

www.revue-eti.net

Editorial du numéro 8	p. 3
Les comités de la 8 ^{ème} édition de la revue eTi	p. 4
Etat de l'art	
Spécification des besoins de performabilité dans les modèles BPMN, <i>Fatima-Zahra Belouadha</i>	p. 7
Recherche	
Le système d'intelligence économique - XEW, <i>A. El Haddadi, A. El Haddadi, B. Dousset & A. Fennan</i>	p. 20
Proposition of a new ex-ante evaluation method for adapting and selecting ERP solutions, <i>A. Khaled & M. Abdou Janati Idrissi</i>	p. 35
Usage et formation	
L'art de cultiver les perles rares: changer l'image véhiculée des professionnels en TI, <i>M.-C. Roy, H. Khechine & D. Ducharme</i>	p. 50
Dossier spécial ArabWIC'2015	
Trust framework for a secured routing in wireless sensor network, <i>O. Hoceini, S. Talbi & R. Aoudjit</i>	p. 61
Translating Discrete Time SIMULINK to SIGNAL, <i>S. Messaoud, N. Saeedloei & S. Shukla</i>	p. 69
Fiche équipe	
Laboratoire d'Analyse des Systèmes, Traitement de l'Information et Management Intégré, <i>N. Berbiche & Jamila El Alami</i>	p. 85
J'ai lu	
Ingénierie des exigences. Méthodes et bonnes pratiques pour construire et maintenir un référentiel de <i>S. Badreau & J.-L. Boulanger</i>	p. 89
Appel aux articles	
Appel aux articles pour la 9 ^e édition	p. 91

Editorial du numéro 8

Les orientations dégagées en 2014 telles que le bilinguisme, l'édition progressive ou la publication en format papier se sont révélées positives, aussi les avons-nous maintenues. Cette 8ème édition a introduit la rubrique **Usages et formation** car nous avons jugé important de soutenir les recherches dans ce domaine.

Le premier article présenté est une revue de la littérature proposée par Fatima-Zahra Belouadha concernant des extensions du standard BPMN qui visent à compléter la modélisation des processus métier par la capture des aspects non fonctionnels, en particulier la performance et la fiabilité. Ces informations sont destinées à être exploitées pour la génération de solutions optimisées.

La rubrique **Recherche** est composée de deux articles ayant pour thème l'informatique décisionnelle :

- Le premier, de Anass El Haddadi *et al.*, présente le système d'intelligence économique XEW (Xplor EveryWhere) qui permet d'extraire et d'analyser des informations concernant le contexte économique d'une entreprise, à partir des diverses sources de l'Internet. XEW offre une large gamme de services allant de l'extraction des informations (sourcing et fusion multimodale, entrepôt de données SQL et NoSQL, etc.) à leur analyse, avec des outils tels que Big Data Analytic.
- Le deuxième article, proposé par Abdelillah Khaled et Mohammed Abdou Janati, décrit l'approche «Systematic EVALuation of ERP Systems» pour l'évaluation et la sélection a priori de solutions ERPs potentielles. SERVALEPS prend en considération non seulement leur adéquation aux besoins de l'organisation, mais également l'effort, le coût et le risque de l'adaptation.
-

Nous avons initié la rubrique **Usage et formation** avec l'article de Marie-Christine Roy, Hager Khechine et Diane Ducharme. Les auteures analysent les perceptions mitigées, voire erronées, de la fonction d'informaticien et dégagent des recommandations pour promouvoir une vision plus réaliste d'une profession marquée par un déficit régulier des effectifs.

La rubrique **Dossier Spécial** est composée de deux articles de la conférence «Women in Computing in the Arab World» organisée par l'université de New York à Abu Dhabi, les 17 et 18 mai 2015. Sa particularité est de promouvoir la contribution féminine arabe en informatique.

Safa Messaoud, Neda Saeedloei et Sandeep Shukla étendent, sur le plan sémantique, le formalisme de Simulink par des mécanismes de gestion asynchrone de la concurrence et la génération de code déterministique *multi-thread*. Ouassila Hocein, Saïd Talbi et Rachida Aoudjit proposent une architecture hiérarchique pour le routage sécurisé dans un réseau de capteurs sans fils. L'objectif est de réduire le coût d'évaluation de la confiance et d'assurer une meilleure sélection de chemins sécurisés qui mènent vers la station de base.

Dans la rubrique **Fiche équipe**, c'est le laboratoire LASTIMI de l'EST de Salé qui est à l'honneur dans cette 8ème édition, présenté par Naoual Berbiche et Jamila El Alami. Quant à la rubrique **J'ai lu**, elle introduit l'ouvrage intitulé «Ingénierie des exigences» de Stéphane Badreau et Jean-Louis Boulanger, édité chez Eyrolles en 2013.

Nous vous présentons nos meilleurs vœux pour l'année 2016 et merci à nos lecteurs et à tous nos partenaires.

Mounia Fredj et Ounsa Roudiès
Rédactrices en chef de la 8ème édition

Les comités de la 8^{ème} édition de la revue eTI

Directrices de publication

- ROUDIES Ounsa EMI, Univ. Mohammed V-Rabat, Maroc
- CHIADMI Dalila EMI, Univ. Mohammed V-Rabat, Maroc

Comité scientifique

- BELLAACHIA Abdelghani George Washington University, USA
- GIRAUDIN Jean-Pierre LIG, Univ. Pierre-Mendès France, France
- ROLLAND Colette Université Paris-1 Sorbonne, France

Rédactrices en chef

- FREDJ Mounia ENSIAS, Univ. Mohammed V-Rabat, Maroc
- ROUDIES Ounsa EMI, Univ. Mohammed V-Rabat, Maroc

Comité de rédaction

- AHMED-NACER Mohamed USTHB, Algérie
- ALIMAZIGHI Zahia USTHB, Algérie
- AMGHAR Mustapha GIE Galileo, Maroc
- BAINA Karim ENSIAS, Univ. Mohammed V-Rabat, Maroc
- BAINA Salah ENSIAS, Univ. Mohammed V-Rabat, Maroc
- BAKKOURY Zohra EMI, Univ. Mohammed V-Rabat, Maroc
- BELOUADHA Fatima-Zahra EMI, Univ. Mohammed V-Rabat, Maroc
- BENHLIMA Leila EMI, Univ. Mohammed V-Rabat, Maroc
- BOUNABAT Bouchaïb ENSIAS, Univ. Mohammed V-Rabat, Maroc
- CHIADMI Dalila EMI, Univ. Mohammed V-Rabat, Maroc
- COHEN Atika ULB, Belgique
- COULETTE Bernard IRIT, Univ. Toulouse Jean Jaurès, France
- DIOURI Ouafae EMI, Univ. Mohammed V-Rabat, Maroc
- ELEULDJ Mohcine EMI, Univ. Mohammed V-Rabat, Maroc
- EL MAGHRAOUI Kaoutar Thomas J. Watson Research Center, IBM, USA
- FREDJ Mounia ENSIAS, Univ. Mohammed V-Rabat, Maroc
- FRONT Agnès LIG, Univ. Grenoble Alpes, France
- GIRAUDIN Jean-Pierre LIG, Univ. Grenoble Alpes, France
- KASSOU Ismail ENSIAS, Univ. Mohammed V-Rabat, Maroc
- MEJRI Mohamed Faculté des Sciences et de Génie, Univ. Laval, Canada
- MOULINE Salma FSR, Univ. Mohammed V-Rabat, Maroc
- OULAD HAJ THAMI Rachid ENSIAS, Univ. Mohammed V-Rabat, Maroc
- RIEU Dominique LIG, Univ. Grenoble Alpes, France
- ROLLAND Colette CRI, Univ. Paris1-Sorbonne, France
- ROUDIES Ounsa EMI, Univ. Mohammed V-Rabat, Maroc
- SALINESI Camille CRI, Univ. Paris1-Sorbonne, France
- TAMZALIT Dalila LINA, Univ. de Nantes, France
- TARI Zahir School of Computer Science, Univ. RMIT, Australie

Comité de lecture

- AJHOUN Rachida ENSIAS, Univ. Mohammed V-Rabat, Maroc
- BAINA Salah ENSIAS, Univ. Mohammed V-Rabat, Maroc
- BAH Slimane EMI, Univ. Mohammed V-Rabat, Maroc
- BOUNABAT Bouchaïb ENSIAS, Univ. Mohammed V-Rabat, Maroc
- CARON Marie-Laurence IAE, Université de Grenoble Alpes, France
- CARTON Sabine IAE, Université de Grenoble Alpes, France
- CHIADMI Dalila EMI, Univ. Mohammed V-Rabat, Maroc
- DIOURI Ouafae EMI, Univ. Mohammed V-Rabat, Maroc
- ELEULDJ Mohcine EMI, Univ. Mohammed V-Rabat, Maroc
- EL MAGHRAOUI Kaoutar Thomas J. Watson Research Center, IBM, USA
- EL KETTANI Mohamed Dafir ENSIAS, Univ. Mohammed V-Rabat, Maroc
- FREDJ Mounia ENSIAS, Univ. Mohammed V-Rabat, Maroc
- GUENNOUN Zouhair EMI, Univ. Mohammed V-Rabat, Maroc
- KASSOU Ismail ENSIAS, Univ. Mohammed V-Rabat, Maroc
- KJIRI Laïla ENSIAS, Univ. Mohammed V-Rabat, Maroc
- MELARD Guy Université Libre de Bruxelles, Belgique
- MESSOUSSI Rochdi FSK, Univ. Ibn Tofaïl, Maroc
- MOULINE Salma FSR, Univ. Univ. Mohammed V-Rabat, Maroc
- ROUDIES Ounsa EMI, Univ. Mohammed V-Rabat, Maroc

Comité logistique

- BENS Aid Hicham, INPT : administration de la plateforme
- MEDDOUN Judith : relecture linguistique
- MOKHTARI Abdelkrim, Univ. Ibn Tofaïl, Maroc : relecture linguistique
- RAPOSO de BARBOSA Appoline : mise en page et infographie

Partenaires



Etat de l'art

Spécification des besoins de performabilité dans les modèles BPMN

Une revue de la littérature

Fatima-Zahra Belouadha

Equipe Siweb, École Mohammadia d'Ingénieurs, Université Mohammed V - Rabat, BP. 765, Agdal-Rabat, Maroc

belouadha@emi.ac.ma

Résumé

La performance du système d'information dépend de l'efficacité opérationnelle des processus métier implémentés, mais aussi de leur capacité à répondre aux besoins non-fonctionnels, en particulier, celui de performabilité (performance et fiabilité). En effet, la phase de modélisation des processus métier est une phase primordiale du cycle BPM (Business Process Management). Néanmoins, la notation BPMN (Business Process Model and Notation) qui est le standard de facto de la modélisation des processus métier, se limite à la modélisation contemplative de leur comportement, permettant aux développeurs de capturer leurs besoins fonctionnels. Dans le cadre d'une approche de modélisation productive, l'intégration des besoins de performabilité dès la conception des processus métier est primordiale pour permettre leur analyse automatique et prédire leur qualité de service. Cet article classe et compare des contributions parmi les plus référencées ou récentes dans ce domaine. L'objectif est de mettre en évidence leurs limites et d'identifier les critères d'une approche efficace.

Abstract

The information system performance depends on the operational efficiency of the implemented business processes, but also on their ability to meet the non-functional requirements, especially those of performability (performance and reliability). Indeed, the modelling phase of business processes is a critical phase within the BPM (Business Process Management) lifecycle. Nevertheless, the notation BPMN (Business Process Model and Notation) which is the de facto standard for business process modelling, is limited to the contemplative modelling of their behavior enabling developers to capture their functional requirements. As part of a productive modelling approach, integrating performability requirements of business processes at design-time is essential to enable their automatic analysis and to predict their quality of service. This article classifies and compares contributions among the most recent or referenced in this field. The aim is to highlight their limitations and to identify the criteria of an efficient approach.

Mots-clés

BPMN, Modélisation des processus métier, Besoins de performabilité, QoS, Annotation

Keywords

BPMN, Business Process modeling, Performability Requirements, QoS, Annotation.

1. Introduction

L'efficacité des processus métier dépend de leur efficacité opérationnelle (capacité de répondre aux besoins fonctionnels), elle-même optimisée par leur efficacité non-fonctionnelle (capacité de répondre aux besoins non-fonctionnels). Rappelons que les besoins fonctionnels expriment ce que le processus métier doit faire alors que les besoins non-fonctionnels expriment la manière de le faire. En fait, il n'y a pas de définition exhaustive des besoins non-fonctionnels dans la littérature. Ils correspondent à titre d'exemple, à des critères tels que la performance, la fiabilité, la sécurité, la disponibilité, la maintenabilité et la portabilité. Nous pouvons constater qu'ils traduisent la ou les politiques adoptées pour implémenter ou configurer un processus métier, y compris les propriétés techniques de son implémentation et le contexte de son utilisation. Ces politiques jouent un rôle crucial lors du développement du système d'information. Elles guident les choix et les décisions prises lors de son implémentation. Comme le confirment Mylopoulos, Chung *et al.* (Mylopoulos, Chung *et al.*, 1992), toute omission ou erreur au niveau des besoins non-fonctionnels est généralement reconnue parmi les erreurs les plus coûteuses et difficiles à corriger, une fois que la mise en place du système d'information est achevée. S'ajoute à cela que la gestion des processus métier (BPM pour Business Process Management) doit de préférence être réalisée au niveau de la conception métier et non pas de l'implémentation technique (Smith *et al.*, 2002). Aussi, la description des propriétés non-fonctionnelles des processus au moment de leur modélisation s'avère-t-elle nécessaire pour contribuer à assurer leur efficacité.

La spécification des propriétés relatives à la performabilité (performance et fiabilité) des processus métier au moment de la conception est en fait importante pour aligner les objectifs de performance et de fiabilité avec les objectifs métier. Elle s'inscrit dans le cadre d'une approche de modélisation productive qui pourrait conduire l'analyse automatique et anticipée des processus. Néanmoins, la notation BPMN (*Business Process Model and Notation*) (OMG-BPMN, 2013), tout comme d'autres langages de modélisation de processus, s'inscrit dans le cadre de l'approche de modélisation contemplative, plutôt que productive (Belouadha, Omrana *et al.*, 2012) ; (Belouadha, Omrana *et al.*, 2014). Elle se limite à la description des besoins fonctionnels des processus métier à des fins de documentation et de communication.

Pour pallier cette limite, des approches d'extension du standard BPMN pour l'intégration des propriétés de performabilité ont été proposées dans la littérature. Cet article présente une revue de ces approches. Il rappelle les concepts et les apports du standard BPMN avant d'exposer les fondements et les principes des approches étudiées et d'en présenter une analyse comparative. L'objectif est d'en dégager les limites et de tirer des conclusions quant aux qualificatifs d'une approche adéquate pour l'extension du BPMN avec des aspects relatifs à la performabilité.

2. Standard BPMN : concepts et apports

La notation BPMN est le standard de facto de modélisation des processus métier. Cette section décrit ses principaux concepts, les illustre par un exemple et présente ses apports.

2.1. Principaux concepts

Le standard BPMN est une notation graphique dédiée à la modélisation des processus métier. Il a été développé à l'initiative de l'organisation BPMI (*Business Process Management Initiative*) en 2004 et adopté comme standard OMG (*Object Management Group*) en 2006. La version 2 du standard BPMN (OMG-BPMN, 2013) considère cinq concepts principaux pour la modélisation des processus : les *swimlanes*, les objets de flux (*flow objects*), les connecteurs (*connecting objects*), les artefacts et les objets de données (*data objects*).

Les *swimlanes* (*pools et lanes*). Ils permettent de créer des partitions d'activités pour un objectif d'organisation des processus.

- Un *pool* constitue un conteneur d'activités réalisées par un participant intervenant dans un processus collaboratif.
- Les *lanes* sont des couloirs qui permettent d'organiser ces activités à l'intérieur du *pool*.

Les objets de flux. Ce sont des nœuds du graphe qui comportent les activités (tâches ou sous-processus), les événements et les branchements (*gateways*).

- Les tâches sont des activités atomiques et les sous-processus sont des activités composées.
- Un sous-processus est noté par un objet graphique qui en représente une vue réduite (*collapsed*) ou développée (*expanded*). À l'encontre de la vue réduite qui cache les détails de modélisation d'un sous-processus, la vue développée montre ces détails en mettant en relief des activités, des branchements, des événements et des flux de séquence. C'est un conteneur d'objets de flux qui peut être constitué de 0 ou plusieurs ensembles de couloirs (*lanes*).

- Un évènement représente quelque chose qui se produit et affecte le cours du processus. Il a souvent une cause et un résultat et peut débuter, interrompre ou terminer un flux.
- Les branchements (*gateways*) sont des éléments de contrôle de flux des activités. Ils sont répartis en branchements exclusif (*exclusive gateway*), inclusif (*inclusive gateway*), parallèle (*parallel gateway*), à base d'évènement (*event-based gateway*) et complexe (*complex gateway*) pour représenter différents branchements conditionnels tels que les fork, join et merge.

Les connecteurs. Ils représentent des flux d'information qui constituent des flux de séquence, des flux de message, des associations ou des associations de données (*data associations*).

- Les flux de séquence enchaînent les activités à l'intérieur d'un couloir (*lane*) pour montrer l'ordre dans lequel elles seront exécutées. Ils lient des activités, des évènements et des gateways et peuvent être conditionnels.
- Les flux de message représentent les messages échangés entre des *pools*.
- Les associations lient des données ou des artefacts à des objets du flux. Les associations de données sont utilisées pour modéliser un flux de données entre des objets de données et des activités ou des évènements (ils peuvent, par exemple, associer des entrées ou des sorties à une activité). Les associations d'artefacts, quant à elles, relient à titre d'exemple, une annotation (texte défini par l'utilisateur) à un élément graphique.

Les artefacts. Ils apportent des informations additionnelles à propos du processus. Ils peuvent être des groupes (*groups*), des annotations textuelles ou encore des artefacts définis au besoin.

- Les groupes sont utilisés pour mettre en relief des ensembles d'activités qui peuvent appartenir à différents *lanes* et/ou *pools*.
- Les annotations textuelles sont associées aux éléments du diagramme BPMN pour les décrire.

Les objets de données. Ces objets désignent généralement soit des données nécessaires pour exécuter les activités, soit des données que les activités produisent (*business documents*, e-mails, etc.). Ainsi, ils peuvent être de types *data inputs*, *data outputs*, *data stores* ou messages.

- Les *data inputs* et *data outputs* représentent respectivement les entrées/sorties d'un processus et peuvent être associés à ses activités par des associations de données.
- Les *data stores* désignent des emplacements où les activités peuvent chercher ou mettre à jour des données stockées (par exemple, des bases de données.).
- Les messages décrivent le contenu d'une communication entre deux participants.

La figure 1 illustre une vue d'un exemple abstrait de modèle de processus métier en BPMN.

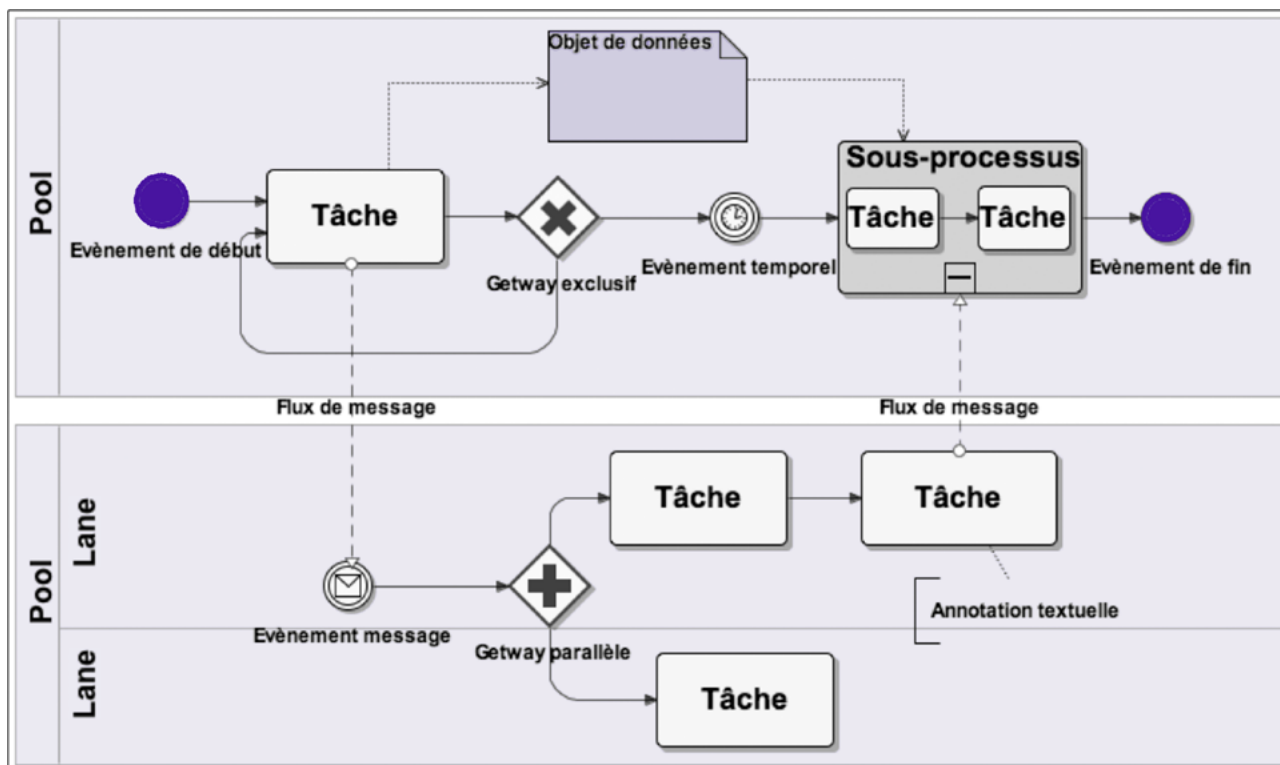


Figure 1. Vue d'un exemple abstrait de modèle de processus métier en BPMN

2.2. Exemple de processus BPMN

La figure 2 illustre un exemple concret de processus modélisé en BPMN. Ce processus d'achat d'articles (produits) est déclenché par le client qui commence par choisir autant d'articles qu'il le souhaite. Le sous-processus *Choisir un article* est en fait un sous-processus *collapsed* qui constitue une boucle consistant en deux tâches : *Sélectionner un article* et le *Rajouter au panier*. Après avoir choisi les articles désirés, le client doit s'authentifier. S'il arrive à le faire avec succès, il devra ensuite payer sa commande, sinon il devra créer d'abord un compte avant de pouvoir effectuer le paiement. La tâche *Créer un compte* a comme entrée les informations du compte à créer (*login*, mot de passe et autres informations y compris le numéro de compte bancaire). Une fois que le paiement est effectué, la date de livraison est retournée et le magasinier de l'agence de vente concernée reçoit la référence du paiement. Suite à cela, il devra préparer la commande. Ensuite, il devra mettre à jour l'inventaire des produits parallèlement à la livraison de la commande. Cette opération est effectuée par le responsable de transport et se concrétisera par l'envoi du bon de livraison au client. Celui-ci, ayant été en attente de la date de livraison, devra signer le bon de livraison une fois qu'il le reçoit à la date prévue. Ainsi, le processus d'achat d'articles prendra fin.

