, "Designing Sustainable Supply Chain Networks under Uncertain Environments: Fuzzy Multi-objective Programming". Journal of Cleaner Production , 2017, doi: 10.1016.

[18] C. Coello Coello, D. van Veldhuizen , G. Lamont, "Evolutionary algorithms for solving multiobjective problems", 2007, Springer, New York.

[19] K. Deb, U. Rao, S. Karthik, "Dynamic multi-objective optimization and decision-making using modified nsga-ii: a case study on hydro-thermal power scheduling". In: Proceedings of the 4th international conference, EMO 2007, vol 4403, pp 803–817.

[20] K. Deb,S. Agrawal,A. Pratap,T. Meyarivan, "A fast elitist non dominated sorting genetic algorithm for multi-objective optimization: Nsga-ii". In: Proceedings of the 6th international conference on parallel problem solving from nature, 2000,vol 1917, pp 849–858.

[21]Y. Jin , J. Branke , "Evolutionary optimization in uncertain environments—a survey". IEEE Trans Evol Comput,2015, 9(3): 303–317

[22] R. Jiao, S. S. Zeng, J. Alkasassbeh and C. Li, "Dynamic multi-objective evolutionary algorithms for single-objective optimization". Applied Soft Computing , 2017,S1568-4946.

[23] M. Greeff, A.P. Engelbrecht, "Solving dynamic multi-objective problems withvector evaluated particle swarm optimization, in: IEEE Congress onEvolutionary Computation, 2008, CEC 2008 (IEEE World Congress onComputational Intelligence), IEEE pp. 2917–2924.

[24] R. Azzouz , S. Bechikh , L. Ben Said , "Multi-objective optimization with dynamic constraints and objectives: new challenges for evolutionary algorithms. In: Genetic and evolutionary computation conference (GECCO 2015)

[25] Y.Wang, B. Li, "Investigation of memory-based multi-objective optimization evolutionary algorithm in dynamic environment". In: Proceedings of the IEEE congress on evolutionary computation, 2009,pp 630–637

[26] C.Li, S. Yang ,"A general framework of multi-population methods with clustering in undetectable dynamic environments". IEEETrans Evolut Comput,2012, 16(4):556–577

[27] L. Cao, L.D. Xu, E. Goodman, S. Zhu,H. Li, "A Differential Prediction Model for Evolutionary Dynamic Multiobjective Optimization". GECCO'18, July 15-19, Kyoto, Japan

[28] M. Cámara J. Ortega, F. de Toro, "Parallel multi-objective optimization evolutionary algorithms in dynamic environments". In: Proceedings of the first international workshop on parallel architectures and bioinspired algorithms, 2008, vol 1, pp 13–20

[29] R.Azzouz, S. Bechikh , LB. Said , "A multiple reference point based evolutionary algorithm for dynamic multi-objective optimization with undetectable changes. In: Proceedings of the IEEE congress on evolutionary computation, 2014, pp 3168–3175

[30] Q. F. Zhang, A. M. Zhou and Y. C. Jin "RM-MEDA: A Regularity Model-Based Multi-objective Estimation of Distribution Algorithm," IEEE Trans. Evol. Comput., vol. 12, no. 1, February 2008.

A behaviorist agent model for the simulation of the human behavior

Hiba Chaher*, Lilia Rejeb**, and Lamjed Ben Said***

Abstract— Recent researches on computational modeling show that emotions have a major influence on human behavior and decision making. Therefore, it is recognized that they are necessary to produce human-like artificial agents. Several computational behavior models have been proposed. However, some of them have incorporated emotion, others have integrated psychological aspects in order to study the human behavior, but few of them have taken into account both of the emotional and the psychological impacts. In this context, we attempt to present an overview of the existent works. Then, we aim to propose a new behavior agent model that integrates both of the psychological and emotional aspects to prove their impacts on the human decision.

Index Terms—Agent Based Model, Decision-making, Emotion, Human behavior, Personality, Simulation

1 Introduction

Cognitive-emotive interactions have been the subject of several research in the past decades. The emotional factor is one of the psychological factors that need great attention when analyzing the human behavior and studying the decision making process. Humans are characterized by complex cognitive emotional and psychological structures. These aspects make the study of human behaviors complicated[1]. However, many researchers in artificial intelligence focused mainly on rational decision-making [2]. However, the human decision is not totally rational [3]. In fact, the integration of the emotional factors is required to simulate human-like behavior[4]. Most research working on human behavior modeling is agent-based models based on the Ortony, Clore & Collins (OCC) model of emotions [5]. Moreover, modeling human-like behavior requires not only the integration of the emotional factor but also the psychological one. Very limited works have adopted these paradigms simultaneously[6], [7]. Consequently, the research on computational modeling of emotions is not advancing compared to other cognitive studies [8]. The main goal of the current research work is to study the psychological emotional dynamics and their effects on human decision making using emotional artificial agents. In this paper, we introduce an agent model that considers emotion and personality in agent decision making. The paper is organized as follow: in section 2, we present the theoretical background of our research work. In section 3, we expose an overview on the major emotional agent-based models that incorporate artificial emotions. Then, in section 4, we describe in details the proposed agent-based model. Then, we give a description of the implementation and tests. Finally, we conclude the paper with a discussion and future scope.

*Hiba Chaher, Phd student , SmartLab Institut Superieur de Gestion, Université de Tunis ,Tunisia. E-mail: chaherhiba@gmail.com
**Lilia Rejeb, Assistant master, SmartLab Institut Superieur de Gestion, Université de Tunis, Tunisia. E-mail: lilia.rejeb@gmail.com.
***Lamjed Ben Said , Professor ,SmartLab Institut Superieur de Gestion, Université de Tunis, Tunisia.E-mail: lamjed.bensaid@isg.rnu.tn

2 Theoretical framework

Our research goal is to build a generic behaviorist computational model that combines emotions and personality in order to analyze the human behavior and the decision making process. Based on these ideas, we looked at research on the emotional and the psychological aspects and how to incorporate them in software architecture. We found that the Ortony Clore and Collins model of emotion (OCC)[5] and the Openness, Consciousness, Extraversion, Agreeableness and Neuroticism model of personality (OCEAN) [9] are the most widely used approaches related to the study of emotions and personality [10] . In the next subsections, we will introduce both of the models in details.

2.1 The Ortony Clore and Collins model of emotions

The Ortony, Clore and Collins (OCC) model of emotions is the most accepted and influential theoretical approach based on appraisal[6] [11]. According to the OCC model, emotions are reactions to perceived stimuli which could be actions, objects or events [5]. In fact, these stimuli are appraised through an appraisal process according to a list of dimensions called appraisal variables. For example, events are evaluated according to the goals of a person. This allows to calculate the Desirability of the events. Joy and Distress are the results of desirable and undesirable events[5][6]. However, appraising an object is related to the agent's preferences in order to decide whether the

object is attractive and appealing or not. The emotion result is either Like or Dislike. Praiseworthiness is used to appraise the actions of agents, on the agent's standards. Generally, praiseworthy actions generate pride and blameworthy actions generate shame, if the action was made by the agent himself [5]. Fig. 1 gives an overview of the OCC approach. The OCC model is considered as one of the most comprehensive models of emotions as it provides a detailed description of the different steps of the emotion generation [12][13]. The emotional agent that we propose incorporates an emotion generation mechanism founded on the OCC model. It integrates also the OCEAN model of personality that will be described in the next subsection.

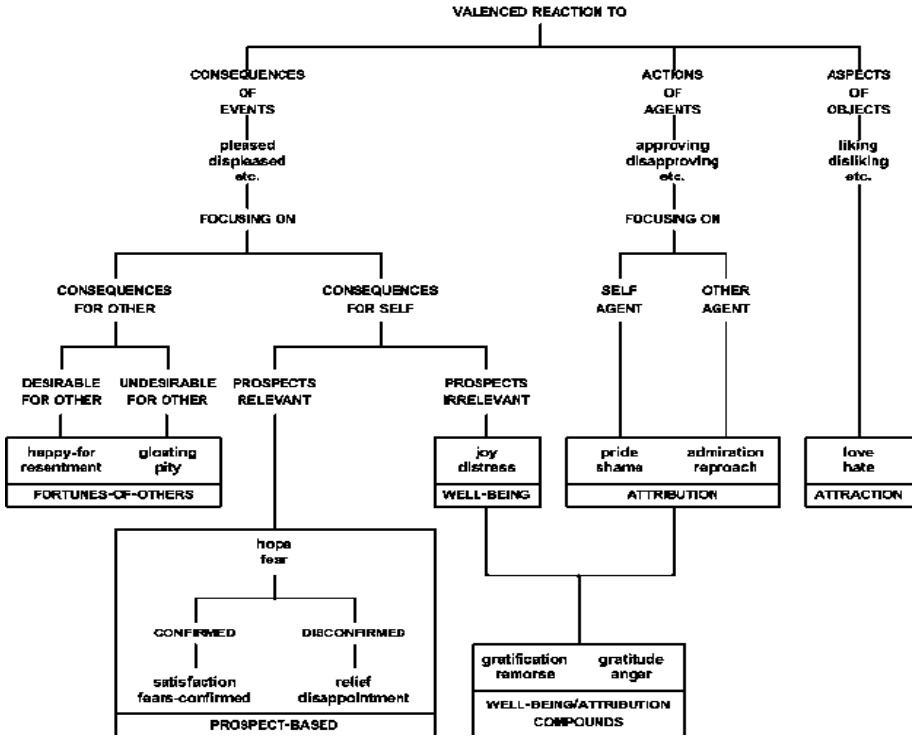


Fig. 1. The OCC model [5]

2.2 The Five Factors model of personality

Researches on human psychology and behavior has demonstrated that personality impacts not only the behavior of an individual, but also the emotional aspect [14]. According to the psychology research, there are many personality models that consist of a set of dimensions, where every dimension is a specific property of the personality [15]. For example, the OCEAN model has five dimensions: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness [9] which are described in Table 1. According to the OCEAN model, we assume that a personality has five dimensions, where each dimension is represented by a value in the interval [0, 1]. These variables can have a value of zero which corresponds to the absence of the dimension in the personality or a value of 1 which corresponds to a maximum presence of the dimension in the personality [9]. The personality P of an individual can then be represented by the following vector[9]:

$$\mathbf{P} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{bmatrix} \forall i \in [0,1] \quad (1)$$

TABLE 1
THE OCEAN MODEL OF PERSONALITY [9]

Factor	Description
Openness	Open mindedness, interest in Culture
Conscientiousness	Organized, persistent in achieving goals
Extraversion	Social behavior and preference for others
Agreeableness	Interactions with others
Neuroticism	Tendency to experience negative thoughts

3 Related Work

The social and computational sciences have seen an explosion of interest in emotions in the last decade. Different research works have tried to reproduce human emotional dynamics in artificial agents [2] . Here, we focus on the computational models of emotion, particularly those that were based on the OCC model. Examples include FLAME (Fuzzy Logic Adaptive Model of Emotions)[17] which uses fuzzy-logic representations to

map events and observations to emotional states, the behavioral multilayered model of Colloc [18], EMIA (Emotional Model for Intelligent Agent) [19] in which fuzzy logic was used to implement emotions. FAtiMA (Fearnot AffecTIve Mind Architecture) is an Agent Architecture with planning algorithms designed to integrate emotions and personality to influence the agent's behavior [20]. Similarly, Bui et al. have implemented another model called ParleE which integrates personality aspects for an embodied conversational agent. The model generates emotion via an event appraisal module based on learning and a probabilistic planning algorithm using six appraisal variables [21]. EMA (EMotion and Adaptation) is a generic domain independent model for creating agents that demonstrate and cope with (negative) affect. The model generates five categories of emotions using the event appraisal mechanism [22]. GEmA is another Generic Emotional Agent which incorporates sixteen emotions by appraising events and actions according to both goals and standards of agents [2]. Moreover, A Layered Model of Affect (ALMA) is a computational model which simulates affective states in virtual humans. The model comprises three layers: an emotion layer, a mood layer and a personality layer [23]. The model of Belhaj et Ben Said [24] analyzed human behavior during emergency situations. A final example is the model of Lejmi and Ben Said [25] which was designed to study the impact of emotions on the employees' decision-making. Since these models have the same theoretical foundation, they have obviously some similar practices in modeling emotions. They are based on appraisal theories that include different steps to generate emotions. However, they differ on the representation of these mechanisms, the number of the appraisal variables used, the methods employed for their computations and the number of emotions type generated. Some of them were theoretic and were not implemented (for example the model of Colloc [18]). Others did not incorporate the collective level and the relation agent-environment, which may generate other emotions (for example the FLAME model[17]. Most of these models were domain dependent, simplistic as they oversimplify the human behavior by using simple scenarios and by representing only few categories of emotions which does not reflect the human real life. Nevertheless, reproducing the human behavior requires also the integration of other factors such as the personality that may impact the decision making process. Although it plays a crucial role in the human behavior, very limited works have integrated it. Only ALMA model, ParleE and FAtiMa have mentioned it but they were too simplistic and they did not integrate all the compounds of the personality model. Moreover, there is always a strong relation between personality and emotions[26]. Consequently, the emotional and the psychological aspects are particularly important in the study of human behavior which was not incorporated in most of these computational works. In response to the limitations of the existent models, we will propose in the next section a multilayered domain-independent behaviorist agent model based on the OCC model and the OCEAN model in order to reproduce human like behavior.

4 Psychological Emotional Modeling Agent (PEMA): The proposed behavior agent Based model

The proposed behaviorist agent based model that we have called PEMA, is based on the OCC model as it deals with all its different steps of the emotion reproduction process. It maps between the OCC model and the OCEAN model in order to study the influence of these factors such as emotion and personality traits on the human's behavior and decision-making process. Agents evaluate what they perceive. Consequently, the resulting emotional state influences then their action selection process. In fact, the proposed emotional agent model involves three components: a Perception module, an Appraisal module and a Cognitive module, Fig. 2 shows the different components of the proposed generic behaviorist agent model. We detail in the following sections these different components.

4.1 Perception module

An agent who is situated in a dynamic environment need to aware of his environment's changes. He is represented by the following components:

-An emotional vector that contains the internal emotional state of the agent (e.g. joy, distress, pride...) according to the emotions categories of the OCC model

-A five-dimension personality vector founded on the OCEAN model of personality

-A goals file that contains all the goals of the agent

- A behavior file that contains a list of actions, reactions, standards and preferences of the agent

-An event file that contains all the events with their degrees and their intensities levels

-A memory file that contains all the perceived data, a history of the past events, actions and objects perceived.

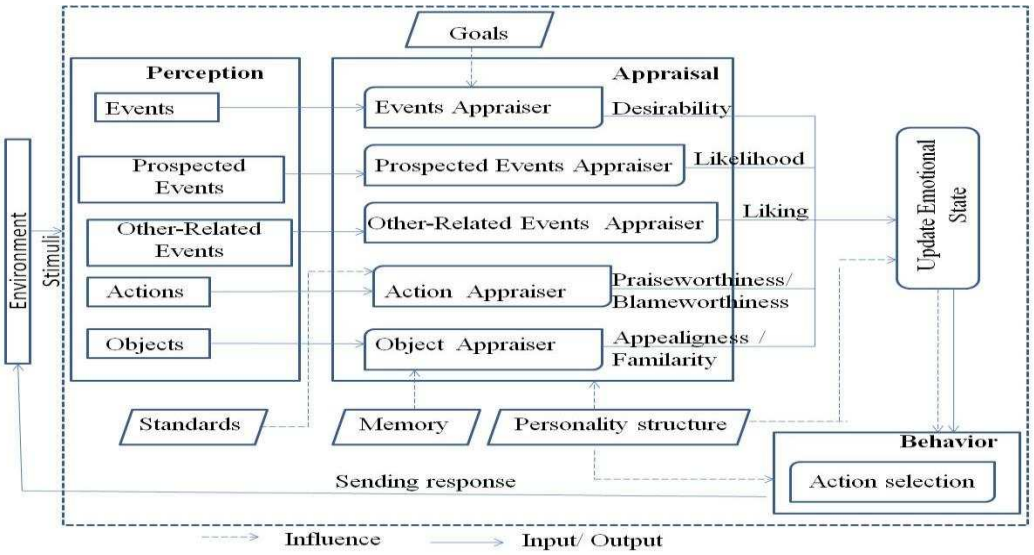


Fig. 2. The behavior agent based model

4.2 Appraisal module

Each category of the perceived data is considered as an input related to a particular appraiser belonging to the three appraisers according to the OCC, which are event appraiser; action appraiser and object appraiser (see Fig.2). Each appraiser uses a set of appraisal variables of the OCC model to calculate emotion intensities. The output is a particular emotion that belongs to one of the emotion categories of the OCC model. To compute the intensities of emotions, Ortony et al. have proposed a list of central and local appraisal variables (Ortony, Clore, & Collins, 1988). In our agent based model of emotions, we used three central variables: Desirability, Praiseworthiness and Appealingness and four local variables: Realization, Likelihood, Liking and Familiarity. These variables are used to update the internal emotional state. As we have noticed in the previous sections, the OCC model represents emotions as valenced reactions to the perception of the world, depicted by couples of (positive/negative) emotions. Indeed, we used all the couples of emotion described in the OCC tree. We summarized all of them in Fig. 1. We describe now the three appraisers:

1. The Self Event Appraiser (SEA) evaluates events that happen to the agent itself. These events are evaluated according to the agent's goals throw the computation of the Desirability of this event in order to generate a Well-Being emotion. The output of this appraiser is then a positive emotion Joy or a negative one Distress since the appraised event is respectively positive or negative. Thus, the Desirability of an event Des (e, g_i, t) corresponds to the contribution or the inhibition of the event on the realization of the current goal of an agent. Before calculating Desirability Des (e, g_i, t), we need to compute the Impact of the event on the goal using two variables called the degree of the event d_e and the Event Intensity Level EIL(e). Then, Desirability is computed based on the impact value of the event (impact (e, g_i) which is in [-1, 1]) and the importance of the current goal itself (import (g_i) which is in [0, 1])[27].

$$Goals = \begin{pmatrix} g_1 \\ g_2 \\ g_3 \\ \dots \\ g_n \end{pmatrix} \forall i \in [1,n],\ g_i \in [0,1] \quad (2)$$

We suppose there are n goals. If we denote the degree of the event defined as $d_e \in$ [-1, 1], the EIL(e) computed for each event e defined in [0, 1] then, the impact of the event on the goal g_i is given by:

$$impact = (e, g_i) = d_e * EIL(e) \quad (3)$$

Then, the Desirability is computed as follows:

$$Des(e, g_i, t) = impact * (e, g_i) * import(g_i) \quad (4)$$

The resulting emotion may be Joy if the event is positive and Distress if the event is negative with the intensities I_{Joy} and $I_{Distress}$ at time t, computed as follows:

$$I_{Joy} = Des(e, g_i, t) \quad (5)$$

$$I_{Distress} = Des(e, g_i, t) \quad (6)$$

2. The Prospected Event Appraiser (PEA) evaluates prospected events that may happen to the agent itself using Desirability (Des (e, g_i , t)) and Likelihood (L) in order to generate Prospect-Based Emotions. Then, when a prospected positive respectively negative event was realized (Realization R =1), Satisfaction respectively Fear-confirmed emotion will be generated. However, if the positive, respectively the negative event, doesn't happen (R= -1), Disappointment, respectively Relief emotion, arises.

3. Other Related Event Appraiser (OREA) computes the Desirability of the events that happen to other agents. The evaluation of these events was performed using two appraisal variables called DesirabilityForOther (DesForOther) and Liking. To assess this Desirability, we endow the agent with the ability to simulate the state of the other with its current goal to compute the Desirability of the event for other. However, Liking was used to compute empathetic emotion between two agents (A_i , A_j).

4. The Action Appraiser (forSelf) (AAS) computes the Praiseworthiness P(a) or Blameworthiness B(a) of the agent's actions using two variables : Action degree d_a which is a signed numeric representation depending on the standards of the agent, Approval Degree Action Ap_a which denotes how strong belief an agent has on the standard is.

5. The Action Appraiser (forOther) (AAO) an agent may be in a situation that may necessitate the help of other agents. Thus, we define a function noted Need Intensity Level NIL (t) as the need intensity of the agent to be saved from the current situation. Then, we measure praiseworthiness PraiseworthyOther(a,t) of accepting the help request or blameworthiness BlameworthyOther(a,t) in case of rejection.

6. The Well-being/Attribution Compound Appraiser (WACA) computes the intensity of the emotions combining the consequences of the event and its agency.

7. The Object Appraiser (OA) evaluates perceived objects using two variables; Familiarity F and Appealingness A.

Table 2 is an overview of all the appraisal variables used by all these appraisers and the emotion triggered.

TABLE 2

DESCRIPTION OF THE APPRAISERS, EMOTIONS TRIGGERED AND FORMULAS RELATED

Appraiser	Appraisal configuration	Formula		Emotions
PEA	$Des(p, g_i, t)>0$	$I_{Hope}(p,t)=D(p,g_i,t)*L(p,t)$	(7)	Hope, Fear
	$Des(p, g_i, t)<0$	$I_{Fear}(p,t) = \lvert Des(p, g_i, t) * L(p,t)\rvert$	(8)	, Satisfaction
	$R=1$	$I_{Satisfaction}(p,t) = I_{Hope}(p,t)$	(9)	Fear confirmed
	$Des(p, g_i, t)>0$ $R=1$	$I_{Fear-confirmed}(p,t) = I_{Fear}(p,t)$	(10)	Disappointment
	$Des(p, g_i, t)<0$ $R=-1$	$I_{Disappointment}(p,t) = I_{Hope}(p,t)$	(11)	Relief
	$Des(p, g_i, t)>0$ $R=-1$ $Des(p, g_i, t)<0$	$I_{Relief}(p,t) = I_{Fear}(p,t)$	(12)	
OREA	$DesForOther(e, g_i, t)>0$	$I_{Happyfor}(A_i, A_j)=DesForOther(e, g_i, t)* L(A_i, A_j)$	(13)	Happy for,
	$DesForOther(e, g_i, t)>0$	$I_{Sorryfor}(A_i, A_j)=\lvert DesForOther(e, g_i, t)* L(A_i, A_j)\rvert$	(14)	Sorry for
	$DesForOther(e, g_i, t)<0$ $L(A_i, A_j) = 0$	$I_{Gloating}(A_i, A_j) = 1 - \lvert DesForOther(e, g_i, t) * L(A_i, A_j)\rvert$	(15)	Gloating
	$DesForOther(e, g_i, t)<0$ $L(A_i, A_j) > 0$	$I_{Pity}(A_i, A_j) = (DesForOther(e, g_i, t)) * (L(A_i, A_j))$	(16)	Pity
AAS	$d_a = 1$	$P(a) = d_a * Ap_a$	(17)	Pride
		$I_{Pride} = P(a)$	(18)	
	$d_a = -1$	$B(a) = \lvert d_a * Ap_a\rvert$	(19)	Shame
		$I_{Shame} = B(a)$	(20)	
AAO	$d_a = 1$	$PraiseworthyOther(a, t) = d_a * NIL(t)$	(21)	Admiration
		$I_{Admiration}(A_i, A_j) = PraiseworthyOther(a(A_j), t)$	(22)	
	$d_a = -1$	$BlameworthyOther(a, t) = \lvert d_a * NIL(t)\rvert$	(23)	Reproach
		$I_{Reproach}(A_i, A_j) = \lvert BlameworthyOther(a(A_j), t)\rvert$	(24)	
WACA		$I_{Gratification} = \max(I_{Joy}, I_{Pride})$	(25)	Gratification
		$I_{Remorse} = \max(I_{Distress}, I_{Shame})$	(26)	
		$I_{Gratitude} = \max(I_{Joy}, I_{Admiration})$	(27)	Remorse
		$I_{Anger} = \max(I_{Distress}, I_{Reproach})$	(28)	Gratitude
OA	$A(o,t)=1$	$I_{Like}(o,t) = A(o,t) * F(o,t)$	(29)	Anger
	$A(o,t) = -1$	$I_{Dislike}(o,t) = \lvert A(o,t)* F(o,t)\rvert$	(30)	Like, Dislike

4.3 The Internal Emotional State update

An Internal Emotional State (IES) changes over time and it is initialized to zero. We define the emotional state e_t as an m-dimensional vector, where m=22, which represents all the emotions types described in the OCC model. The m emotion intensities are represented by a value in the interval [0, 1].

$$e_t = \begin{bmatrix} \beta_1 \\ \dots \\ \beta_m \end{bmatrix}, \forall i \in [1, m], \beta \in [0,1] \quad (31)$$

Besides, we define an emotional state history ω_t that contains all emotional states until e_t, thus:

$$\omega_t = (e_0, e_1, \dots e_t) \quad (32)$$

The update of the intensity of some couple of emotion (well-being emotion) was computed using not only the appraisal formula described in Table 2 but also throws other emotions intensities (prospect emotions) as shown in Table 3.

4.4 Personality

The OCEAN model is widely detailed, we have used it to illustrate the personality's traits of the agent, which was composed by five dimensions Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience. Each personality factor has a weight in determining how something is perceived and what decisions are being taken [15]. For an overview of the relationship between personality traits, goals, standards, agent's preferences and emotion generation, we define the following assumptions inspired from [15], [28]–[33]:

-Openness influences the attitude of someone towards new elements in his environment thus, Like emotion is often triggered with high intensities.

-A conscientious personality tries to achieve all the goals that the agent has set thus, positive emotions (Joy, Hope) are triggered with high intensities than not conscious ones.

-Extraversion includes sociability, assertiveness, and high amounts of emotional thus, Happy for, Sorry for and Pity are experienced with high intensities for extraverted individual than introverted ones.

-Agreeableness defines if someone will adopt or abandon goals in favor of others thus; positive emotions (Pride and Admiration) are triggered for agreeable individual with high intensities than not agreeable ones.

-Neurotic personality experiences often-negative thoughts, which can lead to less ambition thus, negative emotions (Distress, Fear), are triggered with high intensities.

In our model, we suppose that the personality of each agent is stable and constant during his life cycle [34]. It is represented as a vector, which is formed by the five dimensions. The personality P of the agent is as follows:

$$P = [P^{Open}, P^{Consc}, P^{Ext}, P^{Agrea}, P^{Neuro}] \in [0,1] \quad (33)$$

TABLE 3

THE RELATION BETWEEN WELL-BEING EMOTIONS AND PROSPECT EMOTIONS

Emotions (Joy, Distress)	Update configuration	Update condition	Update emotion
	$Des(p, g_i, t)>0$ $I_{Distress}(t) = \lvert I_{Joy}(t) - I_{Distress}(t-1)\rvert$	$I_{Distress}(t) < I_{Distress}(t-1)$	$I_{Distress}(t) = \lvert I_{Joy}(t) - I_{Distress}(t-1)\rvert$
		$I_{Distress}(t) > I_{Distress}(t-1)$	$I_{Distress}(t) = \lvert I_{Distress}(t-1)\rvert$
	$Des(p, g_i, t)<0$ $I_{Joy}(t) = \lvert I_{Distress}(t) - I_{Joy}(t-1)\rvert$	$I_{Joy}(t) < I_{Joy}(t-1))$	$I_{Joy}(t) = \lvert I_{Distress}(t) - I_{Joy}(t-1)\rvert$
		$I_{Joy}(t) > I_{Joy}(t-1))$	$I_{Joy}(t) = \lvert I_{Joy}(t-1)\rvert$
(Joy, Distress) (Hope, Fear)	$Des(p, g_i, t)>0$	$I_{Distress}(t-1) < \lvert I_{Distress}(t-1) - I_{Hope}(t)\rvert$	$I_{Distress}(t) = \lvert I_{Distress}(t-1)\rvert$
		$I_{Distress}(t-1) > \lvert I_{Distress}(t-1) - I_{Hope}(t)\rvert$	$I_{Distress}(t) = \lvert I_{Distress}(t-1) - I_{Hope}(t)\rvert$
	$Des(p, g_i, t)<0$	$I_{Joy}(t-1) < \lvert I_{Joy}(t-1) - I_{Fear}(t)\rvert$	$I_{Joy}(t) = \lvert I_{Joy}(t-1)\rvert$
		$I_{Joy}(t-1) > \lvert I_{Joy}(t-1) - I_{Fear}(t)\rvert$	$I_{Joy}(t) = \lvert I_{Joy}(t-1) - I_{Fear}(t)\rvert$

4.5 Cognition module

As shown in Fig.2, the resulting emotion has a major impact on the agent's action selection process. Choosing an action is considered as behavioral responses sent to the environment throw the exchange of messages. Thus, based on the agent's goals, personality traits and the internal emotional state, the agent selects the appropriate action to be executed and send it to the environment. The action selection mechanism is based on a rewarding system; for each action, a weight W is assigned with a value in [0,1]. Then, the immediate reward is a response that can be either a perceived event or another agent's reaction that will update his internal emotional state.

5 SIMULATION AND RESULT ANALYSIS

5.1 The scenario

The scenario considered for the experimentation included an interaction between a set of agents situated in a college. They are sharing the same goals, which are "Success the year", "Success the exam", "Social growth" and "Appreciation". For that, they may perceive events, actions or objects that facilitate or inhibit the realization of these goals which will generate then different emotions. We noted that each agent has his own personality traits. Consequently, each one of them will behave differently according to his personality structure. Table 4 represents a summary of the events, the emotions triggered, the impacted goals, the actions and the correlation between them and the agent's personality traits. We note that the activation conditions describe when an action must be selected according to the personality structure of an agent which is related to the values of O, C, E, A and N.

TABLE 4

PERCEPTUAL DATA CATEGORIES AND POTENTIAL RESULTING EMOTIONS

Goal	Perceptual data	Emotion triggered	Action	Activation conditions
Success the exam	Bad mark Prospect Bad mark	Distress Fear Fear-confirmed OR Relief (if the prospect undesirable event was not realized)	High effort Asking for help	O>=0.5, C>=0.5 N<0.5
			Asking for help Giving up	0<0.5 C<0.5 N>0.5
	Good mark Prospect Good mark	Joy Hope Satisfaction OR Disappointment (if the prospect desirable event was not realized)	Keep working	
Success the year	Failing the activities Prospect Failing the activities	Distress Fear Fear-confirmed OR Relief (if the prospect undesirable event was not realized)	High effort Asking for help	O>=0.5, C>=0.5 N<0.5
			Asking for help Giving up	0<0.5 ,C<0.5 N>0.5
	Have success the activities Prospect success the activities	Joy Hope Satisfaction OR Disappointment (if the prospect desirable event was not realized)	Keep working	
Appreciation	Negative feedback Prospect Negative feedback	Distress Fear Fear-confirmed OR Relief (if the prospect undesirable event was not realized)	Improve level of competence Asking for help	O>=0.5 , C>=0.5 N<0.5
			Asking for help Giving up	O<0.5, C<0.5 N>0.5
	Positive feedback Prospect Positive feedback	Joy Hope Satisfaction OR Disappointment (if the prospect desirable event was not realized)	Finish the activity	
Social growth	Asking for help (other)	Pride (for self) Admiration (for other)	Accepting help	E>=0.5,A>=0.5
		Shame (for self) Reproach (for other)	Rejecting help	E<0.5 ,A<0.5
	Positive events (for other)	Happy for		
	Negative events (for other)	Sorry for Pity		If liking>0
		Sorry for Gloating		If liking =0
	Receiving an object	Like	Accepting	O>=0.5
		Dislike	Rejecting	O<0.5

5.2 Implementation and results

We choose Java Agent DEvelopment framework (JADE) to implement our model as it is a rich and efficient agent-based simulator. We have conceived a supervisor agent who launches 120 different agents with different personality structures; 30 Opened agents, 30 Conscious agents, 30 Extraverted Agreeable agents and 30 neurotics.

One of the most striking results of this model is that the personality traits and emotional internal state play an important role in defining the agent's behavior. The results focus also on the correlation between the personality structure and the intensity of the emotion triggered. For that, we focused on the behavior of six experimented agents belonging to different personality's structures. These agents are: $Agent_{21}$, $Agent_{32}$, $Agent_{38}$, $Agent_{8}$, $Agent_{29}$ and $Agent_{19}$ (Table 5).

TABLE 5

AGENT'S PERSONALITY TRAITS

Agent	O	C	E	A	N
$Agent_{21}$	0.5	0.77	0.91	1	0.06
$Agent_{32}$	0.5	0.02	0.3	0.39	1
$Agent_{38}$	0.9	0.02	0.19	0.22	0.93
$Agent_{8}$	0.9	0.66	0.02	0.05	0.3
$Agent_{29}$	0.88	0.09	1	1	0.19
$Agent_{19}$	0.09	1	0.9	0.83	0.9

Two couples of charts (Fig.3/ Fig.4) display respectively the evolution of emotions for $Agent_{21}$ and $Agent_{32}$ related to the perception of events in order to highlight their impact on emotion's generation.

1) Appraisal of Self-related Events and Generation of Wellbeing Emotions

As we can observe in (Fig.3 and Fig.4) at the beginning of the simulation, both of the agents have perceived the same negative event (Negative feedback) which elicits Distress for both of them. Then, Joy was generated when Positive feedback was perceived. But this emotion was decreased when an undesirable event such Bad mark was appeared. These negative self events give rise to the emotion Distress. Each event is appraised separately. The intensity of each emotion and the desirability of the causing events are recomputed at each time step of the simulation. Therefore, we find two corresponding value of Joy related to the perception of a positive event such Good mark respectively for $Agent_{21}$ and $Agent_{32}$ at t=26 and t=31. The intensities of this emotion, which corresponds to the absolute value of the desirability of the causing events, will be combined in the Internal Emotional State update phase. In fact, the intensity of emotions may increase or decrease depending on the evolution of the causing events. Nevertheless, despite the fact that all the agents have perceived the same events, each one of the agents has reacted differently. In fact and in case of negative events, $Agent_{21}$ try to do a high effort in order to surpass the situation. However, $Agent_{32}$ was waiting only waiting for help from his entourage. Consequently, the evolution of positive and negative emotions was not the same for both of them. In fact and as we can observe in the end of the simulation, $Agent_{21}$ remained in a good situation and maintained the generation of the positive emotions.

2) Appraisal of Prospected Events and Generation of Prospect-Based Emotions

An Agent may perceive prospect positive events such as Prospect Good mark which elicit Hope emotion respectively for $Agent_{21}$ and $Agent_{32}$ at t=21 and t=27. Then, Satisfaction emotion was generated for both of these agents because this event was confirmed. Besides, they may perceive prospect negative events such as Prospect Failing activities which elicit Fear emotion respectively for $Agent_{21}$ and $Agent_{32}$ at t=51 and t=61. These events generate Relief emotion when they are disconfirmed at instant 65 (Fig.3) or Fear confirmed emotion when they are confirmed at instant 65 (Fig.4).

3) The relation between personality traits and emotion generation

The aim is to put different agents with different personalities in the same situation in order to analyze the impact of the psychological aspect on the agent behavior and on triggered emotion. First, we start to analyze the impact of Consciousness on the generation of Joy/Distress emotions. As we have mentioned in Table 5,$Agent_{21}$, is highly conscious who tries usually to achieve all his goals following an organized and methodical behavior. Consequently, he feels usually positive emotions with high intensities (e.g. Joy, Satisfaction) and negative emotions (Distress, Disappointment) with low intensities compared to other agents with low indicator of Consciousness (e.g.$Agent_{32}$) (Fig.3). In fact and as we have mentioned before, $Agent_{32}$ did never overcome any stressful situations by himself which made him unable to achieve any of his goals. As a conclusion, the absence of Consciousness has usually a negative impact on the realization of the agent's goals and on the generation of the wellbeing emotions.

Besides, agents are able to perceive Other-Related Events (ORE) of nearby agents. A Happy for emotion arises from the appraisal of the events that happened to the other agents such Good mark (other) (Fig.3/t=31,). In fact, results show also that Agreeableness and Extraversion influences also the intensity of Empathetic emotions towards other agents. As we can observe in Fig.3/Fig.4, an Extraverted and sociable agent such $Agent_{21}$feels Happy for emotion with high intensities more than $Agent_{32}$(Fig.4/t=36).

Fig.5 illustrates the evolution of the emotions related to the agent's actions according to two different agents with different level of Agreeableness and Extraversion such $Agent_{21}$ and $Agent_{32}$ as mentioned in Table 5. In fact and in case of a help request sent from the other agents in need, an agreeable cooperative agent such $Agent_{21}$ accepts often to help them (Fig.3/t=35). However, $Agent_{32}$ who has a low level of Agreeableness, refuses usually helping even when he is in a good situation (Fig.4/ t=41). Consequently and as shown in Fig.5, Pride and Admiration emotions related to $Agent_{21}$were experienced with high intensities more than $Agent_{32}$who he feels often ashamed as he has received too many reproach messages for too many times. Therefore, Gratification and Gratitude resulting from the evolution of Joy, Pride and Admiration emotions were maintained for $Agent_{21}$(Fig.6). However, Distress, Remorse and Anger emotions were generated and maintained during the rest of the simulation as shown in Fig.7. Secondly and in order to analyze the impact of the Neuroticism on the generation of Fear emotion, Fig.8 shows that a high neurotic level affects deeply Fear intensity emotion. A neurotic agent such $Agent_{32}$ experiences a strong feeling of Fear more than other agents do, as he was often unable to manage any situation by himself. However, $Agent_{21}$ tries to maintain his emotional stability even in case of a stressful situations (e.g. from t=34 to t= 49) as he was emotionally more stable than $Agent_{32}$ .Results show also that agents can change their behaviors. They can adopt new ones from their environment. For example, $Agent_{38}$ has a much opened, not conscious and very nervous personality. As we can see in Fig.9, as the later he has a low level of Consciousness

and in case of negative event and same as$Agent_{32}$, $Agent_{38}$ was only asking for help from his entourage. Thus and in many cases, he remained in the same negative situation because no one accepted to help him as we can see in Fig.9 (t= 11) which has augmented the Distress emotion. Consequently and as he has an opened personality, he tries often to adopt new behaviors that can help him to avoid the negative events (t= 17).So, he must adopt and learn new behavior which is doing high effort instead of asking for help in order to generate positive emotions. As we can observe, $Agent_{38}$ has been influenced by $Agent_{21}$ and has adopted the latter's reaction as they were friends (Liking ($Agent_{38}$,$Agent_{21}$) =0, 88). Therefore and after adopting $Agent_{21}$'s behavior, positive events were perceived (e.g. Joy emotion (t= 21)). Thus, $Agent_{38}$ will try to keep this behavior in order to remain in a good situation (t= 29). Consequently, asking for help's weight decreased until disappearing (t=17) as the new reaction's weight (Doing high effort) has been incremented according to $Agent_{38}$ reward system.

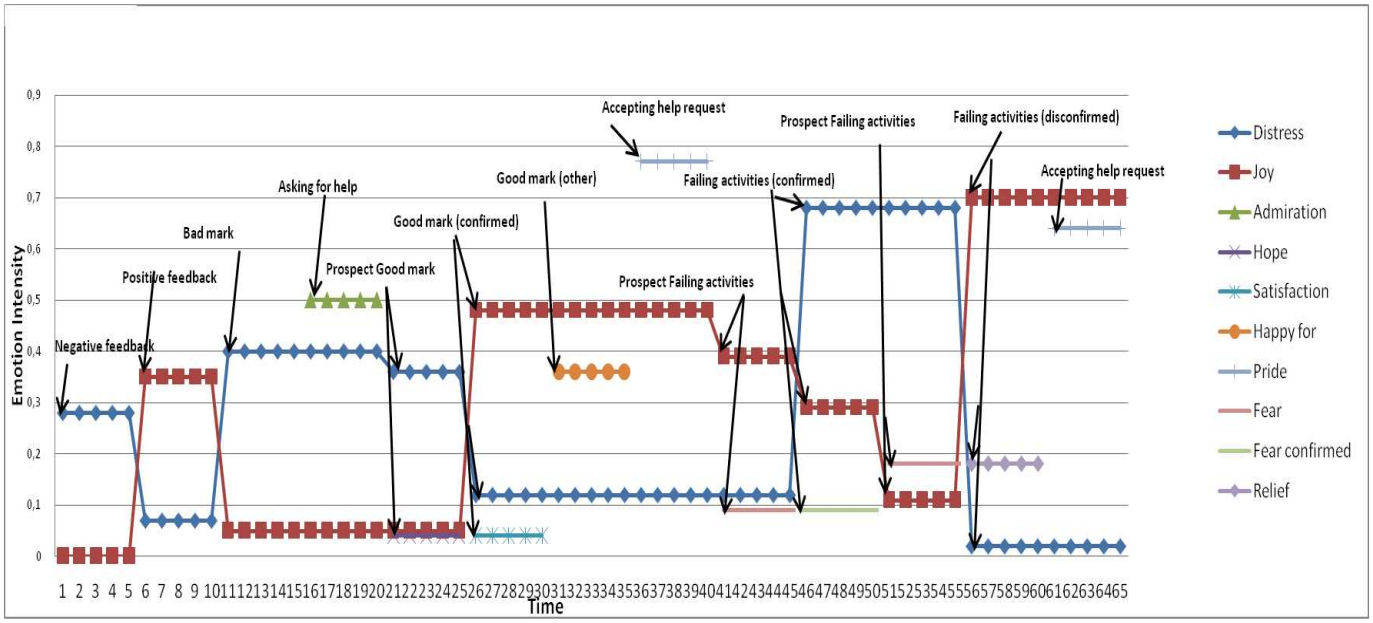


Fig. 3. Evolution of Goal-Based Emotions of Agent $_{21}$

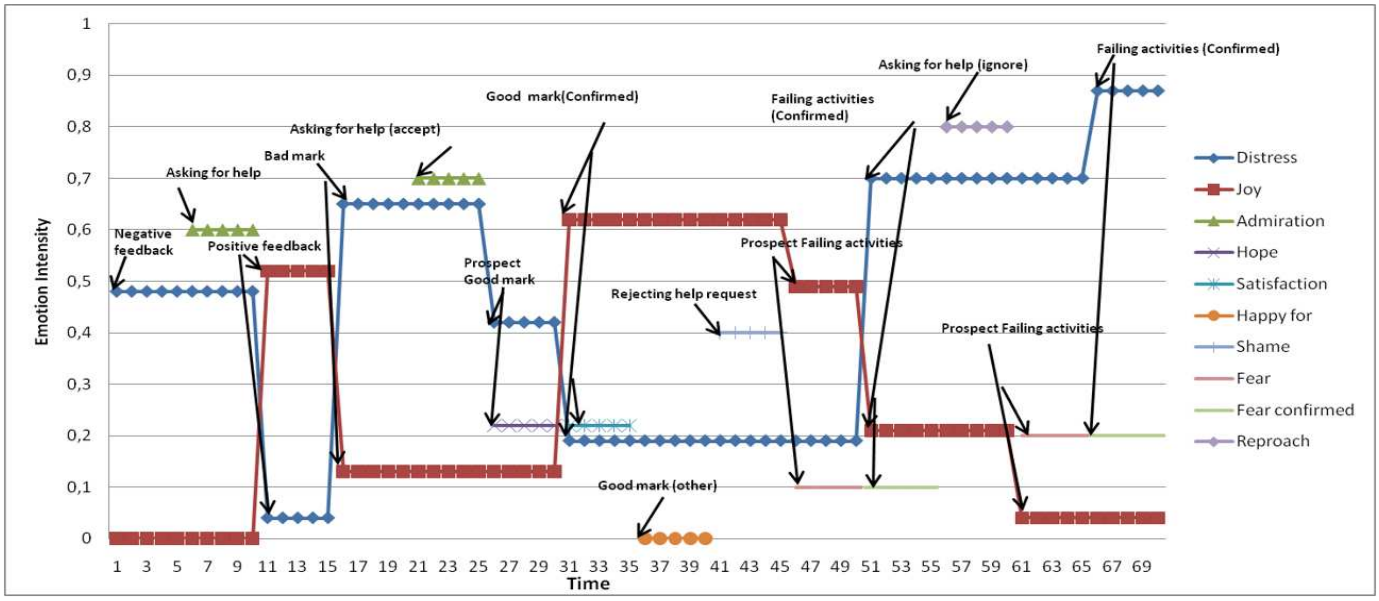


Fig. 4. Evolution of Goal-Based Emotions of Agent $_{32}$

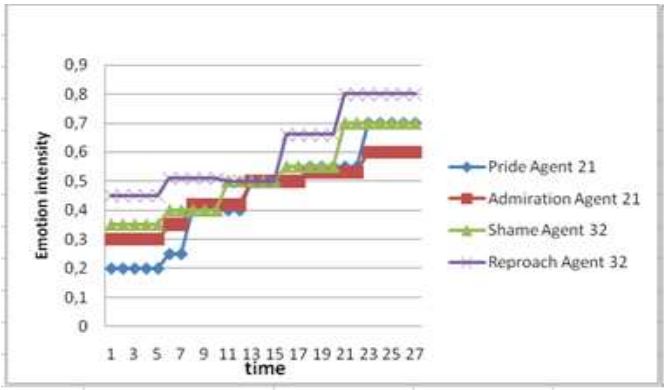


Fig. 5. The relation between Agreeableness and positive emotions

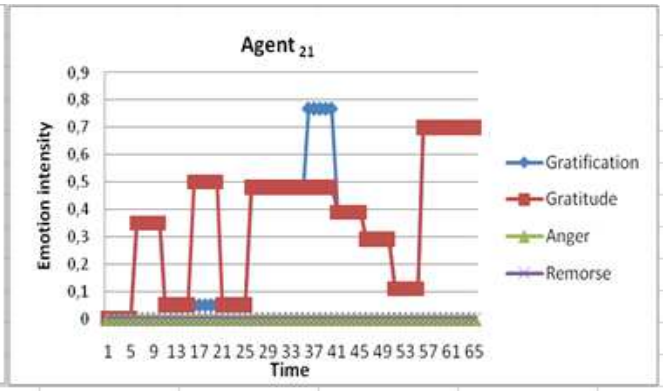


Fig. 6. Evolution of Compound Emotions Agent $_{21}$

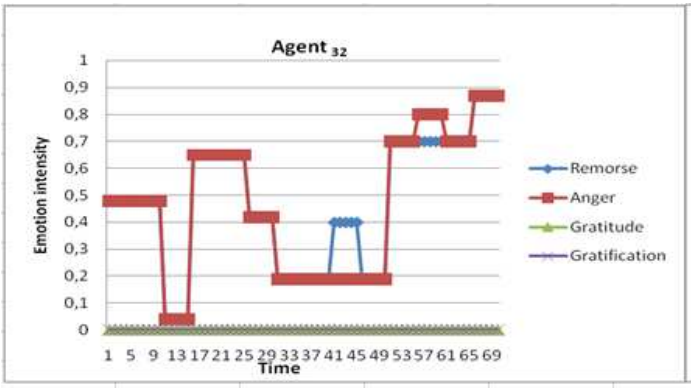


Fig. 7. Evolution of Compound Emotions Agent $_{32}$

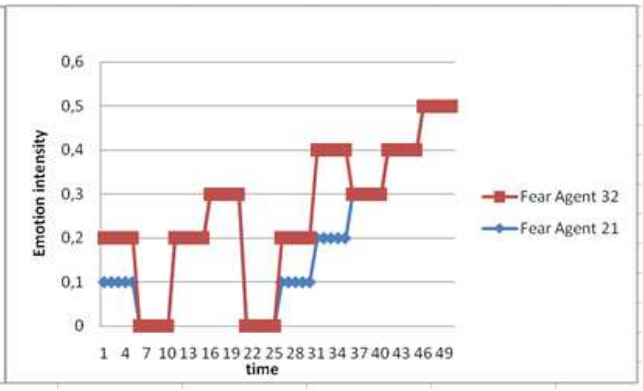


Fig. 8. The relation between Neuroticism and fear emotion

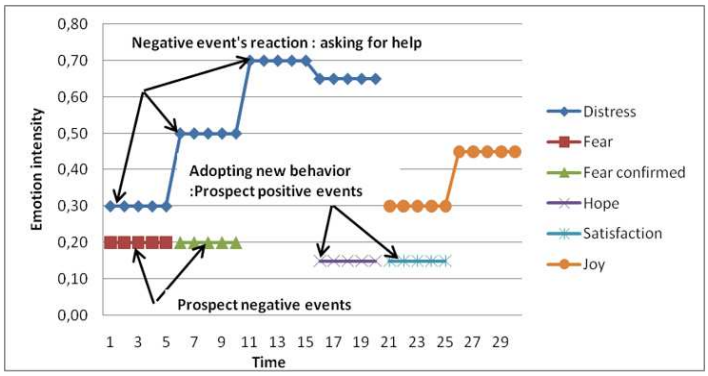


Fig.9. The relation between Openness and the adoption of new behavior

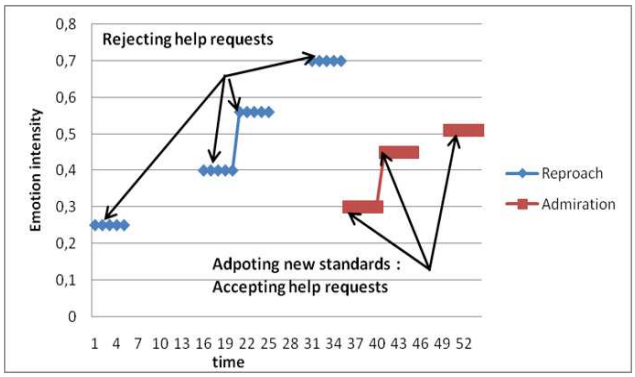


Fig. 10. The relation between Openness and the adoption of new standards

Moreover, opened personalities may adopt also new standards if their own standard does not much with the entourage and inhibit the realization of a social goal (e.g. Social growth).For that, Fig.10 illustrates an example of $Agent_8$who was highly opened but not very agreeable that's why he has received many reproach messages from his entourage as he was not cooperative with the other agents. Therefore and as he has an opened personality, $Agent_8$ tried to change his standards in order to decrease these reproach reactions and to be more friendly with his entourage. Thus, he has adopted new positive standards from other agents e.g. $Agent_{29}$ because he was the most agreeable and extraverted in $Agent_8$'s entourage. As we can observe in Fig.10 and at t= 34, after receiving much reproach messages, $Agent_8$ decided to adopt new standards (t=37).

Finally, results shows that opened personality tend to behave differently towards new objects. For that, Fig.11 describes the behavior of two different agents: $Agent_{29}$ who is much opener than $Agent_{19}$.In fact and when a new object was perceived, $Agent_{29}$ has accepted to discover it which was not the case for the second one $Agent_{19}$. Consequently, the new object becomes more familiar according to$Agent_{29}$ compared to$Agent_{19}$. Then, Like respectively Dislike emotions were generated respectively for $Agent_{29}$ and $Agent_{19}$ as we can observe in Fig.11 bellow.

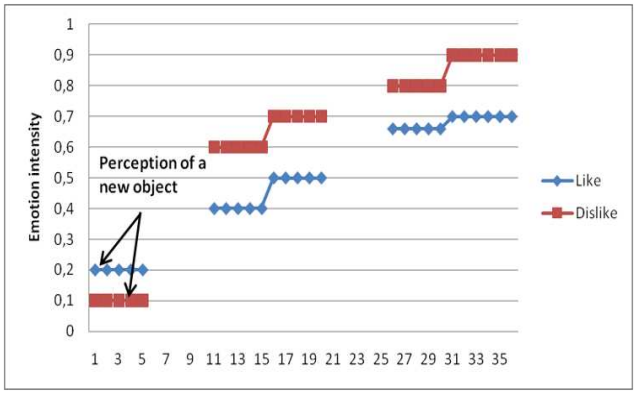


Fig.11. The relation between Openness and the adoption of new preferences

7 Conclusion and Future Works

In this article, after an overview of the related works from the literature, we have introduced a domain-independent computational agent based model which integrates emotions based on appraisal theories of emotions and personality structure based on the OCEAN model of personality. Both of them influence decisions making process on multiple levels. More specifically, the main goal of this model is to highlight the role played by these factors in decision making system where agents have different personalities and emotions in order to simulate human-like behavior. Besides, our implementation results are promising; they indicate a high correlation between the emotional and the psychological factors in human decision making. The appraisal mechanism presented in this paper can be very useful not only for computational emotion modeling researchers but also for cognitive systems researchers who study the effect of emotion on cognition or use emotion in decision systems. In fact, the proposed model will serve to implement the emotion and the personality generation process in any other situation and domain. After that, we intend to implement the effect of emotions on collective behaviors. We also aim to study the interpersonal emotional dynamics through emotional contagion in social complicated phenomenon such the addiction in order to test the impact of such object (the drug substance) on the generation of emotions and on the decision making mechanism. Furthermore, we are planning to run a series of case studies using various age and personality distributions and to compare the results with published data in order to evaluate the validity of the proposed model and simulator.

References

[1] M. Ivanovic et al., "Emotional Agents - State of the Art and Applications," Comput. Sci. Inf. Syst., vol. 12, pp. 47–47, Nov. 2015, doi: 10.2298/CSIS1410260471.

[2] M. Kazemifard, N. Ghasem-Aghaee, and T. I. Ören, "Design and implementation of GEmA: A generic emotional agent," Expert Syst. Appl., vol. 38, no. 3, pp. 2640–2652, 2011.

[3] A. R. Damasio, L'erreur de Descartes: la raison des émotions. Odile Jacob, 2006.

[4] M. Kazemifard, N. Ghasem-Aghaee, and Tuncer Ören, "Agents with ability to understand emotions," in Proceedings of the 2009 Summer Computer Simulation Conference, 2009, pp. 254–260.

[5] A. Ortony, G. L. Clore, and A. Collins, "The Cognitive Structure of Emotions," 1988, doi: 10.1017/CBO9780511571299.

[6] "Affective Computing, Focus on Emotion Expression, Synthesis and Recognition - PDF Free Download," epdf.pub. [Online]. Available: https://epdf.pub/affective-computing-focus-on-emotion-expression-synthesis-and-recognition.html. [Accessed: 17-Jan-2020].

[7] A. Egges, V. Kshirsagar, and N. Magnenat-Thalmann, "Imparting Individuality to Virtual Humans," 2002.

[8] S. Ojha and M.-A. Williams, "Emotional appraisal: A computational perspective," in Fifth Annual Conference on Advances in Cognitive Systems, 2017.

[9] R. R. McCrae and Oliver P. John, "An introduction to the five-factor model and its applications," J. Pers., vol. 60, no. 2, pp. 175–215, 1992.

[10] A. Egges, S. Kshirsagar, and N. Magnenat-Thalmann, "A model for personality and emotion simulation," in International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, 2003, pp. 453–461.

[11] H. Kessler, A. Festini, H. C. Traue, S. Filipic, M. Weber, and H. Hoffmann, "Simplex-simulation of personal emotion experience," Affect. Comput. Emot. Model. Synth. Recognit. -Tech Educ. Publ. Vienna, 2008.

[12] S. Hendrickx, "Enriching the Story by Expressing Emotions," p. 142.

[13] G. L. Clore and A. Ortony, "Psychological construction in the OCC model of emotion," Emot. Rev., vol. 5, no. 4, pp. 335–343, 2013.

[14] A. Bandura, "Social cognitive theory of personality," Handb. Personal., vol. 2, pp. 154–196, 1999.

[15] A. Egges, S. Kshirsagar, and N. Magnenat-Thalmann, "Generic personality and emotion simulation for conversational agents," Comput. Animat. Virtual Worlds, vol. 15, no. 1, pp. 1–13, 2004.

[16] H. J. Eysenck, "Toward a new model of intelligence," Personal. Individ. Differ., vol. 7, no. 5, pp. 731–736, 1986.

[17] M. S. El-Nasr, J. Yen, and T. R. Ioerger, "Flame—fuzzy logic adaptive model of emotions," Auton. Agents Multi-Agent Syst., vol. 3, no. 3, pp. 219–257, 2000.

[18] J. Colloc and C. Bertelle, "Multilayer agent-based model for decision support system using psychological structure and emotional states," Proc. ESM'2004, pp. 325–330, 2004.

[19] S. Jain and K. Asawa, "EMIA: emotion model for intelligent agent," J. Intell. Syst., vol. 24, no. 4, pp. 449–465, 2015.

[20] J. Dias, S. Mascarenhas, and A. Paiva, "Fatima modular: Towards an agent architecture with a generic appraisal framework," in Emotion modeling, Springer, 2014, pp. 44–56.

[21] D. Heylen, M. Poel, and A. Nijholt, "Parlee: An adaptive plan based event appraisal model of emotions," in Annual Conference on Artificial Intelligence, 2002, pp. 129–143.

[22] J. Gratch and S. Marsella, "Evaluating a computational model of emotion," Auton. Agents Multi-Agent Syst., vol. 11, no. 1, pp. 23–43, 2005.

[23] P. Gebhard, "ALMA: a layered model of affect," in Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems, 2005, pp. 29–36.

[24] M. Belhaj, F. Kebair, and L. B. Said, "Agent-based modeling and simulation of the emotional and behavioral dynamics of human civilians during emergency situations," in German Conference on Multiagent System Technologies, 2014, pp. 266–281.

[25] H. L. Riahi, F. Kebair, and L. B. Said, "Agent-based modeling and simulation of the emotional experiences of employees within organizations," in Proceedings of the

Conference on Summer Computer Simulation, 2015, pp. 1–10.

[26] T. H. H. Dang, S. Letellier-Zarshenas, and D. Duhaut, "Comparison of recent architectures of emotions," 2008.

[27] E. J. Mehdi, P. Nico, D. Julie, and P. Bernard, "Modelling character emotion in an interactive virtual environment," p. 9.

[28] C. A. Hutcherson, P. R. Goldin, W. Ramel, K. McRae, and J. J. Gross, "Attention and emotion influence the relationship between extraversion and neural response," Soc. Cogn. Affect. Neurosci., vol. 3, no. 1, pp. 71–79, 2008.

[29] M. Hiebler-Ragger, J. Fuchshuber, H. Dröscher, C. Vajda, A. Fink, and H. F. Unterrainer, "Personality influences the relationship between primary emotions and religious/spiritual well-being," Front. Psychol., vol. 9, p. 370, 2018.

[30] O. Luminet and N. Vermeulen, "Personnalité et psychopathologie cognitive," 2008.

[31] W. Ng, "Clarifying the relation between neuroticism and positive emotions," Personal. Individ. Differ., vol. 47, no. 1, pp. 69–72, 2009.

[32] M. Hansenne, Psychologie de la personnalité. De Boeck Supérieur, 2018.

[33] S. Klamer et al., "Association between neuroticism and emotional face processing," Sci. Rep., vol. 7, no. 1, p. 17669, 2017.

[34] L. Nicholas, Introduction to psychology. Juta and Company Ltd, 2009.

La contribution de la responsabilité sociale de l'entreprise a l'identification organisationnelle : étude exploratoire dans le contexte tunisien

Mohamed Fadhel Djlassi, Doctorant en management à l'ISG Tunis

Chercheur au LARIME – ESSECT – Université de Tunis

jlassi.fadhel@yahoo.fr

Anissa Ben Hassine, Maître des conférences à l'ESSECT

Chercheur au LARIME – ESSECT – Université de Tunis

benhassineanissa@gmail.com

Le résumé

Cet article se propose d'étudier l'effet de la responsabilité sociale de l'entreprise (RSE) perçue par l'employé sur l'identification organisationnelle. Dans ce cadre, une étude qualitative a été menée dans une entreprise multinationale opérant dans le secteur pétrolier en Tunisie (Total Tunisie). 10 entretiens ont été effectués auprès des cadres de ladite entreprise, et une analyse de contenu thématique en a été réalisée. Les résultats font valoir que la perception par les employés de la RSE, orientée aussi bien vers les acteurs sociaux et non sociaux que vers les employés, a un effet significatif et positif sur l'identification organisationnelle alors que les résultats de la RSE orientée vers les consommateurs sont plus mitigés.

Mots-clés : La responsabilité sociale des entreprises, les parties prenantes, la théorie d'identité sociale, l'identification organisationnelle.

***L'auteur correspondant :** Mohamed Fadhel Djlassi, doctorant en management à l'Institut Supérieur de Gestion de Tunis, Tél : (+216) 58 069 305. E-mail :jlassi.fadhel@yahoo.fr

Introduction

Le phénomène de la responsabilité sociale de l'entreprise (RSE) a gagné en importance au fil des années et a fait couler beaucoup d'encre auprès des académiciens et des praticiens. Dans ce cadre, un bon nombre d'études a porté sur les effets de la perception de la RSE sur les comportements et attitudes des clients et des actionnaires (Pettijohn, Pettijohn et Taylor, 2007 ; Stites et Michael, 2011). Néanmoins, rares sont les études qui ont porté sur l'effet interne de la perception de la RSE (c'est-à-dire sur les employés) (Peterson, 2004 ; De Roeck et al, 2016).En outre, les recherches antérieures ont eu tendance à utiliser une mesure globale de la perception de la RSE, sans s'intéresser à la distinction entre les actionsde la RSE orientées envers les différentes parties prenantes, dont la société, les organisations non-gouvernementales, les consommateurs et les employés (Stites et Michael, 2011, Turker 2009a). Enfin, lesétudes effectuées dans ce domaine se sont particulièrement concentrées sur l'impact de la perception de la RSE sur l'engagement organisationnel (Brammer et al, 2007 ; Maignan &Ferrell, 2001 ; Peterson, 2004) et sur la satisfaction au travail (Valentine et Fleischman, 2008 ; Tziner et al., 2011). Il existe, à notre connaissance, peu de recherche ayant étudié l'effet des différentes composantes de la RSE(l'effet différentiel)sur une attitude qui nous semble motrice dans l'explication des comportements au travail, à savoir l'identification organisationnelle (Newman. A, et al., 2015 ; Shen et Benson (2016).

L'identification organisationnelle est, en effet, devenue un pilier de la recherche en management. Elle sert, notamment, à comprendre comment une personne acquiert le sentiment de « faire partie d'une organisation ». D'ailleurs, un nombre croissant de recherches ont tendance à la mobiliser, en tant que ressource inimitable génératrice d'avantage concurrentiel (Cardador and Pratt, 2006, Fiol, 2001). C'est pourquoi il nous semble intéressant d'étudier l'effet de la RSE perçue par l'employé sur l'identification organisationnelle. Pour cela, nous allons tout d'abord présenter le cadre conceptuel de cette recherche en nous intéressant notamment aux deux concepts de base étudiés, à savoir la RSE et l'identification organisationnelle. Suite à quoi nous présenterons notre modèle de recherche et nos propositions que nous validerons dans le cadre d'une recherche exploratoire qualitative dont nous présenterons la méthodologie ainsi que les principaux résultats.

Vers une vision partenariale de la RSE

La plupart des définitions de la responsabilité sociale des entreprises décrivent ce concept comme l'intégration volontaire des préoccupations sociales et écologiques des entreprises à leurs activités commerciales et leurs relations avec leurs parties prenantes. Être socialement responsable signifie non seulement satisfaire pleinement aux obligations juridiques applicables, mais aussi aller au-delà de ce qu'imposent les réglementations en vigueur et investir dans le capital humain, l'environnement et le développement des relations avec les diverses parties prenantes (Freeman,1984).

Plusieurs théories, ou combinaison de théories (Chen et Roberts, 2010), peuvent servir de cadre de référence à la RSE, dont la théorie institutionnelle (Husted et Allen, 2006 ; Campbell, 2007), la théorie basée sur les ressources (Russo et Fouts, 1997), la théorie du contrat social (Donaldson et Dunfee, 2002) ou encore la théorie de la citoyenneté de l'entreprise (Matten et Crane 2005). Mais la théorie des parties prenantes nous semble la plus pertinente dans le cadre de cette recherche dans la mesure où elle classe les activités de l'entreprise en fonction des acteurs (intervenants) envers lesquels l'entreprise est socialement responsable (Shropshire et Hillman, 2007). Notre recherche s'intéresse justement à l'un de ces acteurs, le personnel, que nous

considérons comme l'intervenant le plus important dans une organisation, sachant que la manière dont les pratiques de RSE le touchent, influe sur son travail (Peterson, 2004).

La RSE et les attentes des parties prenantes

Notre conception de la perception de la RSE est inspirée des travaux de Turker(2009a) qui a fait valoir quatre dimensions de la RSE(RSE envers les acteurs sociaux et non sociaux, RSE envers les consommateurs, RSE envers les employés et RSE envers le gouvernement). Nous avons retenu trois de ces dimensions, car la quatrième dimension, RSE orientée vers le gouvernement, est considérée par bon nombre d'auteurs comme une obligation minimale, ne s'apparentant pas à un comportement socialement responsable (Jones, 1980 ; Sims, 2003). Mieux encore, Sims (2003) définit la RSE comme « *des actes qui vont au-delà de ce qui est prescrit par la loi* ». Il nous semble donc que la satisfaction des exigences légales relève plutôt des devoirs que de la responsabilité d'une entreprise.

La RSE visant les acteurs sociaux et non sociaux, a pour cible les parties prenantes secondaires qui ne sont pas directement touchées par les activités de l'entreprise à court terme (Turker, 2009b). Cet auteur la définit comme l'ensemble des actions qui visent essentiellement la communauté, l'environnement et la société (bien-être général et intérêts des générations futures). Plus précisément, elle comprend la participation à la protection de l'environnement naturel et les initiatives visant à rendre la vie meilleure pour les générations futures, à encourager le développement durable et à contribuer à des projets et des activités qui favorisent le bien-être de la société dans sa globalité (Hofman et Newman, 2014).

Quant à la RSE orientée vers les employés, elle fait référence aux pratiques de management équitable et flexible de la part de l'entreprise envers ses collaborateurs, et ce, via l'offre de possibilités de développement professionnel et personnel de l'employé et la satisfaction de ses besoins et aspirations (Turker, 2009b).

Enfin, la RSE qui vise les consommateurs met, pour sa part, l'accent sur la satisfaction de ces derniers et sur l'engagement du respect des droits des consommateurs par la mise en place, par exemple, d'un système de communication externe transparent et responsable (Hofman et Newman, 2014).

L'identification organisationnelle et la théorie de l'identité sociale

L'identification organisationn elle a été définie par Ashforth et Mael (1989) comme une forme spécifique d'identification sociale dans laquelle les individus se définissent par la perception d'unicité (ou le sentiment d'appartenance) à une organisation particulière. Van Dick et al., (2006) ajoutent par ailleurs qu'un haut degré d'identification organisationnelle au niveau des employés est toujours souhaitable et sain pour toutes les organisations et certains auteurs rajoutent que l'attachement cognitif mène à l'attachement émotionnel sous forme de sentiment d'appartenance et d'engagement ; ce qui va conduire davantage à des comportements de nature citoyenne (Brammer et al., 2007 ; Tyler et Blader , 2002 ; Collier et Esteban, 2007).

Quant à la théorie de l'identité sociale (TIS), c'est un cadre théorique qui sous-tend particulièrement la relation entre la perception de la RSE et l'identification organisationnelle (Cropanzano et Mitchell, 2005). Ainsi, avoir une identité sociale particulière, c'est être parmi les membres d'un certain groupe, être comme les autres dans le groupe, voir les choses du point de vue du groupe (Stets et Burke, 2000), donc s'identifier au groupe.

La RSE envers les acteurs sociaux et non sociaux et l'identification organisationnelle

Selon la théorie de l'identité sociale, l'appartenance à une entreprise qui fait des sacrifices pour préserver le bien collectif de la société et pour promouvoir la communauté dans laquelle elle opère est susceptible de générer auprès de l'employé un sentiment de fierté à y travailler et à hisser le niveau de l'identification organisationnelle dans cette entreprise(Turker, 2009b).

Dans ce cadre, De Roeck et Delobbe (2012) ont mené une étude auprès de 155 employés d'une entreprise pétrochimique afin de mieux comprendre pourquoi, comment et dans quelles circonstances, les employés pourraient répondre positivement aux actions de RSE de cette entreprise. Les résultats de cette étude soutiennent que la perception de la RSE, notamment celle orientée vers l'environnement, est positivement liée à l'identification organisationnelle des employés ; une identification qui s'érige en déterminant important des comportements individuels. Les résultats montrent également que la relation entre la perception de la RSE et l'identification organisationnelle des employés est médiatisée par la confiance envers l'organisation. Plus précisément, ces auteurs ont démontré que les initiatives de la RSE destinées à d'autres intervenants (faisant référence dans leur étude à l'environnement naturel) ont le potentiel de faire valoir une position éthique de l'entreprise envers ses membres, ce qui rejoint les travaux antérieurs selon lesquels les actions de RSE sont de nature à véhiculer aux employés certains traits et certaines valeurs organisationnelles, dont la bienveillance et l'honnêteté, qui vont à leur tour les rassurer à propos de la crédibilité de leur entreprise (Hansen et al., 2011). Ces derniers auteurs avancent, dans le même ordre d'idées, que l'identification organisationnelle soutient le processus de la catégorisation de soi auprès des employés, un processus par lequel, ils tendent à adhérer à des entreprises qui partagent leurs valeurs afin de renforcer leur image de soi et donner plus de sens à leur vie.

De même, Jones (2010) a mené une étude auprès des employés d'une entreprise qui exerce dans le domaine du commerce et a abouti au résultat selon lequel l'identification organisationnelle est un processus psychologique qui médiatise la relation entre les attitudes des employés envers les programmes de bénévolat de leur entreprise et leurs comportements. Particulièrement, l'auteur a fait référence au processus de valorisation de soi des employés pour démontrer que la relation entre la RSE, en l'occurrence, le programme de bénévolat et l'identification organisationnelle des employés est partiellement médiatisée par la fierté des employés décrite comme le lien psychologique entre les membres de l'organisation et le concept de soi (Jones, 2010 ; Riketta, 2005).

Suite à ces écrits, nous pouvons donc émettre la proposition suivante :

Proposition 1 : la perception de la RSE envers les acteurs sociaux et non sociaux par les employés, a un effet positif sur l'identification organisationnelle des employés.

La responsabilité sociale envers les consommateurs et l'identification organisationnelle

Les employés d'une entreprise interagissent constamment avec les clients et reçoivent leurs feedbacks au sujet de la perception qu'ils ont du service de leur entreprise, de la qualité des produits et même du rôle plus large de l'entreprise dans l'environnement. Toujours en se référant à la théorie de l'identité sociale, si les employés reçoivent des commentaires positifs de la part des consommateurs sur les pratiques de leur entreprise, y compris les pratiques de la RSE, ils tendent à interpréter ce feedback d'une manière positive, ce qui renforce leur sentiment d'identification organisationnelle (Turker, 2009a). Toujours dans la même veine, Turker (2009b) soutient que, si une organisation induit en erreur ses clients ou bien produit des produits dangereux, les employés peuvent en ressentir de la honte et s'y identifier moins.

Nous pouvons donc dire que si une organisation prête attention à ses clients, les employés ressentent de la fierté à lui appartenir, parce que, selon la théorie d'identité sociale, les membres d'une catégorie sociale sont susceptibles de partager son succès mais aussi, le cas échéant, son échec. En outre, l'étude faite par Farook et al, (2014) a fait valoir que la RSE axée sur le consommateur a un effet positif sur l'identification organisationnelle des employés. Ce qui nous permet d'émettre la proposition suivante :

Proposition 2 : la perception de la RSE envers les consommateurs par les employés, a un effet positif sur l'identification organisationnelle des employés.

La RSE à l'égard des employés et l'identification organisationnelle

La RSE à l'égard des employés se manifeste de diverses manières. Nous pouvons en citer l'amélioration de la flexibilité en milieu de travail, l'octroi d'une rémunération compétitive, le développement d'opportunités de carrière et la mise en avant de la justice organisationnelle (Turker, 2009a). En effet, lorsque les employés ont des perceptions positives quant aux réponses de leur organisation par rapport à leurs besoins et à ceux de leurs collègues, ils sont plus enclins à considérer que l'organisation partage des valeurs semblables aux leurs, et à développer ainsi des niveaux élevés d'identification organisationnelle. De même, en préservant le bien-être des employés, l'entreprise est susceptible d'avoir une réputation externe positive en tant que bon employeur. Cette réputation débouche sur l'amélioration de l'estime de soi des employés et l'accroissement de leur identification organisationnelle (Hofman et Newman, 2014).

De surcroit, l'évaluation que font les employés du niveau de respect avec lequel l'organisation les traite peut influencer leur identification organisationnelle via l'amélioration de leur statut perçu au sein de l'organisation (Tyler et Blader, 2002).

Toujours dans le cadre d'une responsabilité sociale orientée vers les employés, Newman et al, (2016), ont étudié l'influence de trois dimensions de la gestion socialement responsable des ressources humaines (GSRRH), à savoir La GRH orientée loi, la GRH orientée vers les employés et la GRH basée sur la facilitation générale de la RSE, sur le comportement de citoyenneté organisationnelle. Les résultats indiquent que l'identification organisationnelle médiatise totalement l'influence de la GSRRH sur le comportement de la citoyenneté organisationnelle. Ainsi, la GSRRH est de nature à permettre de développer des niveaux supérieurs d'identification organisationnelle.

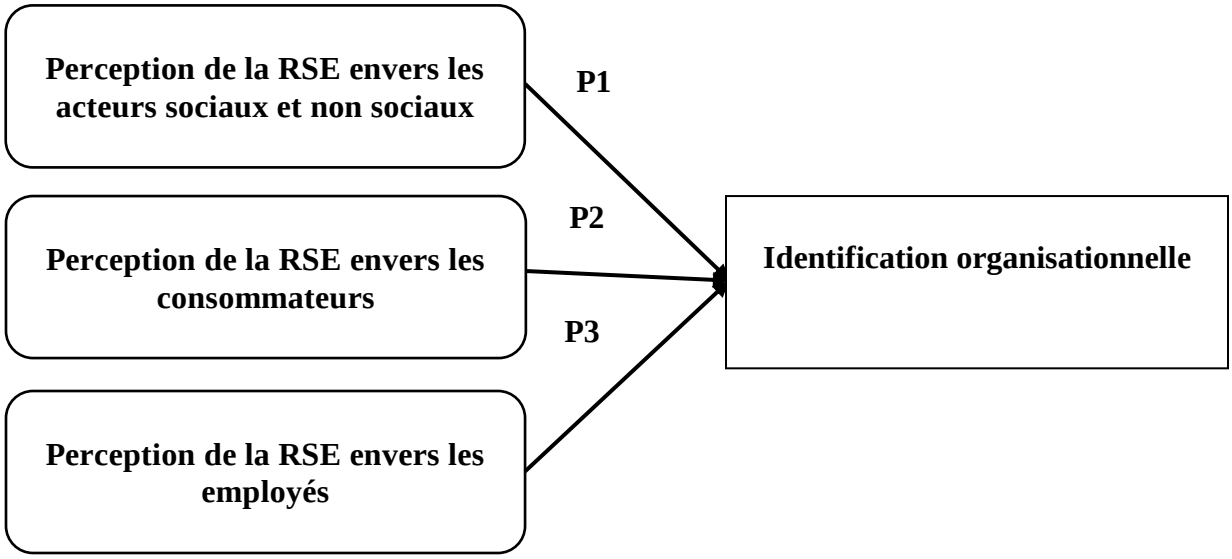
Quant à Shen et Benson (2016), ils ont développé et testé un modèle expliquant les mécanismes par lesquels le management des ressources humaines socialement responsable (MRHSR) affecte le comportement des employés au travail en se focalisant sur deux dimensions (performance à la tâche et comportement d'aide au travail). Leur étude a conclu aux effets du HRMSR sur la performance à la tâche et le comportement d'aide, par le tremplin de l'identification organisationnelle, ce qui corrobore les résultats des recherches antérieures (Carmeli et al., 2007).

L'ensemble de ces constants nous permet d'émettre la proposition suivante :

Proposition 3 : la perception par les employés de la RSE orientée vers ces derniers, a un effet positif sur l'identification organisationnelle.

Suite à ces propositions de recherche étayant la relation entre la perception de RSE orientée vers les trois acteurs d'une part, et l'identification organisationnelle d'autre part, le modèle de notre recherche se présente comme suit :

Figure n°1 : Proposition d'un modèle conceptuel de recherche



La méthodologie de recherche et le cadre opératoire

Pour mener à bien notre travail de recherche, nous avons opté pour une méthodologie qualitative à visée exploratoire. Cette dernière permet une lecture approfondie et variée de l'objet de recherche étudié (Goles et Hirscheim, 2000). A ce propos, Leech et Anthony (2007) ; Mucchielli (1994) et Smith (1983) soutiennent que les méthodes qualitatives, développées en sciences humaines, sont des méthodes qui prospectent, formulent et analysent des phénomènes qui sont peu explorés, ce qui est le cas des relations que nous nous proposons d'étudier.

De même et dans le cadre d'une recherche qualitative, les techniques d'échantillonnage sont moins strictes que celles recommandées dans les études quantitatives. A cet égard, Blanchet et Gotman (1992) affirment que « *la constitution d'un échantillon est souvent guidée par l'obligation de déterminer les personnes dont on estime qu'ils sont capables de fournir des réponses aux questions que l'on se pose* ». Dans ce cas d'espèce d'ailleurs, la définition de la taille de l'échantillon n'est pas déterminée dès la première étape, mais à l'issue du processus de l'enquête. En d'autres termes, c'est la progression des entretiens qui dicte la taille de l'échantillon, tout en obéissant au principe de saturation (Van der Maren, 1996).

Afin de s'assurer de la pertinence de notre modèle de recherche, nous avons choisi de mener des entretiens semi-directifs auprès d'un échantillon d'employés de TOTAL Tunisie, une filiale d'une multinationale française opérant dans le secteur pétrolier. Les entretiens constituent, en effet, l'outil de prédilection des études exploratoires dans la mesure où ils sont eux-mêmes un processus exploratoire permettant diverses pistes de réflexion (Charreire et Durieux, 1999 ; Hollowitz et Wilson, 1983 ; Blanchet, Gotma et De Singly,1992).

Nous avons amorcé nos entretiens par des questions relatives à la perception qu'ont les employés de la RSE pratiquée par leur entreprise (la RSE orientée vers les acteurs sociaux et non

sociaux, la RSE orientée vers les consommateurs et la RSE orientée vers les employés). Puis nous avons essayé d'explorer l'impact que pourrait avoir une démarche de RSE (dimension par dimension) sur l'identification organisationnelle. Les entretiens ont été réalisés en face à face dans les bureaux des interviewés au mois de février 2017. Les employés interrogés nous ont consacré, en moyenne,35 minutes pour répondre à nos questions. Nous avons pu obtenir l'autorisation d'enregistrer nos différents entretiens, ce qui nous a permis de travailler, par la suite, en profondeur à partir des réponses données. Une fois les entretiens enregistrés, ils ont fait l'objet d'une retranscription linéaire et d'une analyse de contenu, sachant que nous avons pu atteindre la saturation de l'information au bout du 10ème entretien. La composition socio-démographique des répondants est la suivante : 3 Hommes et 7 Femmes, cadres moyens et supérieurs, avec une ancienneté moyenne de 9 ans(tableau n°1).

Tableau n°1 : La composition socio-démographique des répondants

Identifiant du participant	Le sexe	Ancienneté au sein de l'entreprise	Durée de l'entrevue
Participant 01	Femme	14 ans	35 mn
Participant 02	Femme	7 ans	37 mn
Participant 03	Femme	2 ans	40 mn
Participant 04	Homme	2 ans	30 mn
Participant 05	Femme	12 ans	38 mn
Participant 06	Femme	9 mois	31 mn
Participant 07	Femme	15 ans	33 mn
Participant 08	Homme	8 mois	27 mn
Participant 09	Homme	10 ans	25 mn
Participant 10	Femme	11 ans	29 mn

Source : Les interviewés

Total Tunisie et la responsabilité sociale de l'entreprise

Filiale du groupe pétrolier Total, Total Tunisie compte 277 collaborateurs et génère plus de 2 000 emplois indirects ; ses activités sont réparties sur 8 sites industriels et 157 stations-service dédiées à la distribution de carburants. En matière de responsabilité sociale de l'entreprise, Total Tunisie réalise des actions socialement responsables envers tous les acteurs, en l'occurrence les acteurs sociaux et non sociaux, les consommateurs et les employés. Ces actions sont présentées au niveau du tableau suivant :

Tableau n°2 : La RSE à TOTAL Tunisie

Les acteurs ciblés	Les actions socialement responsables de Total Tunisie
Les acteurs sociaux et non sociaux	▪ La promotion de l'égalité des chances et la lutte contre la discrimination homme-femme, ont été mises en place dans le cadre du partenariat avec FACE. ▪ La signature de chartes en faveur de l'égalité des chances et la mise en œuvre de partenariats avec le secteur public et les associations. ▪ Un engagement fort envers la diversité, l'intégration des personnes handicapées et l'emploi des minorités visibles. ▪ L'octroi de la possibilité d'une expérience enrichissante en entreprise aux futurs diplômés, leur permettant de valoriser leur savoir-faire et d'accroître leurs chances de succès sur le marché du travail (le « Young Graduate Program », 2014). ▪ La réduction des émissions liées aux installations ; la réduction des émissions liées aux produits (Total Ecosolutions) ;le captage-stockage de dioxyde de carbone (CO2) et le développement d'énergies peu émettrices de gaz à effet de serre.
Les consommateurs	▪ L'instauration d'un climat de confiance avec les clients tout en satisfaisant leurs besoins et en leur offrant des prestations de qualité. ▪ La mise e place d'une écoute personnalisée et des services adaptés à aux besoins des consommateurs. ▪ Le respect des standards de qualité et de sécurité tout au long de son processus de distribution,
Les employés	▪ La mise en place d'outils permettant d'adapter le capital humain aux besoins de l'entreprise. ▪ Mettre en avant des parcours variés et motivants par l'entreprise, et ce, via de nombreux programmes de formation aux employés, afin de leur permettre de progresser en permanence pour évoluer dans leurs métiers. ▪ La mise en place d'une politique de rémunération individualisée, équitable, motivante et diversifiée, tout en favorisant l'équilibre entre vie privée et vie professionnelle, et en encourageant le travail à temps partiel, la flexibilité des horaires ou le travail à temps partagé. ▪ La lutte contre toute forme de discrimination en termes de recrutement et d'évolution, pour le soutien de l'égalité professionnelle et la féminisation de l'encadrement, l'intégration et le maintien dans l'emploi des personnes en situation de handicap.