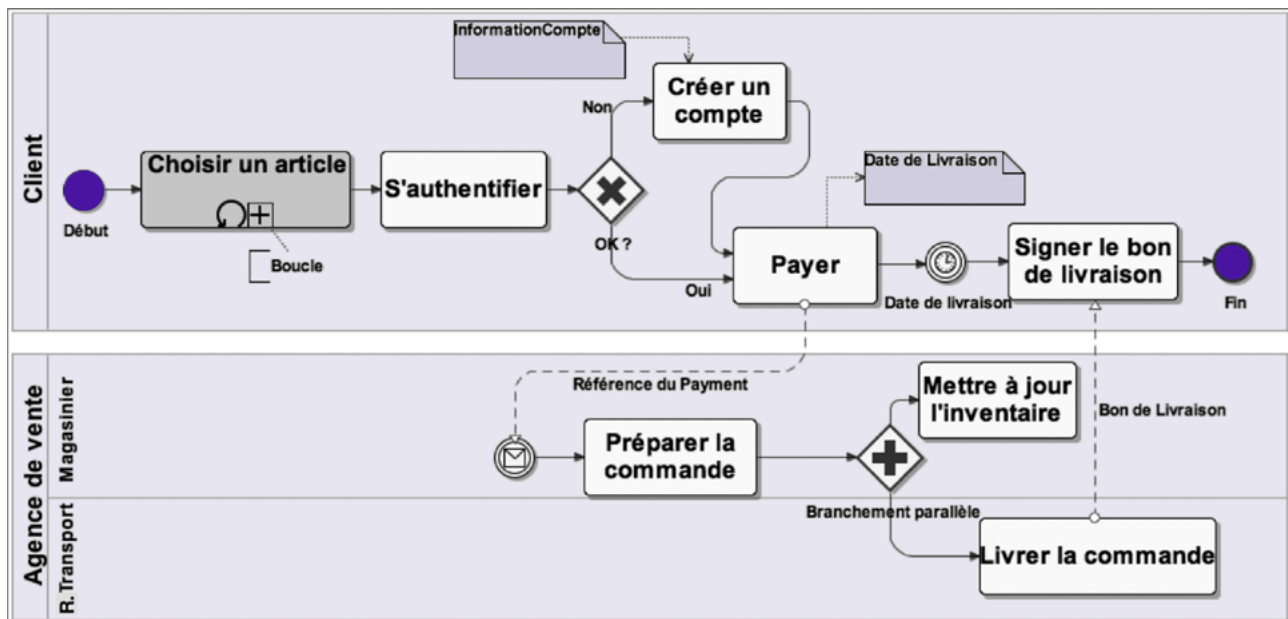


Figure 2. Exemple de processus modélisé en BPMN

2.3. Apports de la notation BPMN

La notation BPMN présente d'importants apports (Owen et Raj, 2003). Parmi lesquels figurent :

La modélisation collaborative. BPMN pallie le problème de la difficulté de communication entre les analystes métier et les experts de technologies de l'information qui avaient à traiter les processus, moyennant des outils et des concepts différents. Il est conçu pour être utilisé par toutes les parties impliquées dans le cycle de vie des processus, à savoir les analystes métier qui modélisent les processus, les experts en sécurité qui définissent les mécanismes adéquats pour les sécuriser, les développeurs techniques qui prennent en charge leur implémentation et l'équipe métier qui gère et contrôle ces processus.

Le mapping automatique à BPEL. BPMN permet la transition directe et standardisée de la conception des processus métier à leur implémentation. Il est automatiquement mappable au standard d'exécution BPEL (*Business Process Execution Language*) (OASIS-BPEL, 2007). En plus, il est compatible avec le langage de requêtes BPQL (*Business Process Query Language*) (Owen et Raj, 2003) qui constitue une interface de gestion standard pour le déploiement et l'exécution des processus métier.

La simplicité et l'expressivité. BPMN est conçu pour être simple et facile à utiliser et à comprendre. Il est cependant expressif et permet de décrire des processus métier complexes. Il propose, entre autres, des notations spécifiques pour décrire en particulier les événements à base de messages et la transmission de messages entre des organisations.

La description consolidée. BPMN représente un processus métier par un seul diagramme. Toutes les spécifications décrivant un processus ne se trouvent pas fragmentées dans plusieurs diagrammes à l'instar de UML (*Unified Method Language*).

3. Spécification des besoins de performabilité dans les modèles BPMN

De nombreux chercheurs se sont intéressés à l'extension des modèles BPMN avec des propriétés de performabilité (Korherr et List, 2007) (Magnani et Montesi, 2007) (Pavlovski et Zou, 2008) (Sampath et Wirsing, 2009) (Saeedi, Zhao *et al.*, 2010) (Bartolini *et al.*, 2013) (Bocciarelli et d'Ambrogio, 2014). Les approches proposées pour l'intégration des propriétés de performabilité dans les modèles de processus BPMN peuvent être classées en quatre catégories selon l'approche d'annotation adoptée : approches orientées artefacts, approches orientées tables, approches orientées annotation textuelle et approches orientées instances UML. Cette section décrit les fondements et les principes de chacune de ces classes d'approches à travers quatre contributions parmi les plus référencées ou les plus récentes dans le domaine. Une analyse comparative de ces contributions est présentée à la fin de la section.

3.1. Approche d'annotation orientée artefacts

La contribution proposée par (Pavlovski et Zou, 2008) constitue une référence dans le domaine. Elle adopte une approche d'annotation fondée sur l'utilisation des artefacts pour annoter les processus avec des propriétés non-fonctionnelles dont les celles de performabilité.

Principe général. Cette contribution est fondée sur deux principes :

- Enrichir les modèles BPMN avec les besoins non-fonctionnels à capturer au moment du développement des systèmes.
- Proposer des éléments de description contribuant à atténuer les risques associés aux contraintes non-fonctionnelles sans toutefois compliquer les modèles de processus.

Aspects considérés. Elle considère des propriétés non-fonctionnelles telles que les politiques de sécurité, les propriétés de performabilité, les politiques organisationnelles et les contraintes réglementaires.

Mécanisme d'annotation. Les propriétés considérées servent à annoter des tâches de processus.

Les auteurs étendent BPMN avec deux nouveaux artefacts (figure 3) :

- *Operating Condition* utilisé pour introduire une contrainte. Il est décrit par un attribut indiquant le type de cette contrainte (contrainte organisationnelle, politique de sécurité, contrainte de performance, etc.).
- *Control Case* utilisé en vue de définir les contrôles métier à mettre en place pour atténuer le risque associé à une contrainte. Il est décrit au moyen d'un ensemble d'attributs catalogués sous format texte sans être introduits dans les diagrammes de processus. Ces attributs sont : le nom de l'artefact *Operating condition* auquel l'artefact *Control Case* est lié, sa description textuelle, la contrainte métier ou technique qu'il exprime, le risque métier qui lui est associé et les contrôles métier à appliquer.

La figure 4 et le tableau 1 montrent l'exemple d'une tâche annotée, tiré d'un exemple donné par (Zou et Pavlovski, 2006). La tâche *FlightReservation* est annotée par une politique de performabilité en vue d'éviter un temps de réponse inacceptable. L'objectif est d'atténuer le risque de perte des clients utilisateurs du système proposant la réservation de billets d'avion.

Cette politique recommande trois mesures (l'utilisation des techniques d'équilibrage de charge (load balancing) et de clustering, l'utilisation d'un cache et la négociation du contrat de la bande passante pour le réseau).

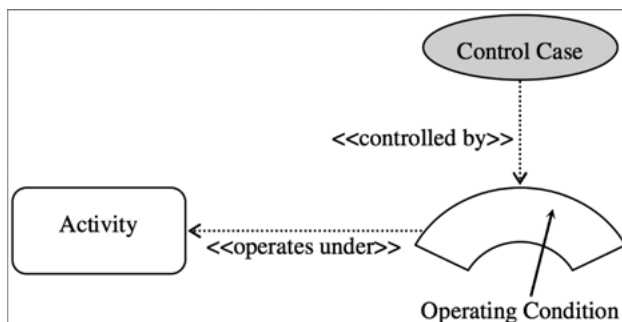


Figure 3. Artefacts *Operating Condition* et *Control Case* (Pavlovski et Zou, 2008)

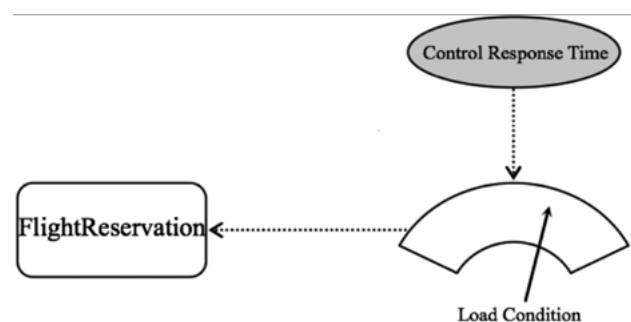


Figure 4. Exemple d'une tâche annotée par une politique de performabilité

Control Case: Temps de réponse	
Operating Condition :	Surcharge due à un accès concurrent (simultané) des utilisateurs.
Description :	Suite à une surcharge due à un accès concurrent de plusieurs utilisateurs à la fois, le temps de réponse du système peut devenir inacceptable. Ce <i>control case</i> définit le temps de réponse maximum des transactions afin d'atténuer le risque de perte des clients.
Contrainte technique :	La durée de la transaction dépend de la bande passante Internet et de l'infrastructure du réseau ISP.
Risque métier :	Satisfaire 90% des utilisateurs en un temps de réponse dépassant 10 secondes exposera l'entreprise à une perte de 20% de ses clients en faveur des entreprises concurrentes.
Contrôles métier :	Pour atténuer le risque, il faut utiliser les techniques d'équilibrage de charge (<i>load balancing</i>) et de <i>clustering</i> , utiliser un cache et négocier le contrat de la bande passante pour le réseau ISP.

Tableau 1. Exemple de spécification des attributs d'un artefact Control Case

Conclusion. Dans le cadre d'une approche de modélisation contemplative, l'approche proposée adopte une représentation simple des besoins non-fonctionnels dont les propriétés de performabilité. Pour fournir plus de détails sur ces besoins, elle les complète par des informations additionnelles. Toutefois, ces informations sont représentées de façon informelle.

3.2. Approche d'annotation orientée tables

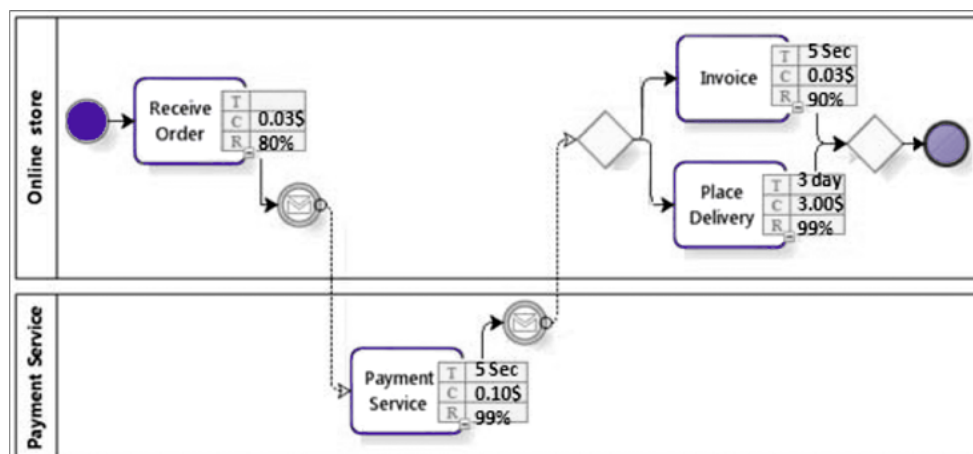
La contribution proposée par (Saeedi, Zhao *et al.*, 2010) adopte une approche d'annotation fondée sur l'utilisation des tables pour l'annotation des propriétés de types temps de réponse, coût et fiabilité.

Principe général. Cette contribution consiste à :

- Décrire les processus BPMN par des propriétés clés de performabilité, souvent exprimées comme préférences d'utilisateur et abordées dans des contrats de service SLA (*Service Level Agreement*).
- Exploiter les propriétés spécifiées pour évaluer la qualité des processus, en utilisant une méthode fondée sur un modèle analytique à base des règles de réduction.

Aspects considérés. Elle considère trois propriétés de performabilité : le temps de réponse, le coût et la fiabilité.

- La notion de temps de réponse réfère à trois composants : le temps de calcul, le temps d'attente (dû par exemple, au délai d'attente d'allocation du processeur selon le rang dans une file d'attente) et le temps de transfert de données qui devient important quand les données transmises sont volumineuses.
- Le coût correspond au coût financier d'installation (*setup*), de déploiement, de gestion, de surveillance et d'exécution du processus.
- La fiabilité est la capacité d'effectuer la tâche attribuée au processus dans une durée donnée. Elle est mesurée en fonction du taux de pannes du système SF (*System Failure rate*) et du taux d'échec du processus PF (*Process Failure rate*).

Figure 5. Exemple d'un processus annoté (Saeedi, Zhao *et al.*, 2010)

Mécanisme d'annotation. Les propriétés considérées sont utilisées pour annoter des éléments de type *SubProcess*, *CallActivity* et *Task*.

Pour annoter une activité, les auteurs utilisent une table qui contient les valeurs correspondant à son temps de réponse, son coût et à sa fiabilité. La figure 5 montre un exemple de processus dont les tâches sont annotées. À titre d'exemple, le temps de réponse de la tâche *Payment Service* est de 5 secondes, son coût est de 0,10 \$ et son taux de fiabilité est de 99%.

Conclusion. Cette approche adopte un mécanisme d'annotation simple pour l'intégration de trois propriétés de performabilité: le temps de réponse, le coût et la fiabilité. Elle s'inscrit dans une perspective de modélisation productive permettant la vérification de la qualité des processus à partir des modèles BPMN. Cette approche a été mise en œuvre à travers l'extension de l'outil *Eclipse Process Modeller*.

3.3. Approche d'annotation orientée instances UML

La contribution proposée par (Bartolini *et al.*, 2013) adopte une approche d'annotation fondée sur un profil UML. Ce profil est utilisé pour l'annotation des processus avec des propriétés non-fonctionnelles dont les propriétés de performabilité.

Principe général. Cette approche permet de surveiller l'exécution des processus en vue de détecter les éventuels défauts de qualité. Elle est fondée sur:

- L'extension Q4BPMN (Quality for BPMN) (Bartolini *et al.*, 2012) qui étend la notation BPMN conformément au métamodèle de propriétés PMM (Property Meta-Model) (di Marco *et al.*, 2011) pour enrichir les spécifications des chorégraphies (celles-ci désignent des interactions entre processus en mettant l'accent sur l'échange entre les participants).
- Les techniques de l'ingénierie dirigée par les modèles pour transformer les modèles Q4BPMN en modèles de qualité conformes à la notation Klaper (*Kernel Language for Performance and Reliability analysis of component-based systems*) et les analyser.

Aspects considérés. Elle considère trois classes de propriétés non-fonctionnelles:

- Les propriétés de performance reliées au nombre de ressources du système, en particulier matérielles.
- Les propriétés de sécurité qui déterminent le niveau de confidentialité (*privacy*) garanti par le système lors de l'échange de données privées.
- Les propriétés de fiabilité qui déterminent le niveau de fiabilité du système.

Mécanisme d'annotation. Les propriétés considérées sont aussi utilisées pour annoter des tâches (*Q4Task*), des acteurs associés à des tâches (*Q4Participants*) et des chorégraphies (*Q4Choreography*).

Des annotations textuelles sont utilisées pour annoter des éléments des diagrammes BPMN avec des instances de classes UML, décrites dans des paquetages qui ne sont pas intégrés à ces diagrammes.

Chaque instance est définie par les attributs suivants:

- *IsHard* pour indiquer si la propriété est obligatoire (contrainte) ou optionnelle (préférence).
- *Metrics* pour spécifier la métrique à laquelle correspond la propriété (par exemple, la durée maximale par défaut).
- *Nature* pour indiquer si la propriété est descriptive (représentant une propriété garantie par le système) ou prescriptive (constituant une propriété requise par le système).
- *PropertyClass* pour spécifier le type de la propriété (performance, sécurité ou fiabilité).
- *Operator*, *Value* et *Unit* pour exprimer une propriété mesurée au moyen d'un opérateur, une valeur et une unité de mesure.

La figure 6 illustre l'exemple d'une tâche annotée par des besoins en performance. La tâche *FlightReservation* est annotée par une propriété qui indique qu'il est préférable (*isHard=false*) que le temps de réponse soit égal à 5000 ms.

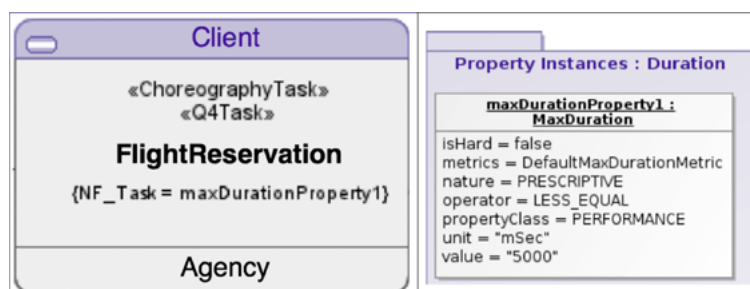


Figure 6. Exemple d'une tâche annotée selon l'extension Q4BPMN

Conclusion. L'extension Q4BPMN s'inscrit dans le cadre d'une approche de modélisation productive visant à détecter les défauts de performabilité et de sécurité à partir des modèles de processus. Elle est implémentée au niveau de l'outil MagicDraw et utilise les outils d'analyse fournis par l'environnement KLAPERSUITE (Ciancone *et al.*, 2011). Néanmoins, cette approche propose une représentation fragmentée des propriétés non-fonctionnelles, en décrivant les modèles BPMN par des instances décrites dans des paquets UML.

3.4. Approche d'annotation orientée annotations textuelles

La contribution proposée par (Bocciarelli et d'Ambrogio, 2014) adopte une approche d'annotation fondée sur l'utilisation des annotations textuelles pour annoter les processus avec des propriétés de performance et de fiabilité.

Principe général. Cette contribution orientée MDA (*Model Driven Architecture*) propose une extension de BPMN, dénommée PyBPMN (*Performability-enabled BPMN*) (Bocciarelli et d'Ambrogio, 2011). Elle est fondée sur deux principes qui consistent à :

- Enrichir le métamodèle BPMN conformément au profil UML MARTE (*Modeling and Analysis of Real-Time and Embedded systems*) (OMG-MARTE, 2011) et ses extensions proposées par (Bernardi, Merseguer *et al.*, 2008) et (Petriu, Bernardi *et al.*, 2008).
- Transformer les modèles PyBPMN en modèles d'analyse de performance et de fiabilité pour prédire la performance des processus et sélectionner les meilleures configurations de services Web pour les implémenter.

Aspects considérés. L'approche proposée permet de spécifier les propriétés de performabilité des processus BPMN. Ces propriétés sont des instanciations de quatre principales métaclasse qui étendent le métamodèle BPMN :

- *PaQualification* permettant de décrire la durée d'exécution (*execution demand*), le rendement (*throughput*) et le temps de réponse.
- *DaError* utilisée pour décrire la fiabilité en termes de latence (temps écoulé entre l'occurrence de l'erreur et sa détection) et de probabilité d'occurrence d'une erreur.
- *DaFault* permettant de spécifier la fiabilité en termes de fréquence d'anomalies (*fault occurrence rate*), de probabilité d'occurrence d'une anomalie (*fault*) et de distribution de probabilité des occurrences des faults.
- *DaFailure* utilisée pour décrire la fiabilité à travers trois attributs : la fréquence des pannes (*failure occurrence rate*), la durée moyenne de fonctionnement sans panne MTTF (*Mean Time To Failure*) et la durée moyenne nécessaire pour la réparation d'une panne MTTR (*Mean Time To Repairation*).

Mécanisme d'annotation. Les propriétés considérées sont utilisées pour annoter les objets BPMN de type *Collaboration* (interaction entre plusieurs processus), *Process* et *task*.

Les auteurs utilisent des annotations textuelles qui spécifient les besoins de performabilité au moyen des attributs/valeurs. L'un des attributs exprime une valeur qui correspond à une fréquence (de panne ou d'exécution, etc.) ou une durée. Les autres attributs fournissent des informations sur cette valeur telles que l'expression de calcul utilisée pour l'obtenir, son origine (calculée, mesurée, requise ou estimée) et son type (maximum, minimum, moyenne, etc.).

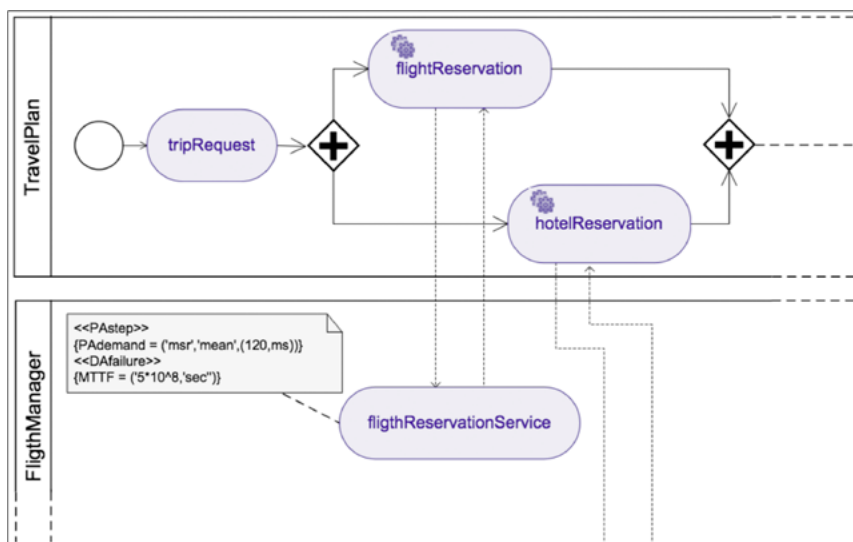


Figure 7. Exemple de processus PyBPMN

La figure 7 montre l'exemple d'une tâche d'un processus, annotée selon *PyBPMN* (Bocciarelli et d'Ambrogio, 2011). Comme spécifié, la durée moyenne d'exécution mesurée de la tâche *flightReservationService* doit être égale à 120 ms et la durée de panne doit être de l'ordre de 5. 108 secondes.