Source : www.total.tn

Résultats et discussion

L'analyse thématique du contenu des entretiens que nous avons réalisés nous a permis de cerner les perceptions de la RSE par nos répondants, et d'évaluer les effets sur leur identification organisationnelle.

Nous nous sommes tout d'abord focalisés sur la représentation que se font les employés de Total Tunisie des pratiques de RSE introduites par leur entreprise et sur la manière dont cette

perceptions' exprime dans leurs propos. Pour ce faire, nou savons demandé auxemployéss'ils ont connaissance des pratiques de RSE réalisées par leur entreprise et s'ils y participent éventuellement. Par la suite, nous nous sommes intéressés à l'impact de la perception de ces actions par les employés interrogés sur leur identification organisationnelle.

La perception par les employés de la RSE orientée vers les acteurs sociaux et non sociaux

De prime abord, les participants ont mis l'emphase sur le caractère distinctif de leur entreprise en prônant son système de valeur. A titre d'exemple le participant n° 5 a affirmé : « *Notre valeur globale c'est le développement durable, et ça nous a été inculqué par nos responsables dans le cadre de la RSE depuis 2 ans, une notion qui est devenue une culture au sein de Total ».*

Plus précisément, et en se basant aussi sur les déclarations d'autres interviewés, nous avons pu constater que les actions de Total Tunisie en matière de responsabilité sociale qui vise les acteurs sociaux et non sociaux, est perceptible par les employés interrogés. Dans ce sens, un des employés (participant 7) affirme que « …. *Depuis longtemps j'entends parler de la RSE et de la notion du développement durable au sein de Total ... ».* Ce volet semble également bien perçu par les répondants car ils s'y identifient et en parlent avec fierté. Ainsi, le participant 5 affirme :**«** *Bien sûr nous visons à respecter l'environnement et d'ailleurs, nous avons des stations photovoltaïques afin de maitriser l'énergie et accéder à une énergie propre et durable, il en va de l'intérêt des générations futures ».*

Nous tenons à signaler par ailleurs que la quasi-totalité des interviewés sont au courant des pratiques de RSE mises en place par l'entreprise pour favoriser l'employabilité. A titre d'exemple, le participant n° 8 déclare que : *Nous faisons des campagnes pour encourager les jeunes en matière d'entrepreneuriat et de création de projets, on leur donne des dons pour s'installer à leur propre compte, Total est considéré comme un partenaire pour ces jeunes, on les accompagne dans leurs projets et on fait un suivi a postériori pour les aider éventuellement ».*

La perception par les employés de la RSE orientée vers les consommateurs

S'agissant de ce volet, les répondants ont peu de visibilité quant aux activités de RSE qui visent les consommateurs. A ce propos, le participant n°10 affirme que « *…vraiment là, je n'ai pas d'idées…* » ; alors que le participant n° 7 avance que « *En fait, j'entends parler de quelques campagnes publicitaires de la sensibilisation de nos clients au niveau des stations et boutiques afin d'éviter les produits périmés et leur assurer une sécurité, c'est tout ».* En revanche, le participant n°1 nous a présenté l'intérêt qu'accorde Total Tunisie à ses clients en affirmant que : *« notre entreprise mène périodiquement des campagnes de sensibilisation pour les clients afin d'adopter de nouveaux comportements au niveau des stations pour éviter les dangers.* ». La RSE de Total envers les consommateurs, comprend également le volet qualité des produits et emballage, ce qui a été constaté d'après les propos du participant n°6 qui affirme que : « *Oui on tient compte de l'intérêt de nos clients au niveau de la qualité de nos produits (lubrifiants et essence), et on veille toujours sur leur sécurité et on prend soin de l'emballage aussi ».*

Suite à ces affirmations, il apparaît que les employés de Total interrogés ne perçoivent qu'en partie certaines actions de RSE qui visent les consommateurs, les résultats pour ce volet ne sont pas significatifs et ne permettent pas de trancher.

La perception par les employés de la RSE orientée vers les employés

Tous les employés interrogés perçoivent bien les actions de RSE effectuées en interne, et notamment le système de valeurs de l'entreprise qui nous a été cité par les répondants :solidarité, écoute, transversalité et équité. A ce propos, le participant n° 1 avance que : « *On ressent la*

responsabilité de notre entreprise en interne, et ce, via un budget énorme pour assurer une sécurité dans le milieu du travail ». Toujours en interne, la RSE se déploie également, selon nos répondants, via des actions de formations ; des formations qui constituent d'après les propos du participant n° 7 : *« L'un des traits saillants caractérisant Total, c'est la formation métier et la formation managériale … notre climat social est serein, et on encourage la responsabilisation de chaque membre », outre « le dialogue et le climat de confiance* » d'après le témoignage du participant n°9. En définitive, nous pouvons conclure que les employés interviewés de Total Tunisie sont conscients de la RSE initiée par leur entreprise, RSE qui vise les acteurs sociaux et non sociaux et des initiatives prises par Total Tunisie et qui leur profitent (RSE orientée vers les employés) ; avec une perception limitée des actions de RSE qui profitent aux consommateurs.

Effet de la perception par les employés de la RSE orientée vers les acteurs sociaux et non sociaux sur l'identification organisationnelle

Notre étude qualitative nous a également permis de démontrer que la perception par l'employé de la responsabilité sociale de l'entreprise orientée vers les acteurs sociaux et non sociaux est un des prédicteurs significatifs de l'identification organisationnelle des employés rencontrés. En effet, un sentiment d'appartenance est perceptible à partir des affirmations des employés interviewés suite aux actions socialement responsables envisagées par Total Tunisie et qui visent les acteurs externes. Dans ce cadre, le participant n°8 affirme : *« Total est une grande maison[1] j'ai des sentiments envers mon entreprise et je veux bien y rester membre et travailler, ils ont investi particulièrement dans le domaine du développement durable en protégeant l'environnement naturel et en se souciant des générations futures ».* Ces propos rejoignent les résultats de la recherche deRodrigo et Arenas (2008) qui précise que si les employés sentent que leur organisation se comporte comme un bon citoyen, ils sont fiers d'en faire partie. Par ailleurs, Bartels (2006) rajoute que le concept de l'identification organisationnelle en tant que sentiment d'appartenance semble être un phénomène résultant d'objectifs communs partagés avec les membres de l'organisation. Dans le même ordre d'idées, le participant n°6 affirme « *…je m'identifie au système de valeurs prôné par mon entreprise, d'ailleurs je les vois en moi ces valeurs-là car il s'agit d'une firme soucieuse des intérêts de tout le monde.*

Ainsi, notre proposition (1) est confirmée d'après les réponses des personnes que nous avons interrogées. Ce résultat peut s'expliquer en partie par la préoccupation croissante par les employés des problèmes environnementaux, et du fait qu'ils sont devenus plus conscients des conséquences de la dégradation de l'environnement et de la communauté (Turker, 2009a).

Effet de la perception par les employés de la RSE orientée vers les consommateurs sur l'identification organisationnelle

Nos résultats ayant démontré que la RSE envers les consommateurs ne parvenait pas à être perçue par les employés, **il ne nous a pas été possible de tester la proposition (2) dans le cadre de cette recherche.**

Effet de la perception par les employés de la RSE orientée vers les employés sur l'identification organisationnelle

A la lumière des discours tenus au sujet de l'identification organisationnelle, nous avons constaté que certains répondants puisent leur identification dans leur forte adhésion et croyances aux

[1] *« Dar kbira » en arabe*

valeurs véhiculées par l'entreprise. Ce sont des valeurs qui alimentent leurs « expériences de socialisation » (Lorence et Mortimer, 1985). Plus précisément, outre le comportement socialement responsable dont fait preuve Total envers les acteurs externes, cette entreprise s'engage dans des initiatives de RSE qui ciblent les employés, des initiatives qui ont débouché sur davantage d'attachement et d'appartenance de ces derniers. Dans ce sens, le participant n° 2 a passé en revue des pratiques socialement responsables notamment dans le domaine de la santé et de l'ergonomie ; à cet égard, il affirme que *: « Total à mon sens est très soucieuse de la santé de ses employés, je vous cite par exemple les actions de sensibilisation de lutte contre les incendies..... Comme je vous ai dit, cet intérêt crée chez moi un sentiment d'appartenance au vrai sens du terme et je me sens un acteur dans la prise de décision ».*

Cette affirmation rejoint les résultats de la recherche de Shen et Benson (2016), qui ont mis en évidence l'efficacité de la mise en œuvre de programmes de RSE, en l'occurrence l'adoption de la RSE qui s'adresse aux employés. ; et qui ont constaté que l'identification organisationnelle médiatise totalement la relation entre le management des ressources humaines socialement responsable d'une part et la performance à la tâche et comportement d'aide d'autre part, cette relation étant entièrement médiatisée par l'identification de l'organisation.

Ces résultats corroborent également les travaux antérieurs en la matière (Carmeli et al., 2007). Ce qui fait qu'il nous semble qu'il y a un sentiment d'appartenance fort à l'entreprise Total, un sentiment éprouvé par la quasi-totalité des employés rencontrés. Ainsi, le fait que l'entreprise Total a toujours traité ses salariés convenablement, s'est répercuté positivement sur leur attachement à l'entreprise ; ce qui rejoint la conclusion de Tyler et Blader(2002) selon laquelle l'évaluation que font les employés du niveau de respect avec lequel l'organisation les traite influence leur identification, et ce, par l'amélioration de l'état perçu au sein de l'organisation.

A ce propos, le participant n° 6 affirme que : « … *mon travail m'importe beaucoup et je ne peux pas m'imaginer dans une autre entreprise ... Je partage avec mon entreprise et mes collègues toutes les valeurs que je vois autour de moi ».*

De même, les employés de Total éprouvent une loyauté envers l'entreprise, et ce, suite aux programmes de RSE qui leur profitent, comme l'avance le participant n° 1 « : *Je me sens engagé dans mon travail et je veille à ce qu'il soit accompli comme il se doit, et je passe même des heures supplémentaires au travail, rien que pour être toujours là pour mon entreprise* ». Ces affirmations s'accordent avec les propos de Van Dick (2001) selon lesquels les individus essayent d'aider l'entreprise dans laquelle ils travaillent, et ce, pour avoir une image de soi positive en tant que membre de l'organisation. Ceci rejoint également la littérature dans la mesure où une forte identification implique davantage de coopération avec les membres de l'organisation mais aussi des efforts supplémentaires pour réaliser les objectifs organisationnels (Dutton et al. (1994).

Conséquemment, ces résultats confirment notre proposition (3).

Conclusion

Cette recherche vise à comprendre la manière dontlesemployésdécriventetperçoivent leurs différentes expériences au sein d'une entreprise qui fait preuve d'un comportement socialement responsable envers toutes les parties prenantes, mais aussi et surtout à saisirl'impactdela perception de la responsabilitésocialedel'entreprisesurl'identification organisationnelle des employés. Pour ce faire, une revue de littérature nous a conduit à émettre trois propositions relatives à l'effet des pratiques de RSE sur l'identification organisationnelle, et ce, conformément à la conceptualisation de la RSE suivant les cibles ou bien les bénéficiaires de Turker(2009a). Nous avons également mobilisé la théorie d'identité sociale comme mécanisme d'explication de ladite relation.

Nous avons choisi de mener une étude qualitative exploratoire, en utilisant un guide d'entretien semi-directif auprès de dix employés travaillant au sein de l'entreprise Total Tunisie. Les résultats de cette recherche montrent que les employés de Total Tunisie perçoivent manifestement les actions de RSE orientée aussi bien vers les acteurs sociaux et non sociaux, mais aussi celles qui visent les employés eux-mêmes. En revanche, les initiatives de RSE qui ciblent les consommateurs, demeurent beaucoup moins perceptibles par les employés. En outre, les résultats viennent confirmer les propositions (1) et (3), selon lesquelles la perception de la RSE (les actions RSE orientées vers les acteurs sociaux et non sociaux, et celles qui visent les employés) a un effet sur l'identification organisationnelle ; avec une infirmation de la proposition (2), du fait que la plupart des employés interviewés n'avaient pas de visibilité sur la RSE orientée vers les consommateurs.

Notre recherche, se voit attribuer des implications théoriques mais aussi managériales.Sur le plan théorique, nous avons mis en avant l'effet de la perception des différentes composantes de la RSE (suivant la théorie des parties prenantes) par les employés sur l'identification organisationnelle. En outre, nous avons mobilisé un cadre théorique (la théorie d'identité sociale), un cadre qui demeure peu utilisé pour expliquer l'effet de la RSE sur les attitudes au travail.

En ce qui concerne les implications managériales, les managers doivent s'atteler à développer des perceptions positives des initiatives de RSEpar les employés, et ce, en développant des émotions et des attitudes positives au travail. De même, les résultats indiquent que différents types d'actions de RSE, influencent différemment l'identification organisationnelle, et dans ces cas les responsables doivent communiquer périodiquement sur les actions de RSE, notamment sur celles qui sont peu perceptibles. De même, les managers sont appelés à engager les membres de l'entreprises dans les programmes de RSE, et ce, pour saisir davantage ces pratiques mais aussi pour développer le sentiment de fierté et d'appartenance à leur organisation.

Notre recherche accuse cependant certaines limites. En effet, les conclusions auxquelles nous avons abouti demeurent préliminaires et spécifiques, ce qui est de nature à limiter la validité externe des résultats. Cette limite est essentiellement afférente au caractère exploratoire de notre étude. En outre, une autre limite de note étude est inhérente au fait que la RSE a été mesurée en fonction de la perception des employés et il est possible qu'il y ait une différence entre les perceptions des répondants et le niveau réel de la participation de l'entreprise dans de telles initiatives. Une dernière limite consiste en le choix d'une seule entreprise pétrolière auprès de laquelle est faite cette étude. Ainsi, d'autres entreprises auraient pu être sollicitées pour enrichir davantage les résultats.

S'agissant des perspectives futures de recherche, nous pouvons proposer l'étude de l'effet de la RSE sur d'autres attitudes au travail, telles que la satisfaction, l'engagement organisationnel, et la confiance organisationnelle, puisque ces attitudes pourraient expliquer les comportements des employés. De surcroît, la mobilisation d'un double cadre théorique basé sur la théorie d'identité sociale et sur la théorie d'échange social, pourrait éclairer davantage la nature des relations entre les dimensions de la RSE et l'identification organisationnelle. De même, et étant donné que notre étude a abouti à des résultats mitigés concernant la perception des initiatives de RSE orientée vers les consommateurs, il serait intéressant de se pencher sur l'effet que pourrait avoir la communication dans ladite relation (Kaddouri et Albertini (2016). Finalement, nous pouvons pousser la réflexion davantage en s'interrogeant sur les variables modératrices qui pourraient expliquer l'intensité de la relation entre la perception de la RSE et l'identification organisationnelle, telles que l'adéquation entre individu et entreprise.

Références bibliographiques

Ashforth, B. E. ;Mael,F. (1989). «Social identitytheoryand the organization»,*The AcademyofManagement Review*, Vol 14, N°*1*, p. 20-39.

Bartels, J. (2006). Organizational Identification and Communication: Employees' evaluations of internal communication and its effect on identification at different organizational levels. University of Twente.

Blanchet Alain ; Gotman Anne ; de Singly François (1992). L'Enquête et ses méthodes : l'entretien, Paris : Nathan Paris.

Brammer S; Millington A.; Rayton, B. (2007), «The contribution of corporate social responsibility to organizational commitment», *The International Journal of Human Resource Management*», Vol.18, N°10, p. 1701-1719.

Campbell, J. (2007), «Why would corporations behave in socially responsible ways? An institutional theory of corporate social responsibility», *The Academy of Management Review*, Vol. 32, No. 3, p. 946-67.

Cardador, M. T.; M. G. Pratt (2006). «Identification Management and Its Bases: Bridging Management and Marketing Perspectives Through a Focus on Affiliation Dimensions», *Journal of the Academy of Marketing Science*, Vol 34, N°2, p. 174-184.

Carmeli, A.; Gilat, G.; Waldman, D. A. (2007). «The role of perceived organizational performance in organizational identification, adjustment and job performance», *Journal of Management Studies*, Vol 44, N°6, p. 972-992.

Charreire S. ; Durieux F (1999). Explorer et tester, cité par Thiétart R.A., Méthodes de recherche en management, Dunod, pp 57-80.

Chen, J.; Roberts, R. (2010), «Toward a more coherent understanding of the organization societal relationship: a theoretical consideration for social and environmental accounting research», *Journal of Business Ethics*, Vol. 97, N°. 4, p. 651-65

Collier, J., & Esteban, R. (2007), «Corporate social responsibility and employee commitment», Business ethics: A European Review, Vol 16, N°1, p. 19-33.

Cropanzano, R.; Mitchell, M. S. (2005). «Social exchange theory: An interdisciplinary review», *Journal of management*, Vol 31, N°6, p. 874-900.

De Roeck, K., &Delobbe, N. (2012), « Do environmental CSR initiatives serve organizations' legitimacy in the oil industry? Exploring employees' reactions through organizational identification theory», *Journal of Business Ethics*, Vol 110, N°4, p.397-412.

DeRoecke, Kenneth, El Ikremi, Assâad, & Swaen, Valérie.(2016). «Consistency Matters! How and When Does Corporate Social Responsibility Affect Employees' Organizational Identification? ». *Journal of Management Studies*, 2016, Vol. 53, No 7, p. 1141-1168.

De Singly, F. (1992). L'enquête et ses méthodes : le questionnaire, Nathan.

Donaldson, T.; Dunfee, T. W. (2002). «Ties that bind in business ethics: Social contracts and why they matter», *Journal of Banking and Finance*, Vol 26, N° 9, p. 853–1865.

Dutton, J. E., Dukerich, J. M., &Harquail, C. V. (1994), « Organizational images and member identification», Administrative Science Quarterly, p. 239-263.

Farooq, Omer, Payaud, Marielle, Merunka, Dwight, (2014). « The impact of corporate social responsibility on organizational commitment: Exploring multiple mediation mechanisms». *Journal of Business Ethics*, Vol. 125, No 4, p. 563-580.

Fiol, C. M. (2001). «Revisiting an identity-based view of sustainable competitive advantage». *Journal of Management*, Vol 27, N°6, p. 691-699.

Freeman, R.E. (1984). Strategic management: a stakeholder approach, Pitman, Boston.

Goles T; Hirschheim R (2000). The paradigm is dead, the paradigm is dead … long live the paradigm: the legacy of Burell and Morgan, Omega, Vol 28, N°3, p. 249-268.

Hansen, S.; Dunford, B.; Boss, A.; Boss, R.; Angermeier, I. (2011). «Corporate social responsibility and the benefits of employee trust: A cross-disciplinary perspective», *Journal of Business Ethics*, Vol 102, N°1, p. 29–45.

Hofman,P.; Newman,A.(2014),«TheImpactof PerceivedCorporateSocialResponsibilityon Organizational Commitmentand theModeratingRoleof Collectivism and Masculinity: EvidencefromChina*»,InternationalJournalofHumanResourceManagement*, Vol, 25, N°5, p. 631–652.

Hollowitz, J.; Wilson, C.E (1983). «Structured Interviewing in Volunteer Selection», Journal of Applied Communication Research, Vol 21, N°1, p. 41-52.

Husted, B.; Allen, D. (2006), «Corporate social responsibility in the multinational enterprise: strategic and institutional approaches», *Journal of International Business Studies*, Vol. 37, N°6, p. 838-49.

Jones, D. A. (2010). « Does serving the community also serve the company? Using organizational identification and social exchange theories to understand employee responses to a volunteerism program», *Journal of Occupational and Organizational Psychology*, Vol 83, N°4, p. 857–878.

Jones, T. M. (1980), « Corporate social responsibility revisited, redefined», *California Management Review*, Vol 22, N°3, p. 59-67.

Leech, Nancy L.; Anthony J. Onwuegbuzie (2007). «An Array of Qualitative Data Analysis Tools: A Call for Data Analysis Triangulation», *School Psychology Quarterly*, Vol 22, N°4, p. 557-584.

Lorence, J. ; Mortimer, J. T. (1 985). «Job involvement through the life course: A panel stud y of three age groups», *American Sociological Review*, Vol 50, N°6, p.618-638.

Maignan, I. ; Ferrell, O. C. (2001). «Corporate citizenship as a marketing instrument-Concepts, evidence and research directions», *European Journal of Marketing*, Vol 35, N°3/4, p. 457-484.

Matten, D.; Crane, A. (2005). «Corporate citizenship: Toward an extended theoretical conceptualization», *Academy of Management Review*, Vol 30, N°1, p. 166–179.

Mucchielli A., (1994). Les méthodes qualitatives. Que sais-je ? », N° 2.591, PUF, Paris.

Newman, A., Miao, Q., Hofman, P. S., & Zhu, C. J. (2016). « The impact of socially responsible human resource management on employees' organizational citizenship behaviour: the mediating role of organizational identification». *The International Journal of Human Resource Management*, Vol 27, N°4, 440-455.

Newman, A.; Nielsen, I.; Miao, Q. (2015). «The impact of employee perceptions of organizational corporate social responsibility practices on job performance and organizational citizenship behavior: Evidence from the Chinese private sector», *The International Journal of Human Resource Management*, Vol 26, N°9, p. 1226-1242.

Peterson, D. K. (2004). «The relationship between perceptions of corporate citizenship and organizational commitment», *Business & Society*, Vol *43, N*°3, p. 296-319.

Pettijohn, C. E.; Pettijohn, L. S.; Taylor, A. J. (2007). «Salesperson Perception of Ethical Behaviors: Their Influence on Job Satisfaction and Turnover Intentions», *Journal of Business Ethics*, Vol.78, N°4, p. 547-557.

Riketta M. (2005). «Organizational identification: a meta-analysis», Journal of Vocational Behavior, Vol 66, N°2, p.358-384.

Rodrigo, P., & Arenas, D. (2008), « Do employees care about CSR programs? A typology of employees according to their attitudes», *Journal of Business Ethics*, Vol 83, N°2, p.265-283.

Russo, M.; Fouts, P. (1997), «A resource-based perspective on corporate environmental performance and profitability», *Academy of Management Journal*, Vol. 40, N°. 3, p. 534-59.

Shen, J., & Benson, J. (2016), « When CSR is a social norm how socially responsible human resource management affects employee work behavior», *Journal of Management*, Vol 42, N°6, p.1723-1746.

Shropshire, C.; Hillman, A. (2007), «A longitudinal study of significant change in stakeholder management», *Business and Society*, Vol. 46, N°1, p. 63-87.

Sims, R. R. (2003). Ethics and corporate social responsibility: Why giants fall. Greenwood Publishing Group.

Smith J.K. (1983), «Quantitative versus qualitative research: An attempt to clarify the issue», *An Educational Researcher*, Vol 12, N° 3, p. 6-13.

Stets, J. E.; Burke, P. J. (2000). «Identity theory and social identity theory», *Social Psychology Quarterly*, p. 224-237.

Stites, J. P. ; Michael, J. H. (2011). «Organizational commitment in manufacturing employees: Relationships with corporate social performance», *Business & Society*, Vol 50, N°1, p. 50-70.

Turker, D. (2009a). «Measuring corporate social responsibility: A scale development study», *Journal of business ethics*, Vol 85, N°4, p. 411-427.

Turker, D. (2009b), «How Corporate Social Responsibility Influences Organizational Commitment», *Journal of Business Ethics*, Vol 89, N°2, p. 189 – 204.

Tyler, T. R.; Blader, S. L. (2002). «Autonomous vs. comparative status: Must we be better than others to feel good about ourselves? », *Organizational Behavior and Human Decision Processes*, Vol 89, N°1, p. 813–838.

Tziner, A., Oren, L., Bar, Y.; Kadosh, G. (2011). «Corporate social responsibility, organizational justice and job satisfaction: how do they interrelate, if at all? », *Revista de Psicología del Trabajo y de las Organizaciones*, Vol *27*, N°1, p. 67.

Valentine, S.; Fleischman, G. (2008). «Ethics programs, perceived corporate social responsibility and job satisfaction», *Journal of Business Ethics*, Vol *77, N*°2, p. 159-172.

Van der Maren, J. M. (1996). *Méthodes de recherche pour l'éducation*. Presses de l'Université de Montréal et de Boeck.

Van Dick, R. (2001).« Identification in organizational contexts: Linking theory and research from social and organizational psychology». *International Journal of Management Reviews*, Vol 3, N°4, 265-283.

Van Dick, R.; Grojean, M. W., Christ, O., &Wieseke, J. (2006). «Identity and the extra mile: Relationships between organizational identification and organizational citizenship behaviour», *British Journal of Management*, Vol 17, N°4, p. 283-301.

www.femmesmaghrebines.com

www.leaders.com

www.total.tn

Optimization of workflow scheduling in an energy-aware cloud environment

Hatem Aziza
LARODEC, Institut Supérieur de Gestion
Université de Tunis
Tunis, Tunisia
hatem.aziza@isg.rnu.tn

Saoussen Krichen
LARODEC, Institut Supérieur de Gestion
Université de Tunis
Tunis, Tunisia
saoussen.krichen@isg.rnu.tn

***Abstract*—Today, we live in an increasingly dematerialized world. The cloud is the generator of the next wave of technologies, the detonator for all cutting-edge developments like artificial intelligence and blockchain. The cloud is mainly composed of datacenters which consume almost 20% of the total energy footprint of the digital world. The host machines that make up these datacenters are becoming more and more powerful and perform calculations more quickly. Indeed, this increase in computing power is accompanied by increased consumption of electrical energy even in the event of inactivity. In this context, we are talking about a very energy-consuming development in terms of electricity consumption and even in terms of CO_2 emissions, the internet of which pollutes 1.5 times more than air transport. This exponential consumption of digital makes it very hard to feed with renewable energies. This problem leads us to focus on solutions that optimize the use of resources in the cloud environment in order to schedule workflows especially since there are still today few techniques to limit energy consumption. This article is part of the resolution of this problem of minimizing energy consumption by datacenters in the cloud environment. For that, we try to study a solution based on the technique Dynamic Voltage and Frequency Scaling (DVFS) which allows to moderate the frequency of the CPU thus influencing the energy consumption. The solution based on DVFS is compared with other techniques in order to show its power to achieve the key objective of this article which is the optimization of energy consumption.**

***Index Terms*—cloud, energy consumption, dvfs, scientific workflows, scheduling**

I. Introduction

Cloud Computing is the combination of various Information Technology(IT) concepts studied and developed since the 60s, associated with certain technologies, as Parallel Computing, Utility Computing, Grid Computing. Cloud is a metaphor for the Internet. Using Cloud Computing, you can access, at any time through any device via the Internet, to data and files that you have downloaded, or software applications you need to use for personal or professional use. Cloud Computing is the development and use of applications accessible only through Internet. Users rely on the Internet to use their software; they are able to access services without installing anything other than a simple Internet browser.

According to the most commonly accepted definition of the US National Institute of Standards and Technology (NIST) [1], Cloud Computing is a practical model, on demand and universal, to establish an access to a shared pool of configurable computing resources (eg, networks, servers, storage, applications and services) that can be rapidly mobilized and made available with minimal management efforts and contacts with the service provider.

Due to the accumulation of demand, the bulk purchase of computer hardware and energy and the reduction of unit costs of labor, the Cloud Computing service providers can achieve significant economies of operating costs and benefit their clients.

Cloud Computing can be used in the same way we use electricity, that is paying only what we consume without even worrying about technical aspects of the functioning of the system.

The main categories of Clouds are public Clouds, private Clouds and hybrid Clouds, and the main types of services provided by these Clouds are software models as a service (SaaS), platform as a service (PaaS) and infrastructure as a service (IaaS). These Cloud Computing services can be provided from datacenters located around the world.

Cloud computing relies on virtualization that offers a huge way to save energy by effectively consolidating virtual machines to use resources optimally. The virtualization enables easy management, optimized and pooled hardware resources and a high elasticity of infrastructure with a view logical rather than physical. Indeed, it is not rare to see a company server running at 15% of its capacity, and consolidate multiple virtual servers on one physical machine to better rationalize resources. More elasticity provides the ability to automatically provisioned resource and to adaptrationally at major mounted load.

Scheduling tasks is a necessary step for routing tasks to the datacenter servers. However, this process is based on algorithms that have an impact on the energy consumption at the cloud level. Indeed, the way in which these algorithms affect the tasks to the virtual machine (VM) influences the number of servers used.

The problem of energy consumption has become one of the major concerns in cloud. Moreover, it is very important to point out that the energy consumption and SLA violation are the two crucial problems in cloud. The cloud provider agrees to respect a set of clauses cited in the SLA contrat. Moreover, this contract specifies the negotiated requirements, in terms

of resources, QoS, minimum expectations and obligations that exist between users and cloud providers.
For that, the cloud provider is always on the horns of a dilemma. On the one hand, the provider has to respect the SLA clauses and especially the QoS which indicates the level of performance, reliability and availability offered by an application or infrastructure supported by the cloud.
To efficiently manage the energy consumed, several techniques exist, including DVFS which has been widely studied to reduce energy consumption in cloud data centers. By dynamically reducing the frequency and the voltage of the processor, DVFS makes it possible to accomplish a task before its deadline with minimal energy consumption.
To address such a problem, this paper focuses on study of an energy-aware workflow scheduling that aims to optimize the energy consumption while respecting the QoS. The solution based on DVFS is compared with other VM allocation policies as Inter Quartile Range, Local Regression, Median Absolute Deviation, and Static Threshold in order to show its power to achieve the key objective of this article which is the optimization of energy consumption. The remainder of this paper is organized into the followings sections. In section 2, we discuss several related work. Section 3 presents a task scheduling environment. Section 4 presents a problem modeling. The energy-aware roblem is formulated in section 5. Experiments and results are reported in section 6.

II. Related work

The problem of energy consumption has become one of the major concerns in cloud. Moreover, it is very important to point out that the energy consumption and SLA violation are the two crucial problems in cloud. The cloud provider agrees to respect a set of clauses cited in the SLA contrat. Moreover, this contract specifies the negotiated requirements, in terms of resources, QoS, minimum expectations and obligations that exist between users and cloud providers.
Calheiros et al. [2] proposed an approach for reducing the power required to execute a set of independent tasks on cloud environment. This approach exploited a novel technical scheduling combined with DVFS capability of modern CPU processors to keep it at the minimum votage level in order to complete tasks execution promptly while meeting deadline. Experiments demonstrated that the proposed solution substantially reduced power consumption compared to existing work.
In [3], C.-M. Wu. et al proposed a scheduling solution for the cloud datacenter with a DVFS technique aiming to increase resource utilization and decrease energy consumption in return. First, the authors provided the feasible combination or scheduling for a job. Secondly, an appropriate voltage and frequency supply for the servers based on DVFS is provided. Using cloudsim toolkit to simulate the cloud environment, they compared their solution with existing ones to show that it can reduce more energy consumption that the others.
Kar et al. [4] provided a green computing approach based on GA in order to reduce time, cost and the environmental pollution. Experiments demonstrated that the proposed solution gives better results than Max-min algorithm.
Tang et al. [5] proposed a energy-aware workflow task scheduling solution with a DVFS technique while meeting the dependency constraints. The proposed algorithm distributed the workflow tasks to appropriate processors aiming to minimize power consumption while meeting deadline and performance. The authors designed five experiments to provide the performance comparisons in different number of processors, various extension ratio, different communication to computation ratio, and different degree of parallelism. These experiments show that the proposed solution outperforms other existing algorithms like HEFT.
In [6], the authors presented a novel approach that can generate energy-aware and soft error resilient schedules for workflow applications. This approach consists of three phases. First, HEFT is applied as an initial scheduling algorithm to generate a feasible schedule under the given completion time and reliability requirements. Next, tasks moved to corresponding nodes with higher energy efficiency. Finally, a reliability-aware DVFS algorithm is designed to minimise the global power consumption. Experiments show that this scheduling solution can optimize energy consumption.

III. Task scheduling environment

As depicted in Figure 1, the provider owns a Cloud with an arbitrary number of datacenters, which are modelled according to their operating systems, processors architecture, hypervisors [7], available network bandwidth, utilization costs, and virtual machines allocation policies. Moreover, The Cloud datacenter comprising n heterogeneous physical nodes. Each node has a CPU, which can be multicore, with performance defined in Millions Instructions Per Second (MIPS). Besides that, a node is characterised by the amount of RAM and network bandwidth.

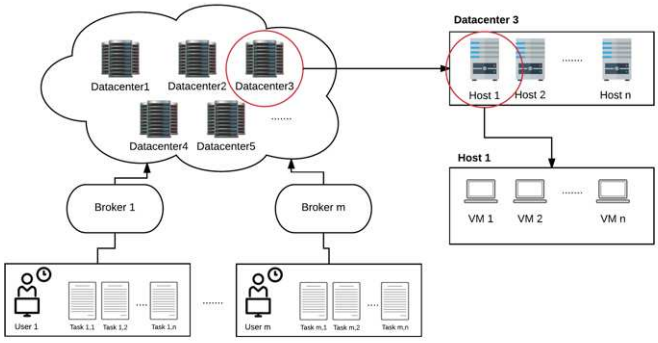


Fig. 1. Task Scheduling environment

Users submit requests for provisioning of m heterogeneous VMs with resource requirements defined in MIPS, amount of RAM and network bandwidth. SLA violation occurs when a VM cannot get the requested amount of resource, which may happen due to VM consolidation.

The resource utilization profile describes the clients' applications and a high-level policy that selects datacenters to deploy virtual machines. This policy is represented by a simulation entity called broker that also defines how Cloudlets will be managed, and in which virtual machine they will be executed. Tasks are modelled using characteristics such as processor cores, size in MIPS, length of input and output files that are transferred between users and providers, and utilization models for CPU, bandwidth, and memory. A virtual machine configuration includes its image size, number of processors, processing capacity in MIPS, amount of RAM and bandwidth, types of hypervisor, and task scheduling policy.

The objectives of a scheduler are mainly:

- Ensure that each task waiting run gets his share of time in the virtual machine.
- Minimize the response time.
- Optimize the use of resources.
- Balanced use of resources.
- Take into account the priorities.

These objectives are sometimes complementary, sometimes contradictory. Increasing the performance compared to any of them may be in detriment of another. It is difficult to create an algorithm that optimizes all criteria simultaneously.

IV. Problem modeling

A. Workflow modeling

Workflow is usually represented using a directed acyclic graph (DAG) which is represented by a pair of vertex (V) and edge (E). Each vertex represents the task while the edge is the dependence link between two tasks. Figure 2 represents the DAG of Montage scientific workflow while figure 3 represents the DAG of cybershake scientific workflow.

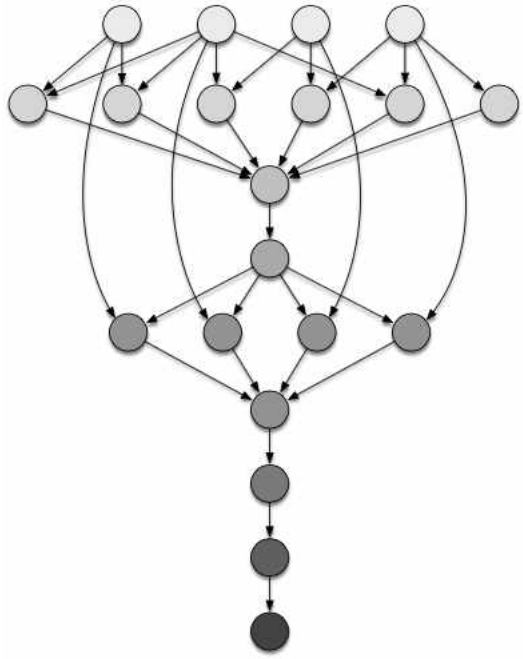


Fig. 2. Montage scientific workflow structure

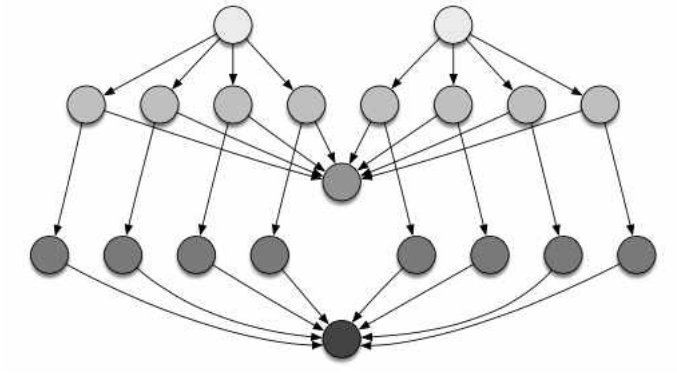


Fig. 3. Cybershake scientific workflow structure

B. Energy consumption modeling

The cloud is mainly composed of datacenters which consume almost 20% of the total energy footprint of the digital world. The host machines that make up these datacenters are becoming more and more powerful and perform calculations more quickly. Indeed, this increase in computing power is accompanied by increased consumption of electrical energy even in the event of inactivity.

In this work, we are talking about the energy consumption by VMs. When the VM is powered on, energy consumption consists of two parts: dynamic energy and static energy. In this work, we will focus on dynamic energy.

The dynamic energy E_{total} is composed of an active energy E_{active} and an inert energy E_{inert} is formulated in equation 1.

$$E_{total} = E_{active} + E_{inert} \tag{1}$$

Each task i making up the workflow consumes this energy during its processing in the VM j for a defined time t_{ij}. In this case, it is active energy since the tasks are running. E_{active} is modeled as follows:

$$E_{active} = \sum_{i=1}^{n} k \times v_{ij}^2 \times f_{ij} \times t_{ij},\ j = 1, \ldots, m \tag{2}$$

where k is a CPU related constant, v_{ij} is the voltage consumed by VM j during the execution of task i and f_{ij} the running clock frequency of task i on VM j.

When the VM is in an inactive state, it still consumes less energy E_{inert} than that consumed in the active state. But this amount of energy consumed is always important, and it has a cost that can become huge after a significant period of inactivity. E_{inert} is formulated in equation 3.

$$E_{inert} = \sum_{j=1}^{m} k \times v_{j_{lowest}}^2 \times f_{j_{lowest}} \times t_{j_{inert}} \tag{3}$$

where $v_{j_{lowest}}$ represents the lowest voltage consumed by VM in idle state while $f_{j_{lowest}}$ is the lowest clock frequency of VM.

V. Problem formulation

We address a Task-scheduling energy aware problem in CC that minimizes the total energy consumption by the Datacenter VMs. The problem consists then to schedule n client tasks on m VM instance types. The objective function of this proposed model can be defined as:

$$Min\ E_{total} \tag{4}$$

VI. Experimentation

we have adopted the CloudSim simulator to simulate a power-aware cloud data center which consists of 200 VMs running on 100 hosts.
We have taken measures on personal computer with the following configuration: Intel Core i5 2.4G.HZ 8GB memory, run on macOS Sierra 10.12.6.
The different experiments are the result of the application of several VM allocation and VM selection policies. These policies are listed in the table I

VM ALLOCATION POLICY	VM SELECTION POLICY	VM MIGRATIONS
DVFS	-	No
Inter-Quartile Range	Maximum Correlation	Yes
Local Regression	Maximum Correlation	Yes
Median Absolute Deviation	Maximum Correlation	Yes
Static Threshold	Maximum Correlation	Yes

TABLE I
VM ALLOCATION AND SELECTION POLICIES

the table gives us shows that in our experiments, we compared the results provided by the VM allocation policies as DVFS, Inter-Quartile Range, Local Regression, Median Absolute Deviation and Static Threshold. We even tried to test a solution without taking into account the reduction in energy consumption.
In these experiments, we focused on several criteria as:

- Energy consumption in kWh
- Number of VM migrations
- Average SLA violation

Besides these criteria, I saw the importance of considering the build total time of the solution by NetBeans IDE. This last criterion is to measure the speed of the implementation of each policy.

The execution of this testbed of tasks with different parameters lead to the results listed in Table II and shown in Figures 4-7.