Conclusion. L'approche proposée adopte une représentation des besoins en performabilité, fondée sur l'extension du standard MARTE. Son utilisation demeure compliquée pour des analystes métier qui ne sont pas familiarisés avec ce standard. Par ailleurs, cette approche s'inscrit dans une perspective de modélisation productive. Elle permet l'analyse et la prédiction de la qualité des processus au moment de la conception ainsi que leur implémentation automatique.

4. Analyse comparative

Avant d'évaluer les différentes approches présentées auparavant, cette section expose d'abord les critères sélectionnés pour l'évaluation de leur pertinence.

4.1. Critères d'évaluation

Trois catégories de critères d'évaluation de la pertinence extraits à partir des approches d'annotation des propriétés non-fonctionnelles sont intéressantes pour analyser les approches étudiées.

Ces catégories de critères se résument comme suit :

- **L'objectif recherché.** Il met en relief trois critères : le domaine ciblé, le contexte d'utilisation et les objectifs spécifiques. Cela devrait permettre d'une part d'évaluer la portée des approches analysées en termes de propriétés ciblées et leur adaptation au contexte des services Web et, d'autre part de distinguer les approches de modélisation contemplatives de celles productives. La prise en considération du critère d'adaptation au contexte des services Web est justifiée par la tendance industrielle. Celle-ci consiste à opter pour le couple BPM/Architecture orientée services (SOA pour *Service Oriented Architecture*), en particulier les services Web et leurs standards W3C comme socle des systèmes d'information adaptables et collaboratifs. En plus, notons que les standards W3C des services Web présentent de nombreux apports par rapport aux modèles sémantiques proposés dans la littérature (Omrana, El Bitar *et al.*, 2013) ; (El Bitar, Belouadha *et al.*, 2013).
- **Le modèle de description utilisé.** Il met en évidence le modèle exploité pour l'extension du BPMN ainsi que le vocabulaire utilisé pour décrire les propriétés de performabilité. L'objectif est d'évaluer les approches étudiées en termes de leur conformité aux standards ainsi que leur niveau d'abstraction.
- **Le mécanisme d'annotation.** Il fait référence à l'objet annoté et aux éléments d'annotation. Ces critères devraient permettre d'évaluer la capacité de couvrir les besoins d'annotation des processus par des propriétés de performabilité, aussi bien que la complexité des modèles obtenus.

4.2. Évaluation des approches étudiées

Le tableau 2 en page suivante donne une synthèse de l'évaluation des approches présentées dans cet article.

Par objectif recherché. A l'exception des deux approches (Pavlovski et Zou, 2008) ; (Bartolini *et al.*, 2013), les contributions proposées dans la littérature y compris celles présentées dans cet article, ciblent exclusivement les besoins de performabilité ou de sécurité et non pas les deux à la fois. En plus, très peu d'approches (Bocciarelli et d'Ambrogio, 2014) sont dédiées à un contexte d'utilisation spécifique qui est l'architecture SOA et les services Web en particulier.

Sur un autre plan, les objectifs spécifiques visés par ces approches varient de la capture des besoins, la prédiction et l'analyse de la qualité, à la surveillance des processus ou encore l'automatisation de leur implémentation fondée sur la découverte des services Web. Elles se répartissent ainsi en approches contemplatives et productives.

Par modèle de description utilisé. Certaines approches analysées s'appuient sur des modèles de propriétés. Notons que la proposition de (Bocciarelli et d'Ambrogio, 2014) est fondée sur un modèle étendu du standard MARTE qui est un standard de description des propriétés non-fonctionnelles dans les systèmes temps réel. Néanmoins, cette approche, qui est dédiée au contexte des services Web, ne s'aligne pas avec les standards de cette technologie

pour prendre en considération les besoins de description des propriétés non-fonctionnelles (dont les propriétés de performabilité) dans ce contexte.

En ce qui concerne le vocabulaire utilisé pour spécifier les propriétés de performabilité des processus, il n'est pas unifié. Cela est dû au non-respect des standards. Notons également que ce vocabulaire reflète le niveau d'abstraction ou de détail de la représentation fournie par une approche. Cependant, les approches proposant des représentations détaillées telles que la contribution de (Bocciarelli et d'Ambrogio, 2014) n'est pas dédiée aux analystes métier. Son niveau de complexité requiert la collaboration avec des experts dans le domaine, voire la connaissance du modèle de description considéré.

	Critères	Pavlovski et Zou, 2008	Saeedi, Zhao et al., 2010	Bartolini et al., 2013	Bocciarelli et d'Ambrogio, 2014
Objectif recherché	Domaine ciblé	Performabilité/ Sécurité : • Performance, réglementation, Organisation, Sécurité...	Performabilité : • Performance, Fiabilité	Performabilité/ Sécurité : • Performance, Fiabilité, Sécurité	Performabilité : • Performance, Fiabilité
	Contexte des services Web	Non	Non	Non	Oui
	Objectifs spécifiques	• Capture des besoins	• Analyse	• Monitoring	• Analyse • Implémentation
Modèle de description utilisé	Modèle			• PMM	• MARTE
	Vocabulaire	• Contrainte • Risque • Contrôle	• Temps de réponse • Coût • Taux de fiabilité	• Valeur • Opérateur • Nature	• Durée • Fréquence • Origine • Type...
Mécanisme d'annotation	Objets annotés	• Tâche	• Tâche, <i>SubProcess</i> , <i>CallActivity</i>	• Tâche, Acteur, Chorégraphie	• Tâche, Processus, Collaboration
	Éléments d'annotation	• Artefacts • Catalogues	• Table de valeurs	• Instances • Attributs	• Annotations textuelles, Attributs

Tableau 2. Synthèse comparative des approches étudiées

Par mécanisme d'annotation. Les approches analysées se restreignent à l'annotation d'un objet BPMN spécifique comme la tâche (Pavlovski et Zou, 2008) ou tentent en plus d'annoter d'autres objets BPMN. Ce choix dépend du niveau d'abstraction de l'approche et ne respecte pas de norme. Dans le cas de la proposition de (Bocciarelli et d'Ambrogio, 2014) dont le contexte d'utilisation est les services Web, nous notons que le mécanisme d'annotation adopté n'est pas conduit par les exigences d'annotation sémantique et de description des propriétés non-fonctionnelles définies par les standards des services Web.

Concernant les éléments d'annotation, certaines approches étudiées utilisent indifféremment des symboles graphiques et/ou des annotations textuelles qui donnent lieu à des représentations informelles. Pour pallier ce problème, d'autres propositions telles que les contributions de (Bocciarelli et d'Ambrogio, 2014) et (Saeedi, Zhao et al., 2010) utilisent des attributs pour décrire les annotations textuelles ou encore une table de valeurs pour décrire les propriétés.

Nous notons également que le fait d'introduire différents symboles et annotations dans les modèles de processus impacte leur lisibilité. Dans ce contexte, la proposition de (Bartolini et al., 2013) prévoit de décrire le détail des annotations par des instances définies dans des paquets UML séparés des modèles de processus. Néanmoins, la représentation proposée s'avère ainsi fragmentée.

5. Conclusions et perspectives

Le standard BPMN est le standard de facto de spécification des processus métier. Il assure la simplicité et l'expressivité de description du comportement fonctionnel des processus ainsi que le mapping automatique avec le langage d'exécution BPEL. En outre, il favorise la modélisation collaborative et garantit la description consolidée (compacte et non fragmentée) des processus. Toutefois, ce standard s'inscrit dans le cadre d'une approche de modélisation contemplative à des fins de documentation et de communication.

Des approches d'extension des modèles BPMN avec les besoins en performabilité ont été proposées dans la littérature pour permettre l'analyse automatique et anticipée des processus. Ces approches peuvent être classées selon l'approche d'annotation adoptée, en approches orientées artefacts, tables, annotation textuelle ou instances UML. Le présent article analyse quatre contributions parmi les plus récentes ou référencées dans ce domaine et dégage leurs limites. L'analyse effectuée permet de conclure que ces approches sont :

- **Spécifiques.** Elles ciblent généralement le domaine de performabilité (Korherr et List, 2007) (Magnani et Montesi, 2007) ; (Sampath et Wirsing, 2009) ; (Saeedi, Zhao *et al.*, 2010) ; (Bocciarelli et d'Ambrogio, 2014) sans s'étendre à d'autres domaines non-fonctionnels tels que la sécurité. En plus, elles ne prennent pas en compte l'aspect sémantique. Aucune des approches ne permet d'annoter les processus avec des propriétés non-fonctionnelles et sémantiques à la fois.
- **Productives, mais axées sur l'analyse plutôt que sur l'implémentation.** Elles s'inscrivent, en général, dans le cadre d'une approche de modélisation qui n'est pas juste contemplative, mais qui conduit surtout l'analyse et la surveillance des processus. Néanmoins, très peu d'approches s'intéressent à l'implémentation automatique des processus.
- **Non conformes aux standards.** Elles ne s'alignent généralement pas avec les standards. Celles utilisées en particulier dans le contexte des services Web ne sont pas conformes aux standards de description des aspects sémantique et non-fonctionnel des services Web.
- **Dédiées exclusivement aux analystes métier ou aux experts.** Elles proposent en général un niveau élevé d'abstraction ou de complexité pour la description des propriétés de performabilité et sont ainsi, destinées soit aux analystes métier, soit aux experts du domaine.
- **Inadaptées au besoin d'évolutivité ou passage à l'échelle (*scalability*).** Elles convergent par l'introduction explicite des annotations de performabilité au niveau du diagramme BPMN. Cela conduit à une densité élevée d'informations au sein du modèle de processus et risque, ainsi, de le rendre illisible et de compliquer son interprétation.

De ces constats, une approche qui contribue à pallier ces limites semble importante. Elle serait encore plus intéressante si elle s'inscrivait dans le cadre d'une approche de modélisation qui pourrait conduire l'analyse, mais aussi l'implémentation automatique des processus. Notons que le couple BPM/SOA consistant à implémenter des processus par des services, en particulier des services Web, est considéré dans l'industrie logicielle comme socle des systèmes d'information adaptables et collaboratifs. Vu l'intérêt de ce couplage, une approche centrée sur le respect des standards W3C des services Web qui normalisent la spécification des aspects fonctionnel, non-fonctionnel et sémantique des services, serait intéressante. Elle devrait permettre de produire des modèles BPMN de processus sémantiques annotés avec les besoins non-fonctionnels tels que les besoins de performabilité et de sécurité, tout en palliant les limites des approches connexes. Cela contribuerait également à l'implémentation automatique de processus sémantiques, performants et sécurisés à travers la découverte et la composition des services Web en alignement avec les standards W3C (Belouadha, Omrana *et al.*, 2010) ; (Omrana, Belouadha *et al.*, 2012). C'est dans ce contexte que nous proposons l'approche IBPM (*Intelligent Business Process Manufacturing*) (Belouadha, 2015). Cette approche est fondée sur eBPMN (*enhanced BPMN*), une extension de BPMN conforme aux standards des services Web. Elle permet à la fois la modélisation des aspects sémantique et non-fonctionnel des processus métier et leur implémentation automatique par réutilisation des services Web.

6. Références

- Bartolini, C., Bertolino, A., De Angelis, G., Ciancone, A., Mirandola, R. (2012, June). *Non-functional analysis of service choreographies*. in *Principles of Engineering Service Oriented Systems (PESOS)*, ICSE Workshop, Zurich, June, pp. 8-14.
- Bartolini, C., Bertolino, A., De Angelis, G., Ciancone, A., Mirandola, R. (2013, March). *Apprehensive QoS monitoring of service choreographies*. in *Proceedings of the 28th Annual ACM Symposium on Applied Computing*, Coimbra, Portugal, March 18 - 22, pp. 1893-1899. ACM.
- Belouadha, F.-Z., Omrana, H., Roudiès, O. (2010). *A MDA approach for defining WS-Policy semantic non-functional properties*. In *International Journal of Engineering Science and Technology (IJEST)*, Vol. 2, Issue 6, juin 2010, pp. 2164-2171.

- Belouadha, F.Z., Omrana, H., Roudiès, O. (2012). *Web services-enhanced agile modeling and integrating business processes*. In E-Business - Applications and Global Acceptance, Prncelo Ifinedo (ed.), InTech, pp. 73-98.
- Belouadha, F.Z., Omrana, H., Roudiès, O. (2014). *Modeling and matching business processes using Web services*. ICMCS'14 International Conference, Marrakech, April 14-16, pp. 617-621.
- Belouadha, F.Z. (2015). *eBPMN pour la fabrication intelligente des processus métier*. Thèse pour l'obtention de doctorat en Génie Informatique, Centre d'études doctorales Sciences et techniques pour l'ingénieur, École Mohammadia d'Ingénieurs, juin 2015.
- Bernardi, S., Merseguer, J., Petriu, D. C. (2008). *Adding dependability analysis capabilities to the MARTE profile*. In Model Driven Engineering Languages and Systems (MoDELS '08), pp. 736-750, Springer Berlin Heidelberg.
- Bocciarelli, P., d'Ambrogio, A. (2011). *A BPMN extension for modeling non functional properties of business processes*. In proceedings of the 2011 Symposium on Theory of Modeling and Simulation: DEVS Integrative M&S Symposium, pp. 160-168, Society for Computer Simulation International.
- Bocciarelli, P., d'Ambrogio, A. (2014). *A model-driven method for enacting the design-time QoS analysis of business processes*. Software & Systems Modeling, Vol. 13 (2), pp. 573-598.
- Ciancone, A., Filieri, A., Drago, M. L., Mirandola, R., Grassi, V. (2011). *KlaperSuite: an integrated model-driven environment for reliability and performance analysis of component-based systems*. In Objects, Models, Components, Patterns, TOOLS'11, Zurich, Switzerland, June 28-30, LNCS, Vol. 6705, pp. 99-114, Springer Berlin Heidelberg.
- Di Marco, A., Pompilio, C., Bertolino, A., Calabrò, A., Lonetti, F., Sabetta, A. (2011). *Yet another meta-model to specify non-functional properties*. In proceedings of the International Workshop on Quality Assurance for Service-Based Applications, pp. 9-16, ACM.
- El Bitar, I., Belouadha, F. -Z., and Roudiès O. (2013). *Review of Web Services Description approaches*. 8th International Conference on Intelligent Systems: Theories and Applications (SITA'13), May 8-9, Rabat, Morocco.
- Korherr, B., List, B. (2007). *Extending the EPC and the BPMN with Business Process Goals and Performance Measures*. In Ninth International Conference on Enterprise Information Systems (ICEIS'07), Madeira, Portugal, pp. 287-294.
- Magnani, M., Montesi, D. (2007). *Computing the cost of BPMN diagrams*. Technical Report. University of Bologna. Available at: <http://www.cs.unibo.it/pub/TR/UBLCS/2007/2007-17.pdf>.
- Mylopoulos, J., Chung, L., Nixon, B. (1992). *Representing and using nonfunctional requirements: A process-oriented approach*. IEEE Transactions on Software Engineering, Vol. 18, N. 6, pp. 483-497.
- OASIS-BPEL (2007). *Web Services Business Process Execution Language*. Version 2.0. Primer, May 9. Available at: https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wsbpel.
- OMG-BPMN (2013). *Business Process Model and Notation (BPMN)*, Version 2.0.2, December 2013. Available at: <http://www.omg.org/spec/BPMN/2.0.2/PDF/>.
- OMG-MARTE (2011). *UML Profile for MARTE: Modeling and Analysis of Real-Time Embedded Systems*, Version 1.1, June, Available at: <http://www.omg.org/spec/MARTE/1.1/PDF/>.
- Omrana, H., Belouadha, F.-Z., Roudiès, O. (2012). *Template-based Matching Algorithm for Dynamic Web Services Discovery*. International Journal of Information and Communication Technology (IJICT), Vol.4, N°. 2, Inderscience Publisher.
- Omrana, H., El Bitar, I., Belouadha, F. Z., and Roudiès, O. (2013). *A Comparative Evaluation of Web Services Description Approaches*. In Proceedings of the IEEE 10th International Conference on Information Technology: New Generations (ITNG '13), April 2013, Las Vegas, Nevada, USA, pp. 60-64.
- Owen, M., Raj, J. (2003). *BPMN and Business Process Management: Introduction to the New Business Process Modeling Standard*. Popkin Software, available at: http://www.omg.org/bpmn/Documents/6AD5D16960.BPMN_and_BPM.pdf.
- Pavlovski, C. J., Zou, J. (2008). *Non-functional requirements in business process modeling*. In proceedings of the fifth Asia-Pacific conference on Conceptual Modelling, Vol. 79, pp. 103-112, Australian Computer Society, Inc.
- Petriu, D., Bernardi, S., Merseguer, J. (2008). *An UML profile for dependability analysis and modeling of software systems*. Technical report, Universidad de Zaragoza, Spain.
- Saeedi, K., Zhao, L., Sampaio, P. (2010). *Extending BPMN for supporting customer-facing service quality requirements*. In Web Services (ICWS'10), pp. 616-623.
- Sampath, P., Wirsing, M. (2009). *Computing the cost of business processes*. In Third International United Information Systems Conference (UNISCON'09), April 21-24, Sydney, Australia. Published in Information Systems: Modeling, Development, and Integration, Lecture Notes in Business Information Processing, Vol. 20, pp. 178-183, Springer Berlin Heidelberg.
- Smith, H., Neal, D., Ferrara, L., Hayden, F. (2002). *The emergence of business process management*. White paper, Computer Sciences Corporation's Research Services., Available at: http://itexpert.ir/download/ebooks/ITExpert.ir_BPM.pdf.
- Zou, J., Pavlovski, C. J. (2006). *Modeling architectural non functional requirements: from use case to control case*. In IEEE International Conference on e-Business Engineering (ICEBE'06), Shanghai, 24 October, pp. 315-322.

Recherche

Le système d'intelligence économique - XEW

Anass El Haddadi

DMI, ENSA, B.P : 3, Al-Hoceïma, Maroc

anass.elhaddadi@gmail.com

Amine El Haddadi

LIST, FST, Ancienne route de l'aéroport km10, Tanger, Maroc

amine.elhaddadi@gmail.com

Bernard Dousset

IRIT, 118 route de Narbonne, F-31062 Toulouse, France

dousset@irit.fr

Abdelhadi Fennan

LIST, FST, Ancienne route de l'aéroport km10, Tanger, Maroc

afennan@gmail.com

Résumé

L'intelligence économique est une activité permanente de management stratégique de l'information, pour une prise de décision collaborative au sein d'une organisation publique ou privée. Cette activité se présente sous forme d'un processus de flux d'information interne et externe, structurée et non structurée, processus supporté par les systèmes d'informations stratégiques et organisationnels. Autrement dit, l'intelligence économique est une cartographie de l'environnement de l'entreprise. Les organisations publiques ou privées ont besoin d'un système d'intelligence économique pour bien se positionner sur le marché, pour anticiper le moindre changement en vue d'y apporter la solution adéquate en temps réel. Le présent article présente le système d'intelligence économique XEW.

Abstract

Competitive Intelligence is a strategic management of information related activity which aims to provide a collaborative decision making in either a private or a public organization. This activity is actually a processing of internal or external, structured or unstructured information flow that is supported by the strategic and organizational information systems. In other words, competitive intelligence is a mapping of the surrounding business environment. Nowadays, every organization needs a competitive intelligence system in order to better its position in the market, or simply to survive, as well as to be able track every single change and to provide the right response to it in a real time scale. This new concept is the subject of the following article.

Mots-clés

Intelligence économique, Système d'Information, Management Stratégique de l'Information

Keywords

Competitive Intelligence, Information System, Strategic Management of Information

1. Introduction

À l'issue des nombreuses analyses stratégiques que nous avons déjà réalisées avec le logiciel Tétralogie et Xplor V1, il est apparu que les utilisateurs finaux des analyses produites veulent, en complément de l'aspect global (vue macroscopique), des analyses microscopiques sur des éléments déjà identifiés (concurrence, marchés, nouveaux produits ou procédés, partenaires potentiels...) ou des analyses leur permettant de découvrir de nouveaux éléments. A posteriori, de nombreux experts ou décideurs sont demandeurs de plus de détails sur les éléments traditionnels de leur environnement. Notamment, pour tout ce qui concerne leur vocabulaire spécifique, les acteurs qu'ils côtoient, les marchés qu'ils convoitent, les alliances qu'ils projettent.

Nous proposons donc de continuer à valider le modèle Xplor d'Intelligence Économique (IE) proposé (El Haddadi *et al.*, 2012) et de compléter les analyses macroscopiques par un modèle avancé Xplor EveryWhere (XEW) en ligne, au coeur de l'information obtenue par des recoupements statistiques, des classifications incrémentales ou des analyses multidimensionnelles. Le but est de privilégier l'extraction d'information en fonction du contexte général et non exclusivement par décryptage du contenu de quelques documents pris séparément. Il devient ainsi possible de retrouver, à partir d'un élément connu (acteur, mot clé), tout ou une partie de l'information qui lui est connexe (équipes, collaborations, concepts, émergences, mots associés ...), et ce par l'utilisation de concepts de filtrage avancé.

Le Système d'Intelligence Économique (SIE) XEW permet d'effectuer des analyses stratégiques sur des corpus d'information issus des sources les plus diverses comme les bases en ligne (publications scientifiques, brevets, portails, annuaires), les Cd, le Web visible et invisible, les news, la presse, les traces de connexions aux sites, les bases internes, les flux RSS, les réseaux sociaux ... et donne la possibilité aux décideurs de mener par eux-mêmes leurs investigations sans la participation d'un analyste confirmé ou d'un expert.

Ses applications sont très diverses : Identification des thématiques et des acteurs du domaine, mise en évidence des stratégies de développement et de coopération, proposition de scénarios d'évolution technologique (innovation), extraction des signaux faibles, consultation des informations à jour et en temps réel, grâce au Web service, remontée d'informations «terrain» lors des salons, des visites en clientèle ou à l'issue des réunions, demande de renseignements spécifiques en urgence qui seront mis en ligne par les Web services dédiés.

La présentation de SIE XEW fait l'objet de cet article. Nous allons aborder le concept d'IE, dans la deuxième section. Puis, nous traiterons des multisources et les types de données pour l'IE dans la troisième section. Ensuite, dans la quatrième section, nous présenterons notre modèle d'IE XEW qui se base sur les modèles MEDESIIE et SITE. Enfin, les deux dernières sections font l'objet de la proposition de notre modèle de SIE XEW avec un exemple d'analyse.

2. L'intelligence économique

L'émergence du concept d'IE et sa compréhension font appel à de multiples champs scientifiques. Cette émergence se rattache aux sciences de la sociologie des organisations, aux sciences politiques, aux sciences de l'information, aux technologies de l'information et à la communication et aux sciences de gestion de l'information et des connaissances. En fait, on peut considérer que sa genèse vient de différentes disciplines de recherche extrêmement vastes et prometteuses.

L'univers de l'IE étant très vaste, très riche en concepts et en domaines d'application, nous allons nous baser sur celle qui est la plus proche à l'analyse massive de l'information pour un eTI, Revue électronique en Technologies de l'Information. ISSN 1114-8802. <http://www.revue-eti.net> Titre de l'article, Auteur1 P1., Auteur2 P1., Revue eTI, Numéro 7, 2013 système d'aide à la décision. Nous retenons pour nos travaux de recherche les trois définitions suivantes: la première définition de l'IE moderne date de 1967 par Harold Wilensky, dans son ouvrage intitulé : «*Organisational intelligence : knowledge and policy in government and industry*». Il définit l'IE comme étant :

“L'activité de production de connaissance servant le but économique et stratégique d'une organisation, recueilli et produit dans un contexte légal et à partir de sources ouvertes”

La définition de l'association professionnelle d'IE (SCIP)¹:

“Competitive Intelligence is a systematic and ethical program for gathering, analyzing, and managing external information that can affect your company's plans, decisions, and operations. Put another way, CI is the process of enhancing marketplace competitiveness through a greater -- yet unequivocally ethical -- understanding of a firm's

¹ SCIP est une association professionnelle à but non lucratif dans le domaine d'IE Créée en 1986, SCIP est basée aux Etats-Unis. Elle comprend près de 6000 adhérents, rassemblés par «chapitres» dans 50 pays. SCIP France a été créée en 1992 par Robert Guillaumot, Yves-Michel Marti, Bruno Martinet et Jean-Pierre Bernat (Elf). Elle est le premier réseau français des professionnels de l'IE et de Gestion des connaissances.

*competitors and the competitive environment. Specifically, it is the legal collection and analysis of information regarding the capabilities, vulnerabilities, and intentions of business competitors, conducted by using information databases and other "open sources" and through ethical inquiry. Effective CI is a continuous process involving the legal and ethical collection of information, analysis that doesn't avoid unwelcome conclusions, and controlled dissemination of actionable intelligence to decision makers."*²

La définition de SCIP n'est pas donc très éloignée de celle de Wilensky. Elles y ajoutent la notion de coordination ou ce qu'on peut nommer l'intelligence collective. Sans oublier que l'IE est à la fois un processus et un produit (Bichard, 2005).

"En tant que processus, l'IE est l'ensemble des méthodes légales et éthiques qu'une entreprise utilise pour mobiliser l'information qui l'aide à réussir dans un environnement mondialisé. En tant que produit, l'IE est l'ensemble des informations sur les activités des concurrents à partir de sources publiques et privées, et son champ d'application et le comportement présent et futur des concurrents, des fournisseurs, des clients, de la technologie, des acquisitions et fusions, des marchés, des produits et services, et de l'environnement des affaires en général."

Avec l'avancement des travaux de recherche, une nouvelle génération de définitions a émergé. De définitions centrées sur la description des processus et des techniques d'IE, elles sont passées à des définitions incluant des objectifs stratégiques puis, depuis peu, à des définitions incluant les fonctions de gestion de connaissance, d'apprentissage collectif, de coopération (Salles *et al.*, 2000), de management stratégique de l'information et des systèmes d'information (Xiaobo et lian, 2010) (Olusoji et Amos, 2010) (Guoliang *et al.*, 2009) (Weidong, 2008) (Ginsburg *et al.*, 2009) (Li et Dong, 2010) (Li *et al.*, 2007) (Chen *et al.*, 2008) (Chen *et al.*, 2009) (Radum, 2006) (liu et Xia, 2009). Cette vision valorise l'intelligence collective en passant d'une entreprise hiérarchisée, isolée face à son environnement, à l'entreprise 2.0 (Jing *et al.*, 2015). Nous retrouvons à partir de ces différentes définitions, un ensemble de fonctions qui sont associées à la notion d'IE, nous le résumons dans le tableau suivant :

Fonctions	Auteurs
Gestion des connaissances	(Salles <i>et al.</i> , 2000)
Intelligence Collective	(Casselman et Samson, 2011)
Management stratégique de l'information	(Xiaobo et lian, 2010) (Olusoji et Amos, 2010) (Guoliang <i>et al.</i> , 2009) (Weidong, 2008) (Ginsburg <i>et al.</i> , 2009) (Li et Dong, 2010) (Li <i>et al.</i> , 2007) (Chen <i>et al.</i> , 2008) (Chen <i>et al.</i> , 2009) (Radum, 2006) (liu et Xia, 2009)
Gouvernance des systèmes d'information	(Chauna <i>et al.</i> , 2011)
Entreprise 2.0	(Jing <i>et al.</i> , 2015)

Tableau 1. Les fonctions associées à la notion d'IE

Pour adapter la notion d'intelligence économique à nos travaux de recherche en Data Mining et visualisation de données, nous définissons l'intelligence économique de la manière suivante :

"L'intelligence économique est à la fois un processus, un produit et un service : en tant que processus l'intelligence économique est le workflow du management stratégique de l'information, pour une prise de décision collaborative. En tant que produit, l'intelligence économique est un système d'information stratégique et organisationnel. En tant que service, l'intelligence économique est une cartographie de l'environnement de l'entreprise."

Selon cette définition, le processus d'intelligence économique de la figure 1 est alors une coordination entre le processus de **management de données** structurées, semi-structurées et non structurées et de la **veille stratégique** basée sur les systèmes d'informations.

Dans la littérature relative à l'intelligence économique, les auteurs (Štefániková, 2014; Sewdass, 2014; Saba, 2014; Nelke, 2012; Colakoglu, 2011) s'accordent sur l'existence d'une étape primordiale consacrée à la compréhension du besoin. Le but de cette première phase consiste à définir un problème décisionnel par un ensemble de scénarios possibles. Elle permet de traduire les orientations stratégiques en axes de développement. Ces axes sont traduits en cibles prioritaires en déterminant les questions opérationnelles sur lesquelles il serait utile de s'informer. Ces cibles peuvent être les signaux forts, les nouveaux centres d'intérêt, les nouveaux produits, les concurrents, les nouveaux acteurs ou les nouveaux procédés.

² <http://competitive-intelligence.mirum.net/business-intelligence/definition-competitive-intelligence.html>, consulté le 18 janv. 11

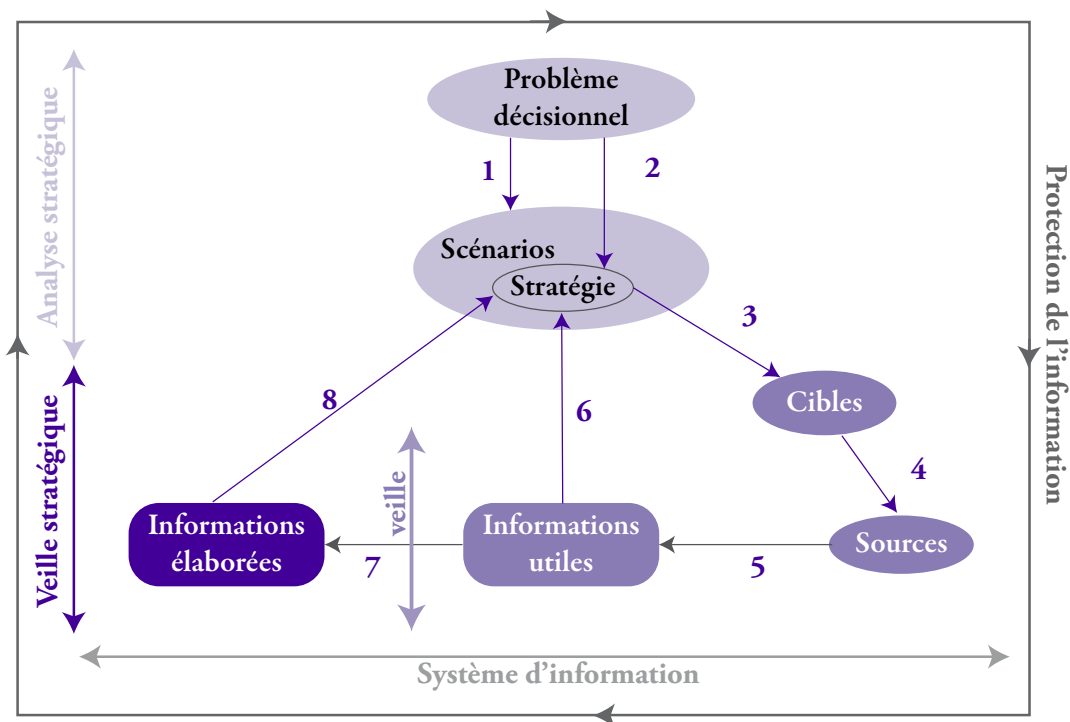


Figure 1. Processus de l'Intelligence Économique

- 1 – Elaborer, 2 – Choisir, 3 – Déterminer & Hiérarchiser, 4 – Identifier & Sélectionner, 5 – Collecter & Evaluer, 6 – Valider & Diffuser, 7 – Analyser & Interpréter, 8 – Valider & Diffuser

Le besoin informationnel déterminé suit la phase de la veille stratégique qui permet d'analyser et de surveiller l'environnement informationnel, les cibles identifiées par la stratégie retenue. En fonction des cibles définies dans la phase précédente, il s'agit de rechercher les informations les plus pertinentes, à l'intérieur et à l'extérieur de l'entreprise. La première étape consiste à identifier les sources d'information qui sont susceptibles de contenir ces informations. À partir de ces sources, les informations utiles sont extraites et évaluées. Il existe deux types d'informations qui sont nécessaires à l'élaboration de la stratégie : l'information utile qui peut être directement utilisée et l'information élaborée qui résulte de l'analyse et du traitement des informations utiles. L'étape de traitement et d'analyse est soutenue par les méthodes et les outils de fouille de masses de données (Big Data Mining) qui permettent de révéler les informations cachées, souvent statistiques, à partir des informations utiles. La dernière phase du cycle consiste à restituer les résultats jugés pertinents aux utilisateurs potentiels qui vont établir et définir les actions stratégiques en se basant sur le degré de fiabilité des résultats obtenus. Cette phase peut engendrer une redéfinition ou un réajustement de la cible, le cycle devient alors itératif. La protection des traitements et des informations est un processus qui s'instaure en parallèle du cycle de l'intelligence économique et s'applique à toutes ces étapes. L'ensemble de ces étapes peut être soutenu par l'utilisation des systèmes d'information. Notre démarche consiste alors à définir un système d'intelligence économique, en se basant sur notre processus d'intelligence économique, dont l'objectif est:

- De supporter les différentes étapes de la démarche en offrant toutes les fonctionnalités de compétitivité, de l'analyse des besoins et de la veille stratégique
- D'extraire et traiter les données structurées, semi-structurées ou non structurées issues de tout type de sources : articles scientifiques, brevets, flux RSS, mails, réseaux sociaux, blogs, forum, etc.
- De répondre aux besoins d'analyse de la démarche de l'intelligence économique en offrant des connaissances synthétiques de l'environnement informationnel de l'entreprise, sur un domaine donné,
- De restituer les résultats sous forme d'un portail Web / Mobile intelligent, offrant une visualisation (reporting) nouvelle génération pour faciliter la prise de décision collaborative.

Afin de répondre à ces besoins, nous proposons dans cet article le prototype d'un système d'intelligence économique qui couple les besoins des étapes du processus d'IE avec les principes et techniques issus du domaine de la découverte des connaissances à partir des masses de données (Big Data Mining). Mais avant de proposer ce type de système, il faut bien comprendre et analyser les différentes sources d'informations.

3. Les multisources et l'Intelligence Économique

Les sources des données sont nombreuses, diverses, et plus ou moins facilement accessibles. Pour l'intelligence économique, il existe plusieurs types de sources à surveiller : les bases de données scientifiques, les bases de brevets, les médias, la presse, les flux RSS, l'Internet, l'Intranet, les forums, les blogs, les wikis, les sites de micro-blogging, les bookmarking sociaux, les mashup, les crowdsourcing, les réseaux sociaux, etc. Nous définissons dans ce qui suit deux types de sources : l'information formelle et l'information informelle.

L'information est dite formelle dès qu'elle est publiée sur support papier, informatique, microfilm, etc. Elle peut être structurée ou non, mais il s'agit dans tous les cas d'une information directement accessible et exploitable. Ce type de sources correspond à l'information blanche.

Les sources formelles sont composées principalement de la presse, la télévision, la radio, les livres, les banques de données et CD-ROM, les brevets, les informations légales, les études réalisées en général par des prestataires publics ou privés, Internet. Ces sources ont l'avantage d'être sûres, assez exhaustives et faciles d'accès. Dans un contexte de veille stratégique, les bases de données les plus consultées sont à dominance scientifique, technologique, réglementaire et se trouvent sur des bases bibliographiques. Parmi les bases de données les plus intéressantes, nous pouvons citer par exemple Factiva dans le domaine économique, Inspec dans le domaine physique et Kompass Europe orientée entreprise, Pascal qui représente une source multidisciplinaire et PubMed spécialisée dans le domaine médical.

L'information informelle est constituée de toutes les informations non formalisées et non disponibles directement. Il est donc nécessaire d'entreprendre des démarches directes auprès des détenteurs supposés de cette information. Ce type de sources correspond à l'information grise. Ces sources peuvent être les expositions et les salons, les fournisseurs, les colloques, les congrès, les clubs. On y échange et communique des informations. L'information qui circule est souvent d'une grande valeur stratégique telle que celles des concurrents à l'occasion de portes ouvertes, ses communications commerciales et financières, ses publications dans des journaux internes, etc. Sans oublier les sources internes de l'entreprise qui représentent plus de 80% des informations de ce qu'un décideur peut chercher comme information alors qu'elle se trouve dans son entreprise, par exemple : des sites personnels, des études de recherche menées par un groupe d'étudiants ou de thésard, etc. ainsi que dans les réseaux personnels dans la limite de la légalité et de la déontologie.

Les sources n'étant pas équivalentes, il s'avère ainsi intéressant de les combiner pour avoir une information plus exhaustive et moins biaisée. Deux bases qui indexent le même journal ne choisissent pas les mêmes articles. L'une va proposer le résumé et de bons mots-clés, l'autre choisira les citations et toutes les adresses. Faut-il les dédoubler ou doit-on conserver toutes les informations complémentaires ? La question reste posée. Le choix est toujours laissé aux utilisateurs, mais nous préconisons de garder toutes les versions disponibles puisqu'une information redondante est moins gênante qu'une information partielle ou manquante.

D'un autre point de vue, compte tenu du fait que les sources sont complémentaires, elles permettent d'une part d'aboutir à un corpus multiformats représentatif du domaine étudié, et d'autre part, d'avoir une vision stratégique, d'un sujet (produit ou procédé), à la fois scientifique (les articles scientifiques) et technologique (les brevets). Cette hétérogénéité des sources permet de : suivre les évolutions de la recherche au service de la technique, mettre en évidence les signaux faibles (par exemple les nouveaux centres d'intérêt), les nouveaux acteurs et les meilleures pratiques.

Une première contribution de notre travail (El Haddadi *et al.*, 2010) a consisté à traiter les différentes sources dans leurs formes natives à la différence de la majorité de logiciels qui imposent aux utilisateurs un format de données propriétaire. Cette restriction engendre une perte de temps considérable de restructuration de données et nécessite parfois une expertise élevée en informatique. Le format natif offre plusieurs avantages à savoir une meilleure réactivité, une mise à jour du corpus facilitée et une préservation de l'ensemble des informations. Néanmoins, afin de s'adapter à quasiment toutes les structures, il est nécessaire d'utiliser les métadonnées qui sont des outils de description des formats dont le principe est de :

- Trouver une technique permettant de différencier les documents les uns des autres (ou les unités textuelles) ;
- Déterminer les balises des champs sémantiques présentes dans la base et leur donner un nom et un sigle standard ;
- Définir leur utilité et leur priorité ;
- Déterminer des techniques judicieuses de découpage permettant d'extraire chaque type d'information.

Notons par ailleurs que plus de 90% des cas rencontrés peuvent être traités sans aucun reformatage (El haddadi *et al.*, 2015). En effet, il est possible de travailler simultanément sur plusieurs formats et sur plusieurs sources en élaborant des règles de correspondance entre les champs utiles à l'aide de métadonnées de second niveau. Celles-ci permettent à la fois d'orchestrer la synchronisation de tous les formats et de les interfacer de façon unique avec les outils d'extraction sémantique. Chaque source a un format spécifique qui lui-même a un descripteur spécifique (métadonnées de premier niveau). Une collection de formats est gérée par un descripteur générique (métadonnées de second niveau : le chef d'orchestre).

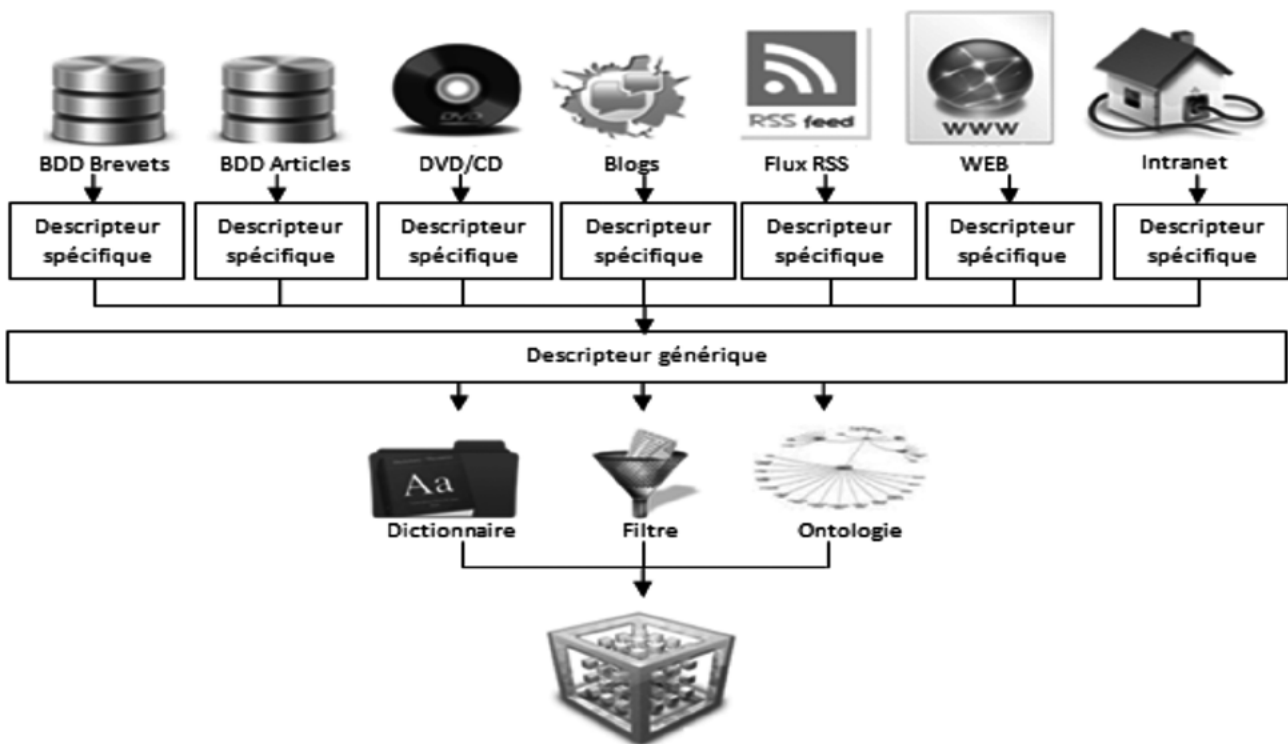


Figure 2. Homogénéisation des sources d'information

Cette vue unifiée associée au corpus ciblé correspond à une représentation logique, structurée, prédéfinie de l'ensemble de ses collections sous forme d'un entrepôt de donnée orientée documents. On a pu automatiser cette étape de collecte et traitement d'informations, grâce aux agents intelligents de crawling et scraping déployés dans notre Système d'Intelligence Économique XEW, qui fera l'objet de la section suivante.

4. Les méthodes de conception des systèmes d'intelligence économique

Depuis deux décennies de travaux de recherche en intelligence économique diverses difficultés se posent toujours lors de la conception des systèmes d'intelligence économique, surtout, que l'on trouve dans plusieurs travaux (Boukrara, 2010; Besson, 2010; Othenin, 2011)

une confusion entre les systèmes d'intelligence économique et les systèmes de business intelligence. Ceci peut s'expliquer par le manque de méthode, la mauvaise traduction du terme anglais «Competitive Intelligence» en Intelligence Économique, ce qui crée une confusion entre l'aspect science économique et l'aspect management de l'information pour la compétitivité d'entreprise. Or, la plupart des sources dans ce domaine sont des constats d'experts et des rapports d'hommes d'État, qui ignorent la dimension théorique du concept d'intelligence économique. Ces travaux de recherche montrent que les modèles d'analyse dans le cadre d'intelligence économique se focalisent sur quatre dimensions clés à savoir (El Haddadi, 2014):

- La dimension environnementale d'une entreprise qui regroupe les éléments pouvant influencer d'une manière directe ou indirecte l'évolution stratégique d'une entreprise. Elle est caractérisée par les partenaires, les concurrents, les marchés, les clients, etc.
- La dimension humaine qui englobe les acteurs impliqués dans une démarche d'intelligence économique, qu'ils soient internes ou externes à l'entreprise. Elle est caractérisée par les réseaux de collaborations, d'interactions et de communications entre les différents acteurs impliqués dans cette démarche.
- La dimension stratégique qui correspond aux différents modèles d'analyse permettant l'élaboration des stratégies d'une entreprise allant de l'identification des objectifs à la décision et à la définition d'actions.
- La dimension technologique qui rassemble toutes les méthodes, outils et techniques utilisées pour répondre aux besoins informationnels dans une démarche d'intelligence économique, par les procédures de recherche, collecte, traitements et diffusion de l'information.