CRITERIA	DVFS [1]	IQR.MC [2]	LR.MC [3]	MAD.MC [4]	NP [5]	THR.MC [6]
Energy consumption (kWh)	154,74	35,31	29,17	34,72	301,35	36,14
Number of VM migrations	0	5391	3859	5218	0	5517
Average SLA violation	0%	9,93%	9,15%	9,39%	0%	0,12%
Total Build Time (seconds)	132	1089	938	961	153	1301

[1] Dynamic Voltage and Frequency Scaling
[2] Inter Quartile Range VM allocation policy and Maximum Correlation VM selection policy
[3] Local Regression VM allocation policy and Maximum Correlation VM selection policy
[4] Median Absolute Deviation VM allocation policy and Maximum Correlation VM selection policy
[5] Non-power aware Datacenter
[6] Static Threshold VM allocation policy and Maximum Correlation VM selection policy

TABLE II
COMPARISON OF VM ALLOCATION AND VM SELECTION POLICIES ON DIFFERENT CRITERIA

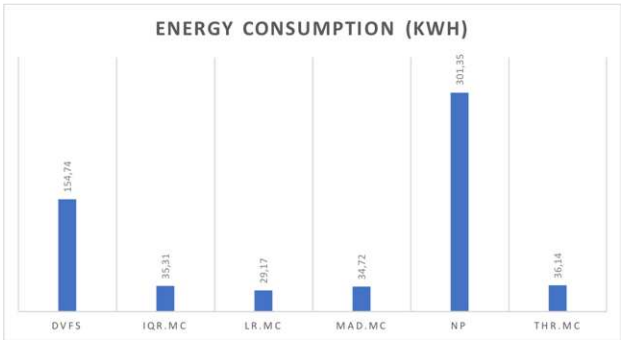


Fig. 4. Energy consumption (kWh)

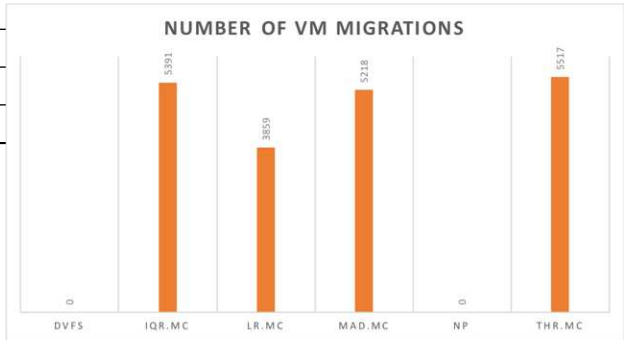


Fig. 5. Number of VM migrations

Fig. 6. Average SLA violation

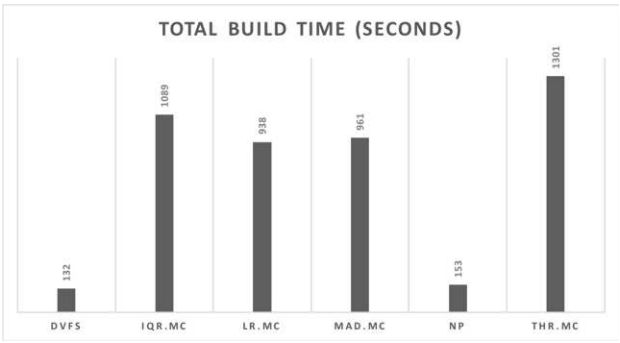


Fig. 7. Total Build Time (seconds)

Comparing all VM allocation policies as DVFS, Inter-Quartile Range, Local Regression, Median Absolute Deviation, Static Threshold and Non-power aware datacenter, we deduce that DVFS is a good policy for optimizing the energy consumed by the data center, especially if we take into consideration other very important criteria such as VM migration and violation of the SLA. Other policies as Inter-Quartile Range, Local Regression, Median Absolute Deviation, Static Threshold consume less energy but in return take a lot of build time and a violation of SLA to varying degrees in addition to the significant amount of VM migration.

VII. Conclusion

In this paper, we studied a solution based on the technique Dynamic Voltage and Frequency Scaling (DVFS) which allows to moderate the frequency of the CPU thus influencing the energy consumption. The solution based on DVFS is compared with other VM policies as Inter-Quartile Range, Local Regression, Median Absolute Deviation and Static Threshold in order to show its power to achieve the key objective of this article which is the optimization of energy consumption.
To simulate these VM allocation policies in combination with VM selection policies, we have adopted the CloudSim simulator. The results show that the solution based on DVFS found a good trade-off between important criteria as energy consumption, number of VM migrations and SLA violation.

References

[1] T. G. Peter Mell, "The nist definition of cloud computing," *National Institute of Standards and Technology*, 2011.

[2] R. N. Calheiros and R. Buyya, "Meeting deadlines of scientific workflows in public clouds with tasks replication," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 7, pp. 1787–1796, July 2014.

[3] X. Wu, M. Deng, R. Zhang, B. Zeng, and S. Zhou, "A task scheduling algorithm based on qos-driven in cloud computing," *Procedia Computer Science*, vol. 17, pp. 1162–1169, 2013.

[4] I. Kar, R. R. Parida, and H. Das, "Energy aware scheduling using genetic algorithm in cloud data centers," *International Conference on Electrical, Electronics, and Optimization Techniques*, 2016.

[5] E. N. Alkhanak, S. P. Lee, R. Rezaei, and R. M. Parizi, "Cost optimization approaches for scientific workflow scheduling in cloud and grid computing: A review, classifications, and open issues," *Journal of Systems and Software*, vol. 113, pp. 1 – 26, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0164121215002484

[6] T. Wu, H. Gu, J. Zhou, T. Wei, X. Liu, and M. Chen, "Soft error-aware energy-efficient task scheduling for workflow applications in dvfs-enabled cloud," *Journal of Systems Architecture*, 2018.

[7] H. A.Sbaa, R. El Bejjet, "Architecture design of a virtualized embedded system," *International Journal on Computer Science and Engineering*, 2013.

[8] X. Xu, J. Wu, G. Yang, and R. Wang, "Low-power task scheduling algorithm for large-scale cloud data centers," *Journal of Systems Engineering and Electronics*, vol. 24, no. 5, pp. 870–878, 2013.

[9] T. Gurout, S. Medjiah, G. Da Costa, and T. Monteil, "Quality of service modeling for green scheduling in clouds," *Sustainable Computing: Informatics and Systems*, vol. 4, no. 4, pp. 225–240, 2014.

[10] C.-M. Wu, R.-S. Chang, and H.-Y. Chan, "A green energy-efficient scheduling algorithm using the dvfs technique for cloud datacenters," *Future Generation Computer Systems*, vol. 37, pp. 141–147, 2014.

[11] pp. 494–495, 2015.

[12] Z. Tang, L. Qi, Z. Cheng, K. Li, S. U. Khan, and K. Li, "An energy-efficient task scheduling algorithm in dvfs-enabled cloud environment," *Journal of Grid Computing*, vol. 14, no. 1, pp. 55–74, 2015.

[13] W. Wu, W. Lin, and Z. Peng, "An intelligent power consumption model for virtual machines under cpu-intensive workload in cloud environment," *Soft Computing*, vol. 21, no. 19, pp. 5755–5764, 2016.

[14] M. Zotkiewicz, M. Guzek, D. Kliazovich, and P. Bouvry, "Minimum dependencies energy-efficient scheduling in data centers," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 12, pp. 3561–3574, 2016.

[15] P. Arroba, J. M. Moya, J. L. Ayala, and R. Buyya, "Dynamic voltage and frequency scaling-aware dynamic consolidation of virtual machines for energy efficient cloud data centers," *Concurrency and Computation: Practice and Experience*, vol. 29, no. 10, 2017.

[16] F. Farhadian, M. M. R. Kashani, J. Rezazadeh, R. Farahbakhsh, and K. Sandrasegaran, "Withdrawn: An efficient iot cloud energy consumption based on genetic algorithm," *Digital Communications and Networks*, 2019.

[17] N. Kumar and D. P. Vidyarthi, "A green sla constrained scheduling algorithm for parallel/scientific applications in heterogeneous cluster systems," *Sustainable Computing: Informatics and Systems*, vol. 22, pp. 107–119, 2019.

[18] A. Beloglazov and R. Buyya, "Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers," *Concurrency and Computation: Practice and Experience (CCPE)*, vol. 24, pp. 1397–1420, 2012.

Combination of personalization parameters based on fuzzy logic in the cloud

Sameh Ghallabi[#1], Fathi Essalmi[#2], Mohamed Jemni[#3], Kinshuk[*4]

[#] *The Research Laboratory of Technologies of Information and Communication & Electrical engineering (LaTICE)*
Higher School of Sciences and Technologies of Tunis (ESSTT), University of Tunis, Tunisia

[1]gallabi_sameh@yahoo.fr

[2]fathi.essalmi@isg.rnu.tn

[3]mohamed.jemni@fst.rnu.tn

[*] *School of Computing and Information Systems*
Athabasca University, 1 University Drive, Athabasca, AB T9S 3A3, Canada

[4]kinshuk@ieee.org

***Abstract*— Integrating e-learning personalization systems in a cloud-computing environment is a relatively recent trend in the technology of enhanced learning. Cloud computing is a new model that allows users to access applications through the Internet. Personalized learning systems use and combine a set of personalization parameters for personalizing courses. Choosing suitable learning objects in accordance with a learner's needs and characteristics using the mentioned system requires precise learner profiles.**

This paper presents an approach for selecting the appropriate and not appropriate learning object to a given learners' characteristic using the cloud-computing technology. The proposed approach is based on fuzzy logic for precisely determine the most appropriate federation of personalization parameters.

Keywords*— *Personalization parameters, Cloud computing, Fuzzy logic.

I. INTRODUCTION

Personalized learning systems combine a set of personalization parameters for personalizing courses [1]. A personalization parameter represents a set of learners' characteristics. For example, the learner's level of knowledge may include the learners' characteristics beginner, intermediate and advanced.

Cloud computing [2] is a model that gives the possibility to access data and services on a remote server. It offers standardized and shared resources. These resources are delivered over the internet. Cloud computing offers several services such as software, platform and infrastructure as a service.

In literature, several approaches have selected and combined the mentioned parameters according to the teachers' preferences in the cloud. For instance, Zaharescu [3], who put forward a new approach that permitted the integration of cloud computing in the education field. Their approach enabled generating a suitable content according to the media preference. Besides, Jeong and Hong [4] proposed a service based personalized learning system in cloud computing environment. In addition, El-Sofany et al. [5] propounded a new approach that integrated cloud computing in e-learning systems. Their approach used the cloud to secure data storage and computing power. Furthermore, Kaewkiriya and Utakrit [6] defined a new model of learning management systems utilizing cloud computing with Web services. The approach allowed users to access various e-learning resources with a low cost.

However, these works have not presented any approach for selecting and combining the precise and accurate composition of personalization parameters. In this paper, an approach to select and combine the precise parameters in cloud computing environment is proposed. The proposed approach allows representing these parameters in a flexible, accessible and available way. It uses fuzzy logic [7] that has given a good response in cloud computing.

The rest of this paper is organized as follows. Section II describes the proposed approach for determining the appropriate combination of parameters. Conclusions and future work are presented in section III.

II. PROPOSED APPROACH

This paper presents an approach, which allows combining set values of the personalization parameters for the personalization of learning objects in the cloud environment with low operation costs. The goal of this approach is to determine the appropriate and not appropriate learning object to a given learner's characteristic. It is based on fuzzy logic. This logic is simple and flexible which allow to reason and combine these parameters based on their specification as linguistic variables. . The proposed approach is illustrated in the following figure :

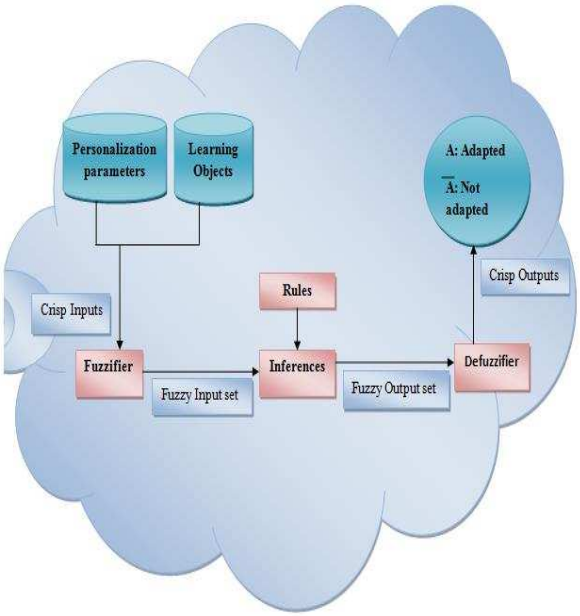


Figure1. Approach architecture

The process of fuzzy logic consists of three stages, which is presented as follow:

The first step is fuzzification, which allow to convert a classical data or crisp set of input data into a fuzzy set using fuzzy linguistic terms, fuzzy linguistic variables and membership functions. This stage represents the used linguistic variables for determining learning objects appropriate to learners' characteristics. Linguistic variables are the input data of the system whose values are words or sentences from a natural language, instead of numerical values. They are decomposed into a set of linguistic terms. In this paper, the used linguistic variables are:

- PP (Personalization Parameters): Each of the personalization parameter constitutes a linguistic variable. For example, level of knowledge is the linguistic variable. The set of terms (values) used to explain the cited above parameter is {beginner, intermediate, advanced} and the set of values used in the literature for the personalization parameter motivation level is {low, moderate, high}.
- LO (Learning Object): The set value of learning object is (simple(easy), hard, very hard).

In addition, in this step, linguistic terms have to be defined by membership functions in order to convert input variables into degrees of membership.

The second step is the inference stage that allows presenting the linguistic rules. These rules are applied to input values in order to generate output values. In particular, a set of IF-THEN rule with condition and conclusion is performed, which represent our reasoning in the determination of learning object adapted with some values of the selected personalization parameter

Finally, in the defuzzification step, the output of the system is the adaptation decision (A). Formally, A is a set of linguistic terms used to determine whether the learning object is adapted or not with some values of the selected personalization parameter

III. CONCLUSION

This paper proposed a new solution that enhances the efficiency within the personalized learning systems. The proposed approach is based on fuzzy logic. It provides a precise and accurate composition of personalization parameters. The main objective of this approach is to offer the most appropriate learning objects to the learners by considering their characteristics with low cost. Future works will focus on implementing and validating the proposed approach.

REFERENCES

[1] Ghallabi, S., Essalmi, F., Jemni, M., and Kinshuk, "Selection and composition of personalization parameters in cloud," International Conference on Advanced Learning Technology (ICALT), pp. 62-64, (2016)

[2] Ghallabi, S., Essalmi, F., Jemni, M., and Kinshuk, "Enhanced Federation and Reuse of E-learning Components Using Cloud Computing," International Conference on Smart Learning Environments (ICSLE), pp. 159-166, (2014).

[3] Zaharescu, E. "'Enhanced Virtual ELearning Environments Using Cloud Computing Architectures," Journal of Computer Science Research and Application, Vol. 2, pp.31-41, (2012).

[4] Jeong, H.Y., and Hong, B. "Service based Personalized Learning System in Cloud Computing Environment," AST/EEC/MMHS/AIA conference program, pp. 95-102, (2012).

[5] El-Sofany, H.F., Tayeb, A. Al., Alghatani, K., and El-Seoud, S.A.,"The Impact of Cloud Computing Technologies in E-learning," International Journal of Emerging Technologies in Learning (iJET), 8, pp. 37 – 43, (2012).

[6] Kaewkiriya, T., and Utakrit, N. "A Model of an e-Learning Management System Based on Cloud Computing and Web Service," Information Technology Journal, 8, pp. 83 – 87, (2012).

[7] Chrysafiadi, K. and Virvou, M. "Evaluating the integration of fuzzy logic into the student model of a web-based learning environment," Expert Systems with Applications, Vol. 39, pp.13127–13134, (2012).

Genetic Algorithms For The Nodes Migration Scheduling Problem.

Fatma Moalla
ISG tunis
University of Tunis
Tunisia
Email: fatma.moalla.fen@gmail.com

Mehdi Mrad
Department of industrial Engineering
College of Engineering
King Saud University
Riyadh,Saudi Arabia
Email: mmrad@ksu.edu.sa

Ali Balma
National Higher Engineering School of Tunis
University of Tunis
Tunis, Tunisia
Email:alibalma05@yahoo.fr

***Abstract*—The Network Migration Scheduling Problem (NMSP) gains nowadays a lot of interest due to its practical relevance in the telecommunication field. The problem is to find the best migration sequence that minimizes the total traffic cost between two networks. In this paper, we propose several genetic algorithms for the network migration scheduling problem. We test firstly the impact of changing the crossover and the mutation operators. Secondly, a new encoding is proposed to study in depth the diversification and the intensification. Finally, we discuss the effect of using a good quality of an initial population. Experiments show that the genetic algorithm proposals were able to provide competitive results.**

***Index Terms*—scheduling, migration, genetic algorithm, encoding, network**

1. Introduction

The evolving expansion of mobile technologies and services and the explosion of the managed data amount force operators to look for profitable solutions to evolve their networks, to optimize investments and consequently migrate to a network supporting the new services. The major challenge for migration is to ensure minimal customer service disruption and to minimize migration costs. This is the backdrop to suggest the use of the network migration-scheduling problem. The problem supposes nodes migration from an outdated network (S1) to a new one (S2). For this purpose, a bridge should separate these networks to ensure the services continuity. The bridge should use minimal capacities. The purpose is to find the migration order that minimizes the cost of boards installed on the bridge. Many practical applications arise such as migrating eNodeBs of a 4G network from an access network to another [1] or migrating virtual machines in geographically distributed cloud data centers [2].Unfortunately, the problem was not fully addressed. Only few works address this variant of problem [1], [2], [3] .

The use of genetic algorithms in scheduling problem is a fertile domain that gains more and more interest due to its applicability in many fields such as manufacturing, transportation, and telecommunications [4]. Among the key factors for the success of the genetic algorithm is its capacity to model a wide variety of problems using tailored operators, encoding and initialization procedures. In this paper, we discuss several genetic algorithms for the network migration scheduling problem. Firstly, we test different versions for the crossover and the mutation operators. Secondly, a new encoding problem based on the bi-partitioning conceptis proposed allowing thereby a better genes interpretation. Finally, we discuss the effect of considering the problem knowledge to get a good quality of an initial population.

2. Genetic Algorithm Proposals For The Migration Scheduling Problem

We consider three different proposals. Each algorithm take into account the problem specifications in order to guarantee effective results.

2.1. A Genetic Algorithm With Combination Of Genetic Operator (SGA)

The SGA algorithm proposes a migration order representation using a permutation encoding. In this vein, specific genetic operators are advised. The use of such an operator is motivated by the fact that the individuals have similar or close fitness values. In our study, we suggest different combinations of crossover and mutation operators in order select the most suitable version. Combining the use of two crossover and mutation operators will accordingly generate four versions to test:

- VERSION I: It is a basic version of genetic algorithm with CX operator and an exchange mutation operator
- VERSION II: It is a basic version of genetic algorithm with CX operator and an inversion mutation operator.
- VERSION III: It is a basic version of genetic algorithm with PMX operator and an exchange mutation operator.
- VERSION IV: It is a basic version of genetic algorithm with PMX operator and an inversion mutation operator.

2.2. A Two-Parts Representation Genetic Algorithm (TPRGA)

The main interest of using such a representation is that the intensification or the diversification strategy could be guided. The proposed encoding proposes two parts to draw the chromosome referring the partitiong concept:

- The first part of the chromosome is represented using the array P which describes the belonging of the nodes to the initial partition.
- The second part of the chromosome is represented using the array O which describes the nodes order in the migration sequence.

The first part uses a binary representation. The second part is a permutation sequence. We apply for each part the specific operators. We adopt conventionally the PMX and the inversion mutation in the evolutionary search. A sensitivity analysis is recommended to determine the appropriate genetic parameters values that generate solutions of good quality. A behavior investigation of these parameters is of great importance especially for large size instances since exact methods failed to solve these instances.

2.3. Genetic Algorithm With Knowledge Based Procedure (KBGA)

Considering knowledge in the representation phase may affect the GA behavior. For the KBGA algorithm, we use knowledge to generate our initial population. The bi-partitioning problem provides a lower bound of good quality by splitting the nodes sets into two disjoint equal size sets. Using such an information in the individuals generation step is of great importance. This allows us to generate an initial population that contains individuals of good quality. A fast running heuristic for the bi-partitionning problem from [5] is used to generate the first individual in the population. The remaining individuals in the population are generated using a knowledge-based algorithm.

3. Experimental Results

For each evolutionary approach version, we run the correspondent algorithm five times and we report the best of all executions.
For the SGA algorithm, we conclude that the PMX operator is more lucky to find a minimal value according to the fitness function since it requires less time . The use of Exchange mutation provides better results in terms of solution quality. So, the best version kept is version IV.
For the TPRGA algorithm, sensitivity analysis is performed on different problem sized instances. Experiments on each parameter delivers guidelines to choose the suitable configuration values. Based on all the experiments, we conclude that high values for the first part crossover are preferable. The performance of the second part operators relies on the effectiveness of the first part. The mutation operator in general has insignificant impact. To bring diversity in the initial population, generating up to the quarter of the population by modifying only the first part values is among the most promising scenarios. The most suitable population size to consider is 75.
Concerning the KBGA algorithm, we conclude that the enhancements made by the KBGA contributes to reach the optimality. Not surprisingly, The success for the KBGA algorithm could be explained by the heuristic used to generate the initial population. This heuristic provides near-optimal results. Generating others neighbors for a solution of good quality may allow finding the optimal solution.
In conclusion, The TPRGA has successfully found satisfactory gap rate. The depth investigation in parameter setting offers acceptable results. Regarding the computational effort, the KBGA and TPRGA have roughly similar CPU consumption.

4. Conclusion

We propose new genetic algorithms for the network migration scheduling problem. First, we propose a simple genetic algorithm with different combinations of crossover and mutation operators. To study in depth the GA behavior, a new representation was considered offering the possibility to understand the conduct of the algorithm and to guide the exploration / exploitation process. Third, we suggest a knowledge-based procedure to generate the initial population for a genetic algorithm. Different proportions were considered to assess the conduct of this evolutionary approach over the generations. Experiments show that the fourth version using the PMX crossover operator and the inversion mutation operator for a simple genetic algorithm is the most performent one. One notes that the new representation was able to produce good solutions. The KBGA successfully reach the optimality for small size instances. However, in term of computational time efforts, the KBGA and TPGA have roughly the same performance.

References

[1] Mrad, M., Balma, A., Moalla, F., Ladhari, T. Nodes Migration Scheduling of Access Networks IEEE Trans. Network and Service Management14(1): 77-90 ,2017

[2] Moalla, F., Balma, A.,Mrad, M., A Rapid Heuristic For The Virtual Machines Migration Scheduling Problem ,in Proceeding of Engineering and Technology PET of The fifth international conference on control and signal processing (CSP'17). Vol.25 pp.44-47, 2017

[3] Moalla, F., Balma, A., Mrad, M., An enhanced evolutionary approach for solving the nodes migration scheduling problem. In Ubiquitous networking: fourth international symposium, unet 2018, hammamet, tunisia,may 2-5, 2018,pp. 112,2018.

[4] Montana, D., Brinn, M., Moore, S., Bidwell, G., Genetic algorithms for complex, real-time scheduling. In Smc98 conference proceedings. 1998 ieee international conference on systems, man, and cybernetics(Vol. 3, p. 2213-2218 vol.3),1998

[5] Wu, J., Jiang, G., Zheng, L., Zhou, S. Algorithms for balanced graph bi-partitioning. In High Performance Computing and Communications, 2014 IEEE 6th Intl Symp on Cyberspace Safety and Security, 2014 IEEE 11th Intl Conf on Embedded Software and Syst (HPCC, CSS, ICESS), 2014 IEEE Intl Conf on (pp. 185-188). IEEE (2014).

Profiling learners' groups in MOOCs

Baraâ JEBALI#1 Ramzi Farhat*2 Mohamed Jemni*3

[1# *2*3] *Research Laboratory of Technologies of Information and Communication & Electrical Engineering (LaTICE)*
University Of Tunis, 5 Avenue Taha Hussein, B.P. 56, Bab Menara, Tunis, Tunisia,
[#I] *IsitCom, University of Sousse*

jebali.baraa@gmail.com
ramzi.farhat@esstt.rnu.tn
mohamed.jemni@alecso.org.tn

Abstract—MOOCs are characterized by their massive and open aspect. Researchers are worried about the learning efficiency in such learning system. They highlight that a course proposed to this large heterogeneous population does not meet obligatory the learners' needs and expectation especially with the lack of support and supervision. Learning personalization is proposed as a solution to adapt learning process to the learners needs. For that, many research works tried to identify the learners' profiles in MOOCS. Experimental results prove that learners' share common features in a significant degree that may unit them in the same learning process. In the following paper, we tried to identify the most determinative features that can profile the learners' groups in MOOCs.

Index Terms—MOOC, learners' profiles, Groups' profiles, personalization.

1 Introduction

Massive Open Online Courses give the opportunity to take courses of prestigious universities without previous requirements. Researchers confirm that MOOCs came with many opportunities, but revealed many new challenges. The open and massive aspects of the MOOCs and low teaching involvement put the learner in an unsupervised learning experience that lacks for support. Even if the founders of MOOCs rely on the collaboration of learners, the open aspect puts the learner in communication with a heterogeneous mass that does not share the same objectives, background, etc. Many researchers are worried about the learning efficiency in MOOCs.

Learning personalization is seen as a potential solution. The personalization takes into account the specific aspects of the learners to adapt learning scenarios. The set of these individual characteristics are modelled in the guise of learner profiles and they are used as personalisation parameters. Therefore, the identification of the specific and determining characteristics of the learner and the development of his/her profile is a very important step before building the learner profile and beginning the process of the learning personalizing. The literature is full of different proposals for modelling learner's profiles [1].

In the context of MOOCs, many researchers highlighted the importance of maintaining the massive aspect of MOOCs. They argue that the mass collaboration on which MOOCs are based is important, but different from the cognitive and motivational support obtained by communicating with a small mass of people who share common interests. The organization of small learning groups in the context of a MOOC is a good strategy. And this sheds the light on the fact that learning personalization features may be defined to identify the appropriate group for the learner and this was defined as the group profile [2].

Due to the information technology advancement, it is easy to collect data about learners using online learning environment such as MOOCS and analyse it. But that is the identification of the determinative features from the several existing features in the literature that have decisive impact on learning process and that can make sense to the explored data and promote the personalization process [3], [4].

Against this context, several quantitative studies have been made to explore the learners' profiles in MOOCs. Other research works defined learners profiles using features defined in literatures.

In this research work, we will begin by investigating the results of the learners' profiles defined by the experimental studies and the theoretical proposed learners' profiles based on the literature. We will identify after the most significant features that allow to group learners in homogeneous groups that can share the same learning experience.

2 Learners' profiing in MOOCs

Most of the researchs investigating the learners' profiles in MOOCs aim to study the learners' behaviour and their preferences in order to improve their learning experience and improve the pedagogical effectiveness of MOOCs. They followed two different strategies to define and identify the profile of the learner attending a MOOC. The first used strategy is the exploration of the mass of learners using their traces and some fulfilled questionnaires in order to identify the profiles of learners are attending MOOCs. These studies [3], [5], [6] [7] generally use a set of measurable features using a set of calculated indicators.They apply after automatic clustering

techniques to identify the topography of the mass. In the second strategy learners' profiles were modeled using a set of features using literatures works[8]–[10].

On the light of the existing works, we found that used learners' features could be divided into three main categories. The first category contains the demographic characteristics such as age region.The second category concerns the learners' knowledge and their learning preferences.The third category includes behavioural characteristics.

4 Learners' groups' porofiling

The experimental results demonstrate that the learners in MOOCs can be divided into harmonious groups that share some common characteristics and take the same learning process. In fact, the resulting groups referring to the chosen sets of features show that we can profile the groups of learners using a set of these features. The overlay of the clustering results with the learners' grade interpreted in [5] shows that learners homogeneous groups have conducts in MOOCs with significant similarity. This fact promotes the collaborative aspect of MOOCs, especially when each group shares an adequate personalized learning process.

The exisiting works confirm that it is possible to compose sub-groups that share some preferences and behaviour. For that, we confirm the idea that proposes personalized learning process to homogeneous learning groups.

Therefore, in the context of learning personalization, the behaviour of the learner is not an objective on it-self but it leads to calculating the learners' preferences, learning objectives, and competencies. For that, we propose to classify the learners' features into main categories: demographic features and learning features. The first category contains the demographic information about the learners. The second category includes informative features about the learners' competencies, about the learners' learning preferences and his learning objectives when attending the MOOC.

TABLE 1
Features defining the learners'groups in MOOCs

category	feature	determinative
demographic features	age	yes
learning features	language	yes
	learning style (active/ reflexive)	No
	learning style (visual/verbal)	yes
	learning objective (certificate importance)	yes
	competencies	yes

In the TABLE 1, we aggregated the used learners' features in the previous works that can be calculated automatically using the learners' traces and we classified them on the light of our main objective to determinative/ not determinative.

5 Conclusion

Learning personalization is proposed as a potential solution to improve learning efficiency in MOOCs. In many researchs, personalization is proposed at the group scale in order to keep the collaborative and massive aspect of MOOCs. In these propositions, researchers proposed to move from the heterogeneous population of learners to homogeneous sub-population. In our work, we investigated the used features to define the learner's profile in the exploratory experimental studies and in theoretical propositions. On the light of our main objective, we identified the determinative features to group learners and propose an adequate learning process. We chosen features that are automatically measurable using the learners' traces. In our future work, we aim to set up a personalization procedure using the identified criterias and to evaluate our choices experimentally.

References

[1] F. Essalmi, L. J. B. Ayed, M. Jemni, S. Graf, and Kinshuk, "Generalized metrics for the analysis of E-learning personalization strategies," *Computers in Human Behavior*, vol. 48, pp. 310–322, Jul. 2015, doi: 10.1016/j.chb.2014.12.050.

[2] S. Jean-Daubias, "Ingénierie des profils d'apprenants," thesis, Université Claude Bernard - Lyon I, 2011.

[3] R. Cabedo, T. C. Edmundo, and M. Castro, "A Benchmarking Study of Clustering Techniques Applied to a Set of Characteristics of MOOC Participants," in *2016 ASEE Annual Conference & Exposition Proceedings*, New Orleans, Louisiana, 2016, p. 26247, doi: 10.18260/p.26247.

[4] S. L. Watson, W. R. Watson, J. H. Yu, H. Alamri, and C. Mueller, "Learner profiles of attitudinal learning in a MOOC: An explanatory sequential mixed methods study," *Computers & Education*, vol. 114, pp. 274–285, Nov. 2017, doi: 10.1016/j.compedu.2017.07.005.

[5] T. A. Hennis, S. Topolovec, and O. Poquet, "Who is the Learner: Profiling the Engineering MOOC Student," *SEFI 44th Annual Conference, Tampere, Finland*, p. 14, 2016.

[6] A.-M. Scott, J. Owers, and V. Kovanovic, "Profiling MOOC course returners: How does student behavior change between two course enrollments?," *Proceedings of the Third ACM Conference on Learning @ Scale (L@S'16)*.

[7] M. Khalil and M. Ebner, "Clustering patterns of engagement in Massive Open Online Courses (MOOCs): the use of learning analytics to reveal student categories," *J Comput High Educ*, vol. 29, no. 1, pp. 114–132, Apr. 2017, doi: 10.1007/s12528-016-9126-9.

[8] G. Paquette, O. Mariño, D. Rogozan, and M. Léonard, "Competency-based personalization for massive online learning," *Smart Learn. Environ.*, vol. 2, no. 1, p. 4, Feb. 2015, doi: 10.1186/s40561-015-0013-z.

[9] H. Hajri, "Personnalisation des MOOC par la réutilisation de Ressources Éducatives Libres," Paris-Saclay, 2018.

[10] F. Clerc, M. Lefevre, N. Guin, and J.-C. Marty, "Mise en place de la personnalisation dans le cadre des MOOCs," in *7ème Conférence sur les Environnements Informatiques pour l'Apprentissage Humain - EIAH'2015*, Agadir, Morocco, 2015, pp. 291–300.

[11] B. Jebali and R. Farhat, "Toward personalization in MOOCs," *2017 6th International Conference on Information and Communication Technology and Accessibility (ICTA)*, pp. 1–6, 2017.

Budget-constrained dynamic Bag-of-Tasks scheduling algorithm for heterogeneous multi-cloud environment

Mouna Karaja, Meriem Ennigrou, and Lamjed Ben Said

Abstract— Cloud computing has reached huge popularity for delivering on-demand services on a pay-per-use basis over the internet. However, since the number of cloud users evolves, multi-cloud environment has been introduced where clouds are interconnected in order to satisfy customers' requirements. Task scheduling in such environments is very challenging mainly due to the heterogeneity of resources. In this paper, a budget-constrained dynamic Bag-of-Tasks scheduling algorithm for heterogeneous multi-cloud environment is proposed. By performing experiments on synthetic data sets that we propose, we demonstrate the effectiveness of the algorithm in terms of makespan.

Keywords—Bag-of-Tasks scheduling, budget-constrained, makespan, multi-cloud environment.

1 INTRODUCTION

Task scheduling, a well-known NP-Complete problem [1], has attracted a lot of attention since it plays a key role to improve the performance of any distributed system. The goal is to schedule users' tasks to Virtual Machines (VMs) by optimizing one or more objectives subject to some constraints. As the number of cloud users evolves, various limitations began to emerge. Thus, inter-cloud appears corresponding to the interconnection of multiple independent clouds in aggregation by a cloud user or a service [2], [3]. Task scheduling in inter-cloud environment is very challenging mainly based on the nature of the inter-cloud to scale dynamically and due to heterogeneity of resources which add more complexity to the scheduling problem [4].

Three mains task types exist in the literature: Single task or independent task, Bag-of-Tasks and workflows described as a directed acyclic graph where tasks are dependent having precedence relationship [5].

This paper is interested on scheduling Bag-of-Tasks (BoTs), a well known model that processes big-data applications supporting embarrassingly parallel jobs where tasks are independent of each other [6]. This model is interested in applications that exist in several domains like business analytics, engineering and science [7]. In this model, data is processed in parallel manner and results are recombined once individual parts have finished. The completion of the entire BoTs is defined by the completion of the last task.

Two main scheduling environment types exist: static and dynamic environment. A static environnement is suitable when all tasks are known in advance before execution. And, a dynamic scheduling environment appropriated when all tasks or machines are not fixed. In such environment, scheduling is done when a task arrives. Two dynamic modes exist: online mode where a task is scheduled to VMs as soon as it arrives and batch mode where tasks are first collected in a set that will be examined for scheduling at predefined times. In this paper, we consider a dynamic batch mode scheduling environment.

In the present work, a budget-constrained dynamic BoTs scheduling algorithm for heterogeneous multi-cloud environment is proposed, aiming to minimize the makespan while taking into consideration budget constaints.

The rest of this paper is organized as follows. Section 2 presents the state of the art. In Section 3 we introduce the scheduling problem by describing the cloud model and providing a mathematical formulation. After that we present the proposed algorithm, in Section 4. Experimental results are discussed in Section 5. Finaly, section 6 ends this paper with conclusions and future directions.

2 STATE OF THE ART

Task scheduling problem is a major research topic for the inter-cloud environment. It has attracted a lot of attention over the latest years. Task scheduling in inter-clouds environment can be seen as a meta-computing scheduling mechanism which, due to heterogeneity of resources, is not a specific problem but a set of problems [8]. The nature of the inter-clouds to scale dynamically and the unpredictable behavior of resources add more complexity to the job scheduling decision [9].

Two main types of inter-clouds exist in the litterature: federation clouds and multi-cloud [10]:

- A federation cloud is a volunteer interconnection of cloud infrastructures between a set of cloud providers in order to share resources among each other.

- *Mouna Karaja, Université de Tunis, Institut Supérieur de Gestion, SMART Lab. E-mail: mouna.karaja.isg@gmail.com.*
- *Meriem Ennigrou, Université de Tunis, Institut Supérieur de Gestion, SMART Lab. E-mail: meriem.ennigrou@enit.rnu.tn.*
- *Lamjed Ben Said, Université de Tunis, Institut Supérieur de Gestion, SMART Lab. E-mail: lamjed.bensaid@isg.rnu.tn.*

- A multi-cloud is the interconnection of multiple independent clouds in aggregation by a cloud user or a service. It does not imply a volunteer interconnection between clouds. Researchers use the word multi-cloud or inter-clouds for this type.

A specific type of multi-cloud is the hybrid Cloud that connects two or more different cloud infrastructures (private and public clouds) for on-demand leasing of public resources [4, 5].

Many algorithms for scheduling BoTs applications in inter-cloud environment have been proposed in the literature. Some authors have been interested on scheduling BoTs by implementing meta-heuristics.

A genetic algorithm for estimating suboptimal sets of cloud resources to execute budget-constrained and deadline-constrained BoTs applications is proposed in [11]. They implement too an agent-based approach for executing these applications by satisfying consumers' budgets and deadlines.

In [12], the problem of scheduling BoTs applications in a heterogeneous multi-cloud environment is considered. The algorithm takes into consideration multi-layered arrivals jobs (local and global) and introduces the notion of critical jobs which interrupt the processes of running jobs. Authors have proposed two meta-heuristic optimization approaches to solve this problem: Simulated Annealing and Tabu Search. These methods aim to minimize cost and maximize the scheduler performance.

The paper [13] is an extension of the previous work [12]. It presented two discrete optimization metaheuristics, Simulated Annealing and Thermodynamic Simulated Annealing approaches, for the scheduling of BoTs applications on a dynamic multi-cloud system with heterogeneous resources. The proposed approaches aimed to minimize cost and optimize performance.

Heuristics have been the subject of several works aiming to schedule BoTs in an inter-cloud environment. The papers discussed below are based on estimation mechanism. In [14], a novel task assignment algorithm named Fully Polynomial time Randomized Approximation Scheme (FPRAS) is presented. Scheduling decisions are made with a non-clairvoyant task assignment; which means, without any prior knowledge of the processing time of tasks.

The papers [15, 16] have treated the problem of scheduling large BoTs onto multi-cloud with no a-prior information about task completion times. They have introduced BaTS, a budget-constrained scheduler aiming to estimate costs and makespan for a given bag-of-tasks on different cloud offerings. The proposed approaches estimate the task execution time and decide which combination of machines would satisfy the budget constraint and optimize the makespan.

Another work [17] aiming to solve BoTs scheduling problem in hybrid cloud presents a dynamic real time approach taking into consideration budget constraints. This approach is composed by a configuration agent that will have the role to estimate the total number of public resources needed to satisfy task's deadline. In addition, a Task Dispatcher Scheduler aims to find the best VM (from either public or private cloud) to each task on the basis of task characteristics.

Another work proposed an autonomic deadline-constrained bag-of-tasks scheduling algorithm for federated clouds [18]. It aims to minimize the execution cost and deadlines. The contribution of this work is that the scheduler includes a runtime estimator based on sampled data to generate the runtime of each task autonomously. The scheduler targets a Hybrid Cloud infrastructure: when the private cloud infrastructure is limited, the scheduler send tasks to the public cloud hired from public providers.

The paper [19] proposed a Distributed Resource Allocation approach (DRA) for cloud federation aiming to minimize marginal cost while improving resource utilization. This approach adopts a pricing strategy when a resource conflict occurs and proposes a task-groping approach aiming to group tasks according to their communication pattern in order to minimize communication cost between clouds.

TABLE 1
COMPARISON OF BOTS SCHEDULING ALGORITHMS IN THE LITTERATURE

Article	Inter-cloud type	Environment	Objective	Constraints
[11]	Multi-cloud	Batch mode	Minimize cost and makespan	Deadline Budget
[12]	Multi-cloud	Batch mode	Minimize cost and makespan	Load balancing
[13]	Multi-cloud	Batch mode	Minimize cost and makespan	Load balancing
[14]	Multi-cloud	Batch mode	Minimize cost and makespan	
[15]	Multi-cloud	Batch mode	Minimize makespan	Budget
[16]	Multi-cloud	Batch mode	Minimize execution time	Budget
[17]	Multi-cloud	Online mode	Minimize starting time	Deadline
[18]	Federated clouds	Batch mode	Minimize execution cost	Load balancing
[19]	Federated clouds	Batch mode	Minimize execution cost	Deadline

3 CLOUD MODEL AND PROBLEM STATEMENT

3.1 CLOUD MODEL

In our cloud model, we consider a multi-cloud system

composed by several clouds. Each cloud is made up of distributed data centers. Each data center is divided into several physical machines (PM) that host plenty of VMs to execute customers' applications (Note that an application is composed by at least one BoTs).

As shown in Fig.1., all customers' BoTs are submitted to the multi-cloud scheduler in the multi-cloud queue. When a BoTs is submitted, the multi-cloud scheduler splits it into several tasks and submits them to cloud schedulers. Each cloud scheduler schedules tasks received from the multi-cloud scheduler to its own VMs.