La prise en compte de l'une ou plusieurs de ces dimensions permet de concevoir plusieurs modèles servant d'analyse (conceptuelle ou pratique) de l'intelligence économique. Dans le cadre de nos travaux de recherche nous avons retenu deux modèles universitaires bâtis sur ces dimensions : le modèle MEDESIIE pour son aspect ingénierie des besoins pour l'IE , le modèle SITE pour son aspect modélisation et conception des SI dans la cadre d'IE.

4.1. La méthode d'analyse MEDESIIE

La démarche d'intelligence économique proposée par (Salles, 2013) dans le cadre du projet MEDESIIE³ est consacrée entièrement à l'ingénierie des exigences pour l'intelligence économique. La méthode MEDESIIE est développée pour analyser les besoins des PME/PMI en matière d'intelligence économique. L'architecture conceptuelle de ce système a été construite selon le référentiel de Seligmann (1989). Maryse Salles a mis un accent tout particulier sur la «façon de penser» en ce qu'elle influe fortement sur la conception du système produit, et sur la «façon d'aider», composant essentiel dans les contextes complexes comme ceux liés à l'aide à la décision. MEDESIIE, propose des modèles pour décrire l'entreprise, sa stratégie, son environnement, son besoin relevant de l'intelligence économique et ses produits et services:

- Un modèle d'entreprise est décrit selon ses différentes fonctions (productive, économique (lien avec le marché), financière et d'innovation (système d'information)). Chaque fonction est elle-même composée d'un ensemble de sous-fonctions de management.
- Un modèle de stratégie est représenté par un ensemble de choix stratégiques structurels et d'axes de développement (recherche d'indépendance, croissance de ses activités, augmentation de ses profits, etc.) ;
- Un modèle d'environnement est décrit dans un premier temps par les fonctions de l'entreprise et des relations qu'elle développe avec son environnement. Dans un second temps, il est décrit par l'environnement exogène à l'entreprise, selon huit grands axes (la géométrie spatiale des marchés, la demande, la technologie, la concurrence, le régime financier, les conditions d'offre, le cadre réglementaire, l'environnement politique et géopolitique).
- Un modèle du besoin sert de cadre à la collecte du besoin, sa formalisation, son analyse et sa validation. Le besoin exprimé est représenté par un ensemble d'unités de besoin. Le modèle d'unité de besoin est décrit en fonction de trois dimensions : le niveau de pilotage de la décision pour lequel l'unité est exprimée (sa valeur : opérationnelle, tactique ou stratégique), la phase du processus de décision associé à l'unité et son contenu informationnel (identification de sa valeur et de sa fonction).
- Un modèle des produits et services, qui présente toute fourniture d'aide à la décision portant sur l'environnement, de l'entreprise qui consiste en un outil de prototypage réalisé en fonction de la satisfaction des besoins recueillis afin d'en définir les coûts, la portée et de pouvoir en évaluer, a priori, les effets.

Le positionnement de MEDESIIE la rapproche des approches d'ingénierie des exigences dirigées par le but. Il s'agit en effet de construire (ou plutôt co-construire) les besoins d'intelligence économique d'une entreprise en relation avec sa stratégie. Pour notre SIE, le modèle d'ingénierie des exigences, proposé dans MEDESIIE, nous aide à définir le modèle des différents critères, indicateurs, contraintes de ciblage et validation des sources d'informations pour répondre à un besoin informationnel, sous forme d'un modèle adaptatif a n'importe quel type de source d'information. Cependant, MEDESIIE ne nous permet pas de comprendre et de modéliser le rôle / comportement des acteurs d'IE.

4.2. Les modèles SITE

Le modèle SITE regroupe les différents modèles d'IE proposés par l'équipe de recherche SITE du laboratoire LORIA de Nancy. La problématique de l'équipe est d'étudier la modélisation et le développement de systèmes d'informations stratégiques dans le cadre de l'intelligence économique.

Ces modèles se basent sur la mise en relation de trois espaces : l'espace des problèmes décisionnels, l'espace des problèmes informationnels et l'espace des médiations qui les relient par l'intermédiaire des échanges entre les deux types d'acteurs (décideur et veilleur).

Les travaux de l'équipe ont en commun la prise en compte de l'utilisateur dans les systèmes d'information. Ils proposent des modèles permettant de définir les différents acteurs, leurs interactions et leurs positionnements dans le processus d'intelligence économique.

Nous retenons trois modèles à savoir : EquAte (Explore Query Analyse Annotate), MEPD (Modèle pour l'explication d'un problème décisionnel), WISP (Model for Information Retrieval query Annotations Based on Expression Levels).

Le modèle EquAte (David, 2002), représente une situation de recherche d'information qui implique les phases cognitives suivantes :

- L'exploration du monde d'information
- L'interrogation de la base d'information
- L'analyse de la base d'information

³ Projet MEDESIIE (Méthode de définition de système d'information pour l'intelligence économique), consacré entièrement à l'analyse du besoin en intelligence économique (i.e.) des PME, et financé par la région Midi-Pyrénées.

- L'annotation fondée sur différentes préférences

Le modèle MEPD (Bouaka, 2010), consiste à définir les différentes facettes d'un problème décisionnel. Il se base sur :

- La modélisation du décideur, par son identité, les traits de sa personnalité, son style cognitif et son expérience,
- La modélisation de l'environnement, par l'environnement immédiat (les clients, les fournisseurs, les concurrents) et global (social, économique, politique, etc.),
- La modélisation de l'organisation, par l'environnement, son signal, les hypothèses que le décideur peut déduire de la détection des signaux recueillis.

Le modèle WISP (Olusoji et Amos, 2010) est associé au modèle MEPD dans lequel il s'intègre en lui empruntant quelques paramètres. Le modèle WISP est un modèle tridimensionnel, multifacettes, qui intègre la notion de point de vue :

- Une dimension analytique qui englobe la compréhension demande-enjeu-contexte (DEC), la définition des indicateurs informationnels et toutes les opérations d'analyse et de création de connaissances qui peuvent être réalisées par l'étude des éléments mémorisés,
- Une dimension méthodologique qui est constituée, à un premier niveau, par les compétences de traduction du problème décisionnel en problème informationnel et à un second niveau par les stratégies de recherche par lesquelles l'information est identifiée et les connaissances sont acquises.
- Une dimension opérationnelle qui correspond à la sélection des plans d'action et à la mise en place des différentes étapes de résolution de la méthodologie associée au modèle WISP.

4.3. Le modèle XEW

En se basant sur le processus d'intelligence économique présenté dans la deuxième section, les meilleures pratiques (best practice) de l'aspect modélisation du rôle / utilisateurs dans un SIC proposé dans les modèles Equate, MEDP et WISP et le modèle d'ingénierie des exigences présenté dans MEDESIIE nous proposons le modèle d'analyse et de conception d'un système d'intelligence économique, nommé XEW.

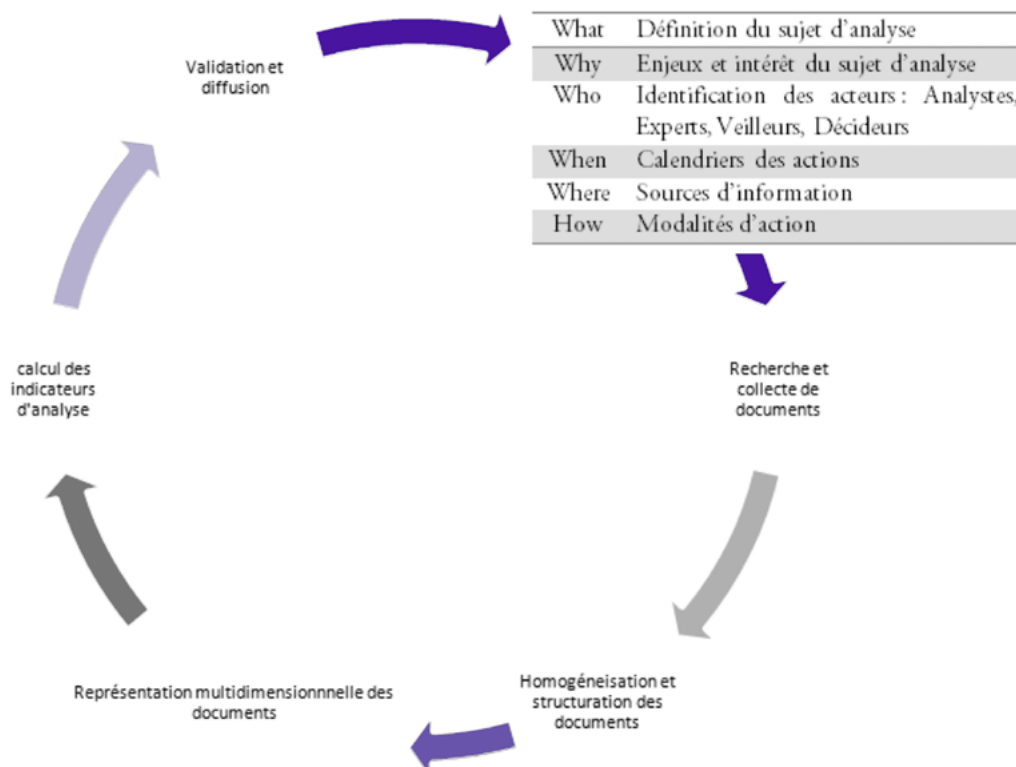


Figure 3. Processus du modèle XEW

XEW présente une capitalisation et une amélioration durable du modèle Xplor que nous avons développé en 2011. Le noyau de XEW se base sur trois modèles :

- XEW Data Analytics (XewDA) : un modèle d'analyse exploratoire multidimensionnelle, qui repose sur l'extraction de la connaissance à partir des grandes masses de données (Big Data) d'une part, et l'étude diachronique et incrémentale de l'évolution de la connaissance.

- XEW Service Quality (XewSQ) : un modèle d'indicateur qui a pour objectif la définition d'un ensemble d'indicateurs, selon les besoins d'analyse, à partir de la représentation multidimensionnelle des documents, dont l'objectif principal est l'évaluation qualitative des différentes interactions dans le système.
- XEW Security (XewSE) : Un modèle de sécurisation qui repose sur le modèle Rôle / Utilisateurs, pour sécuriser les informations et les traitements effectués sur notre système d'intelligence économique.

La première étape de ce modèle est l'analyse des besoins qui consiste à traduire le problème décisionnel en un problème informationnel. Elle est établie à partir du problème informationnel exposé par les décideurs. L'objectif de cette activité est de décrire la démarche de pilotage du processus d'analyse. Nous définissons cette activité selon le principe de 5W- 1H (El Haddadi, 2010): What, Why, Who, When, Where, How (Quoi, Quand, Où, Qui, Comment, Pourquoi).

Ce principe nous permet de décrire le besoin informationnel posé et d'orienter l'analyse. Une fois que l'analyse des besoins est validée, la deuxième étape consiste à rechercher les informations nécessaires pour l'analyse. Cette étape permet de se focaliser, suivant des critères prédéfinis, sur des données supposées à la fois «interprétables» et à fort potentiel informatif. De plus, la préparation des données consiste, dans un premier temps, à les sélectionner en accord avec les objectifs que l'on s'impose, en ayant recours aux techniques de recherche d'information (Büttcher, 2010). Puis, vient l'étape de la fouille de donnée (Data Mining), ou plutôt le Big Data Mining, puisqu'on est submergé par des montagnes de données de plus en plus non structurées, ce qui nécessite le développement d'outils de traitement et d'exploitation de plus en plus efficaces. Une fois l'analyse validée, nous diffusons la connaissance sous forme de sortie graphique dernière génération Web/Mobile, en faisant référence aux dernières avancées dans le domaine de la visualisation des données (Data Visualisation).

L'objectif de nos travaux de recherche est de définir une architecture modulable du système d'intelligence économique XEW, en se basant sur l'ingénierie dirigée par les modèles (MDA). L'approche adoptée, pour le développement du prototype, permet de combiner les techniques de recherche d'information, d'extraction, de la fouille de données (Data Mining) et de la visualisation (Data Visualisation) des grandes masses de données (Big Data). Chacune de ces techniques est vue comme un composant modulable aux fonctionnalités précises et délimitées. Plus simples à concevoir, à développer, plus robustes et testés dans des contextes différents, ces composants peuvent s'assembler de plusieurs manières pour créer ainsi des applications variées et adaptées aux besoins des utilisateurs finaux.

5. Le système d'intelligence économique

Le système d'intelligence économique induit une culture collaborative de l'information accompagnée d'actions offensives, avec le souci constant de protéger le capital informationnel de l'organisation.

On trouve dans la littérature différentes propositions de conception des systèmes d'intelligence économique. Pour Romagni dans son livre «L'intelligence économique au service de l'entreprise», la définition d'un système d'information adapté à la démarche d'intelligence économique est la suivante : *Ensemble organisé de procédures permettant, à tout moment, de donner aux décideurs une représentation de la place de l'entreprise dans son environnement et sur son marché. Il produit de l'information pour assister les individus dans les fonctions d'exécution, de gestion et de prise de décision.*

Il doit :

- Faciliter les décisions, en automatisant un certain nombre d'actions ou en mettant à disposition des décideurs des éléments nécessaires à la prise de décision,
- Coordonner les actions pour le traitement de l'information,
- Stocker de manière durable et stable les informations,
- Traiter les données, c'est la fonction la plus importante du système d'information, car il entraîne la création d'informations directement utilisables par les décideurs.

La plupart des autres propositions (Xiaobo, 2010; Olusoji, 2010; Weidong, 2008; Ginsburg, 2009; Li, 2010; Chen, 2008; Chen, 2009; Radun, 2006; Liu, 2009; Li, 2007) sont d'accord au niveau de la conception et des composants, avec Romagni. Cette conception permet de concevoir des systèmes d'information permettant aux organisations une meilleure gestion de l'information et d'offrir une base de coordination des actions entre les différents acteurs. Cette coordination transversale est soutenue par les motivations suivantes :

- Les objectifs d'une démarche d'intelligence économique sont inter-reliés. Ils ne peuvent être traités séparément.
- Le besoin de partage d'informations entre les différents acteurs,
- Le partage des connaissances acquises au cours d'une démarche,
- L'organisation de l'entreprise de manière fonctionnelle est transversale.

Cette conception doit évoluer d'une architecture verticale vers une architecture transversale qui permet une gestion globale de l'information.

L'architecture transversale est basée sur une architecture modulaire et évolutive structurée autour des projets de l'entreprise. Cette architecture va permettre:

- Une réduction des nombreuses coordinations verticales en diminuant les échelons hiérarchiques,
- Une meilleure surveillance de l'environnement des unités décentralisées,
- Un décloisonnement par une communication transversale,
- Des relations axées sur la complémentarité des métiers,
- Une meilleure adaptation à la dynamique du marché.

La démarche d'intelligence économique s'inscrit alors dans le schéma global du système d'information stratégique de l'entreprise : le système d'intelligence économique XEW. Le modèle IC XEW repose sur une architecture décisionnelle à quatre niveaux :

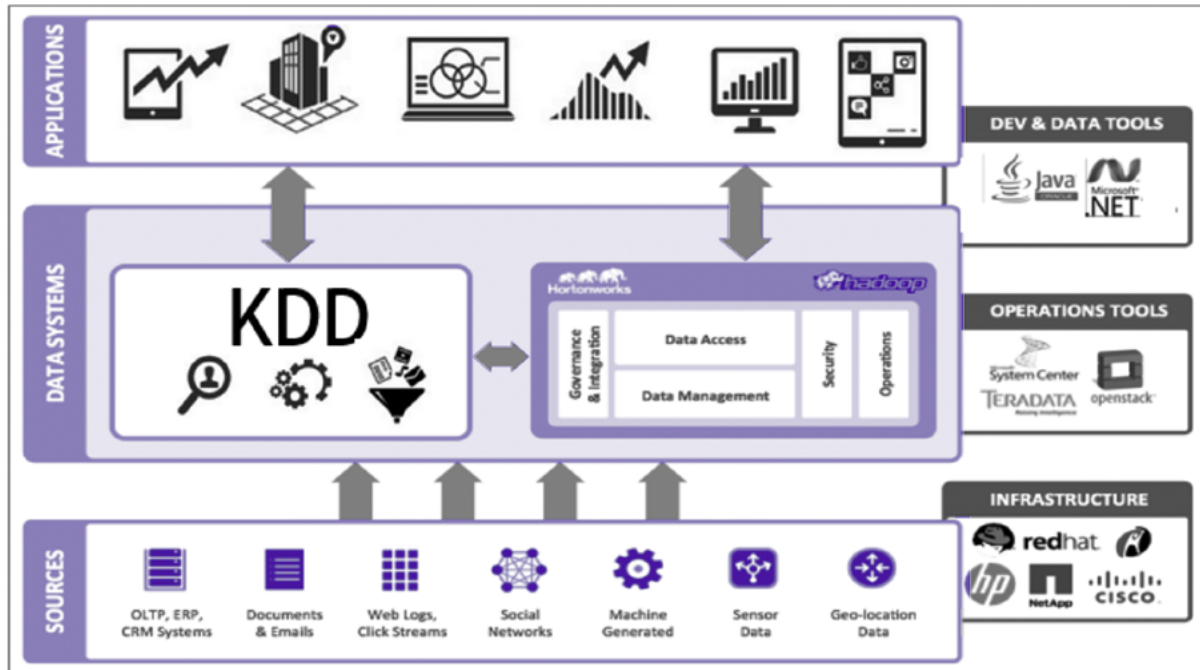


Figure 4. L'architecture XEW

5.1. Service de sourcing XEW (SS-XEW)

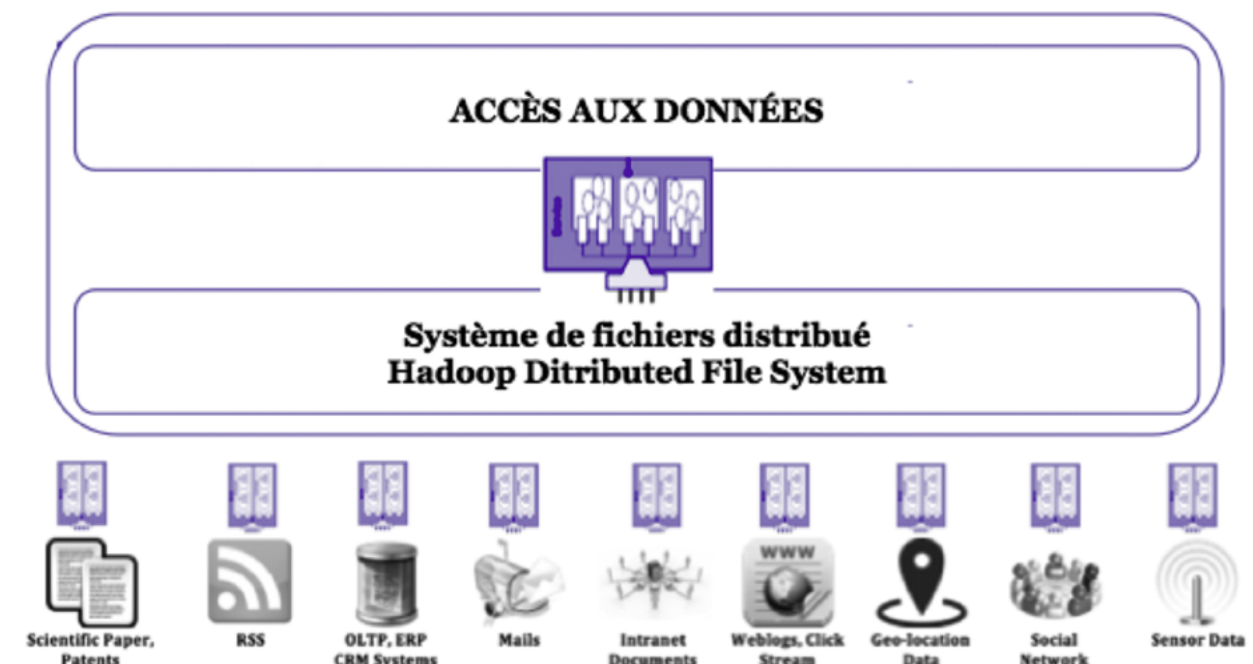


Figure 5. Service de sourcing

Ce service permet la recherche, la collecte, le traitement des données issues de différentes sources, avec la prise en compte de techniques de fusion multimodale pour tenir compte de l'hétérogénéité, l'imprécision et l'incertitude qui entachent les données multisources. Cette prise en compte de fusion assure une maîtrise des connaissances et des informations, et par conséquent facilite amplement une prise de décision. SS-XEW traite l'hétérogénéité des informations, d'un point de vue : contenu sémantique (scientifique, technique, etc.), structurel (fortement structuré (brevet) à non structuré (e-mails)), linguistique (multilinguisme), format du support (Word, html, pdf, etc.), taille : définition de l'unité d'information à analyser (granularité de l'information). L'objectif de ce niveau architectural est de fournir une description complète de l'ensemble du processus de traitement de données issues des différentes origines. Pour cela, les techniques employées s'appuient sur des agents intelligents de crawling et scraping adaptés à chaque source d'information.

5.2. Service Data Warehousing (SDW-XEW)

Ce service est un espace de stockage qui à un premier niveau permet d'offrir une vue unifiée du corpus cible, l'extraction et le stockage des données sources non structurées, semistrukturées et structurées sous forme d'une représentation multidimensionnelle. Le second niveau porte sur les traitements de création de l'entrepôt de données : SQL Classique et NoSQL comme : MongoDB, Neo4j, GraphDB et hbase.

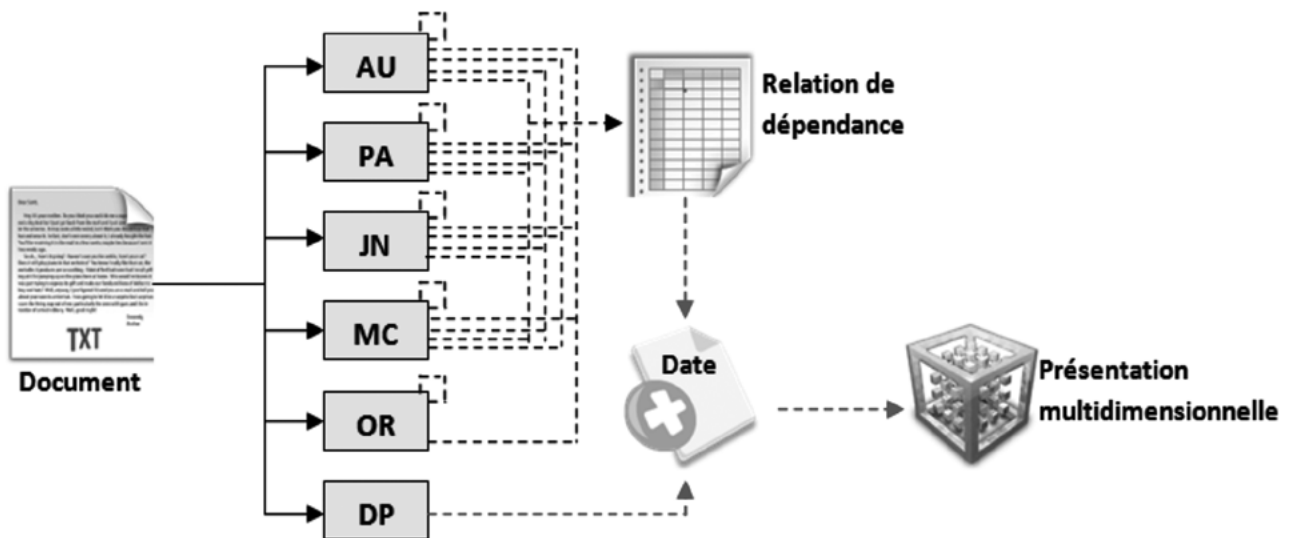


Figure 6. Service Data Warehousing

5.3. Service Big Data Analytic (SBDA-XEW)

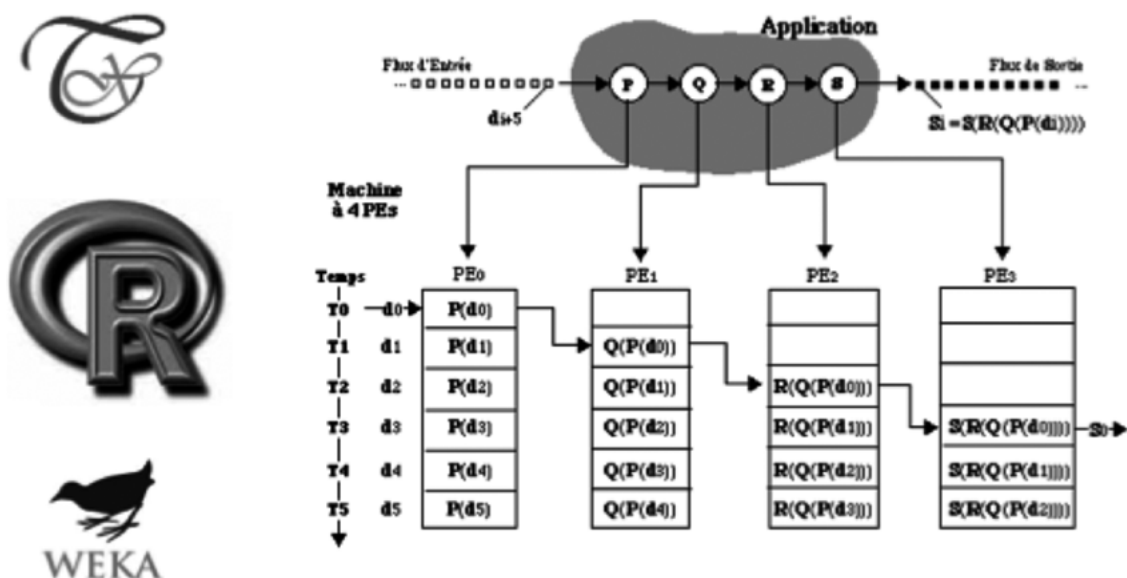


Figure 7. Service Big Data Analytic

Le SBDA-XEW permet d'effectuer des analyses multidimensionnelles, en adaptant les algorithmes de data mining à la problématique de Big Data. En se basant sur le parallélisme des algorithmes développés dans le système Tétraxplor ou les open sources comme Weka et R.

5.4. Service Data Visualisation (SDV-XEW)

L'objectif de ce service est d'offrir aux utilisateurs une interprétation des connaissances claire et compréhensible, grâce à un espace interactif et collaboratif. Il facilite la navigation dans l'environnement relationnel d'un élément connu (entreprise, centre de recherche, chercheur, inventeur, mot clé, etc.). Il devient ainsi possible de retrouver, à partir de cet élément, toute ou partie de l'information qui lui est connexe (alliances, concurrents, équipes, émergences, signaux faibles, etc.). Le SDV-XEW propose des méthodes innovantes de visualisation des grandes masses de données, comme le clustering incrémental, la comparaison par pas de temps (diachronie) et les graphes temporels. La gestion de l'ensemble des services permet de représenter le SIC comme un ensemble de capacités organisationnelles permettant de fournir de la valeur aux utilisateurs finaux sous forme de services. Cette valeur se base sur une approche qualité logicielle d'une part, et des indicateurs de contrôle qualité dans toute la chaîne de traitement.

6. Exemple d'analyse sur les anticorps

L'objectif de cette analyse est d'avoir un état des lieux de la recherche concernant les molécules similaires aux anticorps et notamment les équipes, les stratégies, les alliances, les points forts et faibles des autres entreprises... Nous nous intéresserons donc plus particulièrement à 12 molécules regroupées sous le nom de **Scaffold**. Grâce à l'agent intelligent de Scraping et Crawling de la base de données internationale de brevets (SC-WIPOXEW) de XEW nous avons collecté 3990 brevets.

N° de publication:	WO/2015/171074	N° de la demande internationale.:	PCT/SG2015/050096
Date de publication:	12.11.2015	Date de dépôt international:	05.05.2015
CIB:	C12N 5/071 (2010.01), C12N 1/38 (2006.01)		
Déposants:	NATIONAL UNIVERSITY OF SINGAPORE [SG/SG]; 21 Lower Kent Ridge Road Singapore 119077 (SG). AGENCY FOR SCIENCE, TECHNOLOGY AND RESEARCH [SG/SG]; 1 Fusionopolis Way, #20-10 Connexis Singapore 138632 (SG)		
Inventeurs:	BENNY, Paula-Beth Angelica Tiqui; (SG). RAGHUNATH, Michael; (SG). LANE, Ellen Birgitte; (SG). BADOWSKI, Cedric; (SG)		
Mandataire:	AMICA LAW LLC; 30 Raffles Place #14-01 Chevron House Singapore 048622 (SG)		
Données relatives à la priorité:	61/988,709 05.05.2014 US		
Titre	(EN) METHODS OF PRODUCING TISSUE-MIMETIC CONSTRUCTS AND USES THEREOF (FR) PROCÉDÉS DE PRODUCTION DE CONSTRUCTIONS MIMÉTIQUES DE TISSU ET UTILISATIONS ASSOCIÉES		
Abrégé:	(EN) The present invention relates, in various embodiments, to methods of producing a tissue-mimetic construct having a basement membrane, methods of producing an acellular scaffold containing an extracellular matrix (ECM), methods of producing a scaffold comprising a hydrogel that is enriched in ECM components, methods of treating a condition in a subject in need thereof with a tissue-mimetic construct having a basement membrane, and methods of assessing whether an agent is suitable for administering to a tissue. The invention further relates to tissue-mimetic constructs and scaffolds produced in accordance with the methods of the invention. (FR) La présente invention concerne, selon divers modes de réalisation, des procédés de production d'une construction mimétique de tissu ayant une membrane basale, des procédés de production d'un échafaudage acellulaire contenant une matrice extracellulaire (ECM), des procédés de production d'un échafaudage comprenant un hydrogel qui est enrichi en constituants de l'ECM, des méthodes de traitement d'une affection chez un sujet en ayant besoin à l'aide d'une construction mimétique de tissu ayant une membrane basale, et des procédés permettant d'évaluer si un agent est approprié pour l'administration à un tissu. L'invention concerne en outre des constructions et des échafaudages mimétiques de tissu produits selon les procédés de l'invention.		
États désignés:	AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW. Organisation régionale africaine de la propriété intellectuelle (ARIPO) (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW) Office eurasiatique des brevets (OEAB) (AM, AZ, BY, KG, KZ, RU, TJ, TM) Office européen des brevets (OEB) (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR) Organisation africaine de la propriété intellectuelle (OAPI) (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).		
Langue de publication:	anglais (EN)		
Langue de dépôt:	anglais (EN)		



Figure 8. Exemple d'un brevet issu de la base WIPO

Le SS-XEW produit automatiquement la présentation multidimensionnelle du corpus Scaffolds-WIPO, ce qui nous permet de définir et identifier les différentes corrélations existantes entre les attributs d'un même document et leurs corrélations avec les autres documents du corpus. Cette présentation est analysée ensuite par le SBDA-XEW, puis interprétée par le SDV-XEW. À titre d'exemple on peut bien analyser le réseau des inventeurs, comme nous le montre la figure 11.

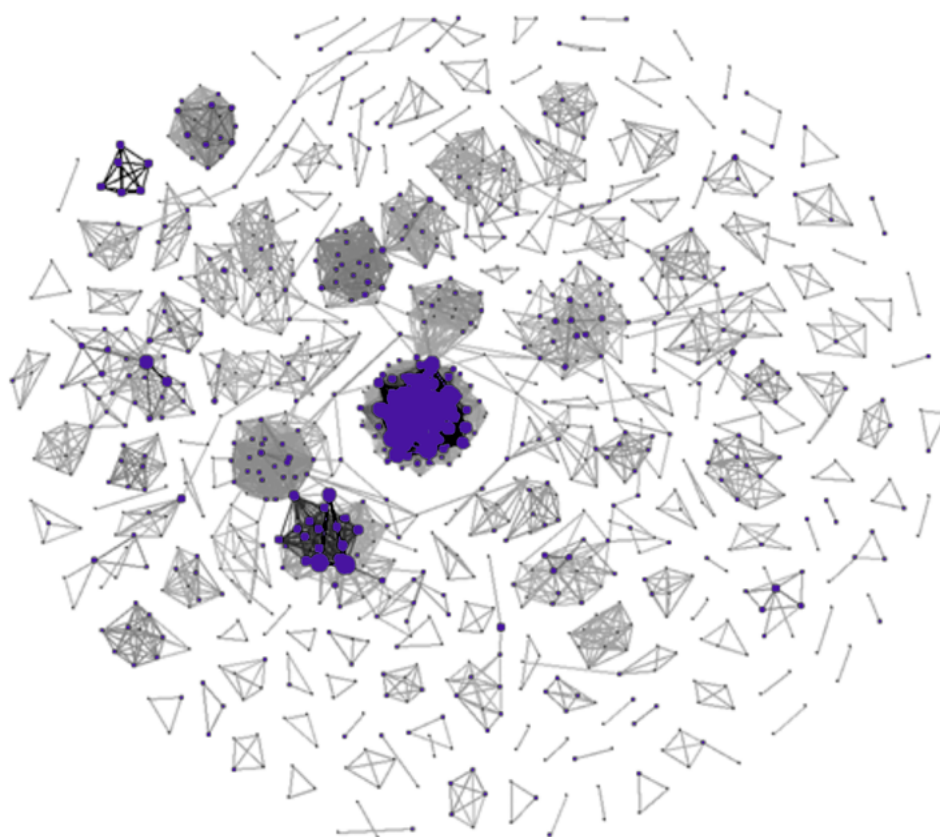


Figure 9. Graphe de clusters de réseau des inventeurs dans le domaine des scaffolds

Ce graphe nous montre deux principaux clusters désignés sur la figure par la grande concentration des noeuds rouge, qui représentent les équipes de recherche les plus importantes dans ce domaine. Cette visualisation macroscopique nous permet d'avoir une vision générale. Dans le but d'étudier les profils des inventeurs, leurs spécialités, leurs évolutions ... SDVXEW permet d'effectuer des analyses détaillées, à titre d'exemple la figure suivante nous montre le top 10 des inventeurs.

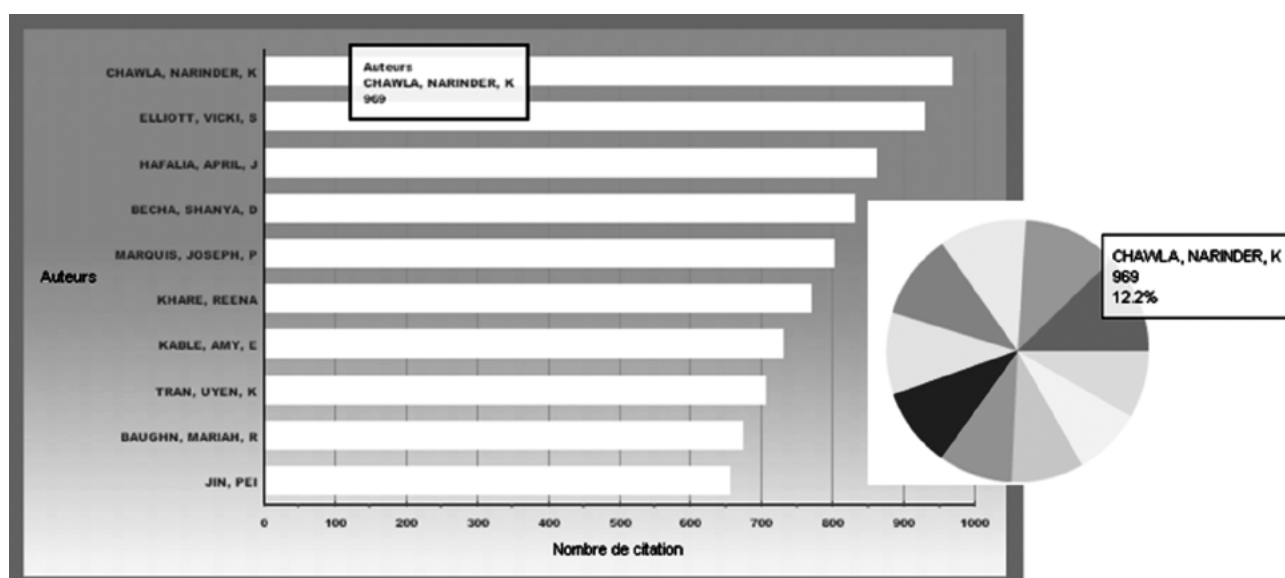


figure 10. Top 10 des inventeurs dans le domaine des scaffolds

La fonction de filtrage de SDV-XEW nous permet de naviguer dans l'environnement de leader, la figure présente toutes les cooccurrences de chercheur Chawla. Son profil, son réseau social, sa spécialité sont visualisés avec un simple clic (figure 11).

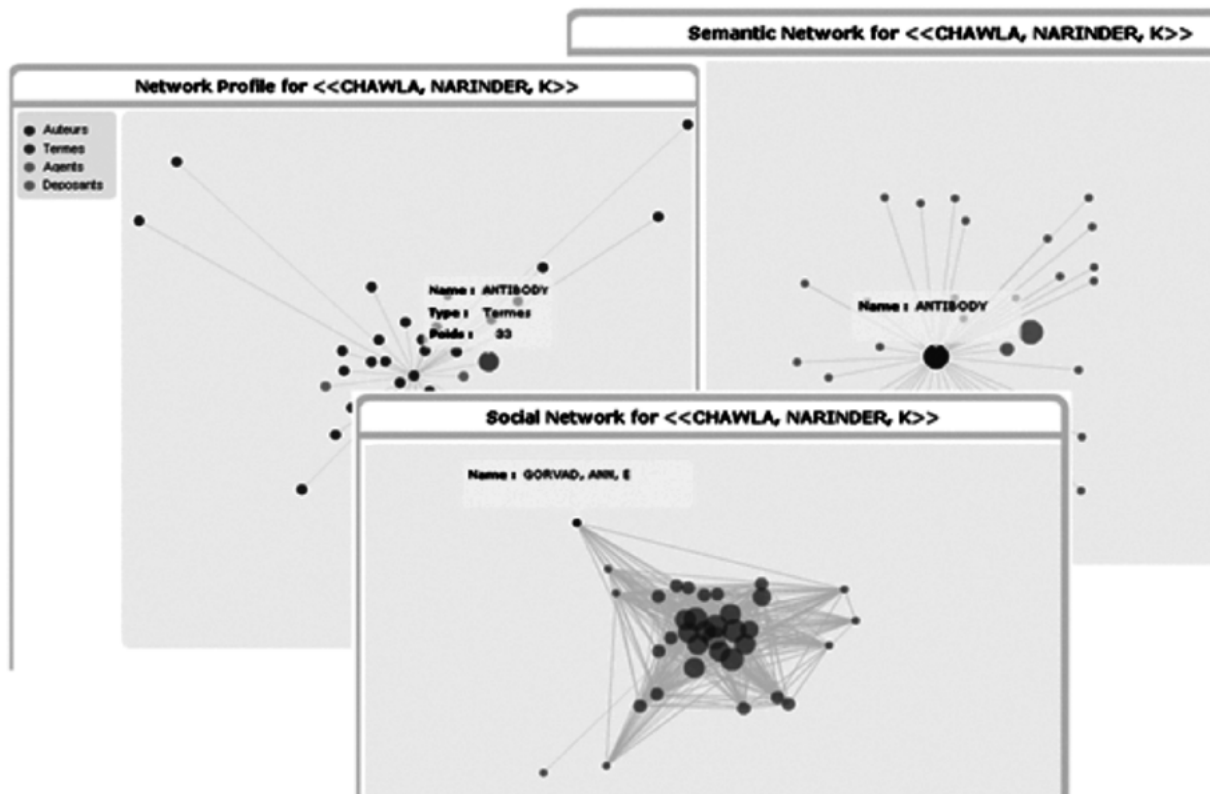


Figure 11. Le profil, le réseau social et le réseau sémantique de Chawla

7. Conclusion

Le SIE XEW se base sur une architecture modulaire des Web services, ce qui nous permet d'automatiser jusqu'à 80% du processus d'intelligence économique, selon l'étude de satisfaction effectuée pour la validation du prototype. XEW prend en considération l'objectif d'interactivité avec les utilisateurs nomades afin de faciliter la navigation dans la connaissance, et d'offrir les bonnes conditions de la veille pour améliorer la compétitivité d'un organisme public ou privé. En plus le service SS-XEW nous a permis d'optimiser le processus de la recherche et collecte d'information dans leurs formats natifs, puisque plus de 90% des cas rencontrés peuvent ainsi être traités sans aucun reformatage, sur l'ensemble des analyses stratégiques effectuées pour nos collaborateurs. Ce qui présente un avantage par rapport aux autres outils existants, vu que la majorité de ces outils demande le reformatage. XEW(<http://www.xploorew.com/XEWProject/pages/login.jsf>) vient compléter l'offre d'outils dédiés à l'intelligence économique, par la possibilité de gérer et d'exploiter l'environnement informationnel, de manière interactive et sécurisée.

8. Bibliographie

- Besson, B., Deschamps, C., Andriamiamina, H., Archanbaud, B., Baret, C., Besnard, M., Blanchet, A., Bouakaz, A., Baptiste, B. (2010). *Méthodes d'analyses appliquées à l'intelligence économique*. ICOMTEC Poitiers.
- Richard J. P. (2005). *De la veille stratégique à la sécurité de l'information*. Décision Informatique, N° 625.
- Boukrara, A., Chalal, R. (2010). *Une approche d'ingénierie des besoins pour la conception des systèmes d'information d'intelligence économique*. Colloque international de Veille Stratégique, Scientifique et Technologique (VSST), Toulouse.
- Büttcher, S., Clarke, C. L. A., Cormack, G. V. (2010). *Information Retrieval: Implementing and Evaluating Search Engines*. MIT Press.
- Casselmann R. M, Samson D. (2011). *Internal, collaborative and competitive knowledge capability*. Proceedings of the 44th Hawaii International Conference on System Sciences (HICSS), pp 1-10.

- Chaunac A., Diakite Z., Naquin p., Wissner m. h. R. (2011). *Le système d'intelligence économique américain*. Dossier de recherche ANAJ-IHEDN, Comité Défense Economique Section Intelligence Economique.
- Chen T., Hou K., Yu K. (2008). *Study on model about competitive intelligence system of enterprise based on data mining under electronic commerce environment*. International Seminar on Business and Information Management (ISBIM), pp 203-206, Wuhan.
- Chen T., Niu X., Yang W. (2009). *The application of Web data mining technique in competitive intelligence system of enterprise based on XML*. Third International Symposium on Intelligent Information Technology Application (IITA), pp 396-399, Nanchang.
- Colakoglu, T. (2011). *The Problematic Of Competitive Intelligence: How To Evaluate& Develop Competitive Intelligence?*. Procedia - Social and Behavioral Sciences, Volume 24, pp 1615-1623.
- David, A., Thiery, O. (2002). *Application of EQuA2te Architecture*. Information and Communication Technologies applied to Economic Intelligence (ICTEI), Ibadan, Nigeria.
- El haddadi, A., Dousset, B., Berrada, I., Loubier, I. (2010). *Les multisources dans un contexte d'intelligence économique*. EGC 2010, Hammamat – Tunisie.
- El haddadi, A., Dousset, B., Berrada, I. (2012). *Establishment and application of competitive intelligence system in Mobile Device*. Journal of Intelligence Studies in Business, Vol. 1, pp 87-96.
- El Haddadi, A., El Haddadi, A. et Fennan, A. (2014). *XEW 2.0, le SIE pour l'analyse de Big Data*. Les Journées d'Intelligence Économique (JIE), Tanger.
- Ginsburg M., Kass A., Yah P. Z. (2009). *Exploring two enterprise semantic integration systems*. Proceedings of the 42nd Hawaii International Conference on System Sciences (HICSS), pp 1-10.
- Guoling S., Xiaorong M., Xu W. (2009). *Designing a network acquisition system of competitive intelligence*. Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing, pp 686-689.
- Jing D., Bo H., Wen L. (2015). *Research on entreprise competitive intelligence development and strategies in the Big Data Era*. Computer and Information Technology (CIT), pp 658-663.
- Li J., Dong Y. (2010). *Post-controlled vocabulary compiling in competitive intelligence system*. International Conference on Information Management and Engineering (ICIME), pp 560-563.
- Li K., Yuchen F., Xiaoke Z., Quan L., Zhiming C. (2007). *Study on competitive intelligence system based on Web*. Workshop Intelligent Information Technology Application, pp 339-342.
- Liu, P., Xia, X.H. (2009). *Web mining technology in competitive intelligence system research*. International Symposium on Information Engineering and Electronic Commerce (IEEC), pp 524-527.
- Moinet, N., Frion, P. (2009). *Knowing is action: from noticing to sense-making*. 3d European Competitive Intelligence Symposium: Competing, Consuming and Collaborating in a Flat World, Stockholm – Suède.
- Nelke, M. (2012). *Competitive intelligence*. Strategic Business Development for Information Centres and Libraries, pp 19-41.
- Olusoji B. O., Amos D. (2010). *AMTEA : Tool for creating and exploiting annotations in the context of economic intelligence (competitive intelligence)*. International Conference on Information Reuse and Integration (IRI), pp 249-252.
- Othenin, G. C., Caron, C. et Guillemette, M. C. (2011). *When competitive intelligence meets geospatial intelligence*. Proceedings of the 44th Hawaii International Conference on System Sciences (HICSS), pp 1-10.
- Radun v. (2006). *The internal and external communication of intelligence and the competitive intelligence process: the case of the higher educational sector in Serbia*. Technology Management PICMET Proceedings.
- Saba, M., DE Rémur, D. B. et Gerbaix, S. (2014). *ICT implementation. Going beyond expectations? An essay of interpretation through competitive intelligence*. International Strategic Management Review.
- Salles M., Clermont Ph., Dousset B. (2000). *Une méthode de conception de système d'IE*. Colloque IDMME'2000, Canada.
- Seligmann, P.S., Wijers, G.M., Sol, H.G. (1989). *Analysing the structure of I.S. methodologies, an alternative approach*. Proceedings of the 1st Dutch Conference in Information Systems, Amersfoort, The Netherlands.
- Sewdass, N., Du Toit, A. (2014). *Current state of competitive intelligence in South Africa*. International Journal of Information Management, Volume 34, Issue 2, pp 185-190.
- Štefániková, L., Masárová, G. (2014). *The Need of Complex Competitive Intelligence*. Procedia - Social and Behavioral Sciences, Volume 110, pp 669-677.
- Tan, A.H, Ong, H.L., Pan, H., Ng, J., Li, Q.X. (2004). *Towards personalised Web intelligence*. Knowledge and Information Systems 6(5), pp 595-616.
- Thakor, A. V. (2013). *Strategic information disclosure when there ; is fundamental disagreement*. Journal of Financial Intermediation, volume 24, Issue 2, pp 131-153.
- Weidong G. (2008). *Establishment and application of competitive intelligence system in china*. Second International Symposium on Intelligent Information Technology Application (IITA), Volume 1, pp 279-283.
- Xiaobo T., Lian L. (2010). *A integration method of competitive intelligence and knowledge management system for corporate decision-making (WiCOM)*. pp 1-4.