In our model, we go with the assumption that a VM can execute only one task on a time and tasks are not preempted.

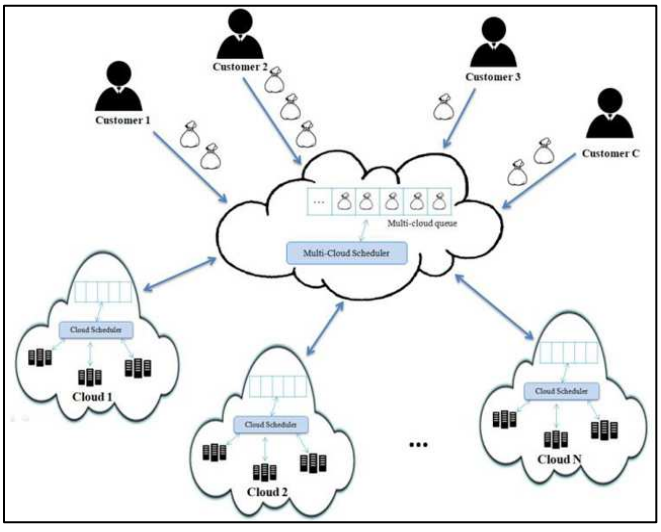


Fig.1.: Cloud model

3.2 PROBLEM DESCRIPTION

Consider an Inter-Cloud composed by N clouds {C_1, C_2, … C_N}. Each Cloud is composed by M types of virtual machines {VM_1, VM_2, …, VM_M}. Each VM is caracterized by an execution speed CPU_M expressed in mips (millions of instructions per seconde) and a cost $cost_M$ expressed on \$/h.

An inter-cloud queue is composed by X BOTs {BOT_1, BOT_2, …, BOT_X} where each BOT is associated to a budget $Budget_X$. An arrival time Arr_X is assigned to each BOT which contains T independents tasks {t_1, t_2, …, t_T} and each task has a number of instructions Ins_T .

In our model, we consider a heterogeneous multi-cloud environment in which each cloud has different Virtual Machine types. That's why; the execution time of tasks differs in each cloud. Note that each task will be executed by one and onlyone cloud at a time.

Let $Exec_{Tijz}$ the execution time of task t_i of BOT_j on the VM_z where i{1, …, T}, j{1, …,X} and z{1,..., M} defined as follows:

$$Exec_{Tijz} = Ins_{Tij} / CPU_z \quad (1)$$

Let $Cost_{Tijz}$ the execution cost of the task t_i of BOT_j on the VM_z defined by equation (2).

$$Cost_{Tijz} = (Exec_{Tijz} / 3600) \times cost_z \quad (2)$$

Consider $Cost_j$ the execution cost of the BOT_j defined by the equation (3).

$$Cost_j = \sum_{i=1}^{T} Cost_{Tijz} \quad (3)$$

The problem consists of scheduling BOTs on Virtual Machines belonging to different clouds while minimizing the completion time $Cmax_j$ where j{1, …,X} and taking into consideration budget constraints .

Let $Cmax_j$ the completion time of the last task in progress belonging to the BOT_j defined by:

$$Cmax_j = MAX\ (C_{Tij}) - Arr_j \quad (4)$$

Where C_{Tij} is the completion time of the task i ∈ BOT_j.

4 PROPOSED ALGORITHM

Solving the problem of scheduling BoTs in a multi-cloud environment can be modeled in a hierarchical manner structured in two levels. The upper level is the multi-cloud scheduler, which has as input the tasks of cloud customers working as a meta-scheduler distributing the arrival tasks to the clouds. The lower level is the cloud scheduler that schedules the tasks received from the multi-cloud scheduler to its own resources.

The scheduling at the multi-cloud level is done at predefined times. Meanwhile, arriving BoTs are putted on the Inter-cloud waiting list and will be scheduled in the next scheduling step.

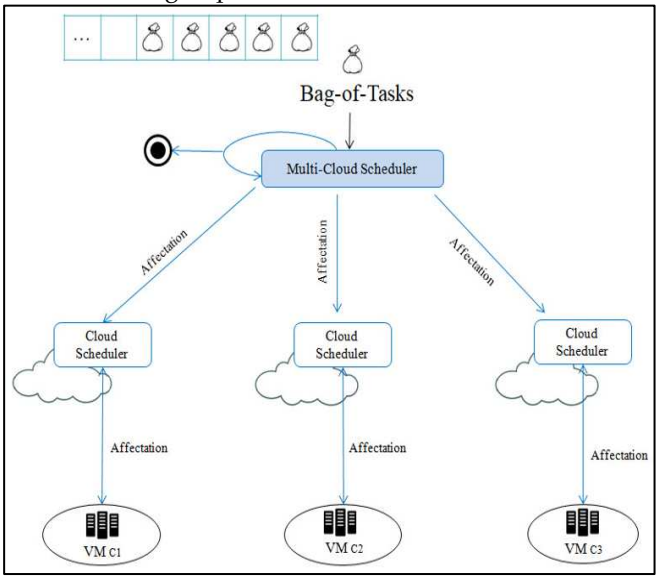


Fig. 2.: Modeling of the proposed algorithm (version 1)

Fig.2. illustrates the first version of the proposed BoTs scheduling algorithm aiming to minimize the makespan while taking into consideration budget constraints.

The scheduling process is as follows: At each scheduling step, the multi-cloud scheduler treats BoTs one by one. It begins by generation a solution that corresponds to the assignment of tasks to clouds randomly. Each cloud scheduler, schedule received tasks (considered as a part of the multi-cloud solution) to his own VMs and calculate the makespan and the cost. Then, it transfers obtained values to the multi-cloud scheduler.

The latter collects parts of the solution from clouds and evaluates it in terms of makespan and cost. If the solution satisfies budget constraint, it will be saved.

This solution undergoes variation mechanisms using a local search in order to minimize the makespan. This process is repeated until a certain number of iterations. The output of this algorithm is a solution with minimum makespan and that respects budget constraints.

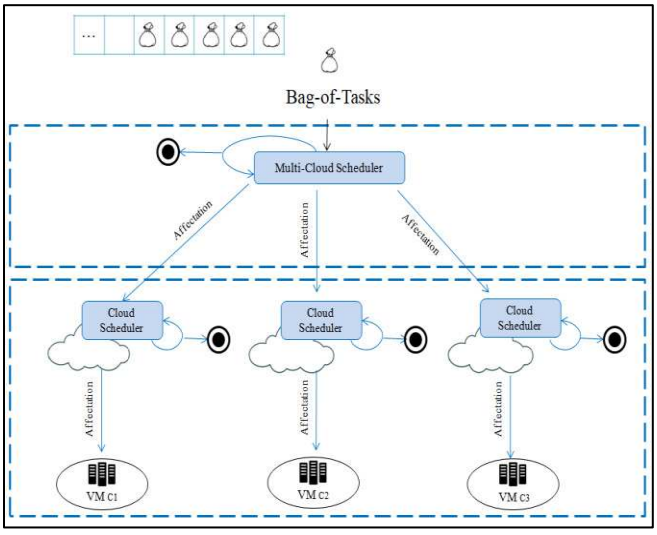


Fig. 3.: Modeling of the proposed algorithm (version 2)

Fig.3. illustrates the second version of the proposed BoTs scheduling algorithm. In this version, the makespan is optimized on two levels: in the multi-cloud scheduler such as in the first version in order to optimize the makespan of the whole BOT. And in the cloud scheduler, where the completion time of the received tasks from the multi-cloud is optimized.

A cloud-level solution is an assignment of tasks to virtual machines. The cloud scheduler evaluates the solution in terms of makespan and execution cost. This solution undergoes variation mechanism applying a local search algorithm. This process is repeated until a well-determined number of iterations. The solution with minimal makespan is saved and sended at the end of the algorithm to the multi-cloud scheduler.

The latter collects parts of his solution from clouds and evaluates it in terms of makespan and cost. If the solution satisfies budget constraint, it will be saved.

This solution undergoes variation mechanisms using a local search in order to minimize the makespan. This process is repeated until a certain number of iterations.

5 EXPERIMENTAL RESULTS

5.1 EXPERIMENTAL DATA SET

In the literature, there are no benchmarks to reflect the workloads of a multi-cloud environment. The researchers are supposed to:

- Either, use some real traces available on the internet and unfortunately that do not prove the efficiency of the algorithms. Or,
- Use statistical models to transform, using a sampling mechanism, a real workload into a synthetic workload but that are limited in terms of data and unfortunately are not feasible in the case of a multi-objective algorithm.
- Or also generate randomly synthetic instances that meet their needs. In this paper, we will opt for this choice.

TABLE 2
BOT INPUT DATA

Problem instance size	20 BOTs
Arrival time of BOT	0, 15, 30, …, 285 minutes
Budget of BOT	0.75, 1.5, 2.25, …, 15$
Nbr. of tasks per BOT	100, 200, …, 2000 tasks
Nbr. of instructions per task (in million of instructions)	5000, 7500, ..., 50000 mi

Table 2 illustrates the BoTs input data. For our experimental study, we randomly generated a problem instance containing 20 BoTs with different sizes (from 100 to 2000 tasks with an increment of 100 tasks) with budgets (ranging from 0.75$ to 15$ with an increment of 0.75$). Note that the minimum budget is associated to the BOT with the minimum number of tasks (for example, the BOT containing 100 tasks is associated with the budget 0.75 $; the BOT containing 200 tasks is associated with the budget 1.5 $; ...; the BOT containing 2000 tasks is associated with the budget $ 15).

Each BoT arrives at the inter-cloud waiting list at a different time ranging from 0 to 285 minutes with 15 minutes increments. The arrival time is assigned to BoTs randomly.

Each task in a BOT contains a different number of instructions assigned randomly (using a uniformly distributed random number) from the available sizes ranging from 5,000 to 50,000 million of instructions.

TABLE 3
VIRTUAL MACHINES INPUT DATA

Cloud 1 (8 VMs)		
Nbr. Of VM	Speed (mips)	Cost ($/h)
2	250	0.1
2	500	0.3
2	750	0.5
2	1000	0.7

Cloud 2 (16 VMs)		
Nbr. Of VM	Speed (mips)	Cost ($/h)
3	250	0.1
3	500	0.3
3	750	0.5
3	1000	0.7
2	1250	0.9
2	1500	1.1

Cloud 3 (24 VMs)		
Nbr. Of VM	Speed (mips)	Cost ($/h)
4	250	0.1
4	500	0.3
4	750	0.5
4	1000	0.7
4	1250	0.9
4	1500	1.1

Table 3 illustrates the VMs input data. For our experimental study, we consider in our model 3 clouds with 8, 16 and 24 VMs respectively. The VM speed is ranging from 250 to 1500 milion of instructions per seconds (mips) with an increment of 250 mips. And the VM cost is ranging from 0.1 to 1.1$ with 0.1$ increments. Note that to the VM with lower speed is associated the cheapest price and vice versa as shown on tables.

TABLE 4
ALGORITHM PARAMETERS

Nbr. Of iterations of the multi-cloud scheduler algorithm	30
Nbr. Of iterations of the cloud scheduler algorithm	20

For the experimental study, as shown in table 4, we have fixed 30 the number of iterations for the multi-cloud scheduler algorithm and 20 iterations for the cloud scheduler algorithms. Those numbers are chosen after several tries.

5.2 RESULTS AND DISCUSSION

Table 5 illustrates the comparison between the two versions of the BoTs scheduling algorithm in a multi-cloud environment in terms of makespan.

The assignment of the arrival time to different BoTs is done randomly. As shown in the table, the first BoTs that arrives to the multi-cloud scheduler is the BoTs named BOT07 at 0 seconds and the last one is the BoTs named BOT05 at 17100 seconds.

As cited in the table 4, BoTs are named as follows: BOT01, BOT02, …, BOT20. The name of the BoT indicates the BoTs size (BOT01 represents a set of 100 tasks, BOT02, represents a set of 200 tasks, …, and BOT20 represents a set of 2000 tasks).

We notice that the version 2 of the algorithm outperforms the version 1 for all BoTs size in the instance.

TABLE 5
COMPARISON IN TERMS OF MAKESPAN

	Arrival Time	Version 1	Version 2
BOT07	0	2810	1030
BOT08	900	5770	4730
BOT14	1800	10470	8520
BOT02	2700	10830	8610
BOT10	3600	13990	11010
BOT17	4500	20130	16810
BOT12	5400	24680	20460
BOT13	6300	29680	24720
BOT20	7200	37240	32060
BOT15	8100	42920	37330
BOT04	9000	44200	38310
BOT16	9900	50190	43960
BOT18	10800	56390	49860
BOT11	11700	60170	53260
BOT01	12600	60080	53130
BOT19	13500	65600	58650
BOT06	14400	67190	59590
BOT09	15300	69550	61380
BOT03	16200	70260	61580
BOT05	17100	71060	62480
Makespan Instance		71060	62480

As we observe, the completion time of the problem instance composed by 20 BoTs with differents size is 71060 seconds for the version 1 (the equivalent of 19 hours 44 minutes and 20 seconds) and 62480 seconds for version 2 (the equivalent of 17 hours 21 minutes and 20 seconds) which proves the effectiveness of the algorithm.

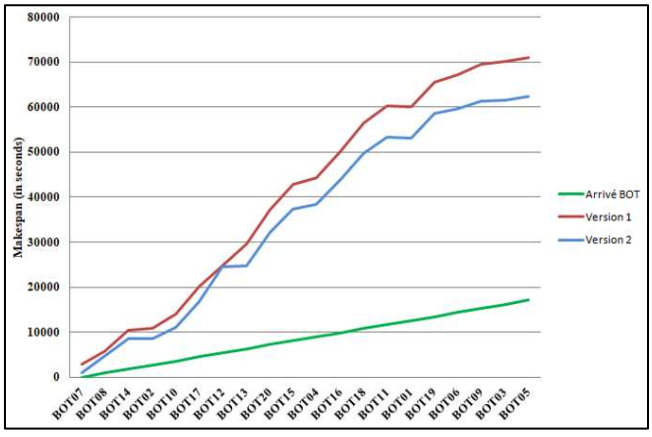


Fig. 4.: completion time of different BoTs in the instance

Since our algorithms are Budget-constrained, we have chosen execution cost as a parameter measure to evaluate its effectiveness. The comparison between the two versions of the algorithm in terms of execution costs is presented in Table 6.

TABLE 6
COMPARISON IN TERMS OF COST

	Budget (in $)	Version 1 (cost in $)	Version 2 (cost in $)

BOT07	5.25	2.34	4.081
BOT08	6	2.481	4.44
BOT14	10.5	4.596	7.894
BOT02	1.5	0.853	1.067
BOT10	7.5	3.221	5.699
BOT17	12.75	5.266	9.787
BOT12	9	3.884	6.721
BOT13	9.75	4.170	7.352
BOT20	15	6.237	10.959
BOT15	11.25	4.707	8.544
BOT04	3	1.375	2.328
BOT16	12	5.144	9.033
BOT18	13.5	5.425	9.912
BOT11	8.25	3.530	6.183
BOT01	0.75	0.576	0.737
BOT19	14.25	5.859	10.352
BOT06	4.5	2.05	3.461
BOT09	6.75	2.896	5.108
BOT03	2.25	1.075	1.786
BOT05	3.75	1.626	3.017
Cost instance		67.311	118.461

As shown in Fig.5., both algorithm versions managed to find an execution cost for each BoTs size that respect BoTs budget.

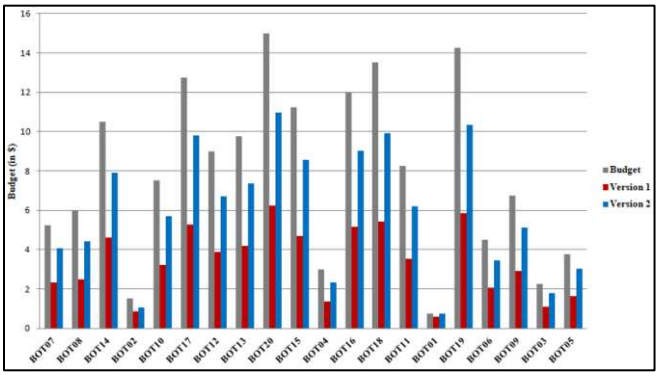


Fig. 5.: Execution cost of different BoTs in the instance

If we compare the figures Fig. 4 and Fig. 5., we observe that the version 2 is more efficient to find minimal makespan. But, the version 1 found a minimal cost. Indeed, this is expected since the most performing VM is the more expensive.

6 Conclusion

In this paper, we have presented two versions of a budget-constrainted dynamic task scheduling algorithm for heterogeneous multi-cloud environment. The optimization level is the main difference between the two versions. In version 1, the inter-cloud scheduler is the only responsible to minimize BoTs' makespan. However, in version 2, the cloud scheduler is also responsible to optimize the makespan of the BoTs' subsets received from the inter-cloud scheduler. The result signifies that the performance of the version 2 is better than that of version 1 in terms of makespan.

As the completion time and the execution cost are two conflicting objectives, our future works will focus on proposing a budget-constrained multi-objective task scheduling algorithm.

References

[1] N. Patil, D. Aeloor, A review – Different scheduling algorithms in Cloud Computing environment, 11th International Conference on Intelligent Systems and Control (ISCO), 2017.

[2] B. Kezia Rani, B. Padmaja Rani, A. Vinaya Babu, Cloud Computing and Inter-Clouds - Types, Topologies and Research Issues, 2nd International Symposium on Big Data and Cloud Computing (ISBCC'15), Procedia Computer Science 50, pp. 24 – 29, 2015.

[3] N. Grozev, R. Buyya, Inter-Cloud architectures and application brokering: taxonomy and survey, Software – Practice and experience, 44: pp. 369–390, 2012.

[4] S. Sotiriadis, N. Bessis, N. Antonopoulos, Towards inter-cloud schedulers: A survey of meta-scheduling approaches, 2011 international Conference on P2P, Parallel, Grid, Cloud and Internet Computing, 2011.

[5] I. Gupta, M.S. Kumar, P.K. Jana, Compute-Intensive Workflow Scheduling in multi-cloud environment, International Conference on Advances in Computing, Communications and Informatics, 2016.

[6] W. Cirne, D. Paranhos, L. Costa, E. Santos-Neto, F. Brasileiro, J. Sauve, F. Silva, C. Barros, C. Silveira, "Running bag-of-tasks applications on computational grids: the mygrid approach", International Conference on Parallel Processing, pp. 407–416, 2003.

[7] M. HoseinyFarahababy, Y.C. Lee, A.Y. Zomaya, "Non-clairvoyant assignment of Bag-of-Tasks applications across multiple Clouds", 13th International Conference on Parallel and Distributed Computing, Applications and Technologies, 2012.

[8] S. Sotiriadis, N. Bessis, N. Antonopoulos, Towards inter-cloud schedulers: A survey of meta-scheduling approaches, 2011 international Conference on P2P, Parallel, Grid, Cloud and Internet Computing, 2011.

[9] N. Bessis, S. Sotiriadis, V. Cristea, F. Pop, Modelling Requirements for Enabling Meta-Scheduling in Inter-Clouds and Inter-Entreprises, Third International Conference on Intelligent Networking and Collaborative Systems, 2011.

[10] B. Kezia Rani, B. Padmaja Rani, A. Vinaya Babu, Cloud Computing and Inter-Clouds - Types, Topologies and Research Issues, 2nd International Symposium on Big Data and Cloud Computing (ISBCC'15), Procedia Computer Science 50, pp. 24 – 29, 2015.

[11] J.O. Gutierrez-Garcia, K.M. Sim, GA-based cloud resource estimation for agent-based execution of bag-of-tasks applications, Information Systems Frontiers 14: pp. 925-951, 2012.

[12] I.A. Moschakis, H.D. Karatza, A meta-heuristic optimization approach to the scheduling of bag-of-tasks applications on heterogeneous clouds with multi-level arrivals and critical jobs, Simulation Modelling Practice and Theory, 2015.
[13] I.A. Moschakis, H.D. Karatza, Multi-Criteria scheduling of Bag-of-Tasks applications on heterogeneous interlinked Clouds with Simulated Annealing, The Journal of Systems and Software, 2014.
[14] M. HoseinyFarahababy, Y.C. Lee, A.Y. Zomaya, Non-clairvoyant assignment of Bag-of-Tasks applications across multiple Clouds, 13th International Conference on Parallel and Distributed Computing, Applications and Technologies, 2012.
[15] A.M. Oprescu, T. Kielmann, H. Leahu, Budget estimation and control for Bag-of-Tasks scheduling in Clouds, Parallel processing letters, Vol. 21, No. 2, pp. 219-243, 2011.
[16] A.M. Oprescu, T. Kielmann, H. Leahu, Stochastic Tail-Phase Optimization for Bag-of-Tasks execution in Clouds, IEEE/ ACM fifth International Conference on Utility and Cloud Computing, 2012.
[17] O.C. Marcu, C. Negru, F. Pop, Dynamic Scheduling in Real Time with Budget Constraints in Hybrid Clouds, Economics of Grids, Clouds, Systems, and Services: 12th International Conference GECON 2015, pp. 18–31, 2016.
[18] Y. Lee, K. Huang, M. Shieh, K. Lai, Distributed resource allocation in federated clouds, Journal of Supercomputing, Vol. 73, Issue 7, pp. 3196–3211, 2016.
[19] V. Peláez, A. Campos, D.F. Garcia, J. Entrialgo, Autonomic Scheduling of Deadline-Constrained Bag of Tasks in Hybrid Clouds, International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS), 2016.

Mouna Karaja is currently pursuing his PhD on computer sciences from the Institut supérieur de gestion de Tunis, University of Tunis. Her research interests are scheduling problems, cloud computing, evolutionary algorithms and optimization.

Meriem Ennigrou received his PhD in computer science from the Institut supérieur de gestion de Tunis, University of Tunis in 2010. She is an Assistant Professor and Research member of SMART Lab. with University of Tunis, Tunisia. She is currently working on scheduling problems, evolutionary algorithms and Vehicle routing problems.

Lamjed Ben Said received the B.Sc. degree from the Université de Tunis, Tunisia, in 1998, and the M.Sc. and Ph.D. degrees from the University of Paris VI, Paris, France, in 1999 and 2003, respectively, all in computer science. He is currently a Professor of Computer Science with the Université de Tunis, where he is the Head of the SMART Laboratory and the Dean of Institut supérieur de Gestion (Université de Tunis). Author and co-author of 130 journal/patent/international conference/chapter book papers. His current research interests include multiagent systems, multicriteria decision making, evolutionary computation, and behavioral economics.

An investigation of AI in games: educational intelligent games vs non-educational games

Sanda Hammedi[1], Fathi Essalmi[1, 2], Mohamed Jemni[1], Alaa A. Qaffas[3]
[1]Research Laboratory of Technologies of Information and Communication and Electrical Engineering (LaTICE), University of Tunis, Tunisia
[2]Assistant Professor Management Information System Department College of Business University of Jeddah
[3]Head of Management Information System Department College of Business University of Jeddah, Saudi Arabia
Sanda.hammedi@yahoo.fr
fathi.essalmi@isg.rnu.tn / feessalmi@uj.edu.sa
mohamed.jemni@fst.rnu.tn
aaqaffas@uj.edu.sa

***Abstract*— Attracting and motivating the player is considered as one of the major reasons of the success of games. In fact, games are the form of art which is characterized by attractiveness. Artificial Intelligence (AI) is used in games to have an interesting and interactive experience of the player. However, the objectives of increasing the attractiveness with AI techniques differ between games for pleasure and games for educational contexts. While non-educational intelligent games focus mainly on the entertainment side, intelligent educational games dive into deep issues. Various AI techniques are used in games such as pathfinding algorithms, decision trees, intelligent narrative technologies and intelligent agents. In one hand, this paper presents a literature review about non-educational intelligent games and intelligent educational games. The paper classifies the presented intelligent games according to the AI techniques used in each game. In the other hand, a discussion is conducted to observe the limitations and the future directions concerning the use of AI in games.**

Keywords- Intelligent games; Educationonal intelligent games; Artificial Intelligence techniques ; non-educational intelligent games.

I. INTRODUCTION

Games are generally developed for entertainment and fun. Playing games is widespread between all categories of people. They are not only dedicated for children, but also a variety of games are addressed for older people. The number of people playing games is continuously growing up. It is predicted that it will reach 2.7 billion gamers in the world by 2021 [1]. This huge number presents a catalyst to develop games with purposes. In this context, games with educational purposes are designed to make learning more enjoyable. Educational games have a great impact in improving player's memory capacity and enhancing his/her skills in problem solving and practical thinking as well as gaining more knowledge and to obtain new skills in various fields such as in computer science field and medical field [3, 4, 5].

Over the last years, artificial intelligence techniques took place in many applications and computer systems. Artificial Intelligence (AI) represents machines that mimic human mind intelligence [14]. These machines become able to do tasks which are normally specific to human abilities such as learning, searching, solving problems and performing thinking tasks, etc [2, 14]. Even in some cases we can talk about a superhuman abilities achieved by AI such as AlphaGo which is famous by defeating one of the top human Go player.

AI is adopted in games in order to make the player more interested and motivated. Many AI techniques are used in games such as adding smart behavior to the non-player characters (NPCs) and controlling it, using of intelligent algorithms and decisions trees. AI techniques used in games are not necessarily visible to the player, such as data mining and procedural-content generation.

The paper is structured as follow: Section 2 deals with the review of relevant literature regarding non-educational intelligent games and intelligent educational games. Section 3 discusses the future directions of AI techniques in educational games. Finally, Section 4 summarizes the paper.

.

II. LITERATURE REVIEW

A. Non-educational intelligent games

Born [14] states that AI has a huge impact in today's world regarding the usefulness of integrating AI in machines. The emergence of AI techniques in games has a great effect in motivating the player and engaging him/her. Many studies highlight the impact of using AI in games and claim that the intelligent side of the game makes it more attractive, exciting and challenging [12, 15, 16].

AI-based games or also called intelligent/smart games are games in which AI techniques are used. Various AI techniques have been used in games in order to attract the player or to ameliorate the player experience in the game. The table1 presents examples of non-educational intelligent games and classify them according to the used techniques.

TABLE 1. EXAMPLES OF NON-EDUCATIONAL INTELLIGENT GAMES

Intelligent game	Description	techniques
Intelligent game system [6]	This study aims to make board and tabletop games intelligent. It consists of an intelligent game system based on the use of electronic materials such as sensors and controller. These materials work cooperatively; The sensors collect object data from intelligent game piece objects and transmit the gathered information to a controller. Other treatment take place to finally prepare and transfer a changing image to a projector .	Use of sensors and a controller
Car racing game [7]	In this study, two version of A* algorithm are used to solve pathfinding problem in car racing game. A third algorithm is also proposed dealing with obstacles avoidance. These algorithms succeed in finding the path in racing game and have a great effect in enhancing the performance and decreasing the lap times comparing to the original A* algorithm.	Pathfinding algorithms
Car race game [8]	In this research, Dynamic pathfinding algorithm and A* algorithm are combined and used in car race game. Results show that the combination of these two algorithms is more efficient than using only the A* algorithm in avoiding obstacles and attaining the finish line.	A* algorithm in combination with dynamic pathfinding algorithm
Military Chess game [9]	In one hand, in order to increase the level of accuracy of predicting enemy chess pieces, Monte Carlo algorithm is used. In the other hand, Q-learning algorithm is used in the interpretation of the player's chess pieces. Findings prove the efficiency of this dynamic evaluation approach comparing with static evaluation strategy.	Monte Carlo algorithm, Q-learning algorithm
Sports game [15]	This study focuses on adding commentary to a sport game. For this goal, an AI system for information retrieval is implemented. This system proposes automatically stories as live comments to tell during the game.	Information retrieval and Machine learning technique

From the table above, a variety of AI techniques are used in games. Some limitations in the presented studies can be figured out. For example, in [8], it is proved that the combination between the dynamic pathfinding algorithm and A* algorithm gives promising results in car racing games without trajectory obstacles or with stable obstacles, while in trajectories with dynamic obstacles, the proposed approach doesn't work properly under certain conditions. In [15], the developed intelligent system for sport game offers only offline commentary, online commentary are not supported.

The next subsequent section deals with some examples of intelligent educational games.

B. Intelligent educational games

Adopting game elements and putting them in a way to serve educational needs, allow the development of educational games which may engage learners to discover more knowledge and to obtain new skills that encompass problem ponder and solving skills.

AI techniques have been used in games based learning in order to improve players' engagement and learning outcomes. Hence, games deal intelligently with issues for learning purposes and make the player entertained and enlightened in the same time. In this context, Weng and his colleagues [10] prove that involving assessments in an intelligent game to do by student present a motivational method of learning since students do not feel stressed as when taking tests in classroom. Lester and his colleagues [12] affirm that intelligent game-based learning environments are continuously growing up to take more large standing of educational settings in all fields and for all ages.

Table 2 displays examples of intelligent educational games and their pedagogical objectives with a brief description and the used techniques for each one.

TABLE 2. EXAMPLES OF INTELLIGENT EDUCATIONAL GAMES

Intelligent educational game	Pedagogical objective	Description	AI techniques
QuizMASter game [10]	Doing assessement	A quiz game based learning aiming to provide adaptive tests for students is developed. Intelligent agents are involved is this game to accomplish some tasks such as generating player's profile, determining player's level of knowledge, etc, in order to give an adaptive assessment for each player.	Multi-agent system

Mobile educational game [11]	Doing activities	This study describes the use of agents in an educational game. The developed game can generate a series of activities dedicated for employees. Each implemented agent has specific tasks and in some cases they cooperate with each others.	Multi-agent system
Crystal Island [12]	Microbiology	An intelligent game-based learning dedicated for microbiology students is developed. AI techniques are used such as intelligent tutoring systems and intelligent narrative technologies	AI-based narrative features and intelligent tutoring system
An intelligent game based learning [13]	Improving player's sign language skills	An intelligent game based learning dedicated for hearing impairment people is implemented for learning sign language. Researchers prove the effectiveness of the developed intelligent game in improving player's sign language skills	Expert system

In the table above, various intelligent educational games are displayed. It is important to mention that despite of the impact of adding an intelligent side in games, there are a limited number of intelligent games with learning purposes. Numerous AI techniques are involved in games. Using intelligent agents has a good impact in solving limited resources problem. Furthermore, it makes the modification of the game more flexible, that improving the game and adding some functions only requires the modification of the agent [11].

In [13], the experimentation proves benefits of the developed intelligent game-based learning system in attracting students, and shows the facility of using it as well as its impact in increasing student's skills in American Sign Language. However, the presented system displays only a limited number of vocabularies comparing to the big quantity of vocabularies presented in school curriculum. In addition, the cost-effectiveness is related to the quantity of the presented vocabularies, that's why, it will be higher by augmenting the number of vocabularies.

III. DISCUSSION

Despite the contribution of AI in games, it is not yet been widely investigated. We noticed that few intelligent games are developed in the medical field. For example, the approach developed in [13] which aims to help hearing impairment people in learning sign language, can be applied in a similar context presented in [17, 18] to help brain-injured rehabilitation patients as they have similar objectives.

New scope concerning AI in games is about artificial general intelligence. It refers to the creation of an AI agent able to solve any mental tasks that human can perform, such as communicating in natural language, predicting future events, making judgment, using different skills to attain a specific goal , and so on. In the same context, we can talk about general game AI. It refers to (1) the development of AI methods that can be used in any game that the method is applied to. (2) The development of method able to do different tasks (modeling, testing…). (3) The development of method that fit to a large variability among people (playing style, preferences…). It is important to mention that talking about generalization of concepts doesn't mean that methods dealing with only one task are useless since they are necessary as a proof of concept, but there is an insisting need to validate a case study in a more general context [19].

IV. CONCLUSION

Intelligent non-educational games and intelligent educational games are presented in this paper. Some limitations are mentioned and future directions are discussed.

A lot of attention should be given to AI issue regarding its great role in computer application. Various AI techniques are involved in games. Many studies prove the efficiency of the use of AI in games as well as those with educational purposes. Among the most used AI techniques in games, we can cite Monte Carlo Search Tree, A* algorithm, intelligent narrative technologies, intelligent tutoring systems and intelligent agents. Some future directions of research are displayed such as general game AI.

REFERENCES

[1] Statista (2017), "Number of active video gamers worlwide from 2014 to 2021". Retrived from: https://www.statista.com/statistics/748044/number-video-gamers-world/

[2] Millington, I. (2019). AI for Games. CRC Press.

[3] Khenissi, M. A., Essalmi, F., & Jemni, M. (2013, October). A learning version of Pacman game. In Information and Communication Technology and Accessibility (ICTA), 2013 Fourth International Conference on (pp. 1-3). IEEE

[4] Hammedi, S.,Essalmi, F., and Chang, M, "Healthy Kidney: An educational game fr health awareness," International Conference on Computers in Education (ICCE), 2017

[5] Khenissi, M. A., Essalmi, F., Jemni, M., & Chang, T. W. (2015). Measuring learners' working memory capacity from their interactions within educational game. In Emerging Issues in Smart Learning (pp. 233-237). Springer Berlin Heidelberg

[6] Maharbiz, M. M., & Jaqua, S. (2019). Intelligent game system for putting intelligence into board and tabletop games including miniatures.U.S. Patent No. 10,265,609. Washington, DC: U.S. Patent and Trademark Office.

[7] Wang, J. Y., & Lin, Y. B. (2012). Game ai: Simulating car racing game by applying pathfinding algorithms. *International Journal of Machine Learning and Computing*, *2*(1), 13.

[8] Sazaki, Y., Primanita, A., & Syahroyni, M. (2017, July). Pathfinding car racing game using dynamic pathfinding algorithm and algorithm A∗. In *2017 3rd International Conference on Wireless and Telematics (ICWT)* (pp. 164-169). IEEE.

[9] Xiaochuan, Z., Wanwan, W., Qin, L., Tianao, W., & Hao, S. (2019, June). The Design and Realization of Dynamic Evaluation Strategy of Pieces in Military Chess Game System. In *2019 Chinese Control And Decision Conference (CCDC)* (pp. 6287-6292). IEEE.

[10] Weng, M. M., Fakinlede, I., Lin, F., Shih, T. K., & Chang, M. (2011, July). A conceptual design of multi-agent based personalized quiz game. In 2011 IEEE 11th International Conference on Advanced Learning Technologies (pp. 19-21). IEEE.

[11] Lu, C., Chang, M., Huang, E., & Chen, C. W. (2011, March). Architecture and collaborations among agents in mobile educational game. In 2011 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops) (pp. 556-560). IEEE.

[12] Lester, J. C., Ha, E. Y., Lee, S. Y., Mott, B. W., Rowe, J. P., & Sabourin, J. L. (2013). Serious games get smart: Intelligent game-based learning environments. AI Magazine, 34(4), 31-45.

[13] Kamnardsiri, T., Hongsit, L. O., Khuwuthyakorn, P., & Wongta, N. (2017). The Effectiveness of the Game-Based Learning System for the Improvement of American Sign Language Using Kinect. Electronic Journal of e-Learning, 15(4), 283-296..

[14] Born, R. (2018). Artificial intelligence: The case against. Routledge.

[15] Subramaniyaswamy, V., Logesh, R., & Indragandhi, V. (2018). Intelligent sports commentary recommendation system for individual cricket players. *International Journal of Advanced Intelligence Paradigms*, *10*(1-2), 103-117.

[16] Wimaladharma, S. T. C. I., Sampath, A. G. A., Sampath, J. D. B., & Sapumohotti, C. H. V. (2019, May). A Game-Based Driving Learning System for Sri Lankan Driving Learners to Enrich the Awareness of Road Rules. In 2019 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS) (pp. 195-199). IEEE.

[17] Putnam, C., Cheng, J., Lin, F., Yalla, S., & Wu, S. (2016, May). 'Choose a Game' Creation and Evaluation of a Prototype Tool to Support Therapists in Brain Injury Rehabilitation. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (pp. 2038-2049).

[18] Lange, B., Chang, C. Y., Suma, E., Newman, B., Rizzo, A. S., & Bolas, M. (2011, August). Development and evaluation of low cost game-based balance rehabilitation tool using the Microsoft Kinect sensor. In 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 1831-1834). IEEE.

[19] Yannakakis, G. N., & Togelius, J. (2018). Frontiers of Game AI Research. In Artificial Intelligence and Games (pp. 279-291). Springer, Cham.

Personalization of MOOCs for increasing the retention rate of learners

Kaouther Kaabi[1], Fathi Essalmi[1, 2], Mohamed Jemni[1], Alaa A. Qaffas[2]
Research Laboratory of Technologies of Information and Communication and Electrical Engineering (LaTICE), Tunis national higher school of engineering (ENSIT), Tunisia[1]
Management Information Systems Department, College of Business, University of Jeddah, Saudi Arabia[2]
Kawther.kaabi@hotmail.com
fathi.essalmi@isg.rnu.tn
mohamed.jemni@fst.rnu.tn
aaqaffas@uj.edu.sa

***Abstract*—Nowadays, MOOCs is interesting for the success of learning through open educational resources (OER). However, the evolution of MOOCs has some challenges; the first one of them is related to the high dropout rate of students, which is caused by the lack of personalized interactivity. In this context, the personalization of MOOCS may help for increasing the retention rate. Therefore, this study examines how the personalization in MOOCs can affect students' expectations for staying connected to MOOCs. We propose a new approach that improves the retention rate in MOOCs by personalizing the content based on optimized selection of the personalization characteristics, based on a classification algorithm. This approach is validated by experiments to test its success when applied to many combinations of strategies and learners characteristics.**

Keywords-OER ;personnalization; classification; MOOCs; characteristics; retention rate.

I. INTRODUCTION

Recently, a movement of Open Educational Resources (OER), is witnessing the emergence of MOOCs around the world, mainly based on the notion of term connectivisim coined by [1, 2] as part of a network and the digital world. In addition, researchers are focused on open education resources and its impact on openness in education. With a particular focus on new trends, one the emerging phenomenon of massive open online courses (MOOCs), which are changing the fundamental foundations of education systems worldwide and forcing teachers to rethink the way teaching is currently exemption [3].

Despite the doubling of the number of MOOC users worldwide in recent times, with a total number of students enrolled in a course of more than 35 million according to statistics [3]. Also, the success of students has been defined in various ways, thus, increased retention becomes the goal of many quality assessments and institution improvement efforts in MOOCs. For this reason, the topic personalizing the MOOC to improve student retention rates is receiving the most research interest. This has motivated researchers and practitioners to emerge a new approach to solve the problem of student dropout. The personalization of MOOCs was introduced in the years after the use of courses, such as ranking and badges [4]. It then aroused significant interest in industry, the academic world with the sustained promise, motivation and promotion of the personalization [5, 6].

The goal of personalization in MOOCs is that increased the quality of learning in online education. And we focus on their problem of failure rates and drop-out. Research on these teachings showed that learners needed a personalized device and in particular a personalized follow-up. Similarly, initial MOOC observations indicate that about 10% of learners go to the end of the course [7, 8]. In particular, this study uses the parameters of personalization which are widely known and used in [9], and the purpose of learners' classification aims to increase retention rates in MOOCs and solve the massively problem of learners by regrouping them to similar groups. We apply the unsupervised classification algorithm K-Means to grouped learners according to similarities and differences in accordance with the parameters of personalization. The results of this algorithm allow to aggregate the data into coherent groups of learners.

The rest of the paper is structured as follows: Section 2 presents a literature review regarding the parameters personalization in MOOCs. Section 3 presents the challenges in MOOCs and the proposed prospects and impact of combining personalization parameters on learners of MOOCs. Section 5 presents the obtained results. Finally, section 6 concludes the paper with a summary of the findings and potential research directions.

II. LITERATURE REVIEW

A. Personnalisation in MOOCs

Our interest is to apply personalization parameters in MOOCs, forth more, in open education increases given its ability to capture and maintain the attention of students, which is retention for student success in educational environments. [10] They are focused to the low completion rate in MOOCs. Therefore, they conducted a study to examine learner engagement and disengagement patterns with the MOOCs course and, in turn, suggested that MOOCs offer adaptive content or assistance to learners based on their needs.

The personalization of learning in MOOCs arises with acuity considering the massiveness of the learners who

subscribe to it. Studies have identified practices to improve the quality of instruction in MOOCs, including support for personalized learning [11], but with no clear guidance on how to integrate them into a MOOC others describe more or less operational approaches to the design of adaptive MOOCs [12] and still others offer personalization solutions that have not yet been tested in MOOC platforms and that learners' data will be shared between organizations to allow personalization in MOOCs.

Others researchers described the adaptive in open learning as a process where learning content is delivered to learners in an adaptive way, appropriate content is delivered to learners appropriately: knowledge, preferences and other characteristics of learners [13]. Despite the apparent need for personalized learning in MOOCs, [14] points out that the existing MOOC courses are not even halfway through in the implementation of personalization. However, without personalization, learners can reduce their participation and eventually drop a MOOC course, which is one of the biggest concerns of MOOCs [15].