Proposition of a new ex-ante evaluation method for adapting and selecting ERP solutions

Abdelilah Khaled

Equipe TIME, ENSIAS, Université Mohammed V de Rabat - Madinat Al Irfane, Maroc.
khaled@igf.finances.gov.ma

Mohammed Abdou Janati Idrissi

Equipe TIME, ENSIAS, Université Mohammed V de Rabat- Madinat Al Irfane, Maroc.
janati@ensias.ma

Résumé

Acquérir une solution ERP est une activité à la fois risquée, fastidieuse et complexe pour toute organisation. Il a été largement rapporté que le choix inapproprié d'une solution ERP est l'une des principales causes derrière l'échec de sa mise en œuvre. Dans cet article, notre objectif est d'élaborer une nouvelle méthode d'évaluation ex-ante d'adaptation et de sélection des solutions de type ERP. Cette méthode, dénommée SEVALERPS (Systematic EVALUation of ERP Systems), est destinée à aider les organisations, plus particulièrement les grands comptes, à mieux adapter et à évaluer des solutions ERP potentielles, afin d'en choisir celles qui réalisent les meilleurs compromis de leurs exigences. Afin de gérer les différents aspects d'évaluation des solutions, la méthode SEVALERPS se base sur plusieurs techniques mathématiques. Il s'agit de la programmation linéaire à variables binaires pour le choix des meilleurs scénarii d'adaptation des solutions ERP, de la technique MACBETH pour exprimer, sous une forme quantitative ou qualitative, les préférences de l'équipe de sélection sur les critères d'évaluation, et de l'intégrale discrète de Choquet pour gérer les interdépendances qui pourraient exister entre les critères d'évaluation.

Abstract

Acquiring an ERP system is one of the most risky, tedious and complex decision making activities for any organization. It has been widely reported that selecting an inappropriate ERP is one of the major reasons for its implementation failure. In this paper, we will develop a new ex-ante evaluation method for ERP system tailoring and selection. The proposed method, called SEVALERPS (Systematic Evaluation of ERP Systems), is basically elaborated to help organizations, especially large ones, to better customize and evaluate the potential ERP solutions in order to choose the ones that meet their best requirements' tradeoff. SEVALERPS relies on many sound mathematical techniques to handle various evaluation aspects: 0-1 linear programming to choose the best ERP tailoring scenarios, MACBETH to express the preferences of the selection team in both a qualitative and quantitative way, and the discrete Choquet integral to address interdependencies that might exist among evaluation criteria.

Mots-clés

Évaluation ERP, Couverture fonctionnelle, Sélection ERP, Adaptation ERP, MACBETH, Intégrale de Choquet, Programmation linéaire à variables binaires

Keywords

ERP Evaluation, ERP Functional coverage, ERP Selection, ERP Tailoring, MACBETH, Choquet Integral, 0-1 linear Programing.

1. Introduction

The industrial concept of Enterprise Resource Planning (ERP) denotes an integrated commercial off-the-shelf software package which streamlines and encompasses a large spectrum of business processes within organizations (Chofreh *et al.*, 2014). Since the appearance of these solutions on the market in the 1990s, they have known a spectacular development. Indeed, organizations have become aware of the strategic competitive advantages that could be bestowed from such technology, if its implementation is successfully done. They aim to sustain their market share in a highly severe competition by adopting the main worldwide best practices generally embedded in these ERP solutions. Basically, these packaged solutions are destined to cut redundant costs, raise quality, receive pertinent information in a timely manner and improve customer satisfaction (Martin *et al.*, 2014).

Historically, ERP systems were destined to large organizations that were willing and able to spend tens or hundreds of millions of dollars on an integrated software system (Katerattanakul *et al.*, 2014). However, over the past few years, the ERP market has shown signs of saturation as large companies almost completed their significant ERP implementations. The time was ripe for the main vendors to steadily maintain the challenge of Small and Medium Enterprises (SMEs) (Consulting, 2013). On the one hand, the vendors have started turning their marketing sights towards the mid-market and redesigning their strategies to match its requirements. On the other hand, SMEs would unfavorably disrupt their business activities if they failed to upgrade their information systems to communicate with their supply chain partners or with those of their corporate headquarters (Pérez-Salazar *et al.*, 2013).

However, despite the pace with which numerous organizations have embraced and implemented ERP packages, it should be noted that not all ERP implementations have given satisfactory results. Actually, since the dawn of ERP projects, it has been widely reported that many of them fail to yield the expected payoffs (Katerattanakul *et al.*, 2014). More than half of these projects have been reported in the literature to be a pure loss, as they often fail to reach their objectives (Poba-Nzaou and Raymond, 2011), let alone the large scale investments engaged by their adoption, which tend to be a very heavy burden on the organizations' budgets. Thereby, given their alarmingly low level success rate, ERP projects are now regarded as highly risky activities that might jeopardize the very existence of even large organizations; the case of smaller sized enterprises is more problematic. In fact, SMEs have limited budgets and little experiences in addressing and overcoming these matters (Chofreh *et al.*, 2014).

But, surprisingly, an extensive part of the academic literature has dealt mainly with ERP implementation and post-implementation issues and has skipped the pre-implementation stage of these solutions, especially the way they are selected (Addo-Tenkorang and Helo, 2011); (Zeng and Skibniewski, 2013). Actually, the major purpose of ERP selection activity is to identify the alternative that would best cover the business requirements of a given organizational context. In light of this, the choice of the appropriate ERP solution is obviously one of the most critical success factors on which relies its future implementation. This seems to explain why the research that has previously been done to address the high failure rate of ERP projects hasn't yet brought satisfactory results. Given that organizations are still ill-equipped to select the most promising solutions which could fit their business requirements, we believe that it is highly justifiable to express more interest in developing selection methods and models geared towards ERP solutions.

According to (Pérez-Salazar *et al.*, 2013), ERP evaluation and selection have become increasingly difficult for the decision makers due the hundreds of software products available in the marketplace. Other authors argue that methodologies developed for custom commercial off-the-shelf solutions may not all be applicable to ERP ones (Yıldız and Yıldız, 2014). Organizations will have to be armed with the right tools to avoid serious errors and make purchases that will generate good returns (Munkelt and Völker, 2013).

The selection of an ERP solution is basically a typical MCDA (Multi Criteria Decision Aid) problem (Pérez-Salazar *et al.*, 2013). This is due to its propensity to operate under multiple, often conflicting criteria and the discrete decision space in which the decision is made. By applying MCDA methods, justifiability and accountability will be improved considering that they are regularly seen as the pre-requisites of a complex and risky purchasing decision (Razmi and Sangari, 2013). Taking this into account, we argue that there is a key feature that should be also considered when it comes to ERP selection. Actually, ERP solutions, unlike custom packaged software, are customizable ones, giving them more flexibility to best meet the specific requirements of each organization. In this regard, ERP candidates should be undoubtedly evaluated according to their anticipated fitness with the business requirements instead of with their current one. In other words, the selection team must be able to determine the best tailoring actions that could resolve, in an optimal way, the mismatches relating to each ERP candidate solution. This resolution should be done by seeking, within limited total cost of ownership, a trade-off between maximizing the functional fitness of each candidate solution and minimizing its non-adaptation risk. And it is only after determining how each ERP candidate solution should be customized that organizations should, in order to choose their best solutions, conduct global evaluations by considering both alternatives' functional and non-functional criteria. Indeed, the selection criteria should go beyond functional aspects to encompass strategic, economic and technical ones.

Despite their importance, the selection methods and approaches proposed in the literature still fail to deal with this aforementioned key feature. With this end in view, in this paper, we will develop a new ex-ante evaluation method

that aims to help organizations, especially large ones, to customize the ERP candidate solutions at hand, and choose among them those that could best meet their business requirements. This method, denoted by SEVALERPS (Systematic EVALuation for ERP Selection), is designed to tackle the current shortcomings underlined above. Actually, SEVALERPS development addresses the two following research questions:

- Given that ERP packages provide numerous options to customize them, how should the selection team choose the best tailoring actions that could offset the functional coverage improvement against the non-adaptation risk reduction?
- Due to the heterogeneous kinds of evaluation criteria and their interdependencies, which systematic evaluation processes and models should the selection team adopt to evaluate locally and globally the ERP candidate solutions against these criteria?

In the remainder of this paper, we underline in a nutshell some of the most common criticism made against the evaluation methods and approaches proposed in the literature relating to ERP selection. We introduce the evaluation process and model of SEVALERPS method, and discuss how the developed method addresses the two aforementioned research questions. A public administration case study is detailed to illustrate how SEVALERPS could be applied in a real context. Indeed, we elaborate, in more detail, the background of using SEVALERPS in this specific case study by focusing on its application scenarios. Analyses of the obtained results are also given. Finally, we present some perspectives for our future research.

2. Previous ERP selection approaches in a nutshell

Over the past two decades, both researchers and practitioners have progressively started to develop methodologies, approaches and methods to assess the fitness of some software packages with given organizational contexts. This growing body of academic literature witnesses the importance of the pre-implementation stage of software adoption, usually referred to by the headings of package procurement, acquisition or selection (Addo-Tenkorang and Helo, 2011).

Our literature review and classification of the previous ERP selection methods and approaches have revealed that the past research addresses, in general, the following aspects: evaluation and selection processes, evaluation models and techniques, selection criteria and automation tools supporting the proposed methods. However, when it comes to their scope, these previous works still remain general, even if their authors pretend that they are especially developed to handle the special case of ERP systems. Indeed, in these methods, neither the evaluation models nor the selection processes take into account that, unlike ordinary software packages, ERP ones are customizable solutions (Please refer to our literature reviews presented in our previous works: (Khaled and Janati-Idrissi, 2011); (Khaled and Janati-Idrissi, 2012)). In this regard, we argue that mismatch handling activity geared towards enhancing the global coverage of the required functionalities should be at the heart of any ERP evaluation approach.

Furthermore, the analytic hierarchy processes (AHP) (Saaty, 2013) and the weighted scoring method (WSM) are the most widely used multi-criteria evaluation models to determine the global scores of candidate packages. Nonetheless, AHP and WSM rely on an additive aggregation model which assumes that evaluation criteria must be totally independent. Actually, such assumptions are often unrealistic in practice (Sen *et al.*, 2009); (Pérez-Salazar *et al.*, 2013).

Finally, a mandatory feature of package software selection methods is the ability to deal with both quantitative and qualitative data when it comes to evaluating candidate packages on selection criteria. In this regard, a great deal of the existent methods uses numerical models that transform qualitative judgments, voided out by the selection team, into quantitative ones. The final score of each ERP candidate is obtained through compensating mutually its preference scores, basically heterogeneous, on the various criteria. However, so as to obtain valid results, the numerical representations relating to these criteria must be at least commensurate among each other in order to ensure their aggregation significance (Bustinzá *et al.*, 2013). In practice, this assumption seems to be ignored by many of the evaluation methods proposed so far. A systematic procedure to elaborate commensurate representations based on the qualitative judgments of the selection team should then be considered when conducting package evaluation activity.

The next section of this paper presents the SEVALERPS method which is mainly developed to address the highlighted limitations of the previous methods.

3. Presentation of SEVALERPS method

SEVALERPS relies on a systematic six stage process that defines the activities that should be carried out to evaluate ERP candidate packages. The flow chart of this process is depicted in Fig. 1. It comprises the following stages: requirements and criteria gathering, candidate searching and screening, functional gap analysis, mismatch

handling, elementary and global evaluations. The remainder of this section describes, in more details, each stage of SEVALERPS's process and discusses how the proposed method overcomes the shortcomings highlighted in the literature review section.

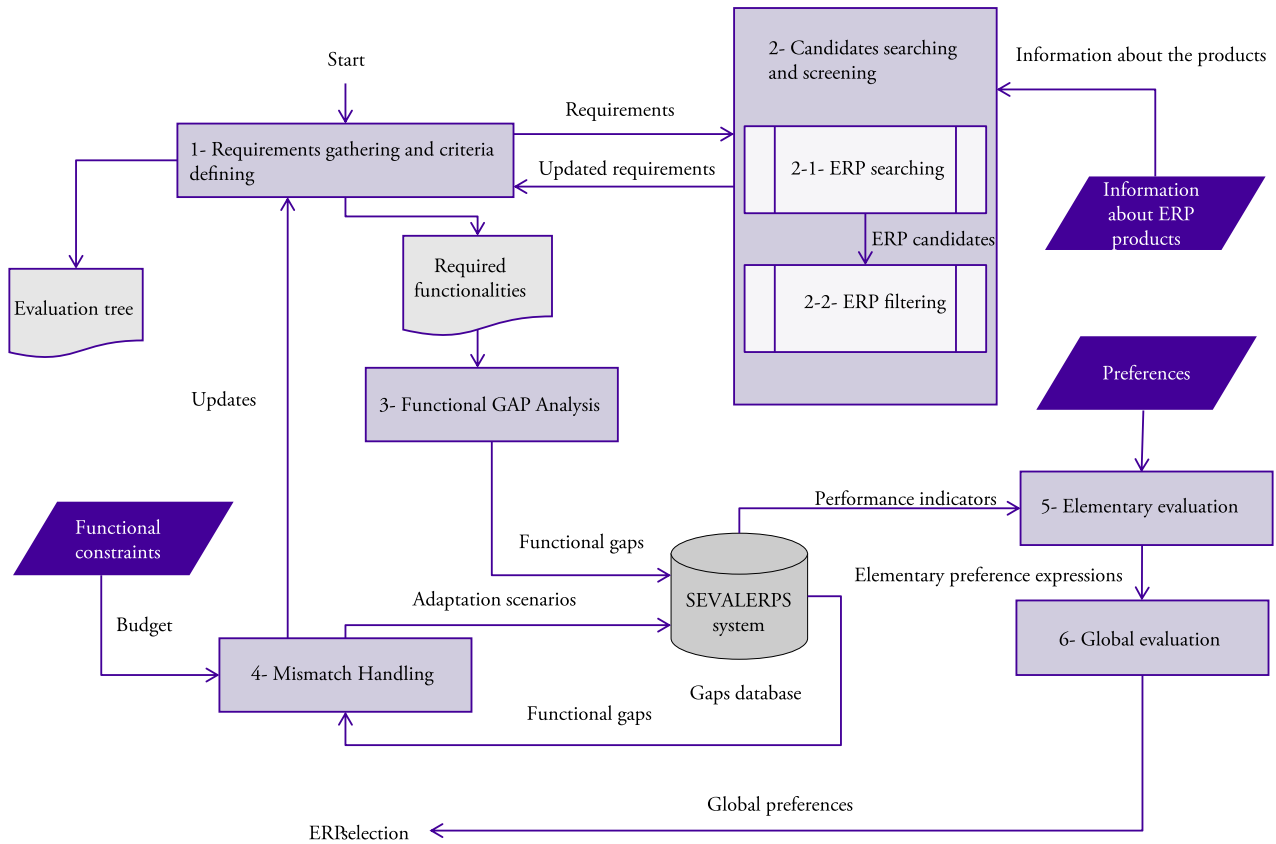


Figure 1. SEVALERPS evaluation flow chart

3.1. Requirements and criteria gathering

First of all, SEVALERPS is a requirement-driven selection method that enables organizations to rely on both functional and non-functional requirements to define their evaluation criteria, which will be used to evaluate and then to choose the appropriate candidate package. In this regard, numerous methods have been proposed in the literature to support requirements gathering (Zoukar *et al.*, 2013). However, these methods could be classified into two groups of strategies that could be adopted to manage requirement specification.

By adopting the first strategy, the selection team defines all its requirements from scratch based only on the organization business needs without considering what the available ERP solutions on the market could really offer. Even if this strategy would be ideal for the organization, it requires both time and effort and would lower the chances to find a solution that could satisfy all the specified requirements.

The second strategy suggests a solution-driven paradigm to define these requirements. Even though it requires less time and effort, it is considered less flexible when it obliges the organization to adopt the predefined business logic of a given ERP system.

Each strategy has its strengths and weaknesses. Nonetheless, when it comes to ERP packages, choosing one of them must take into account that these systems, unlike in-house ones, are not developed from scratch. This means that requirements specification should be kept at a moderate abstraction level in order not to discard all available ERP options.

For this reason, SEVALERPS uses mainly the first strategy to capture the organization's high-level requirements and to point out some candidate solutions that would satisfy them. The second strategy is also used during the candidate package examination so as to enrich and to update the initial requirements obtained by the first strategy. To support the selection team in its functional requirement specification, SEVALERPS suggests using UML's use case formalism to describe, in a higher abstraction level, what functionalities are expected from the future solution. Furthermore, characteristics that define how the future system must operate could also be taken from some standard quality models such as ISO-9162 (ISO9126, 2001).

3.2. Candidate searching and screening

In this stage, a wide market searching activity geared towards looking for ERP potential solutions and inquiring about them should be immediately carried out so as to detect the most promising ones and to narrow their spectrum. Basically, this searching and filtering activity relies on the criteria defined in the previous stage. Besides, many information sources relating to ERP solutions could be used in this stage. They include internet, benchmarking studies conducted by consulting firms, white papers published by ERP editors, functional and technical brochures of some solutions, specialized conferences and also requests for proposals (RFP) that could be launched by the organization.

3.3. Functional gap analysis

The third stage concerns detecting and assessing mismatches relating to the ERP candidate solutions with reference to the organizational requirements. That's why the selection team should appraise to what extent a given ERP candidate package would satisfy a required functionality $f_j \in F' = \{f_1, f_2, \dots, f_j\}$. As a result, we define a satisfaction function, denoted by SAT, which shows initial satisfaction levels given by ERP solutions. This function is illustrated in (1).

$$SAT: ERP \times F' \rightarrow [0,1] \quad (1)$$

$$(ERP, f_j) \rightarrow a_{ij}$$

3.4. Mismatch handling

Unlike ordinary packages, ERP solutions are customizable ones. This feature must be taken into account when it comes to their evaluation. For this reason, ERP evaluation shouldn't be undertaken based on the current functional coverage of these solutions, but rather on their anticipated fitness resulting from tailoring them, in order to handle the mismatches detected in the previous stage.

According to (Brehm *et al.*, 2001), there are nine tailoring patterns that identify adaptation strategies that the selection team could apply to handle identified mismatches. These tailoring patterns are: configuration, bolt-on, screen mask, extended reporting, user exit and programming. For each ERP candidate package, the choice of the optimal mismatch handling scenario depends on three main factors: maximizing anticipated functional coverage, reducing non-adaptation risk and respecting the financial resources' limit.

This end in view, we introduce the concept of adaptation strategies, denoted by $S_{ijk}, k \in IN$, relating to a given ERP_i candidate product and a required functionality, denoted by f_j . Each adaptation strategy S_{ijk} aims to improve the ERP_i functional coverage of f_j from a_{ij} to b_{ijk} . Furthermore, the potential implementation of S_{ijk} incurs to the organization a non-adaptation risk, denoted by r_{ijk} and induces an additional adaptation cost, denoted by c_{ijk} .

However, at most, only one adaptation strategy S_{ijk} must be chosen to handle a given mismatch relating to ERP_i and f_j . For this reason, we consider a binary unknown factor x_{ijk} to mention whether an adaptation strategy is chosen or not. Table 1 sums up the parameter set used in SEVALERPS mismatch handling model.

To determine which adaptation strategies are included in the optimal tailoring scenario of a given ERP candidate package, binary values relating to x_{ijk} must be assigned. For this purpose, SEVALERPS relies on the linear optimization model of (2) to determine these values.

In this optimization system, the objective function O_i makes a tradeoff between functional coverage improvement and its underlying non-tailoring risk, through the use of the multiplication operator. In this regard, the more the adaptation risk is important, the less is the value of O_i ; and, conversely, the more the functional coverage's improvement is important, the more is the value of O_i .