B. Parametrs of personnalization

Personalization parameters include a set of learners' characteristics. Like the model Felder-Silverman learning style (includes the learners' characteristics in four dimensions: Active/Reflective, Sensing / Intuiting, Visual / Verbal, and Sequential / Global).

In order to ensure the adaptive e-learning system according to his/her profile, a set of learner characteristics should be determined. According to [9] this is a constraint for providing effective e-learning experience and for rationalizing the personalization needs of the pedagogues, and its personalization parameters.

The learning personalization from the parameters that one proposes [5] parameters personalization defines some divergent characteristics the learning needs: e-learning style Felder [6] was the subject treated by the researchers for a long time ago. This problem was solved by the e-learning systems as an attempt to give the best support for the student. This systems e-learning is based on a set of personalization parameters. [7] They proposed a system that parameters personalization defines some characteristics the learning needs: e-learning style Felder, the four dimensions Learning Style (Sensing/intuiting, visual/verbal, Active/Reflective, Sequential/Global), level of knowledge ..., and our objective is to minimize complexity of parameters in order to achieve a strategy for the system.

III. CHALLENGES AND PROSPECTS

A. Challangesin MOOCs

The retention of students at the level of the course, program or degree has been a timeless concern of educators. For example, the following reasons were the main reasons for their abandonment according to [16] there is a lack of time and student support also motivation problem.

The goal of personalizing MOOCs is that in the early days of distance and online education drop-out and failure rates were very high. Research on these teachings showed that learners needed a personalized device and in particular a personalized follow-up. Similarly, initial MOOC observations indicate that high drop out of learners go to the end of the course [8, 9]. This type of situation covers situations that we can problematize in research. In this number of students who give up, there are of course curious. However, we think that these students are not the most numerous and that dropouts occur because the device is not suitable for these students. Thus providing a personalized device could help them maintain their motivation and learning

B. Prospects

We proposed an approach that improves the quality of retention in MOOCs by optimizing the selection of personalization parameters and combining an appropriate personalization strategy, based on a classification algorithm K-Means. The proposed approach selects the best strategy to use for personalizing a particular course according to two criteria: The selection of the learning objects: we choose the personalization parameters which have the highest number of learning objects that fits with their characteristics. Then, The classification of students for the assessment of the personalization parameter: we aim to reduce the total groups of learner's assessment. Thus, we get more precise results without getting the learner demotivated.

This approach aims to increase retention rates of learners in MOOCs by personalization of parameters, with a minimum effort of teacher and student for the appropriate personalization strategy to consider for a given course.

IV. PROPOSED APPROACH

A. Principe Approach

To allow a personalization in MOOCs to achieve optimal retention, we propose to set up the following personalization cycle: And the next figure 1 that presents the personalization cycle:

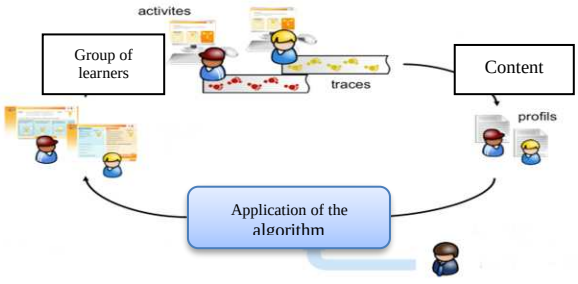


Figure 1 : Perasonalization cycle

The principle is simple: within a MOOC, learners will play games, during which all their actions will be traced. Throw to the traces generated by these interactions with the platform, a learner profile can be generated for everyone. The proposed approach in charge of the MOOC will for its part define a pedagogical strategy. This educational strategy will make it possible to automatically determine for each learner new activities and new paths, according to the information contained in his profile. Then, the cycle will be able to begin again, since new traces will be generated by the learners when they realize these games.
For each learner, the implementing application can thus build activities that correspond to its characteristics (learner profile) according to the teacher's vision (pedagogical strategy) and in the context of a given session.

B. Case of study

In order to help the teacher select the best combination of parameters for the personalization of particular course in MOOCs, we will present in this study first, an approach that analyses personalization parameters and extract the best one for the course. This approach is based on the k-means algorithm .~~And~~ in order to ameliorate the effectiveness of using games in learning, we have integrated learning strategies into gaming scenarios by applying unsupervised learning algorithm. In fact, this algorithm classifies students to organize what they have learned during the game learning.
Before starting the development of web application, we have to study first many works in the literature concerning the educational games to have an idea about how to implement and design this learning tool. Figure 2 summarizes the different steps followed to develop web application using.
The architecture of application is presented in the figure 2. It is composed by the following elements:
Platform interface (MOOCs): It is the unique element to which the Learners can get access. It is a simple interface that doesn't contain any complicated instructions.
Traces of learners: the Database stores all the information about the traces of learners and his personalization parameters and users.
K-Means Algorithm: all the results are generated in this part of the system by apply the k-means algorithm. It request information from Database when it's needed.
Classification learners in similar groups: The K-means algorithm applied for students' classification based on their traces to get students in similar groups.

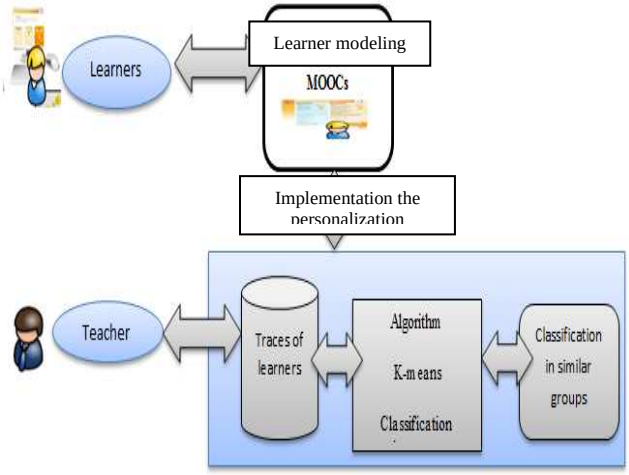


Figure 2: Personalization in MOOCs

In this study we presented an approach that improves the quality of retention in MOOCs by optimizing the selection of personalization parameters and combining an appropriate personalization strategy, based on a classification algorithm. The proposed approach selects the best strategy to use for personalizing a particular course according to two criteria: The selection of the learning objects: we choose the personalization parameters which have the highest number of learning objects that fits with their characteristics. Then, The classification of students for the assessment of the personalization parameter: we aim to reduce the total groups of learner's assessment. Thus, we get more precise results without getting the learner demotivated.

V. METHOD

In this paper, we present the design game and vision of the system implementation be discussed by getting a clear idea of what the goals and components should be and what methods will be used to achieve these goals. In order to help the teacher select the best combination of parameters for the personalization of particular course, we propose in this study, an approach that classifies students through possible combination of personalization parameters. Furth more, we conduct experiments in order to validate our assumptions; we validate it by designing a process of system which supports the teacher's decision about the appropriate personalization strategy to consider for a given course in MOOCs. We validate it by a web application system called system for optimization the strategy personalization parameters which supports the teacher's decision about the appropriate personalization strategy to consider for a given course.

To evaluate the effect of personalization MOOCs on students' perceived preferences for web application, an experiment was conducted in a secondary school in Tunisia. The students used a newly developed personalization MOOCs for 3 days to learn database courses. Personalization in MOOCs is considered as one of the most suitable environments for learning because it provides games. In this

context, the games elements are implemented in our MOOCs. For instance, each time the students finish a game, they traces will be saved in database. This learning with game can make them keep track of their progress and be motivated to go to the next learning, they will be more active.

A. *Participants*

The participants in this study were secondary school students specializing in computer science and they were all between 15 and 18 years of age. In addition, all of these participants have not previously used the MOOCs and this is an opportunity to find out.

B. *Procedure*

In the beginning, the instructor gives a short introduction to the students concerning the purpose of this experiment and the purpose of using the MOOCs, since they are the first time to use. In particular, we have explained the role of each element of the game and how it is implemented. This helped the learners to understand the purpose of the experiment. They then answered the FSLSM questionnaire to determine their personalization parameters. After that, students used the MOOCs to learn the computer course in game mode, this helped them learn and more active. Finally, the students answered a questionnaire and they play games.

VI. RESULTS

To validate this work, learners have participated in the experiment to test their knowledge and level of satisfaction using different technologies such as the technology acceptance modem (TAM), pre and post-test, and questionnaires.

As a result, this designed learning tool is very efficient when it comes to learning and keeping learners satisfied while learning. For validating the system created, a number of learners participated in an experimentation to test their level of satisfaction while using the system by answering to the Technology Acceptance Model (TAM). The result obtained showed that this system could really help users in the personalization process. As a conclusion, this designed web system is very efficient to select the personalization strategy, to reduce the teacher's efforts of creating extra learning objects and to reduce the time spent and for a principle purpose to increasing retention rate for learners.

VII. DISCUSSION

This section examines the results obtained affirmed that learners have a positive attitude towards using versions of learning a course with played. These results are similar to the results of recent studies [17,18]; by personalize learning games according to the characteristics of the learners.

The results obtained and their similarities to other research work concerning the preferences for using the Web application, including open educational games for regrouping the diversity of learners.

The uses of personalization parameters and the k-means algorithm are implemented to classify students and consequently reduce the problem of massive learners. This allows using parameters of personalization in massive open courses to increase the quality of learning in MOOCs.

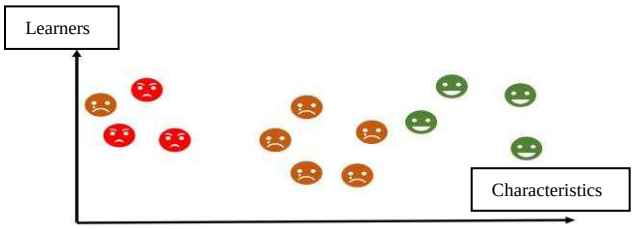


Figure 3: Regroupement of learners by K-Means Algorithm

In figure 3 we visualize that the learners are grouped with the similarity of their characteristics. These results aim to optimize personalization strategies are aided for learning success, to improve the retention rate, we will solve those problems by grouping learners.

VIII. CONCLUSION

This study investigated the impact of different personalization parameters on learners' perceived preferences for increasing their retention within personalized MOOCs. The obtained results, collected from students from a public secondary school in Kairouan from Tunisia, showed that learn with play in MOOCs may affect students' more engagement with it.

The findings of this study for a strategic personalization is offered to individualize interactions with students; the pedagogical conditions likely to favors the personalization of learning in different contexts related to the control of the learner as well as his training and his learning. In addition, optimized personalization strategies in MOOCs are helped for the success of learning, in order to help the teacher to select the best combination of parameters for the personalization of a particular course. We presented an approach that analysis and give a solution for every possible combination of personalization parameters for to find optimal retention in MOOCs, that for extracts the best one for the course.

This study presented a solution to the central question: How to optimize the retention rate of learners in MOOCs according to parameters of personalization? A process is presented for the personalization of an e-learning system together with an evaluation of personalization parameters, and architecture design of personalized learning strategy. By aims to design a web-application tool for assisting the teacher's decision about the personalization strategy to use for his/her course. To achieve our main goal, various steps were taken into considerations which are:

- Extract the meaning of the concept e-learning in particular, the adaptive learning in MOOCs.

- Make a literature review in e-learning personalization systems with identifying the used personalization parameters for each one.
- Extract the meaning of the retention rate in MOOCs with identifying some of the existing methods of solving optimization problems in the field.
- Create an approach based on classification algorithm for the teacher to guide his/her decision about the personalization strategy to use for his/her course.

REFERENCES

[1] Conole, G. (2014). The use of technology in distance education. In Anderson, T. & Zawacki---Richter, O. (Eds.), OnlineDistanceEducation:Towardsaresearchagenda. AU Press, Athabasca University.

[2] Bell, F. (2011). Connectivism: Its place in theory---informed research and innovation in technology---enabled learning. InternationalReviewofResearchinOpenandDistanceLearning, 12(3), 98---118.

[3] Jemni, M, Kinshuk and kouthair khribi : Provides a comprehensive overview of open educational resources and MOOCs, as well as their emerging pedagogical perspectives Open Education Ressources ; From OERs to MOOC , Springer-Verlag Berlin Heidelberg ,(2017)

[4] https://www.class--central.com/report/moocs--2015--stats

[5] Essalmi, F., Jemni Ben Ayed, L., &Jemni, M, (2007).A multi-parameters personalization approach of learning scenarios. In The 7th IEEE international conference on advanced learning technologies, Niigata, Japan, 90–91.

[6] F. Essalmi, L. J. B. Ayed, M. Jemni, S. Graf, et al.(2015).Generalized metrics for the analysis of e-learning personalization strategies. Computers in HumanBehavior, (310–322).

[7] R. M. Felder and L. K. Silverman, (1988).Learning and teaching styles in engineering education. Engineering education, vol. 78, (pp. 674–681).

[8] Khalil, H. & Ebner, M. (2014). MOOCs Completion Rates and Possible Methods to Improve Retention - A Literature Review. In Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2014 .

[9] Essalmi, F., Jemni Ben Ayed, L., Jemni, M., Kinshuk, Graf, S, (2010). A fully personalization strategy of E-learning scenarios. Computers in Human Behavior 26(4), 581–591 (2010)

[10] Mystakidis, S., & Herodotou, C. (2016). OpenQuest: Designing a Motivational Framework for MOOCs Instruction. in Europe, 141.

[11] Vaibhav, A., & Gupta, P. (2014, December). Gamification of MOOCs for increasing user engagement. In MOOC, Innovation and Technology in Education (MITE), 2014 IEEE International Conference on (pp). IEEE.

[12] Kizilcec, R. F., Piech, C., and Schneider, E. (2013, April). Deconstructing Disengagement: Analyzing Learner Subpopulations In Massive Open Online Courses. In *Proceedings of the third international conference on learning analytics and knowledge*: pp. 170-179. ACM.

[13] Alaa Qaffas, Lei Shi, Alexandra I. Cristea , Jonathan G. K. Foss and Dana Al Qudah : Un environnement d'apprentissage électronique adapté et social: une étude de cas dans Topolor, (2013).

[14] Henning J, Morton J, Hla T, Meers J. Mortality rates adjusted for unobserved deaths and associations with Newcastle disease virus serology among unvaccinated village chickens in Myanmar. Prev Vet Med. 2008;85:241–252. doi: 10.1016/j.prevetmed.2008.01.014

[15] Bagnol B. Good Practices for Family Poultry Production. Advocate Gender Issues: A Sustainable Way to Control Newcastle Disease in Village Chickens. International Network for Family Poultry Development. GPFPP Note No 03. Rome, Italy: Food and Agriculture Organization of the United Nations; 2012.

[16] Kay, J., Reimann, P., Diebold, E., Kummerfeld, B.: MOOCs: so many learners, so much potential. IEEE Intell. Syst. 28(3), 70–77 (2013)

[17] Stevanović, N. (2014) Effects Of Motivation On Performance Of Students In MOOC. SINTEZA 2014 Internet and Education. DOI:10.15308: pp. 418-422

[18] R. M. Felder and L. K. Silverman, (1988).Learning and teaching styles in engineering education. Engineering education, vol. 78, (pp. 674–681).

[19] Frankola, K. (2001). Why online learners dropout. Workforce, October 10, 53-63. [viewed19/11/2005]

[20] Mohamed Ali Khenissi, Fathi Essalmi, Mohamed Jemni , (2012). Toward the personalization of learning games according to learning styles.

[21] Mohamed Ali Khenissi, Fathi Essalmi , Mohamed Jemni , Kinshuk ,Sabine Graf , Nian-Shing Chen ,(2016).Relationship between learning styles and genres of games. Computers in Human Behavior.

Use of Machine Learning techniques in Financial forecasting

M. Becha, O. Dridi, O. Riabi, Y. Benmassaoud

Abstract— Financial institutions play a crucial role in the acceleration of the economic stability and sustainability. In fact, one of their major sources of gain consists of debts provided to their customers in the form of loans. However, the risk that this service carries can lead to financial instability within these institutions. To avoid such incidents, an adequate credit approval process is needed through different approaches to forecast and assess delinquency and avoid falsification. The aim of this paper is to deal with credit risk management through the development of a model that classifies potential borrowers as good or bad credits. Furthermore, the paper reflects the importance of using machine learning techniques in the financial field. Historical observations of good and bad credits for a number of customers based on mixed attributes were used to train and test our data by carefully applying data preparation techniques and analysis (Univariate, Bivariate using ANOVA and Chi-Square tests), and KNN algorithm. Consequently, an accuracy test led us to an optimal K based on significant independent variables.

Index Terms—Financial forecasting, credit risk management, machine learning

1 INTRODUCTION

One of the major players that is closely related to economic growth worldwide is the financial institutions and the services that they provide. In fact these institutions play a big role as intermediaries which is considered as an accelerator of economic stability and sustainability.

Therefore, the stability and sustainability of financial institutions themselves is considered as one of the main factors to economic growth.

Based on the fact that debts are considered the major group of financial institutions' services and are a major source of gain for each institution as well, it is important to understand that credit risk has an important influence and requires a special concentration. In fact, Credit risk management is a major determinant of an institution's performance including profitability and sustainability. Consequently, it is a systemic issue which requires special attention from each institution that deals with credit giving as it causes an important threat to the process of reaching the goals of the institution [2]

In fact, Credit creation is the main income generating activity for banks. However this also carries many major risks for both parties: the borrower and the lender. For example the fact those banks with high credit risk have high bankruptcy rates which discourage customers to make deposits. On the other hand the risk related to the inability of the borrower to fulfil his obligations and paying back their loan, which is the part that we will be focusing on in our project.

Through effective and proper credit risk management, banks not only support the sustainability and profitability of their operations but also positively contribute to the economy's growth and efficient allocation of capital within it. In fact, these institutions have structured and well-studied credit approval process which is personalized from one lender to another includes a well-established procedure for comprehensive credit appraisal.

Weak credit risk management, however, will cause an important increase in the level of non-performing credits leading to financial instability and negative variations in the health of the loan portfolio which in turn the institution's performance.

Consequently, more accurate assessments and forecasts of delinquencies and defaults reduce the likelihood of such false incidents. Finally, given the size of this part of the industry more accurate forecasts would improve macro prudential policy decisions, and reduce the likelihood of a systemic shock to the financial system.

Different approaches to this issue have been made, however, one of the most powerful and efficient ways to avoid defaulted and sleeping accounts is the prediction of the ability of the customers to meet the expectation and manage their accounts efficiently and provide return to the bank. In fact by assessing whether an individual is a good candidate or not, based on a number of attributes, we can filter out of the applicants that are considered a bad match and only keep the credit worthy individuals, reducing significantly the level of delinquency, and increasing the revenue of financial institutions.

In our paper we will be dealing with the credit card

- *M. Becha, Tunis Business School, University of Tunis, Tunisi. E-mail: malikabecha@outlook.fr.*
- *O. Dridi, Department of IT, Tunis Business School, University of Tunis, Tunisia. E-mail: olfa.dridi.bz@gmail.com*
- *O. Riabi, Tunis Business School, University of Tunis, E-mail: riabi.oumayma@gmail.com*
- *Y. Benmassaoud, Tunis Business School, University of Tunis, E-mail: yasmine.bnm5@gmail.com*

***Please provide a complete mailing address for each author, as this is the address the 10 complimentary reprints of your paper will be sent*

*Please note that all acknowledgments should be placed at the end of the paper, before the bibliography (**note that corresponding authorship is not noted in affiliation box, but in acknowledgment section**).*

risk management through the classification of potential customers as good or bad fits.

2 PROPOSED APPROACH

We will classify the creditworthiness of potential borrowers for a sample of 1000 individuals based on a group of attributes that describe the lender characteristics that influence their ability to pay back the loan and avoid default and delinquency.
The performance indicator in this project is the class of the potential borrower (Good Credit/Bad credit). In order to determine this classification, we will be opting for the K-nearest neighbor algorithm (KNN) by dividing the dataset into a training dataset and test data set. The use of KNN is based on the fact that it is a simple yet effective method of classification based on majority voting of the selected neighbors given the value of k.

2.1 Data collection

In order to apply the classification analysis, we have opted for a credit card risk management related database that deals with 1000 observations, based on 20 attributes divided into categorical and numeric variables, based on which a classification of the credit worthiness of each individual is forecasted and evaluated.
The data set was taken from [1] for one financial institution' potential borrowers' records.

2.2 Feature selection

In order to better understand the data provided in the dataset, we will be using:

1. Frequency Table and histograms to identify the different classes and the distribution of the categorical variables.
2. Measures of central tendencies to determine the distribution of numerical variables.

We are concerned in building a classifier that performs an accurate classification task. This section aim to determine the variables that have an impact on the dependent variable.
We perform a bivariate analysis to select the attributes that have a significant relationship with the dependent variable, and to check whether a correlation exists between the attributes themselves. If so, we decide either to drop one of them or to create a new variable that summarizes the information between the correlated variables.
As we have variables of a different type, we use different significance tests.

Correlation Coefficient: is a statistical measure that calculates the strength of the relationship between the relative movements of two numerical variables.

Chi Square Test: It is an independence test that evaluates whether or not two categorical variables are related. The frequency of each category for one nominal variable is compared across the categories of the second nominal variable.

Null hypothesis:
X and Y are independent: $\mu_X = \mu_Y$
Alternative hypothesis:
X and Y are dependent: $\mu_X \neq \mu_Y$

ANOVA Test: It tests the significance between the dependent variable and the independent variable. It considers the variation of the numerical variable means among the different categories of the categorical variables.
Null hypothesis: There is no significant difference among the groups.
The alternative hypothesis: there is at least one significant difference among the groups.

2.3 Model Fitting

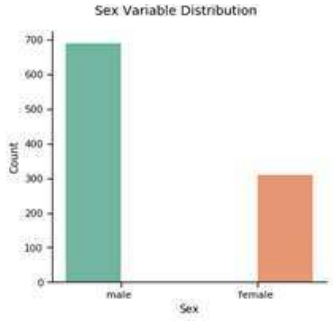
Before fitting the model, a common practice is to split the historical available date for the credit risk into two subsets, naming the training set and the test set. While the training set contains the records to be used to estimate each individual's class and it is counted for 99% of the random observations, the test set serves to evaluate its accuracy which will be applied for the rest 1% percent of the dataset.
The accuracy test should provide a reliable indication of how well the model is likely to perform classification on new data.

3 DATA PRE-PROCESSING

The data set contains 1000 records with respect to 16 variables. We have 5 quantitative variables and 11 qualitative ones. The data summarizes bank information (e.g. Credit history, Checking account), loan information (Credit amount, Purpose), and socio demographic information (Age, Sex) of the borrower.

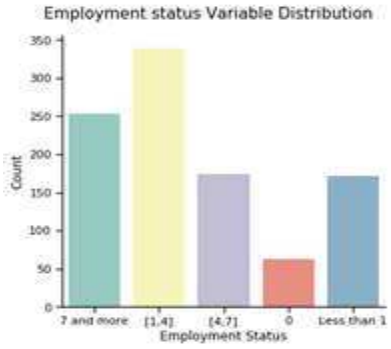
Univariate Analysis:

Sex variable: 69% of the customers we have in the dataset are male (690 out of 1000)

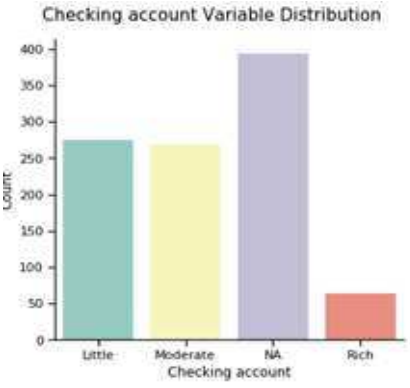


Employment Status Variable: Employment status is a four-class categorical variable. The following histogram indicates that the

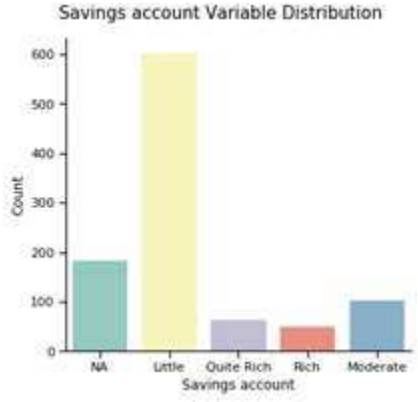
majority of the customers who are currently working have been employed from 1 to 4 years.



Checking Account Variable: Following is the distribution of the checking account variable. 39.4% of the customers do not have a checking account. 27, 4% are considered to have a little amount in their checking accounts. 26, 9% belongs to the moderate class as they have an amount between 0 and 200 in their checking account.



Savings Account Variable: The majority of the customers (60, 3%) tend to have less than 100 in their savings account.



Purpose Variable: The purpose is a categorical variable with 10 classes. The above frequency table illustrates that 28% of the customers are willing to get the loan in order to have a new radio or TV. However, 23, 4% want to buy a new car.

Credit History Variable: Credit history variable has 5 classes. 53% of the customers have existing paid credits. However, 29, 3% have existing credits in other institutions that are not yet paid.

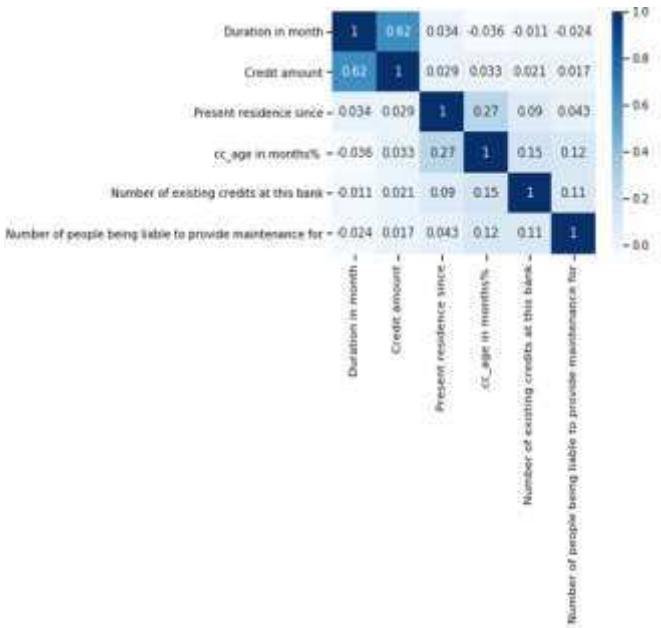
Following are the central tendency measures related to the numerical variables. The average amount of loan is 2,319 however the maximum amount of load equals 18,424. The average age of the credit cards equals 33 months.

The customers have on average 1 existing credit at that same bank. The mean duration of the credits is 18 months

	Credit amount	cc_age in months%	Number of existing credits at this bank	Duration in month
count	1000.0	1000.0	1000.0	1000.0
mean	3271.3	35.5	1.4	20.9
std	2822.7	11.4	0.6	12.1
min	250.0	19.0	1.0	4.0
25%	1365.5	27.0	1.0	12.0
50%	2319.5	33.0	1.0	18.0
75%	3972.2	42.0	2.0	24.0
max	18424.0	75.0	4.0	72.0

Bivariate Analysis:

Following is the correlation matrix (heat map) of the numerical independent variables. We notice that there are no strong correlations between these variables. The correlation is weak and insignificant for most of the variables.



Y vs Credit Amount: for the Anova output between the dependent variable and credit amount variable, the null hypothesis of this test states that the credit amount means are equal, which means there is no significant relationship between the Good/Bad class and the amount of the credit.

P-value = 0.0000009 is much lower than our significance level 5%. Meaning that we have enough evidence to reject the null hypothesis. Therefore, ther is a significant relationship between Y and Credit amount.

Y vs Duration in month: The null hypothesis states that average value if the duration in month is the same in the good and the bad classes. We have P-value lower than our significance level. Therefore, there is a significant relationship between Y and the duration in months.

Y vs number of existing credit: We performed the Anova test to test for the significance of the relationship between Y and the number of existing credits in the same bank. The P-value is higher than 5%. We fail to reject the Null hypothesis, there is no significant relationship between Y and the number of existing credits in the same bank.

*Y vs number of People being liable to provide maintenance for & Present residence since:*P-value is higher that our significance intervals. So, we fail to reject the null hypothesis, so we are 95% sure that:

- There are no relationship between Y and the number of people being liable to provide maintenance for.
- There are no relationship between Y and Present residence since variable.

Y vs number of cc age in month: The test is testing the variation of the mean age In month among the two classes of the dependent variable. The P-value is lower than 5%. So, we reject the null hypothesis. There is a significant relationship between Y and the credit card age in months.

4 The KNN algorithm

In order to apply the KNN algorithm on our dataset we divide the latter into a training seta and a test set.
Trainings Set: This set includes 700 observations of potential borrowers classified as "Good" or "Bad" credits.
As the information that we have for each potential borrower includes both qualitative and quantitative variables, we have transformed the dataset into the following by turning categorical attributes into ordinal one so that we can apply our algorithm.
Test set: Composed of 300 observations. This is an example of the test set composed of unclassified observation which we are going to determine the classes that each one of them belongs to:

5 Results and Interpretations

First, we calculated the distance between each training row and testing row using the Manhattan method. Then, we identified the nearest neighbors depending on the value of K chosen. In our case we opted for K values that are odd, which is due to the fact that using pair nearest neighbors to forecast the class of a given example can lead to a case of confusion of the class' result. We have opted for 4 values of k: 3,5, 9 and 11. Accuracy test was conducted to evaluate how well our data was classified. Different K neighbors yielded to the following accuracy results:

- 3NN showed the classification accuracy equal to: 63%

```
In [53]: KNN_Class(3)
[[ 31  62]
 [ 48 159]]
              precision    recall  f1-score   support

         bad       0.39      0.33      0.36        93
        good       0.72      0.77      0.74       207

    accuracy                           0.63       300
   macro avg       0.56      0.55      0.55       300
weighted avg       0.62      0.63      0.62       300
```

- 5NN showed a classification accuracy equal to: 67%

```
In [54]: KNN_Class(5)
[[ 35  58]
 [ 41 166]]
              precision    recall  f1-score   support

         bad       0.46      0.38      0.41        93
        good       0.74      0.80      0.77       207

    accuracy                           0.67       300
   macro avg       0.60      0.59      0.59       300
weighted avg       0.65      0.67      0.66       300
```

- 9NN showed a classification accuracy equal to: 70%

```
In [65]: KNN_Class(9)
[[ 33  60]
 [ 31 176]]
              precision    recall  f1-score   support

         bad       0.52      0.35      0.42        93
        good       0.75      0.85      0.79       207

    accuracy                           0.70       300
   macro avg       0.63      0.60      0.61       300
weighted avg       0.67      0.70      0.68       300
```

- 11NN showed a classification accuracy equal to: 72%

```
In [66]: KNN_Class(11)  #BEST
[[ 33  60]
 [ 24 183]]
              precision    recall  f1-score   support

         bad       0.58      0.35      0.44        93
        good       0.75      0.88      0.81       207

    accuracy                           0.72       300
   macro avg       0.67      0.62      0.63       300
weighted avg       0.70      0.72      0.70       300
```

In order to create an efficient and effective classification and thus forecasting model for the credit risk management for financial institutions, we opt for a value of k that shows the highest level of efficiency which is in our case k = 11.

6 Conclusion

The classification of the potential borrowers to good or bad credits provides financial institutions with an assessment of the creditworthiness of their clients. This is one possible measure to use in order to prevent defaulted and delinquent accounts and maintain the sustainability of the institution through increasing its growth and profitability thus the stability of the economy.

With a high accuracy our classification allowed us to classify the test sample efficiently for an optimal K= 9 based on the most significantly impactful variables including : The credit amount, Duration in months, number of people being liable to provide maintenance for, duration at present residence, credit history, purpose of the credit, savings account, employment status, housing status, property (Collateral).

Since in our case we are classifying individuals as bad or good credits, it is important to note that we opt for a prudent classification, meaning that for a given observation, when in doubt, it is more fitting to classify it as a bad credit than a good credit. This is mainly due to the fact that the risk of the default outweighs the opportunities and the advantages related to a possible good credit. Here we can say that the financial institutions are risk averse.

Although the KNN is one method to use in order to determine, however it still considered as a lazy learner, meaning that it doesn't actually learn from the training set but uses the training set itself to determine the class of the test set. This can cause a problem of generalization, not being robust to noisy data, in addition to the fact that changing the level of k changes the resulting prediction of the class. Which is why using the decision tree for this case can be a more effective way to classify new observations although it can be a bit time consuming. Also, since KNN algorithm lazy learning method it is challenging to apply it in many relatively complex applications such as dynamic web mining for a large repository. To solve this problem, building a faster KNN model is a very interesting approach that is based on using cosine similarity as distance approach, as it is fast, simple, and has a comparatively better accuracy than other distance metrics.

References

[1] The data set http://www.Kaggle.com

[2] I. Bose, R.. K. Mahapatra , "*Business data mining – a machine learning perspective", Information &Management,* vol. 39, no 3 pp.211-225, (2011)

[3] C. Jiang, Z. Wang, R, Wang, Y. Ding, "*Loan default prediction by combining soft information extracted from descriptive text in online peer-to-peer lending", Annals of Operations Research,* vol. 266, pp. 511-529, (2018)

Etat de la recherche sur la veille stratégique et l'intelligence économique dans le milieu académique Tunisien

Mots clés

Etat des lieux, recherche, Veille stratégique Intelligence économique gestion des connaissances, technologies numériques.

Résumé

En faisant l'État des connaissances dans le domaine de la veille stratégique et l'intelligence économique, nous nous proposons de rechercher le maximum d'informations disponibles concernant ce domaine et à en faire une synthèse. L'objectif de ce travail est de combler un vide auquel font face les chercheurs académiques, intéressés par des problématiques liées à ce champ de recherche. La rareté des méta-analyses de la revue de littérature et des études rétrospectives constituent un obstacle à la cumulativité de la recherche et à la capitalisation des connaissances actionnables issues des différentes investigations.

Nous constatons souvent, que la plupart des chercheurs, se rabattent sur les démarches exploratoires, prétendant, et souvent à tort, que le phénomène étudié est nouveau. Mais la réalité est que les recherches antérieures ne sont ni suffisamment visibles ni assez bien documentées pour permettre de progresser dans le confirmatoire. La rupture perçue dans cette dynamique de la recherche pourrait laisser supposer que la Veille Stratégique (VS) et l'Intelligence Economique (IE), n'ont pas réussi à s'imposer en Tunisie comme un axe fort pour l'amélioration de la compétitivité des produits tunisiens. Le constat partagé est celui de l'effritement de la recherche locale et l'absence d'une communauté homogène capable d'offrir aux praticiens une alternative sérieuse à des recettes d'experts venus d'ailleurs.

Pour apporter des éléments de réponse à ces suppositions, nous procéderons comme suit :

Nous définissons les deux concepts clés du titre, ensuite nous dressons une rétrospective de la recherche publiée par des auteurs tunisiens depuis 1996 dans les proceedings de deux colloques internationaux spécialisés VSST (Veille Stratégique et Technologique) et SIIE (Systèmes d'Information et Intelligence Economique). Cette base de données est exploitée pour analyser la situation de la recherche en matière de VS et IE en Tunisie, et ce, en nous basant sur les problématiques récurrentes, les résultats et voies futures dégagées par les chercheurs. Le classement des recherches recensées dans le cadre de cette étude rétrospective, selon leur ancrage disciplinaire a pour but de permettre d'établir une cartographie permettant de mieux visualiser les orientations et préoccupations sociales des chercheurs dans ce domaine et pourquoi pas, en faire une représentation commune pouvant aider à donner du sens à ce qui se fait. L'éventuelle émergence d'axes interdisciplinaires fédérateurs serait une contribution pertinente de ce travail. Ces résultats préliminaires, basés sur une analyse de contenu, feront l'objet d'une consultation d'experts pour validation.

1) **Les concepts clés**

La Veille stratégique comme pratique de surveillance des environnements immédiats et lointains au niveau micro économique est un système d'information externe de l'entreprise.

Les informations pertinentes pour l'aide à la prise de décision stratégiques ne sont pas les informations structurées produites par les unités internes mais celles de type « signal faible » (Igor. Ansoff, 1975) qui alerte sur des ruptures ou changements pouvant générer des menaces ou des opportunités pour une entreprise. Comme l'a fait remarquer F. Aguilar (1967), ce sont ces « bribes » d'informations qui arrivent aux décideurs par différents canaux notamment oraux qui ont de la valeur pour l'action stratégique.

Aujourd'hui, grâce aux progrès du numérique, ce type d'information est devenu plus disponible et mieux accessible. Les Big Data constituent un gisement précieux pour les décideurs désireux d'améliorer leur positionnement stratégique soit en local, régional ou à l'international.

La Business intelligence analytique et l'Intelligence artificielle faible (pour l'instant), constituent autant d'outils susceptibles de faciliter l'accès à l'information stratégique et son déploiement à des fins de compétitivité.

Au niveau macro-économique, l'intelligence économique dont la finalité première est la protection du patrimoine national, et sa promotion au niveau international, peut aussi, dans le cadre d'une stratégie du numérique, capitaliser sur les progrès réalisés dans ce domaine pour favoriser l'accès des entrepreneurs aux informations et connaissances utiles pour leur pérennité. Le patriotisme économique Allemand et le sens japonais de la primauté de l'intérêt national ont fait de ces deux pays les leaders de la robotique et de l'IA faible.

Depuis les travaux de Jacobiak (1999), la veille scientifique et technique ou veille technologique associée à la R&D et donc à l'innovation industrielle, a évolué dans plusieurs directions selon les ancrages disciplinaires des laboratoires de recherche qui en ont fait leur axe de recherche. Tous ces laboratoires se retrouvent autour de l'information comme ressource fondamentale : les ingénieurs des structures de R&D dans l'industrie, les chercheurs en sciences de l'information, les chercheurs en sciences de gestion (gestion des ressources immatérielles), et les informaticiens.

Les objets de recherches, différent selon les objectifs, et les pratiques de veille et IE sont analysés sous l'angle de la discipline d'appartenance du chercheur :

1. L'ingénieur industriel serait attentif à la traduction de l'innovation tout au long du processus de veille technologique ;
2. Les spécialistes des sciences de l'information, accorderaient une attention particulière aux sources et canaux de transmission de l'information ;
3. Les gestionnaires, se placeraient sur l'une ou l'autre dimension de leur discipline : la dimension organisationnelle et managériale ou la dimension artéfactuelle ;
4. Les informaticiens privilégieraient l'aspect ingénierique et artefactuel (analyse des données, conception et modélisation des systèmes d'information).

Un autre concept lié à l'information, unit les chercheurs de ce domaine. Il s'agit du concept d'incertitude environnementale perçue (Daft, R.L. et Lengel R.H., 1984).

L'incertitude peut correspondre à des situations de surabondance d'informations, d'insuffisance des informations disponibles ou d'ambiguïté.

En focalisant leur attention sur les sources susceptibles de leur permettre de combler leur déficit informationnel, les dirigeants sont supposés adopter un comportement informationnel qui consiste à surveiller les sources de changements en vue de collecter, traiter et exploiter les informations pouvant les aider à rendre intelligible leur environnement : « l'information est utile pour diminuer l'incertitude dans la prise de décision » (E. Pateyron, 1998 : 19), elle permet de réduire l'ignorance et l'ambiguïté qui entourent souvent le processus de décision.

2) Le contexte tunisien

La Tunisie est un pays considéré en voie de développement et en éternelle transition politique et économique. Cette situation et de nature à générer une forte incertitude environnementale perçue. Mais le statut de consommateurs et pourvoyeurs de data gratuites du numérique, de laissés pour compte de l'industrie de la data ne permet pas à ces pays de profiter comme il se doit des opportunités que les surpuissances mondiales GAFAM (Google, Apple, Facebook, Amazone Microsoft) et BATX (Baidu, Ali Baba, Tencent et Xiaomi) ont capitalisées en leur faveur grâce à leur dimension visionnaire. L'Europe, mais à un degré moindre, ne parvient pas non plus avec les CNILS (Commissions Nationales Informatique et Liberté) à faire émerger une industrie européenne de la data (L. Alexandre (2017), alors que la réduction de l'incertitude environnementale perçue dans un monde globalisé, ne semble plus possible aujourd'hui sans stratégie dans l'industrie de la data. La limite de l'attention humaine trop sollicitée de nos jours, face à la multiplicité des sources de data rend incontournables le recours aux technologies numériques : les medias (images,vidéos, audios, podcasts, médias sociaux, comme facebook, twitter, youtube, instagram) les clouds publics et privés, le Web, l'IOT, les bases de données traditionnelles et modernes. Faute de compétences en matière de littératie informationnelle, l'incertitude environnementale ne peut que s'accentuer.

Mais que révèlent les recherches menées dans le contexte tunisien en matière de pratiques de VS et d'IE ? Existe-t-il au niveau des institutions publiques et/ou au niveau des entreprises un usage courant de ces sources et/ou une conscience des enjeux d'une possible sous-utilisation des technologies numériques ?

A partir d'une cartographie de la recherche en VS &IE en Tunisie, nous concevrons le questionnaire destiné aux chercheurs/experts identifiés à travers l'analyse de contenu. La méthode Delphi qui sera mobilisée recommande de ne pas aller au-delà de 18 experts (Paliwoda, 1983, cité par Okoli & Pawlowski, 2004). Nous veillerons cependant à faire en sorte que les quatre disciplines mentionnées ci-haut soient représentées au prorata du volume de la recherche publiée et recensée à l'étape de l'analyse de contenu.

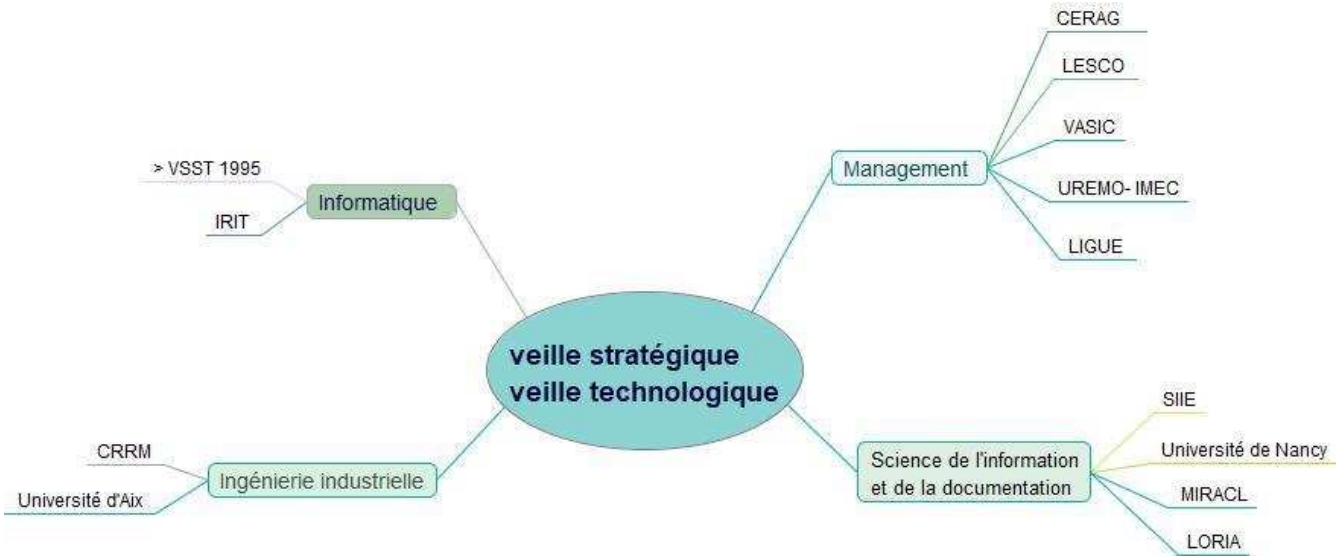


Figure 1 cartographie

3) Méthode de collecte des données

Lors du processus d'exploration nous sommes basés sur des données collectées à partir des communications publiées dans le cadre de la conférence SIIE, des communications publiées lors du séminaire VSST, des revues scientifiques, des mémoires de recherche et des thèses, en tenant compte des critères suivants : la nationalité tunisienne des auteurs, leurs affiliations dans des laboratoires de recherche tunisiens et des problématiques émergentes dans un contexte tunisien.

Toutefois, la base de données collectée est le résultat d'une recherche individuelle qui ne peut prétendre à l'exhaustivité. D'ailleurs, nous avons été confrontés à l'absence et/ ou à l'insuffisance de certaines données. En premier lieu, l'indisponibilité de toutes les communications publier lors de la conférence SIIE et VSST, qui se sont limitées seulement à quatre éditions de SIIE à savoir les éditions des années 2008, 2009, 2010,2012, et seulement cinq éditions du séminaire VSST soient les éditions des années 1998, 2008, 2001,2004,2007et 2010. En second lieu, l'absence d'une base de données recensant des thèses et des mémoires soutenues. Notre collecte s'est basée seulement sur les mémoires et les thèses disponible à la bibliothèque universitaire de l'école supérieure de commerce. En somme nous avons pu obtenir 57 références réparties sur quatre sources à savoir : SIIE, VSST, revues scientifiques, mémoires et thèses

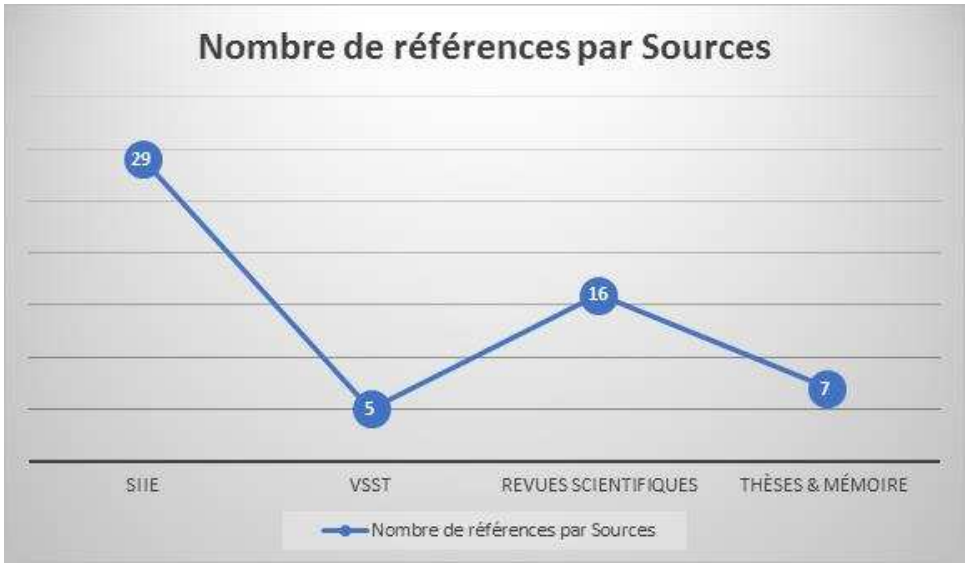


Figure 2 : Nombre de référence par source

L'écart entre le nombre des références collectées par source n'est pas dû à un hasard, cela peut être expliqué tout d'abord par la disponibilité et par l'accessibilité aux données comme mentionné ci-dessus, et par la nature de la source. Par exemple le nombre le plus élevé des publications recueillies est celui lors des conférences SIIE (29 publications), cela peut être expliquer tout d'abord par le positionnement géographique de cette conférence : la Tunisie ce qui encourage les chercheurs tunisiens à y participer et à publier, ensuite par le nombre des éditions réalisées : la conférence est à sa huitième édition, et enfin le spectre élargi de la conférence qui touche plusieurs domaines, ce qui donne la possibilité aux chercheurs de différents disciplines d'y participer.
Par contre, le domaine de recherche de séminaire VSST se base strictement sur la veille stratégique, scientifique et technique. Notamment VSST est une conférence qui n'a jamais eu lieu en Tunisie elle se tient souvent à son siège situé à Toulouse, et elle s'est déroulé une fois au Maroc, ce qui explique la dominance des publications françaises et le nombre réduit des publications tunisiennes.

4) Les articles par domaine

Grâce à la collecte des données primaires, soumise à un traitement automatique à l'aide de NVIVO 11 sur les 57 Références collectées, nous avons pu les repartie sur les 4 domaines mentionnés ci-haut dans la cartographie à savoir : le management, l'informatique, l'ingénierie industrielle, et les sciences de l'information et de la documentation
La figure suivante permet de visualiser cette répartition :

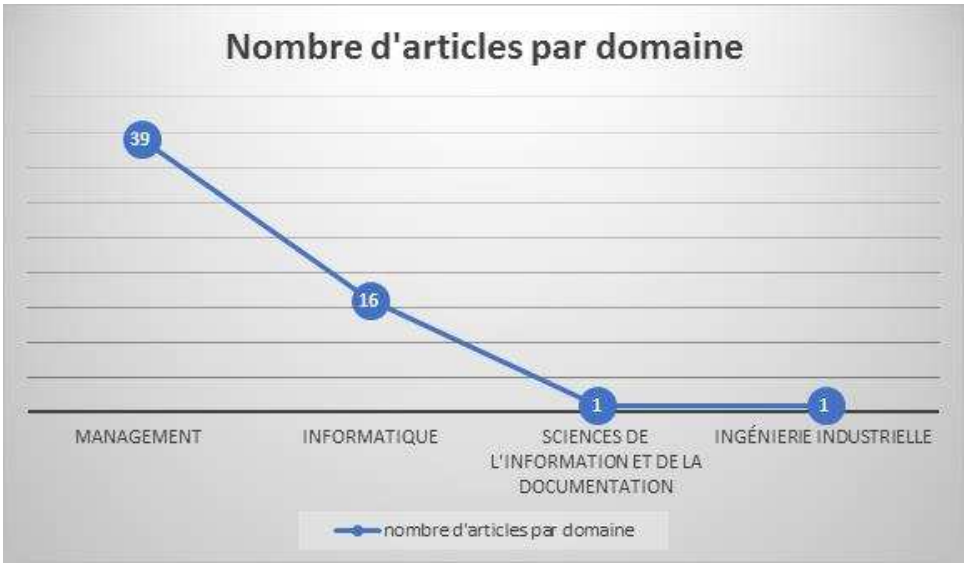


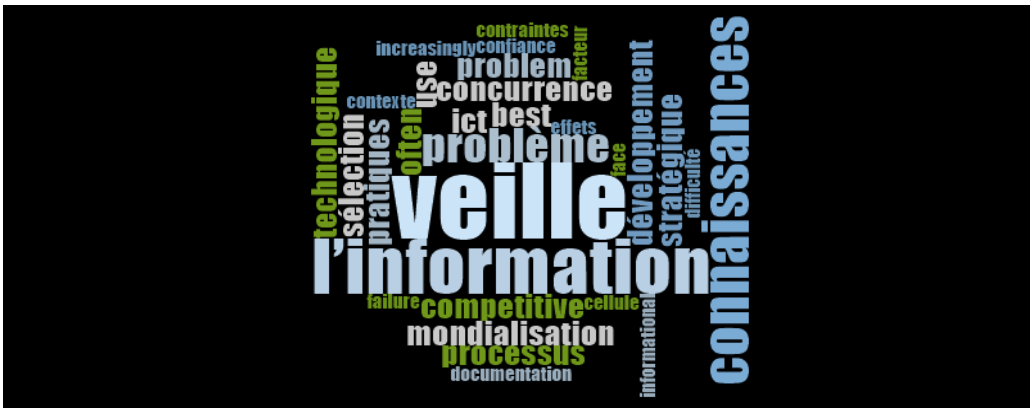
Figure 3 : Répartition des articles par domaine de recherche

Le domaine de management totalise le plus grand nombre des articles trouvés, 39 articles, alors que le domaine de la science de l'information et l'ingénierie industrielle ne comporte qu'une seule publication chacun. Ce qui nous amène à constater que les axes de recherche les plus traités se focalisent essentiellement sur les problématiques liées à la veille stratégique, qui englobe entre d'autres la veille commerciale, la veille marketing etc. C'est dans cette perspective, que nous avons opté pour l'analyse des occurrences de concepts les plus fréquents pour dégager les problématiques les plus traitées et les perspectives de recherche les plus citées.

5) Les occurrences de concepts les plus fréquents

Nous avons soumis au logiciel Nvivo 11 plus un corpus de texte et nous avons défini les nœuds permettant de trier et d'éliminer les redondances, pour repérer les concepts les plus fréquents dans les problématiques et les perspectives de recherche des articles trouvés. L'intérêt du nuage de mots est de pouvoir visualiser l'étendue lexicale et les termes les plus fréquents.

Les nuages de tags suivants permettent une représentation visuelle des mots clés les plus cités dans les problématiques de recherche selon les différentes sources :



Nuages de tags des problématiques de recherche (SIIE)

Parmi les mots clés les plus cités dans les problématiques de recherche traitées dans les articles publiés par les conférences SIIE : veille, information, connaissances, veille technologique, compétitivité, concurrence, ces mots se rapportent essentiellement au domaine du management. D'autres mots clés nous renvoient aux problématiques de la pratique de la veille tels que : processus, pratique, confiance, facteurs. De ce fait, les problématiques les plus traitées portent sur la veille stratégique et la veille technologique comme déterminants et facteurs de succès pour la compétitivité des entreprises dans différents contextes affectés par la mondialisation. La connaissance qui découle des pratiques de veille semble aussi occuper une place importante dans les préoccupations des chercheurs. Le Knowledhge management étant un prolongement de la vielle et de l'intelligence économique.

Nuages de tags des problématiques de recherche (VSST)

Les mots-clés les plus cités dans les problématiques de recherche traitées lors des séminaires VSST s'orientent vers le domaine du management cela peut être expliquer par l'aspect moin généraliste et plus pointue de colloque, spécialement dédié à la Veille Scientifique et Technique VSST. Les problématiques les plus traitées se distribuent comme on le voit dans le nuage autour de la notion centrale de système. Le champ lexical comporte des notions fréquentes comme Intelligence, démarche, modélisation, méthode, l'amorçage. Les problématiques de la mise en place des dispositifs de veille stratégique semble être centrales dans chez cette communauté scientifique.

Nuages de tags des problématiques de recherche (RS)

Les mots-clés les plus fréquents dans les problématiques de recherche évoquées dans les revues scientifiques sont plus vastes, traitant plusieurs domaines : informatique, management. Les problématiques traitées ne sont pas seulement liées à la veille stratégique mais aussi à la

veille technologique. Les articles de revues scientifiques semblent aller au-delà la problématique des facteurs de succès, et déterminants de la bonne pratique de la veille, pour se focaliser de plus en plus sur les dispositifs et les technologies numériques déployées lors de la mise en place du dispositif de veille.

Nuages de tags des problématiques de recherche (mémoires et thèses)

Les 7 mémoires et thèses traités dans le cadre de cette étude préliminaire, sont ceux disponibles à l'école de commerce de Tunis, de ce fait il n'est pas surprenant que les mots clés soient limités au domaine du management. Les problèmes liés à la veille stratégique et veille technologique sont les plus traités dans les mémoires et thèses.

6) Les perspectives de recherche

L'analyse des corpus relevés dans les différents travaux, a été faite manuellement étant donné le nombre relativement limité des corpus. Notre lecture des passages relatifs aux perspectives nous fait découvrir une tendance faible à traiter des problématiques liées à l'intégration de la veille dans les Systèmes d'information stratégiques ou décisionnels et la prise en compte des progrès techniques réalisés. Un seul article dans notre corpus à traité et envisage d'aller plus loin dans le sujet. Les big data offrent une opportunité aux spécialistes de la veille et de l'IE pour utiliser les données semi-structurées et non structurées qui sont les plus pertinente pour l'anticipation. L'IA forte permet quant à elle de traiter plus rapidement des documents sources de plus en plus variés et volumineux.

Conclusion

Rappelons au terme de ce travail de recherche que nous sommes partis sur un objectif de combler un vide auquel font face les chercheurs académiques, qui sont intéressés aux problématiques liées à VS et IE. Face à un déficit de traçabilité de la recherche dans ce domaine les chercheurs ont tendance, à travers leurs publications de se rabattre sur la facilité perçue de l'exploratoire. L'apport de notre contribution se situe essentiellement à ce niveau. Grâce à la collecte des données primaires, soumise à un traitement automatique à l'aide de NVIVO 11 plus sur les 57 Références collectées nous avons pu répartir les publications sur 4 domaines de recherche à savoir : management, informatique, ingénierie industrielle, et science de l'information et de la documentation. Et afin de mettre en exergue les occurrences de concepts les plus fréquents, nous avons soumis au logiciel Nvivo 11 plus un corpus de

texte et nous avons défini les nœuds permettant de trier et d'éliminer les redondances, pour repérer les concepts les plus fréquents dans les problématiques et les perspectives de recherche des articles trouvés. Nous constatons à travers notre étude préliminaire que les recherche en Tunisie porte essentiellement sur les déterminants de succès de la mise ne place de la veille, et que la pratique de la veille est occasionnelle. Nous pouvons constater que les décisionnaires adoptent toujours le mode commande de la veille et non pas le mode alerte, de ce fait, nous pouvons conclure que la pratique de la veille n'est pas toujours inscrite dans la culture de l'entreprise. D'un autre côté, en examinons les occurrences de concepts les plus fréquents dans les perspectives de recherche nous remarquons qu'elles tiennent en compte l'évolution technologique et l'importance de l'intégration des nouvelles technologies dans la structuration de la veille. D'ailleurs depuis que la veille stratégique s'est dotée de l'internet, cet outil lui offre un potentiel qui lui permet d 'anticiper et de bénéficier d 'une information numérisée et pertinente en vue de produire des connaissances et d'assister les décisions performantes. Malgré la conscience des chercheurs de l'importance des nouvelles technologies dans la pérennité de la veille mais la recherche en Tunisie est encore loin de la tendance actuelle des occidentaux qui émergent déjà vers l'utilisation de l'intelligence artificielle dans la pratique de la veille. Cette conclusion fera l'objet d'une consultation d'experts pour validation.

Références

Aguilar, F.J. (1967) - *Scanning the Business Environment,* New York : Macmillan, 239p.

Alexandre, L. La guerre des intelligences : Comment l'Intelligence artificielle va révolutionner l'éducation. Editions JC-Lattès, Paris, 2017

Ansoff, H.I (1975), Managing Stratégic Surprise by response to weak signals, *California Management Review*, vol. 18, n° 2, pp. 21-33.

Daft R.L., Lengel R.H. (1984)- Information Richness : A New Approach to Managerial Behavior and Organization Design. *Research in Organizational Behavior*, 6, pp. 191-233.

Pateyron, E. « La veille stratégique » – Paris : Economica, 1998. – 212 p. – ISBN : 2- 7178-3695-0.

Jakobiak, F. (1991) - *Pratique de la veille technologique,* Les Editions d'Organisation, 232p.

Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: an example, design considerations and applications. Information & Management, 42, 15-29.

Vers une nouvelle approche de processus de systèmes de recommandation

Imen Gmach, Administrative Sciences Department, Imam Abdulrahman university ,Community College Arabie Saoudite, Dammam ihgmmach@iau.edu.sa

Lofi khrifish, Research Unit Business Management, Faculty of Economic Science and Management , Sfax University, Sfax,Tunisie, lotfi.khrifech@gmail.com

Résumé —
Cet article fait le point sur un état de l'art des systèmes de filtrage informationnel. Il présente diverses techniques de filtrage évoquées dans la littérature dont principalement le filtrage à base de contenu, le filtrage collaboratif et les filtrages hybrides. Il est question d'étudier les différentes étapes du processus de recommandation. L'article présente également la notion de prédiction dans le filtrage informationnel. Et il se termine pour une proposition d'un processus de recommandation en valorisant le lien les différentes étapes pour améliorer la recherche d'information personnalisée.

***Mots-clés* — recherche d'information, approches de filtrage, filtrage collaboratif, agent de filtrage, systèmes de recommandations.**

I. INTRODUCTION

Ce travail s'inscrit dans le contexte d'une recherche en cours d'évolution sur l'apport des méthodes de filtrage pour améliorer le processus de recherche d'informations (RI) et remédier aux carences liées au profilage de l'usager d'un système RI.

Le volume d'information présent sur le Web et dans les systèmes d'information ne cesse de croitre en acquérant une sémantique de plus en plus complexe. Ainsi, retrouver précisément l'information recherchée sur le Web devient une tache complexe car, pour une requête particulière, soumise à un moteur de recherche d'information du Web, donnera un grand nombre de réponses dépassant les limites cognitives de l'usager mais proportionnellement avec l'expansion du contenu mis à disposition.

Par conséquent, il est devenu nécessaire de proposer des approches et des outils permettant de cibler au mieux les réponses fournies à chaque usager, afin que les réponses soient plus proches à ses attentes et à ses souhaits rapprochant ainsi son profil cognitif. Ceci a l'avantage de répondre aux critères de la facilité d'utilisation ou d'utilisabilité requis aux systèmes d'informations [10] en permettant à l'utilisateur d'être facilement mis en présence de l'information pertinente qui l'intéresse et le concerne personnellement. Dans ce contexte, le système de filtrage informationnel se présente comme une alternative pour une solution pertinente au problème posé.

Une méthodologie avec des outils associés pour filtrer des informations sur des ressources est donc nécessaire afin de ramener l'ensemble des ressources qui sont à l'échelle de l'Internet vers la dimension cognitive de l'usager : humainement gérable. Dans cette considération, les systèmes de filtrage d'informations sont des exemples pour les solutions cherchées.

Dans la littérature grise, un système de filtrage d'information « achemine des documents qui se présentent vers des groupes de personnes, en se basant sur leurs profils à long terme » et « élaborés à partir de données d'apprentissage » [12].
Le filtrage d'information est l'expression utilisée pour décrire une variété de processus se rapportant à la fourniture de l'information adéquate aux usagers qui en ont besoin [10].
Le terme « filtrage » est souvent interprété comme l'élimination de données indésirables sur un flux entrant, plutôt que la recherche de données spécifiques sur ce flux. Beaucoup sont les recherches effectuer sur le processus de la recommendation ,la majorite se sont interesses sur l etape de la recommendation sans tenir compte de la prerecommendation. D une part et en se basaant sur les recherché de (Tarek Abdelatif) qui a montre que la recommendation suit un processus a trois

etapes a savoir :1) L evaluation et le classement des alternatives de recommendations 2) collecter plus de detalils sur le besoin de l utilisateur afin de lui proposer une liste de souhaits qui repond a son besoin 3)comparer les differentes alternatives afin d efectuer le choix final.De meme (Adomavicuis and Tuzhilin 2005)qui ont propose un modèle de recommandation a trois étapes a savoir comprendre l'utilisateur 2)proposer des recommandation 3)impact de la recommandation propose
D'autre part,(Anna Casali et Armando),(Monali Gandhi et al)ont montré que le système multi agent permet d'aider l'internaute à avoir une meilleure recommandation qui répond a ses besoins .
La pluralité des travaux de recherches se sont focalises a répondre a la question comment aider l'utilisateur à accepter la recommandation proposes peu sont les recherches qui se sont intéressés a étudier qu est ce au on doit recommander? Et à qui?
Ainsi donc, nous pouvons résumer le but de cette article est de voir Comment pouvons-nous formaliser les étapes préalables a la recommandation?
Pour ce faire il s'avère nécessaire de résoudre les questions suivantes :
Pouvons-nous parler d'une phase antérieur à la recommandation?
-Comment s'articule les étapes de la phase antérieure à la recommandation?

Le but de cet article est de présenter un « état de l'art » sur le filtrage de données et d'informations ainsi que de proposer un processus de recommandation produit/services.D' ou on soulève la question de recherche suivante:

Cette article s'articule autour de trois grandes thématiques (ou parties) : la première présente une introduction générale sur le rapprochement entre les systèmes d'information et les systèmes de filtrage ; la seconde énonce les diverses approches de filtrage et la dernière est consacrée pour présenter notre approche « processus de recommandation produit services».

I. ETAT DE L'ART

Un système de recherche d'information a pour but « La mise à la disposition des documents pour l'utilisateur afin de satisfaire son besoin en informations »[10].

Quant à un système de filtrage d'information « amène des documents à des groupes de personnes en se basant sur leur profils sur le long terme » [12] et il est « fondé à partir des données d'apprentissage ».

L'accès à l'information pertinente se fait à travers le système de filtrage d'information, à l'encontre des outils de recherche d'informations qui ne permettent pas cette formulation ou adéquation systématique par rapport au besoin informationnel de l'utilisateur [10].
Le tableau en ci-dessous (cf. Tab.1) résume les critères de divergence entre un système de filtrage et un système d'information :

	Systèmes d'information	Systèmes de filtrage
Approche	Trouver l'information recherchée	Filtrer l'information non désirée
Analyse du contenu	Utiliser des mots clés (termes ou multitermes)	Analyser pour le filtrage basé sur le contenu et non pour le filtrage collaboratif
Livraison	Corpus statique à la demande	Flux de données dynamiques
Fonctionnalités	Non personnalisée, passif, à court terme, non adaptatif aux changements	Personnalisé, filtre sur les données entrantes et de manière dynamique, à long terme, s'adapte aux changements du profil utilisateur
Persistance	Des besoins à court terme	Des intérêts à long terme

Tableau 1 **: Recherche d'information versus filtrage d'information [9]**

A. Le filtrage d'information

Le filtrage d'information est décrit comme [1] étant une variété de processus ayant pour but la mise à disposition des informations à des

personnes en adéquation avec leurs centres d'intérêts.
Donc le filtrage permet de sélectionner l'information pertinente à partir d'un flux important d'informations selon le profil utilisateur.

Les objectifs majeurs du filtrage sont :

- Eviter la surcharge en informations pour des utilisateurs : le filtrage permet de sélectionner l'information pertinente pour l'utilisateur et ceci à partir d'un flux important de données en se basant sur son profil.
- Rechercher l'information pertinente : le filtrage permet de rechercher l'information pertinente pour l'utilisateur.
- Contrôler l'accès à l'information : le filtrage permet l'autorisation de l'information utile en bloquant l'information inutile.

Dans ce contexte, il existe deux méthodes de filtrage :

- **Le filtrage actif :**

Il permet de chercher l'information pertinente pour l'utilisateur. En effet reposant sur cette technique, le filtrage collecte explicitement les préférences de l'utilisateur et ceci par l'intermédiaire des questionnaires, des votes et des commentaires exprimés par l'utilisateur vis-à-vis des ressources [14,6].
Il est vrai que cette technique parait simple néanmoins, elle possède des inconvénients. En premier lieu vient la forte implication de l'utilisateur dans le processus : étant donnée qu'il devra répondre à plusieurs questions afin d'affiner et de personnaliser le contenu, ce qui affectera l'utilisabilité du système d'information.
Prenons comme exemple le domaine du tourisme, les pionniers sont Triplehop de Trip Matcher [13], dans ce cas l'utilisateur à l'aide d'un formalisme bien déterminé exprime ses besoins, ses attentes et les contraintes, l'approche effectue alors une correspondance entre les préférences de l'utilisateur et les objets du catalogue qui doivent être décrite avec le formalisme.

- **Le filtrage passif :**

Cette approche vient remédier aux problèmes de la forte implication de l'utilisateur dans les systèmes de filtrage pour le recueil explicite des préférences.
Le filtrage passif permet de sélectionner l'information pertinente à partir d'un flux important d'information selon le profil de l'utilisateur [7,2] : l'approche de filtrage passif modélise l'utilisateur par l'intermédiaire de l'analyse de son comportement compte tenu du contenu informationnel mis à sa disposition.
Le service Amazon est un bon exemple, il reconnait directement les identités des utilisateurs et leur recommande des livres qui pourraient les intéresser sans être craint à leur demande sur l'expression de leur besoin.

B. Les différentes approches de filtrage:

Ils existent trois grandes familles de systèmes de filtrage d'informations [3] :

1. Le filtrage à base de contenu

C'est la proche la plus ancienne dans les systèmes de filtrage d'information basé sur le contenu des documents. Elle trouve ces racines dans le monde de la recherche d'informations.

Cette technique appelé aussi filtrage thématiques, cognitif ou encore intelligent s'appuie sur le contenu des documents (thèmes abordés ou encore les items) afin de les comparer à un profil composé généralement des thèmes abordés par l'utilisateur.
En effet, le filtrage à base de contenu est fondé sur une comparaison des thèmes abordés par les documents avec ceux qui intéressent l'utilisateur. Il convient alors de déduire que chaque utilisateur possède un profil qui décrit ces centres d'intérêts. Ce profil peut contenir une liste de thèmes que l'utilisateur aime ou pas [6].
Cette approche est similaire au fonctionnement d'une requête de recherche d'information. Le profil utilisateur occupe le rôle d'une requête de recherche afin de filtrer les documents en adéquation avec le besoin de l'utilisateur [9].
L'avantage majeur du système de filtrage à base de contenu est l'utilisation des techniques de l'intelligence artificielle pour parfaire le filtrage, mettre à jour le profil et coupler les profils avec les documents. Autrement dit, l'utilisateur d'un tel système ne dépond pas des

autres utilisateurs, mais il peut bénéficier des recommandations des autres même s'il est le seul inscrit et à communiqué peu d'information sur ces centres d'intérêts.

Toutefois le filtrage à base de contenu possède certaines limites à savoir : le démarrage à froid, prenons l'exemple d'un utilisateur qui trouve du mal à exprimer ces préférences ou encore à définir ces thèmes pour que le système calcule son profil.
L'effet d'entonnoir, qui restreint le champ de vision des utilisateurs, est en perpétuelle évolution ou en restriction avec l'évolution des profils utilisateurs calculés.

1. Le filtrage collaboratif

Le premier système de filtrage collaboratif est le système de Tapestry [8].
Le filtrage collaboratif tente de pallier les insuffisances du filtrage à base de contenu en se fondant sur une évaluation des contenus par une communauté d'utilisateurs ayant des profils similaires.

Donc, le filtrage collaboratif exploite les évaluations des documents faits par l'utilisateur dans le but de recommander ces mêmes documents à d'autres utilisateurs sans analyse de leur contenu [8,13].
Plus précisément, le système de filtrage collaboratif est un processus de recherche et d'exploitation des corrélations qui peuvent exister entre utilisateurs ou encore entre des items qu'ils ont appréciés [7]. Ainsi, le filtrage collaboratif utilise principalement une matrice dont les lignes correspondent aux utilisateurs et les colonnes aux ressources. Chaque cellule de la matrice correspond à une note adéquate de l'utilisateur à la ressource. L'effort est construit sur comment prédire les notes attribuées par l'utilisateur à chaque ressource pour pouvoir recommander par la suite les meilleures ressources (ou notes prédites).

L'architecture générale du système de filtrage collaboratif repose sur deux fonctionnalités :

- Le calcul de la prédiction de l'évaluation d'un document par l'utilisateur ainsi que le calcul des proximités entre les utilisateurs ;
- La mise à jour continue des profils de l'utilisateur et ce au fur et à mesure de la collecte de leur évaluations (ou données utilisateurs).

Il est vrai que le système de filtrage collaboratif vient remédier aux insuffisances du filtrage à base de contenu, car il possède plusieurs avantages et principalement le fait qu'un utilisateur peut se servir des documents déjà évalués et traités par d'autres utilisateurs dans le système lors de leurs activités antérieures de recherche d'informations. Cette approche résout le problème de filtrage à base de contenu étant donnée qu'il devient possible de traiter n'importe quel type ou forme de données et de diffuser des informations non similaires à celles déjà proposées.
Toutefois, les personnes qui ont un gout ou une préférence peu fréquente risquent bien de ne pas recevoir de propositions, puisque le système ne pourra pas filtrer les données : un minimum d'informations est requis sur les utilisateurs (données de profils) avec des données collectées à partir de leurs avis (données d'apprentissage).
Cette approche se limite aussi pour le démarrage à froid [8], à savoir qu'un nouvel utilisateur dans le système débute avec un profil vide et donc le système devra le constituer à partir de données manquantes. Ce profil vide devra passer par une période d'apprentissage avant de refléter concrètement les préférences de l'utilisateur. Tout au long de la période d'apprentissage, le système ne pourra pas filtrer efficacement le profil et les données de l'utilisateur.

2. Le filtrage Hybride

C'est la combinaison du filtrage par le contenu et du filtrage collaboratif.
Le filtrage hybride tente de tirer profit de toutes les approches de filtrage en combinaison.

Cette approche vise à combiner les approches entre elles afin d'améliorer la qualité. En effet, les ressources seront donc acheminées vers d'autres utilisateurs en employant les critères de filtrage collaboratif pour les profils et le filtrage à base de contenu pour les contenus [3, 4, 14].

Dans la littérature, plusieurs approches d'hybridation ont été proposées et qui recommandent l'utilisation de deux ou plusieurs méthodes de personnalisation dans un même système [11]. Toutefois, le choix de types d'hybridation ne se fait pas de manière

triviale, il dépendra du domaine et de l'information disponibles (cf. **Tableau 2**).

Approches de filtrages	Avantages	Limites
Filtrage à base de contenu	- Répondre aux intérêts à long terme des utilisateurs en employant des techniques efficaces dans le domaine. - Intègre la sémantique.	- Difficulté d'indexer les documents multimédias (images, son…). - Problème de seuillage. - Non prise en compte des critères de pertinence outre que le thème abordé à savoir : qualité scientifique du contenu, fiabilité de l'information… - Non prise en compte des nouveaux thèmes - Effet d'entonnoir - Effet de masse
Filtrage collaboratif	- Simple. - Appréhende tout type de documents (texte, images, son….) - Ne filtre pas les documents traitant les nouveaux thèmes.	- Subjective - Indépendance du domaine et de l'information existante. - Difficulté de quantifier les critères de performance - Le système de notation est grossier.
Filtrage Hydride		- Dépond tes techniques sous jacentes.

Tableau 2: Les avantages et les limites des différentes approches de filtrage

Une revue de la littérature noua a permis de classifier les divers approches avec les références correspondant (cf. **Tableau 3**).

Approches	Réferences
Filtrage à base de contenu	[Pazzani, 99] [Balabanovic, 97] [Deshpande, 04] [mooney, 00]
Filtrage collaboratif	[Linden, 03] [Balabanovic, 97] [Goldberg, 92] [Herlocker, 99] [Resnick, 94] [Sarwar, 01]
Filtrage hybride	[Burke, 07] [Berkovsky, 08]

Tableau 3 : **combinaison approches/références**

Les deux approches principales de filtrage d'informations à savoir le filtrage à base de contenu et filtrage collaboratif donnent des résultats intéressants. D'une part, le filtrage collaboratif, en se basant sur les opinions de groupes d'utilisateurs similaires, recommandent aux utilisateurs certains documents qui n'ont pas encore consultés. D'autre part, le filtrage à base de contenu en se basant sur le profil thématique de l'utilisateur évalue si un texte ou un document est important ou non pour l'utilisateur. Ces deux approches semblent néanmoins importantes. En effet, une approche hybride combinant les opinions de l'utilisateur et les goûts ou encore les préférences thématiques d'un utilisateur, parait une méthode pertinente pour fournir des recommandations.

Il est à noter sur les techniques de filtrage qu'elles ne tiennent pas en considération un certain nombre d'informations qui expliquent le jugement effectué par l'utilisateur. Ceci nécessite la prise en compte de diverses sources d'informations sur l'utilisateur (ces préférences, ces centres d'intérêts, son environnement social, sa culture, ses traits de personnalités,…) et leurs corrélations avec les documents qui lui sont proposés ou encore les prédictions.

Dans ce qui va suivre nous allons exposer le concept de prédiction et les diverses techniques de calcul.

B. Les agents intelligents des filtrages

Actuellement, des recherches sont dirigées vers les systèmes « intelligents » ou encore appelé « automatique ».

On appelle agent intelligent un système qui prend les décisions à une situation donnée sans intervention humaine, c'est-à-dire l'agent intelligent agira automatiquement sans l'aval humain.

Les orientations actuelles visent à l'intégration de ces agents dans le domaine de recherche d'informations et de filtrage collaboratif :

- les agents constituent leurs bases de données en se fondant sur leur observations du comportement de l'utilisateur et ainsi ils agissent suite à une équation de recherche effectuée par l'utilisateur,
- l'agent effectue sa sélection des documents qui répondent à la requête utilisateur, il classe les documents par

ordre de pertinence et puis les présentent,

– l'agent observe le compte de l'utilisateur dans le but d'ajuster son équation puis de décider de lui envoyer ou non une liste de documents profilés [7].

II. APPROCHE PROPOSEE DU PROCESSUS DE RECOMMENDATION

Plusieurs sont les recherches qui se sont intéressés a comprendre l'utilisateur afin de construire son profil et par la suite lui recommender des produits et services.

Certes, il est primordial de comprendre les préférences utilisateurs et d'étudier leur profil toutefois il s'avère qu'étudier et savoir quoi recommender a cette utilisateur est une étape primordial préliminaire a la recommandation, c'est l'objet de notre article

En effet, la collecte de des données est relatifs à l'utilisateur ou au produit dépond des sourcesd'informations et des objets intéressants utilisés dans le système. Or ces informations. Or ces informations ne sont pas tous facile à collecter, c'est le rôle des systèmes de recommendation

Notre travail, se divise en 5 étapes (FIGURE1)

Etape 1: collectionner ou encore rassembler les diverses types de produits

Etape 2: Ayant une liste de produit diversifies, nous commençons par filtrer les produits par catégories (Catégorie A, Catégorie C,…..)

Etape 2.1:Ensuite, dans la liste de divers groupes de catégories nous essayons de raffiner le filtrage par marque et notoriété

Etape 2.2: Dans cette même base de données, appliquer un autre filtrage par prix (Opportunités ou pas)

Etape 2.3: Obtenir une liste de produit a recommender

Etape 3: Sélectionner les préférences des utilisateurs pour cette base de données obtenues

Etape4: Organiser et structurer les recommendation selon la matrice produit/préférence utilisateur

Etape 5:Decider de recommander tel produit a tel utilisateur

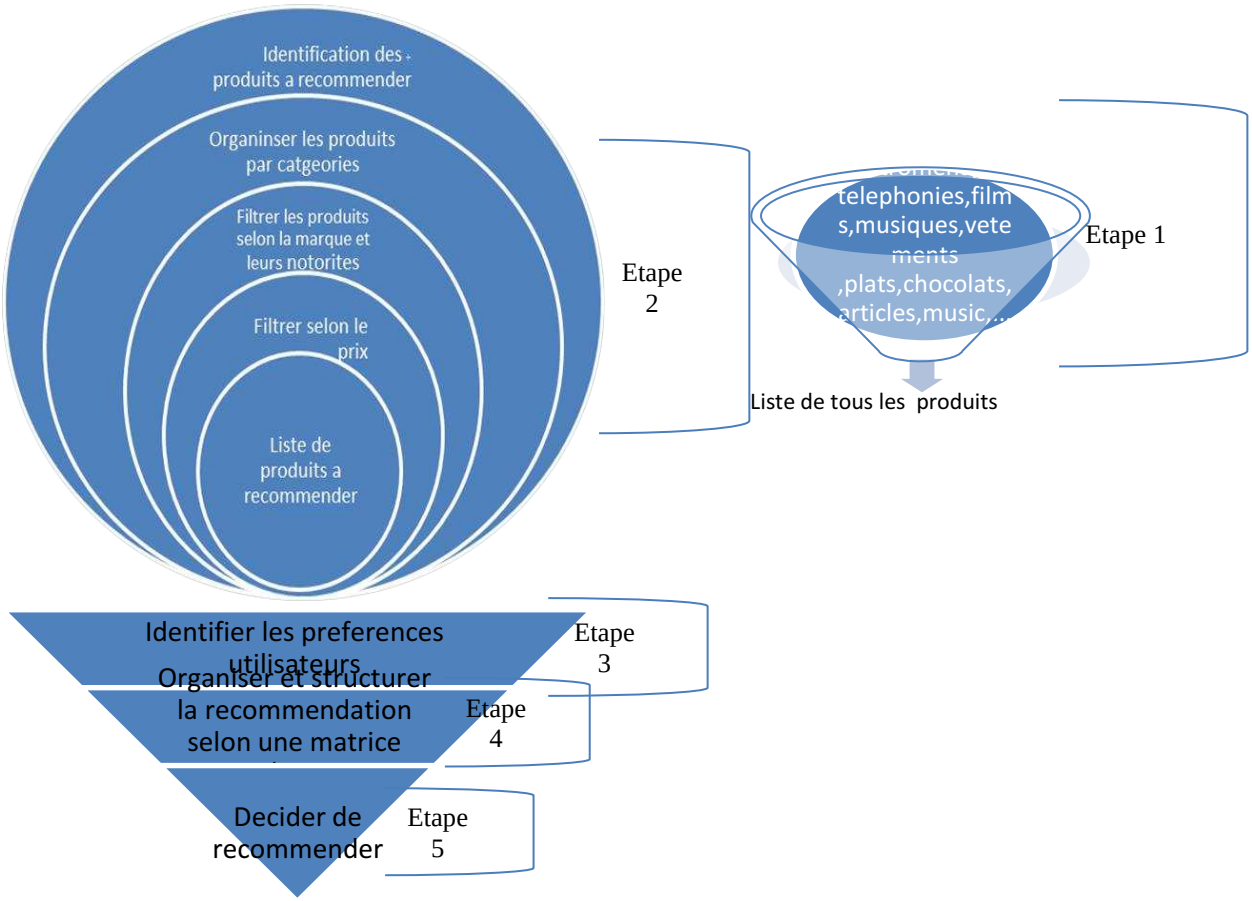


Figure 1:Procesus de recommandations de produits /services

III. DISCUSSION:

En se basant sur les diverses recherches antérieures sur les systèmes de recommandation nous
La majorité des recherches se sont focalises sur l'adoption et l'acceptation de la recommandation ou encore sur l'impact de la recommandation
Or il s'avère nécessaire de s'intéresser à l'étape de la pré recommandation vu que les consommateurs n'ont pas des informations complètes relatifs au produit ou services sur le net.et que la majorité des internautes s influencent par les produits aimes ou achètes par leurs amies.
Dans cette étude nous avons analysé les différents types de systèmes de recommandation ainsi que les différentes techniques de filtrage existantes dans la littérature, à savoir les systèmes à base de contenu, les systèmes hybrides et les systèmes collaboratifs sur lesquelles se base notre étude.
Nous avons par la suite étudié le processus de recommandation et ses différentes étapes.
Nous nous sommes base sur les travaux de (Tarlatif et al)et celui de afin de construire notre processus qui se décompose d'une phase de prerecommandation c'est la nouveauté propose dans cette étude.
En effet nous nous sommes focalises sur l'étape qui précédé la recommandation tel qu'il faut bien étudier cette étape afin de savoir bien choisir que recommander selon des critères que nous avons soulevé pour construire notre processus de recommandation.

Conclusion

La recherche d'information et le filtrage informationel (sur le contenu et/ou sur le profilage de l'utilisateur) consiste à concevoir, modéliser et à mettre en œuvre des systèmes permettant le traitement des flux d'informations au fur et à mesure de leurs arrivées dans le but de sélectionner et de présenter les documents pertinents. Cependant, la recherche d'information et le filtrage ont un lien fort ensemble pour compléter la chaine de traitement associant les documents pertinents et les préférences de l'utilisateur.

Dans cette étude, la recherche d'informations a été observée pour inclure trois fonctionnalités principales, à savoir :
(i) l'organisation et l'indexation des contenus (collection de documents), (ii) la requête de l'utilisateur et ses préférences intrinsèques, et (iii) la méthode utilisée pour effectuer la comparaison entre l'indexation de contenus et la requête de l'utilisateur associant ses préférences. Par l'aspect sur le filtrage d'informations, la complexité se situe dans l'association qui traite les contenus entrants, les profils intégrés qui sont dans la requête, et la méthode utilisée pour effectuer le rapprochement entre les documents et les profils. Aux travaux de recherche actuelle, ce rapprochement s'inspire beaucoup plus des techniques classiques utilisées en recherche d'information. Par conséquent le filtrage d'information à associer sera fortement lié au concept de l'utilisateur en activité de recherche d'informations. Tout en améliorant les méthodes et la nature des résultats proposées à ce dernier, répondre à ses besoins (en termes d'information) et à ses attentes (en termes de satisfaction dans ses préférences) reste une problématique complexe à résoudre pour y contribuer.
Intrinsèquement, le filtrage d'information est un domaine de recherche en plein essor pour améliorer l'extraction des flux de données entrants, des informations utiles pour le compte de l'utilisateur.

La modélisation de l'utilisateur est par conséquent, le point central pour l'accès adaptatif et personnalisé à l'information. Dans le but de rendre cet accès efficace et pertinent, le système doit connaitre l'utilisateur à travers la description de son profil et devra évoluer avec ses activités afin de refléter ces centres d'intérêts courants aussi bien les changements dans ses préférences : des problèmes sont à résoudre et un certain nombre de contraintes restent liées à la prise en compte des données relatives au contexte de la recherche d'informations. La contextualisation est ainsi observée dans la dynamique des préférences de l'utilisateur.

IV. RÉFÉRENCES

[1] Nielsen J. Usability Engineering, chapter" What is Usability", Academic Press, pages 23-48., Cambridge, MA, 1993.

[2] Robin D. Burke." Hybrid web recommender systems. In Peter Brusilovsky", Alfred Kobsa, and Wolfgang Nejdl, editors, The Adaptive Web, Methods and Strategies of Web Personalization, volume 4321 of Lecture

Notes in ComputerScience, pages 377-408. Springer, 2007.

[3] Herlocker J.-L., Konstan A.-J., Borchers A., Riedl J.," An Algorithmic Framework for Performing Collaborative Filtering", Proceedings of the 22nd International ACM Conference on Research and Development in Information Retrieval (SIGIR'99), , p. 230-237. USA,1999

[4] Lieberman H., Letizia: "An agent that assists web browsing", Proceedings of the 14thInternational Joint Conference on Artificial Intelligence (IJCAI-95), p924-929,Canada, 1995.

[5] Lang K., NewsWeeder:" Learning to Filter Netnews", Proceedings of the 12th International Conference on Machine Learning (ICML'95), CA, p. 331-339,USA, 1995,

[6] Salton G., Buckley C., "Term Weighting Approaches in Automatic Text Retrieval",Information Processing and Management, vol. 24 (5), , p.513-523, 1988.

[7] Gong, S. ; Ye, H. et Dai, Y,"Combining singular value decomposition and item-based recommender in collaborative filtering". In Proceedings of the 2009 Second International Workshop on Knowledge Discovery and Data Mining, , Washington, DC, USA. IEEE Computer Society, pages 769–772. (WKDD'09).

[8] Goldberg, D; Nichols, D. ; Oki, B. et Terry, "Using collaborative filtering to weave an information tapestry". Communications of theACM, 35(12):61–70. 1992

[9] Berrut, C, Denos N.. «Filtrage collaboratif». In Assistance intelligente à la recherche d'informations, Hermes Science Publications, p. 30., 2003 En ligne. <http://www-mrim.imag.fr/publications/2003/CBOO 1/berrut03 b. pdf>

[10] Belkin,N J., Bruce Croft W.. «Information filtering and information retrieval: two sides of the same coin?». Commun. ACM. vol. 35, no 12, p. 29-3, 1992. En 1igne. <http://doi.acm.org/10.1145/138859.138861 >.

[11] Nguyen A.-T., Denos N., Berrut C., « Modèle d'espaces de communautés basé sur la théorie des ensembles d'approximation dans un système de filtrage hybride », Actes de la 3ème Conférence en Recherche Information et Applications (CORIA'06), p. 303-314, Lyon, France, 2009.

[12] CROFT W.B., « Knowledge-based and Statistical approaches to Text Retrieval »,IEEE EXPERT, vol. 8, n° 2

[13] Balabonovic M.,Shoham Y., " Fab: content-based, collaborative recommendation",Communication af the ACM,vol 40,n°3,P 66-72,1997.

[14] Xiayuan .S , Taghi M.K., « A survey of collaborative filtring techniques ».Hindawi Publishing Corporation,advances in artificial intelligence, vol 2009,P 19,2010.

[15] Pazzani M.,"a framework for collabaortif, content-based, and demographic filtering",artificiel intelligence rev,P. 393-408,Dec 2019.

[16] Sarwar B., Karypis G., Konstan J.,Reidl j., " Item-based collaborative filtering recommendation algorithms", Proc of the 10 international conference on world Wide Web,P 285-295,2001.

[17] Linden G. ,Brent S., York J, "Amazon.com recommendations: Item-to-item collaborative filtering" ,IEEE internet computing , Vol 7, N°1,p76-80,2003.

[18] Linerman,H.,Maulsby,D.; "Instructible agents:software that just keeps getting better",IBM Systems Journal, Vol. 35, num 3 et 4,pp. 539-556.1996.

[19] Ardissno,L.,console,L., "An adaptative system fot the personalized access to news".AI Communications,num 14,pp 129-147,2001..

[20] Miiyahara,K., Pazzani,M.J., "Collaboartive filtring with simple Bayesian classifier ". Proc 6th Pacific RIM Internationnal Conference on Artificial Intelligence,Australia,pp 679-689,2015.

[21] Aggarwal,C.C,J.L.,Wolf,J.L.,WU,K.,YU, P.S.. " Horting hatches an egg: anew graph-theoretic approach to collaborative filtering". Proc.knowledge Discovery and Data Mining, pp 201-212,1999.

[22] Brajnik,G., Tasso,C. “ A flexible tool for developing user modeling applications with non monotonic reasoning capabilities”. International Journal of Human-Computer Studies, num 40,pppp 31-62,1994.

Evolutionary algorithms for solving the multi-location and multi-item transshipment problem

Noomen Selmi, Mohamed Hmiden and Lamjed Ben said

SMART Laboratory, ISG, University of Tunis, 41 Liberty Avenue Bouchoucha, 2000 Tunis, Tunisia.
noomen.selmi@gmail.com

Abstract— Transshipment is items transfer between locations belonging to the same echelon level. It is adopted to improve inventory systems performance measures. It consists to derive the optimal replenishment decision variables while a transshipment policy is practiced. A huge body of works has addressed this topic where several configurations are considered and many approaches are adopted. A few of them has interested in the multi-item and multi-location configuration because of its complexity. We focused in this paper on this complex configuration and we propose two evolutionary algorithms PSO and DE to resolve the studied problem and we compare between them. Also, we study the impact of a number of problem parameters on the inventory system performance measures.

Index Terms— multi-item, multi-location, transshipment problem, Particle swarm optimization, Differential Evolution.

1 INTRODUCTION

Inventory management aims to satisfy demands while optimizing inventory performances which are generally expressed in terms of profit or cost functions. When companies evolve in an uncertain and competitive environment, this objective becomes hard to be achieved. Consequently, many inventory flexibility techniques are practiced in order to cope with these constraints and they help companies to achieve their primary objective. The substitution and the transshipment are two well-known inventory flexibility techniques largely practiced in industry. The first consists to replace the unavailable items by alternatives ones having the same functionalities. However, the second consists to transfer the items from locations in excess to ones in need. These two flexibility techniques help companies to improve their fill rate and to reduce, simultaneously, their inventory cost. The transshipment is hugely practiced in many domains as spare parts and fashion items sold in several locations. Generally, these kinds of items are replenished from suppliers and transshipment is practiced to serve customers requiring, from a location, items which are out of stock. For example, when a customer looks for a specific item from '*Zen la Soukra*' and this item is out of stock in this location, then the required item could be transshipped from '*Zen Manar 2*' where the item is in excess. So, the customer demand is satisfied, the profit in '*Zen la Soukra'* is improved and the inventory cost of *'Zen Manar 2'* is reduced. The transshipment problem, which consists to derive the optimal replenishment quantities where transshipment policy is adopted, has been studied. So, many configurations, parameters and approaches are investigated, and many transshipment policies are identified. Paterson et al. [1] overviewed works related to transshipment problem and they proposed a classification based on a set of criteria linked to: replenishment parameters, transshipment policies and environmental parameters. They identified two classes of works according to the transshipment policies: the reactive transshipment and the proactive one [2]. The reactive is triggered once the demands are observed and the locations in need and other in excess are identified. In contrast, the proactive transshipment is started before the realization of demand and it aims to redistribute stock in order to avoid a possible shortage. Two streams of works are identified according to number of locations involved: the two-location and the multi-location transshipment problem. The first stream adopts exact methods in order to derive the optimal inventory decisions. Krishnan and Rao [3] are the first that studied the two-location transshipment problem. They developed a single period model aiming to minimize an inventory cost function expressed in terms of holding and shortage costs. Since that, many research's focused on this stream and investigated many configurations. Among recent ones, we quote Olsson [4] who considered the transshipment lead times for two-location inventory system adopting continuous reviewing. Yao et al. [5] studied the two-location

transshipment problem with a single replenishment, at the beginning of a season, while reactive transshipment is practiced during the season. Feng et al. [6] studied the two-location transshipment problem in a competitive context and with dynamic demand information's. The second stream of works focused on the multi-location configuration and it adopts simulation-based methods to derive approximate solutions [7],[8]. The meta-heuristic is a simulation-based optimization approach which is widely adopted to resolve complex problem as the multi-location problem considering uncertain demands. Miao Z. [9] resolve the transshipment problem with fixed schedules with a genetic algorithm. Hochmuth and kochel [10] resolve the multi-location transshipment problem with many realistic parameters with particle swarm optimization (PSO) Algorithm. Danloup [11] compared the performances of two meta-heuristics applied for the transshipment problem: Local Neighborhood Search and genetic algorithm. All the mentioned works above, focused mainly on the single item configuration. Few of works have interested in the multi-item configuration. They investigated in the two-location network with periodic reviewing or the multi-location configuration with continuous reviewing.

In this paper, we focus on the transshipment problem for multi-location and multi-item configuration considering uncertain demands. We aim to resolve the studied problem with meta-heuristic approach. We compare between the performances of two algorithms, Particle swarm optimization (PSO) and Differential Evolution (DE), according to a set of criteria.

Our contributions here are threehold: first we studied the multi-location and multi-item transshipment problem considering periodic review and we formulate the studied problem, second we resolve the problem with PSO and DE and we compare between their performances and we study the impact of transshipment and uncertainty on the inventory system performance.

The rest of the paper is organized as follows: in the second section, we present the studied problem and its formal model. The third section is dedicated to the meta-heuristics algorithms PSO and DE and their application to resolve the studied problem. The fourth section is related to experimentation.

2. PROBLEM DESCRIPTION

We study an inventory system composed of N locations selling many products. At the beginning of the period, these locations are replenished from a common supplier and over the period demands are observed and satisfied. At the end of the period, a location L_i could be in excess related to product P_k and in need for Pm. However, location L_j could be in need for P_k and in excess for Pm. Transshipment, from Locations L_i to L_j of P_k units, corrects P_k shortage at L_j and it reduces the P_k holding cost at L_i. Here, we consider the fixed transshipment cost. The goal is to determine the transshipped quantities between locations and the replenishment quantities of products at each location optimizing a profit function. In order to introduce our studied problem, we present an illustrative example, shown by Figure 1, of three locations L_1, L_2 and L_3 selling two products P_1 and P_2. P_1 is in need (-3) at L_1, in excess (+4) at L_2 and in need (-2) at L3. However, P_2 is in excess (+4) at L_1, in excess (+3) at L2 and in need (-6) at L_3. Shortages at L_1 and L_3 could be corrected by transshipping items from L_1 (P_1) and L_2 (P_1, P_2).

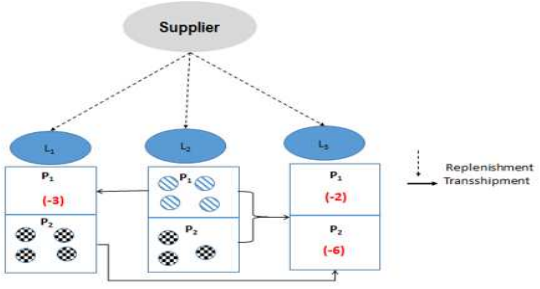


Fig. 1. Illustrative example of three locations.

2.1 Notations

Throughout this paper, we adopt the following notations

- Indexes :

i index of location $i \in \{1,2,\dots,N\}$.

k index of product, $k \in \{1,2,3,\dots,K\}$.

- Parameters:

P_i^k The unit selling price of the product k at location i.

r_i^k The shortage unit cost of product k at location i.

s_i^k The salvage unit cost of product j at location i .

c_i^k The replenishment unit cost of product k at location i .

tc_{ij}^k The transshipment unit cost between locations i *and* j for product k

D_i^k The demand of the product k at location i.

$f(D_i^k)$ The probability density function of the demand of the product k at location i.

$X = (X_1, X_2, \dots, X_N)$ the replenishment quantity vector where X_i is the replenishment quantities at location i. $X_i = (Y_i^1, Y_i^2, \dots, Y_i^K)$ is the products quantities at location L_i. Y_i^k is the replenishment quantity at location L_i for product k.

T_{ij}^k Transshipment quantity between location i and j related to product.

Performance measure functions

$\pi(X, D)$ The total profit function

$TC(X, D)$ The total Inventory cost

$TR(X, D)$ The total revenue generated before transshipment execution.

$TP(X, D)$ The total transshipment profit.

2.2 Problem formulation

$$\pi(X,D) = TR(X,D) - TC(X,D) + TP(X,D) \quad (1)$$

$$TR(X,D) = \sum_{i=1}^{N}\sum_{k=1}^{K} p_i^k \times \min\{D_i^k, Y_i^k\} \quad (2)$$

$$TC(X,D) = \sum_{i=1}^{N}\sum_{k=1}^{K} c_i^k \times Y_i^k + \sum_{i=1}^{N}\sum_{k=1}^{K} s_i^k \times \left(Y_i^k - D_i^k\right)^+ + \sum_{i=1}^{N}\sum_{k=1}^{K} r_i^k \times \left(D_i^k - Y_i^k\right)^+ \quad (3)$$

$$TP(X,D) = \sum_{i=1}^{N}\sum_{\substack{j=1 \\ i\neq j}}^{N}\left(\sum_{k=1}^{K}\left(s_i^k + P_j^k - tc_{ij}^k\right) \times T_{ij}^k\right) \quad (4)$$

$$\sum_{i=1}^{N} T_{ij}^k \leq \left(D_j^k - Y_j^k\right)^+ \quad \forall j \in \{1,2,\dots,N\}, \forall k \in \{1,2,\dots,K\} \quad (5)$$

$$\sum_{j=1}^{N} T_{ij}^k \leq \left(Y_i^k - D_i^k\right)^+ \quad \forall i \in \{1,2,\dots,N\}, \forall k \in \{1,2,\dots,K\} \quad (6)$$

3. DE AND PSO FOR THE TRANSSHIPMENT PROBLEM

Here, we resolve the studied problem descripted and formulated above by two Evolutionary algorithms, Differential Evolution (DE) and particle Swarm optimization (PSO), that confirmed their efficiency in resolving complex problem.

3.1 Differential evolution

Differential evolution (DE) was introduced for the first time by Storn and Price, in 1997 [12], as a stochastic and population-based optimization algorithm. DE was proved to be the fastest evolutionary algorithm (EA) and it was used for solving nonlinear optimization problem over continuous spaces. DE has been shown having a good convergence and very simple but very powerful for optimizing continuous functions. As many others evolutionary algorithms, DE uses three operations: mutation, crossover and selection that guide the individuals of (population) to move toward a global optimum. DE was used to solve diverse optimization problems and it has proved his performance. The DE algorithm results depends enormously of the mutation strategy and the control parameters: the population size (NP), the crossover operator (CR) and the differential weight factor (F).

- **DE Structure**

The general structure of DE algorithm is composed by the steps shown in figure 3. The steps are executed sequentially till stop continuation is met. These steps are:

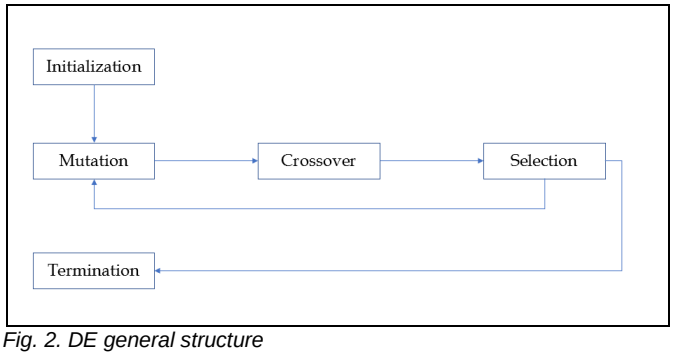


Fig. 2. DE general structure

- *Initialization*: DE starts with initialization of the population by producing NP individuals in the problem space domain. Each individual X_i is presented by a vector of D values (each value associated to one variable of the problem dimension):

  $X_i = \{\boldsymbol{x_i^1}, \boldsymbol{x_i^2}, \dots, \boldsymbol{x_i^N}\}$, i ∈ {1, 2, …., NP}. Where N is the number of locations.

  $X_i^j = \left(x_i^{j1}, x_i^{j2}, \dots, x_i^{jK}\right)$, where k is the number of items

- *Mutation* **:** After the initialization and in each generation g, DE uses the mutation operation to create at mutant vector $V_{i,g}$ associated to each target vector $X_{i,g}$ (X_i at generation g), $V_{i,g} = \{\boldsymbol{v_{i,g}^1}, \boldsymbol{v_{i,g}^2}, \dots, \boldsymbol{v_{i,g}^D}\}$ with i ∈ {1, 2, …., NP} and g ∈ {1, 2, …., G}, (G is the maximum number of generation. There are different DE variants identified according to mutations operation formula. Here, we are limited to the DE best/1 having the following mutation formula:

$$V_{i,G} = X_{best,G} + F * \left(X_{r1,G} - X_{r2,G}\right) \quad (7)$$

- *Crossover operation*: The crossover step, as presented by the formula below, is used to introduce some diversity in the population during each generation in order to look for the optimum. In this phase DE produce, at each generation g, a trial vector $U_{i,g}= \{\boldsymbol{u_{i,g}^1}, \boldsymbol{u_{i,g}^2}, \ldots.., \boldsymbol{u_{i,g}^D}\}$ associated to each individual $X_{i,g}$. The binomial Version of the crossover operation is presented below:

$$u_{i,g}^j = \begin{cases} v_{i,g}^j & if\ rand < CR\ or\ j == Rand \\ x_{i,g}^1 & otherwise \end{cases} \quad (8)$$

CR represents the crossover parameter (CR ∈ [0.1]), Rand is a random integer value with Rand∈ [0.D] and rand is a random real (rand ∈ [0.1]).

- *Selection*: During this step, at the generation g DE have to decide which vector to keep between the pair X_g and U_g depending on their fitness values, for the generation g+1.

$$X_{g+1} = \begin{cases} U_g & if\ fitness(U_g) > fitness(X_g) \\ X_g & otherwise \end{cases} \quad (9)$$

This selection formula is in case of maximization problem.

The algorithm restarts the cycle from the mutation phase until the stop condition is reached.

- **DE parameters**

The main DE parameters are:

- The Cross-over probability (CR ∈ [0.1]).
- The differential weight factor (*F*)
- The population size (NP).

- **DE algorithm:**

Algorithm 1 DE (BEST/1)

1. P (NP) ← Initialize Population
2. Evaluate each individual of P by **algorithm 2**
3. **While** (Stop-condition not met) Do
4. **For** i= 1 to NP
5. Radom choose {individual$_1$ and individual$_2$} from P
6. Individual$_{best}$ ←Look for the best individual from P
7. R ← RandomInteger [0, N*k]
8. **For** j = 1 to (N*K)
   a. CrossoverProbability ← random()
   b. **If** (crossoverProbability < CR or R== j)
      Candicate$_j$ ← Individual$_{best,j}$ + W * (individual$_{1,j}$ – individual$_{2,j}$)
   c. **Else**
      Candidate$_j$ ← individual$_{i,j}$
9. **End for**
10. Evaluate candidate fitness by algorithm 2
    If (Candidate Fitness > Individual$_i$ Fitness) **Then**
    Individual$_i$ ← Candidate
11. **End for**
12. **End while**

- **Expected profit algorithm:**

Algorithm 2

1. Expected_profit←initilize to zero
2. X inventory vector
3. **For** i= 1 to Number of Simulation (NS)
   D← generate_Demand ()
   Expected _profit← Expected _profit + $TP(X, D)$
4. **End for**
5. Expected _profit← Expected _profit /NS
6. **Return** Expected _profit

3.2 PSO algorithm:

The Particle swarm optimization (PSO) is one of the most known population-based algorithms proposed by Eberhart and Kennedy in 1995 [13]. Since its appearance, several improvements were introduced to the basic version and was used for solving global optimization problems. The basic operation of Particle Swarm Optimization consists in searching for the optimal solution in a search space D (number of dimension) where each particle 'i' is characterized by its position (Xi) and its velocity (Vi) which are represented as follows:

$X_i= \{\boldsymbol{x_i^1}, \boldsymbol{x_i^2}, \ldots.., \boldsymbol{x_i^N}\}$, i ∈ {1, 2, …., NP}. Where N is the number of locations.

$X_i^j = \left(x_i^{j1}, x_i^{j2}, \ldots, x_i^{jK}\right)$, where k is the number of items

Each particle must also keep track of its best previous position $P_{besti} = (p_{i1}, p_{i2}, p_{iD})$ as the best among all the particles of the population $P_{gbest} = (p_{g1}, p_{g2}, p_{gD})$. At each iteration, each particle in the population adjust its velocity and calculate its new positions vector using the best fitness in the population according to the following two formulas:

$$v_{id} = wv_{id} + c_1 r_1 (p_{id} - x_{id}) + c_2 r_2 \left(p_{gd} - x_{id}\right) \quad (10)$$
$$x_{id} = x_{id} + v_{id} \quad (11)$$

Where c1 and c2 are the acceleration constants and w is the inertia weight parameter. r1 and r2 are two random generated numbers in the interval [0, 1] according to the uniform law.

- **PSO algorithm:**

Algorithm 3 PSO

1. P (NP) ← Initialize Population
2. **While** (Stop-condition not met) Do
3. Update pBest of each particle
4. Update gBest of the population
5. **For** i= 1 to NP
   a. Candidate← Particle$_i$ (copy Particle$_i$ in Candidate)
   b. Calculate new_Velocity for Candidate according to (10)
   c. Calculate new_Location for Candidate according to (11)
   d. Evaluate Candidate fitness by algorithm 2
   e. **if** (Calculate Fitness > Particle$_i$ Fitness) Then
      Particle$_i$ ←Candidate
6. **End for**
7. **End while**

4. Experimental Results

In this experimental section, we focuce on three different problem configurations of transshipment problem identified according to the number of items. We are interested in four locations selling 2, 4 or 8 items. Characteristics of the studied configurations are presented in table 1.

Location		L_1	L_2	L_3	L_4
P_1	$P_i^!$	15	15	14	9
	$s_i^!$	4	4	6	4
	r_i^k	8	5	4	5
	c_i^k	10	10	9	9
	D_i^k	u[12, 26]	u[10, 20]	u[5, 36]	u[15, 35]
P_2	P_i^k	17	17	18	11
	s_i^k	5	5	5	4
	r_i^k	6	6	6	6
	c_i^k	12	12	13	11
	D_i^k	u[13, 26]	u[12, 20]	u[4, 36]	u[16, 35]
P_3	P_i^k	15	15	14	9
	s_i^k	4	4	6	4
	r_i^k	8	5	4	5
	c_i^k	10	10	9	9
	D_i^k	u[12, 26]	u[10, 20]	u[5, 36]	u[15, 35]
P_4	P_i^k	17	17	18	11
	s_i^k	5	5	5	4
	r_i^k	6	6	6	6
	c_i^k	12	12	13	11
	D_i^k	u[13, 26]	u[12, 20]	u[4, 36]	u[16, 35]
P_5	P_i^k	15	15	14	9
	s_i^k	4	4	6	4
	r_i^k	8	5	4	5
	c_i^k	10	10	9	9
	D_i^k	u[12, 26]	u[10, 20]	u[5, 36]	u[15, 35]
P_6	P_i^k	17	17	18	11
	s_i^k	5	5	5	4
	r_i^k	6	6	6	6
	c_i^k	12	12	13	11
	D_i^k	u[13, 26]	u[12, 20]	u[4, 36]	u[16, 35]
P_7	P_i^k	15	15	14	9
	s_i^k	4	4	6	4
	r_i^k	8	5	4	5
	c_i^k	10	10	9	9
	D_i^k	u[12, 26]	u[10, 20]	u[5, 36]	u[15, 35]
P_8	P_i^k	17	17	18	11
	s_i^k	5	5	5	4
	r_i^k	6	6	6	6
	c_i^k	12	12	13	11
	D_i^k	u[13, 26]	u[12, 20]	u[4, 36]	u[16, 35]

Table 1 configurations data.

We note here, that in addition to the uniform distribution of demand mentioned in table 1, we used the normal distribution with the same parameters.

4.1 Algorithms Parameters Tuning

In this subsection, we tune the parameter setting to be used for the two meta-heuristics (PSO and DE) based on the problem configuration 4X4.

- **DE/Best/1 Tuning**

We study, here, the parameters of DE to determine which specific configuration of parameters feet better with our problem. The DE/best/1 parameters are the crossover parameter (CR), the differential weight (F) and the population size.

- *CR Tuning*: We tested, here 9 different values, from possible crossover (CR) domain [0, 1], starting from 0.1 and adding at each step 0.1 to stop at 0.9.

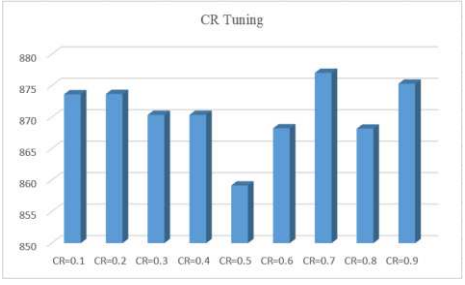


Fig. 3. CR Tuning

We notice that the impact of Crossover parameter (CR) on system performance is very remarkable. The best performance is for CR = 0.7.

- W Tuning: We tested here 8 different values, from possible differential weight domain [0, 2], starting from 0.4 and adding at each step 0.2 to stop at 1.8. We decided to eliminate two non-significant intervals ([0, 4[and]1.8, 2]) from our tuning tests for differential weight parameter and to use discrete values.

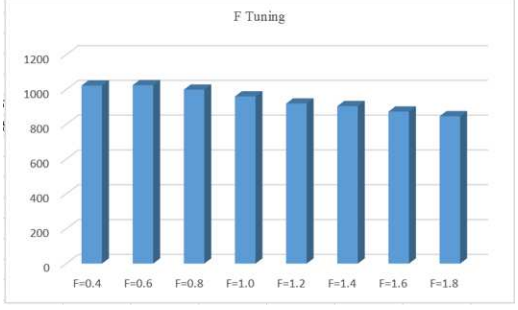


Fig. 4. W Tuning

The previous differential weight tuning histogram shows a slight advance of the value (0.6) compared to (0.4) and (0.8), and it will be used as the suitable F value for the DE

algorithm during the rest our experimentation.

- Nunmer of Population Tuning: in order to identify the best population size, we compare between these systems derived according to their NP belonging to {40, 50, 60, 70, 80, 90, 100}.

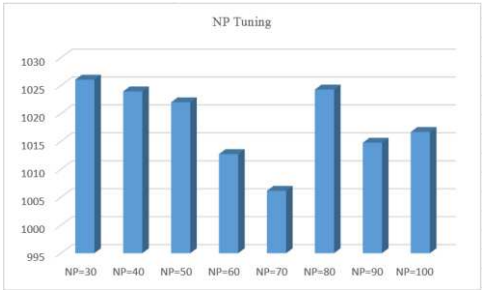


Fig. 5. NP Tuning

The previous NP tuning histogram shows an advance of the value (30) compared to the others.

- **PSO TUNING**

The PSO is influenced by several control parameters (inertia weight: w, acceleration coefficients: c1 and c2, the number of iterations or generation) that we must tune, to improve the performance of the algorithm.

- **W TUNING**

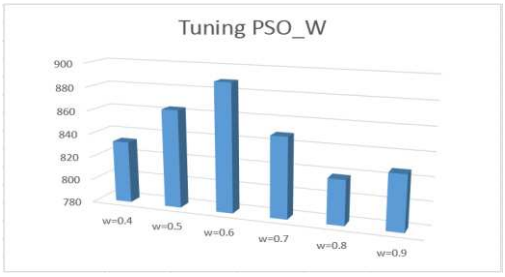


Fig. 6. W Tuning

The previous W tuning histogram of the PSO metaheuristic shows that the system performance reaches his best results when using the w value of 0.6 compared to {0.4,0.5,0.6,0.7,0.8,0.9}, and so it will be used as the suitable w value during the rest our experimentations with PSO.

- **ACCELERATION COEFFICIENTS C1 AND C2 TUNING**

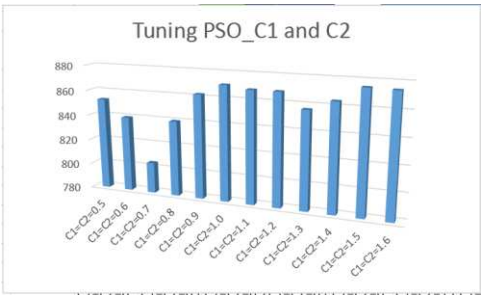


Fig. 7. C1 and C2 Tuning

The previous tuning result histogram of Acceleration coefficients C1 and C2 of the PSO metaheuristic shows that the system performance reaches his best results when using the value of 1.5 compared to {0.6,0.7,0.8,0.9,1.0,1.1,1.2,1.3,1.4, 1.5,1.6}, and so it will be used as the suitable value (C1=C2=1.5) during the rest our experimentations with PSO.

- **NUMBER OF GENERATION TUNING**

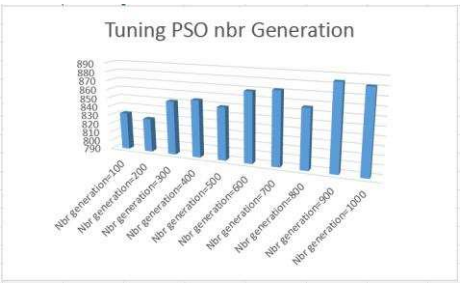


Fig. 8. Number of generation Tuning

A result of tuning the number of generation needed for the PSO shows that the higher system performance is reached at the value of 900 generations.

TUNING RESULTS

As we found previously, the DE/Best/1 and PSO parameter resulted of the tuning phase and that give the best results in our problem configuration and will be used in the all rest of our experimentations, these parameters are summarized in the Table 2 below:

TABLE 2 LIST OF PARAMETERS USED FOR PSO AND DE/BEST/1

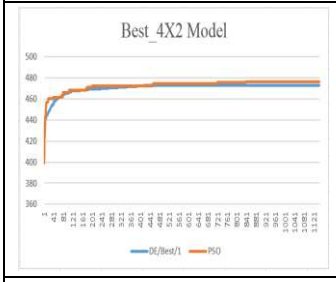
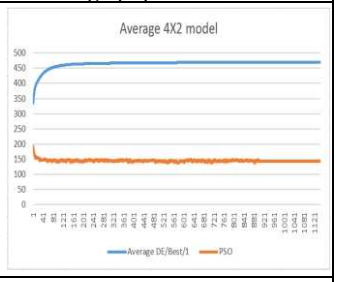
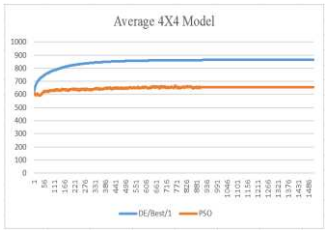
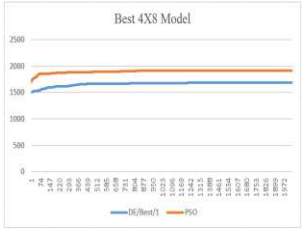
DE/Best/1 Tuning Results		PSO Tuning Results	
Parameter	Best Tuned value	Parameter	Best Tuned value
CR	0.7	W	0.6
F	0.6	C1=C2	1.5
NP	30	NP	30*
N. generation	Stop stability condition*	N generation	900

4.2 DE AND PSO COMPARISON

We treated here 3 different problem configurations (3 models) and we present below the results of DE/Best/1 and PSO showed in the table below. This table contains respectively the results of the best and the average system fitness value realized by each algorithm corresponding to the three configurations mentioned.

BEST AND AVERAGE RESULTS OF DE/BEST/1 AND PSO SYSTEM FITNESS VALUE

TABLE 3: DE/BEST/1 AND PSO RESULTS

Best individual results	Average population results
	
	
	

The curves show that PSO always takes the top compared to DE in all models in the case of the best individual results, which proves the advance of PSO compared to DE. On the other hand, the average value achieved by individuals of PSO population is always lower than the others of DE.

4.3 SYSTEM PERFORMANCE STUDIES

We study in this section the impact of two different parameters (fixed cost transshipment and the demand uncertainty) on the system performance. We chose also to use two demand types distribution (uniform demand distribution and normal demand distribution). We present below the results of these experimentations:

DEMAND IMPACT ON SYSTEM PERFORMANCE

We present below the results of the study of the impact of the variation of normal and uniform demand distribution on the system performance:

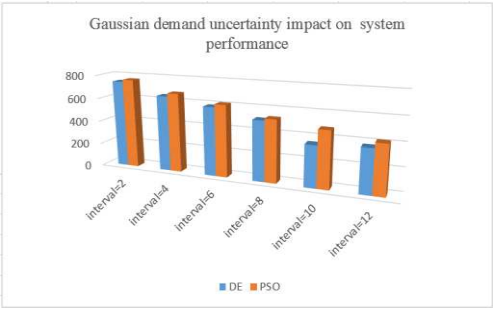


Fig. 9. Impact of normal demand on system performance

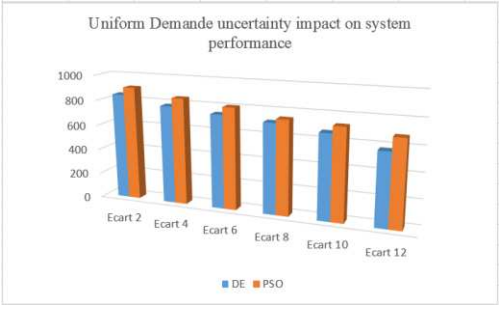


Fig. 10. Impact of uniform demand on system performance.

Figure 9 and Figure 10 show that the impact of normal and uniorm demand uncertainty on system performance is very remarkable. the more the uncertainty of the demand increases the more it causes a decrease of the system performance and affects the result of the fitness function regardless of the algorithm. we also notice a rapid decrease in results in the case of normal demand faster than in the case of uniform demand, this proves that our system depends directly on the nature of demand and its domain. We also note that the performance of the PSO is always better than that of the DE even with the variation of the uncertainty of the demand in both normal and uniform cases.

5. CONCLUSION

In this paper, we studied the multi-location and multi-item inventory management considering lateral transshipment. we proposed two evolutionary algorithms (DE/best/1 and PSO), to resolve the studied problem. The model considered many items which could be transshipped between locations. We proposed a simulation algorithm to derive the expected profit value of a replenishment vector. our experimental study shows that the PSO algorithm performs better results than those of DE. We studied the uncertainty demand effect on the system performance. We noted that the transshipment impact on the system performance is more significative for high level of demand uncertainty. Our actual research could be extended by:(1) considering the fixed cost of transshipment and (2) integrating a substitution flexibility technique allowing to customer to replace their first choice by an alternative one when it is out of stock.

REFERENCES

[1] Paterson, C., Kiesmüller, G., Teunter, R., and Glazebrook, K., Inventory models with lateral transshipments: A review, European Journal of Operational Research, 210, pp. 125–136. 2011.

[2] Seidscher, A., & Minner, S. A Semi-Markov decision problem for proactive and reactive transshipments between multiple warehouses. European Journal of Operational Research, 230 (1): 42-52. 2013.

[3] Krishnan, K.S. and V. R. K. Rao. Inventory Control in N Warehouses. Journal of Industrial Engineering, 16, 212-215. 1965.

[4] Olsson F. Emergency Lateral Transshipments in a Two-Location Inventory System with Positive Transshipment Leadtimes. European Journal of Operational Research. 2015;242(2):424-433. https://doi.org/10.1016/j.ejor.2014.10.015

[5] Yao, D., Zhou, S. X. and Zhuang, W. 'Joint initial stocking and transshipment - asymptotics and bounds', Production and Operations Management 25(2), 273–289. 2016.

[6] Pingping Feng,Feng Wu, Richard Y. K. Fung, Tao Jia, Wei Zong. The order and transshipment decisions in a two-location inventory system with demand forecast updates. Computers & Industrial Engineering 135. DOI: 10.1016/j.cie.2019.04.043. 2019.

[7] KÖCHEL, P. Retrospective optimisation of a two-location inventory model with lateral transshipments. In Proceedings of the 2nd International Conference on Traffic Science – ICTS '98, 129-139. 1998.

[8] Kochel, P. & Nielander, U. Simulation-based optimisation of multi-echelon inventory systems, International Journal of Production Economics 93-94:505-513, DOI: 10.1016/j.ijpe.2004.06.046. . 2005

[9] Miao Z., Fu K., Fei Q., Wang F. Meta-heuristic Algorithm for the Transshipment Problem with Fixed Transportation Schedules. In: Nguyen N.T., Borzemski L., Grzech A., Ali M. (eds) New Frontiers in Applied Artificial Intelligence. IEA/AIE 2008. Lecture Notes in Computer Science, vol 5027. Springer, Berlin, Heidelberg. 2008.

[10] Hochmuth, C. A. and K¨ochel, P.. How to order and transship in multi-location inventory systems: The simulation optimization approach. International Journal of Production Economics, 140:646–654. 2012.

[11] N. Danloup, H. A. (2018). A comparison of two meta-heuristics for the pickup and delivery problem with transshipment. Computers & OR 2018 volume 100, 155-171.

[12] Price, R. S. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization, 11:341 -359. 1997.

[13] Eberhart, R. C., & Kennedy, J.. A new optimizer using particle swarm theory. In *Proceedings of the sixth international symposium on micro machine and human science*(pp. 39–43), Nagoya, Japan. Piscataway: IEEE. 1995.