In addition, the first constraint of this optimization system ($\sum_k x_{ijk} \leq 1$) indicates that, at most, only one adaptation strategy must be chosen. The second constraint ($\sum_{j,k} x_{ijk} c_{ijk} \leq \text{cost}_i$) indicates that the elementary adaptation costs' sum mustn't exceed the budget limit, cost_i , relating to ERP_i .

$$(\forall i) \begin{cases} \max(O_i) \\ O_i = \sum_j |a_{ij} \neq 1| w_j (b_{ijk} - a_{ij}) (1 - r_{ijk}) x_{ijk} \\ (\forall j | a_{ij} \neq 1) \sum_k x_{ijk} \leq 1 \\ (\forall j, k | a_{ij} \neq 1) \sum_{j,k} x_{ijk} c_{ijk} \leq \text{cost}_i \end{cases} \quad (2)$$

Parameter	Description
$ERP_i, i=1 \dots I$	ERP products from which the organization has to choose its solution
$f_j, j=1 \dots J$	Required functionalities from ERP products
w_j with $\sum_j w_j = 1$	f_j 's importance weight describing its importance in achieving organizationnal goals
$S_{ijk}, k=1 \dots K$	Tailoring strategies related to f_j and ERP_i . These strategies are destined to handle the identified mismatches between the required functionality and the one proposed by the ERP. These mismatches are identified in the Functional Gap Analysis stage of SEVALERPS's process.
$a_{ij} \in [0,1]$	Initial functional coverage relating to f_j and ERP_i
$b_{ijk} \in [0,1]$	Anticipated functional coverage relating to f_j and ERP_i after applying S_{ijk}
$r_{ijk} \in [0,1]$	Tailoring risk relating to S_{ijk}
$c_{ijk} \in \mathbb{R}$	Tailoring cost relating to S_{ijk}
$cost_i \in \mathbb{R}$	Budget limit allowed for ERP_i tailoring
$x_{ijk} \in \{0,1\}$	Decision binary unknown factor to mention whether the adaptation strategy S_{ijk} is chosen ($x_{ijk}=1$) or not ($x_{ijk}=0$).

Table 1. Parameters of SEVALERPS model

By solving this linear 0-1 programming system, we define, for each ERP_i , the optimal tailoring scenario expected under a specific budget constraint. By adopting this mismatch handling model, we provide an answer to our first research question. Finally, based on the values of the aforementioned adaptation model, we define in Table 2, in order to evaluate ERP candidate packages in the next stages of SEVALERPS, a set of performance indicators relating to the anticipated functional coverage, the non-adaptation risk, the adaptation cost and the non-guaranteed functional coverage of the ERP candidate packages.

Performance indicators	Description
Anticipated Functional coverage (ERP_i) = $\sum_j w_j \max(\sum_k b_{ijk} x_{ijk}, a_{ij})$	It represents the new functional coverage relating to ERP_i after its adaptation.
Non-adaptation risk (ERP_i) = $1 - \frac{\sum_{j,k a_{ij} \neq 1} \Gamma_{ijk} x_{ijk}}{\sum_{j,k a_{ij} \neq 1} w_j \Delta_{ijk} x_{ijk}}$ With $\Delta_{ijk} = (b_{ijk} - a_{ij}) \dots$ and $\dots \Gamma_{ijk} = w_j \Delta_{ijk} (1 - r_{ijk})$	It represents the global risk associated to the all adaptation strategies.
Adaptation cost (ERP_i) = $\sum_{j,k a_{ij} \neq 1} c_{ijk} x_{ijk}$	It represents the global cost induced by the chosen adaptation strategies
Non-guaranteed functional coverage (ERP_i) = $\sum_{j,k a_{ij} \neq 1} w_j \Delta_{ijk} x_{ijk} \Omega_{ijk}$ With $\Omega_{ijk} = \begin{matrix} 0 & \text{if } S_{ijk} \equiv & \text{customization} \\ 1 & & \text{Otherwise} \end{matrix}$	It represents the functional coverage's part that the organization would lose immediately after the ERP_i version update.

Table 2. Performance indicators of SEVALERPS

3.5. Elementary evaluation

In this stage, the selection team has to express quantitatively their preferences relating to the ERP candidate packages on the lower criteria of the hierarchical evaluation tree, defined in stage 1. The preference expression relating to the upper criteria is defined, in the next stage of SEVALERPS, as the aggregation of the preferences defined on their sub-criteria.

In SEVALERPS, we use the multi-criteria Measuring Attractiveness by a Categorical Based Evaluation Technique (MACBETH) (Bana e Costa *et al.*, 2011) to handle the elementary evaluation of candidate packages. MACBETH is mainly developed in the context of multi-criteria decision aid, and it is based on sound mathematical foundations. MACBETH introduces the concept of cardinal scales to define the curve of the selection team preferences with reference to the alternatives. The main reason behind choosing this technique is that it requires only verbal judgments that qualify the attractiveness differences between every two single actions to quantify their underlying attractiveness.

MACBETH comprises seven predefined judgments which describe the difference of attractiveness between every two candidate products: No difference of attractiveness, Very weak difference of attractiveness, Weak difference of attractiveness, Moderate difference of attractiveness, Strong difference of attractiveness, Very strong difference of attractiveness, Extreme difference of attractiveness. MACBETH introduces two reference actions: SUP and INF. Those actions denote respectively the best and the worst potential actions relating to the evaluation criterion.

In the context of elementary evaluation of ERP candidates, an illustrative example of a judgment matrix relating to the comparison of three ERP systems (ERP A, ERP B and ERP C) with respect to the security criterion is illustrated in Fig. 2. The MACBETH interval scale is obtained thanks to the M-MACBETH software that supports MACBETH method.

	sup	ERP A	ERP B	ERP C	inf	Current scale	
sup	no	moderate	positive	positive	positive	1.00	extreme
ERP A		no	moderate	positive	positive	0.73	v. strong
ERP B			no	mod-strg	positive	0.45	strong
ERP C				no	weak	0.18	moderate
inf					no	0.00	weak
							very weak
							no

Figure 2. Judgment matrix of MACBETH

For more details about using MACBETH in ERP selection, the reader is asked to refer to our previous work (Khaled and Janati-Idrissi, 2011b).

3.6. Global evaluation

The global preference score relating to an ERP candidate product is defined through aggregating preferences values along the evaluation criteria tree. Based on these scores, recommendations about optimal solutions are given with respect to their decreasing ranking order. Accordingly, the best candidate is the one that has the highest score. The originality of SEVALERPS is its introduction of a new aggregation function which is based on the discrete Choquet Integral (Choquet, 1953). In fact in order to represent the interdependencies among criteria, SEVALERPS relies on the criteria coalitions' importance concept.

Indeed, for a set of criteria, denoted by $N, (1_x, 0_{N-x})$ represents a binary vector which has the value of 1 on the criteria belonging to X and the value of 0 on the criteria belonging to $N-X$. The set of $\{ \forall X \subseteq N | (1_x, 0_{N-x}) \}$ represents the all possible coalitions of the criteria belonging to N . In order to assign an importance value to each coalition, SEVALERPS uses MACBETH to determine these values on interval scales. To illustrate this concept, we suppose that the selection team has to define the importance weights relating to the coalitions of the three following criteria: security, portability and extensibility. These importance weights are needed, for instance, to judge the quality of an ERP candidate package that has respectively the following three scores on those criteria: 5%, 35% and 45%.

If we suppose that the qualitative judgments of the evaluation team relating to the importance of differences between each two couples of the 3-uplets criteria coalitions are given in Fig. 3; then the strategic importance of each coalition is defined in the scale column illustrated in the same figure.

	upper	(1,1,0)	(0,1,1)	(1,0,1)	(0,1,0)	(1,0,0)	(0,0,1)	lower	Current scale
upper	no	very weak	weak	mod-strg	moderate	strg-vstr	v. strong	extreme	1.00
(1,1,0)		no	very weak	very weak	moderate	strong	strg-vstr	v. strong	0.86
(0,1,1)			no	very weak	weak	moderate	mod-strg	strong	0.73
(1,0,1)				no	weak	moderate	moderate	strong	0.70
(0,1,0)					no	very weak	weak	strong	0.51
(1,0,0)						no	very weak	moderate	0.35
(0,0,1)							no	weak	0.24
lower								no	0.00

Consistent judgements

Figure 3. Attractiveness' differences among criteria's coalitions

Accordingly, we denote by $\mu(X) | X \subseteq N$ the importance function that assigns weights to each coalition of the X criteria belonging to N as it is illustrated in (3).

$$\forall X \subseteq N \mu(X) = \text{SCALE}_{\text{MACBETH}}(1_X, 0_{N \setminus X}) \tag{3}$$

In order to extend the definition of μ from $\{0,1\}^n$ to $[0,1]^n$ (n is the number of elements within N), we interpolate the μ function within the $[0,1]^n$ domain. According to (Grabisch, 2006), the discrete Choquet integral is the only valid linear interpolator of such functions, called capacities. The Choquet integral relating to a μ capacity is defined in (4).

$$\left\{ \begin{array}{l} C_\mu(X) = \sum_1^n x_{\sigma(i)} [\mu(A^{\sigma(i)}) - \mu(A^{\sigma(i+1)})] \\ X = (x_1, x_2, \dots, x_n) \in [0,1]^n \end{array} \right. \tag{4}$$

σ is a n -permutation that ranges the elements of X as follows:

$$x_{\sigma(1)} \leq x_{\sigma(2)} \leq \dots \leq x_{\sigma(n)}$$

$$\left\{ \begin{array}{l} A^{\sigma(i)} := \{\sigma(i), \dots, \sigma(n)\} \\ A^{\sigma(n+1)} = \emptyset \end{array} \right.$$

For instance, in our previous example, by considering the importance values obtained through MACBETH as a capacity in the Choquet integral, we compute the following aggregated score of Fig. 3. This score is assigned to the preference vector of (5%, 35%, 45%) given above.

```
> Quality <-c(0.05,0.35,0.45)
>
>
> mu <- capacity(c(0,0.35,0.52,0.24,0.86,0.70,0.73,1))
>
> mu
      capacity
{ }          0.000000
{1}          0.350000
{2}          0.520000
{3}          0.240000
{1,2}        0.860000
{1,3}        0.700000
{2,3}        0.730000
{1,2,3}      1.000000
>
>
> Choquet.integral(mu,Quality)
[1] 0.293
```

Figure 4. Global scores computed by Choquet integral

Hence, by using this systematic evaluation process and model, SEVALERPS provides a quantitative approach to comparatively evaluate ERP candidate packages, which represents our answer to the second research question. For further information about using Choquet integral in evaluating ERP packages, the reader is asked to refer to our previous work: (Khaled and Janati-Idrissi, 2011a); (Khaled and Janati-Idrissi, 2011b).

4. Case study

In this section, in order to validate SEVALERPS, we present its experimentation through a case study relating to an organization belonging to the public administration in Morocco. Due to confidentiality concerns, we won't disclose the name of the organization and we denote it by "X".

4.1. Background

In this case study, our goal is to experiment to what extent SEVALERPS could predict and assess the risks relating to choosing one alternative among the others. Indeed, organization "X" aimed to acquire an ERP system to manage its public accounting process. This project was started in 2009 and resulted in choosing a solution that hasn't given satisfactory outcomes. For this reason, by applying SEVALERPS on this case study, we underline the SEVALERPS added value and see whether our method, applied in 2015, would have recommended the same risky solution chosen by organization "X".

The ERP system which should be chosen has to manage mainly the following aspects:

- Public Expenditure management;
- Public Income management;
- Public Accounting management;
- Public Debt management.

More precisely, the selected ERP system should cover the modules depicted in Fig.5.

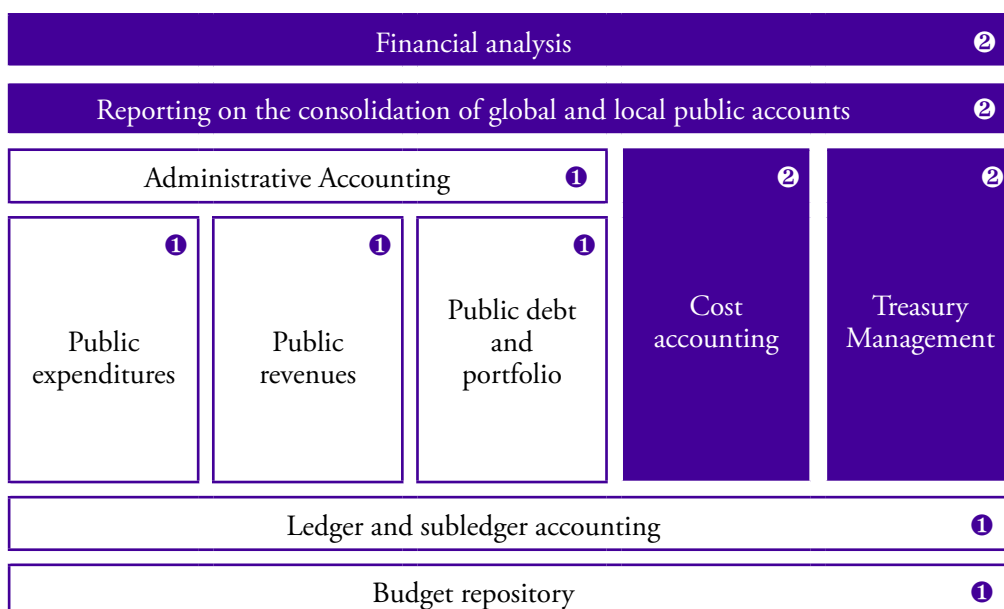


Figure 5. Functional project's scope

These modules represent the functional scope of the acquisition project. Furthermore, the modules tagged with number "1" represent the mission critical functions that must be covered immediately by the adopted system, whilst the ones tagged with number "2" are not urgently needed and could be implemented in the medium term. Hence the implementation strategy adopted by organization "X" is based on progressive acquisition of the modules described in the project's scope. In addition, it is worth mentioning that upon the completion of the implementation project, organization "X" aims to interoperate its ERP system by exchanging financial data with other information systems belonging to other public departments, such as the tax and the custom ones.

Organization "X" has to choose among three market leader ERP systems, which we denotes here by "Solution 1", "Solution 2" and "Solution 3". These solutions have respectively the following three initial acquisition costs: \$5.8 million, \$4 million, and \$2 million. It should be noted that based on the organization financial resource constraint, the selection team of organization "X" has basically a tendency to choose solution 2.

In this overall context, it should be noted that the selection team has already defined the required functionalities relating to each module as well as the evaluation criteria tree which will be used to judge the potential ERP solutions. Accordingly, in this case study, SEVALERPS is applied starting from the third step of the evaluation process described in Fig.1. Besides, four main criteria categories are taken into account by the selection team in order to evaluate the candidate products: Functional, Technical, Strategic and Financial.

4.2. Application of SEVALERPS

SEVALERPS was applied, in this case study, with regard to three different scenarios:

- First evaluation scenario (short term), in which SEVLAEPRS is applied to recommend an ERP solution based only on the functionalities relating to the first module of Fig. 5;
- Second evaluation scenario (medium term), in which SEVALERPS is applied to choose an ERP solution based on the functionalities comprised in both the first and the second sets of modules of Fig. 5;
- Third evaluation scenario (long term), in which SEVALEPRS is applied to evaluate the three ERP solutions by taking into account the fact that the whole information systems of public departments should interoperate with each other.

4.2.1. First scenario

For each of the three potential ERP solutions, Table 3 shows the number of natively covered functionalities (before tailoring), the number of detected mismatches, and the one relating to the critical functionalities not covered by the standard releases of the ERP solutions.

	Solution 1	Solution 2	Solution 3
Natively covered functionalities	150	110	94
Detected mismatches	25	65	81
Critical functionalities	11	30	41

Table 3. Functional coverage of the three ERP candidate packages

In the SEVLAEPRS mismatch handling stage, the selection team has identified several tailoring strategies for each of the ERP potential solutions. By solving the 0-1 linear programming systems of (1) relating to these solutions, the impacts of the best tailoring scenarios on their performances are shown in Table 4.

	Solution 1	Solution 2	Solution 3
Anticipated functional coverage	10%	16%	5%
Number of tailoring strategies	18	53	46
Non-adaptation risk	30%	35%	32%
Total cost of ownership (% tailoring cost)	\$5,8 million (24%)	\$4,2 million (56%)	\$3,7 million (49%)

Table 4. Impact of the best tailoring strategies

In order to assign global evaluation scores to the three ERP solutions, the selection team has to consider interdependencies that might exist among the three criteria of functional coverage, tailoring risk and total cost of ownership. These three criteria share the strategic importance of 75%, whilst the criterion of technical performance was considered as independent and has the strategic importance of 25%. In order to define the importance weight of each coalition relating to the three interdependent criteria, SEVALERPS suggests using MACBETH to represent them on an numerical [0,1] scale. As a result, the aggregation of the three ERP solutions' preference values through the discrete Choquet integral provides the global scores presented in Table 5.

Evaluation criteria	Strategic Importance	Solution 1	Solution 2	Solution 3
Anticipated functional coverage	Coalitions	90% (+10%)	88% (+16%)	48% (5%)
Tailoring Risk		70%	65%	68%
Total cost of ownership		30%	50%	60%
Aggregated score by Choquet integral	75%	0,616	0,6309	0,5664
Technical performance	25%	0,91	0,89	0,35
Global preference score		0,69	0,70	0,51

Table 5. Global preference scores

In this scenario, SEVALERPS has recommended solution 2 as the best one, followed by solution 1 and solution 3. However, we notice that solution 1 and solution 2 have almost the same scores. Even if solution 1 provides more anticipated functional coverage than solution 2, its higher total cost of ownership has downgraded its ranking order.

4.2.2. Second scenario

Similarly to the first scenario, Table 6 describes the standard functional coverage of the three ERP solutions before tailoring them.

	Solution 1	Solution 2	Solution 3
Natively covered functionalities	250	209	175
Detected mismatches	45	86	120
Critical functionalities	25	55	97

Table 6. Functional coverage of the three ERP candidate packages

In this second scenario, solution 3 was discarded because some mismatches relating to a set of critical functionalities couldn't be resolved by the proposed tailoring strategies. Indeed, the number of tailoring strategies (60) is lower than the number of critical functionalities that must be covered through tailoring (97).

Contrary to the recommendation of the first scenario, the aggregated scores obtained for this scenario (see Table 7) show that, in the medium run, solution 1 is more suitable than solution 2. In fact, even if solution 1 remains a bit more costly than solution 2. The tailoring risk of solution 2 is too high to promote its selection. For this reason, the global score of solution 1 is significantly more important compared to the one relating to solution 2.

Evaluation criteria	Strategic Importance	Solution 1	Solution 2
Anticipated functional coverage	Coalitions	85% (+15%)	84% (+35)
Tailoring Risk		72%	35%
Total cost of ownership		50%	52%
Aggregated score by Choquet integral	75%	0,6754	0,4742
Technical performance	25%	0,91	0,89
Global preference score		0,69	0,73

Table 7. Global preference scores

4.2.3. Third scenario

In this last scenario, the selection team of organization "X" has decided to use the RatIop approach (Elmir and Bounabat, 2012) which is geared towards the assessment of the interoperability maturity level among many information systems. In this case study, the interoperability feature is considered as a sub criterion of the technical performance.

According to (Elmir and Bounabat, 2012), it evaluates the compatibility of these systems based on the external interfaces involved within business processes. RatIop takes into account the following three operational aspects:

- Interoperation potentiality (PI);
- Interoperation compatibility (DC);
- Operational performance (PO).

The key performance indicator defined by RatIop to evaluate interoperability is defined in (5).

$$\text{RatIop} = (\text{PI} + \text{DC} + \text{PO}) / 3 \quad (5)$$

Actually, the chosen ERP system must interoperate with ten other information systems, as it is shown in Fig. 6.

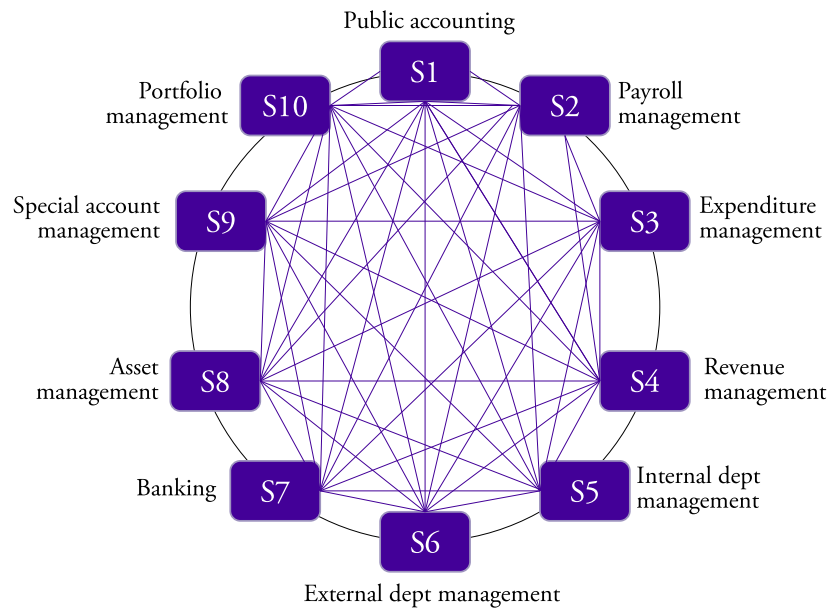


Figure 6. Interoperability among the ten information systems (Elmir and Bounabat, 2012)

According to a previous study conducted by (Elmir and Bounabat, 2012), the levels of interoperation relating to solution 1 and solution 2 are respectively 0.75 and 0.62. Based on these values, the global preference scores of solution 1 and solution 2 are presented in Table 8.

Evaluation criteria	Importance	Solution 1	Solution 2
Anticipated functional coverage	Coalitions	85% (+15%)	84% (+35)
Tailoring Risk		72%	35%
Total cost of ownership		50%	52%
Aggregated score by Choquet integral	75%	0,6754	0,4742
Technical performance	25%	0,83	0,75
Global preference score		0,71	0,54

Table 8. Global preference scores

As result, we conclude that in order to ensure better communication among the information systems and the future solution, it is advised to choose solution 1.

4.3. Analysis and discussion

By analyzing the results obtained in the three aforementioned scenarios, we can see as it is shown in TABLE 9 that the two last scenarios recommend clearly to choose solution 1, whilst the first scenario is barely making a distinction between solution 1 and solution 2. With regard to solution 3, it is either ranked in the last position or eliminated from further consideration.

	Solution 1	Solution 2	Solution 3
Scenario 1	0,69	0,70	0,51
Scenario 2	0,73	0,58	-
Scenario 3	0,71	0,54	-

Table 9. Summary of SEVALERPS evaluation

In the first scenario, we can see that the total cost of ownership has pledged for the selection of solution 2 with a difference of \$1.6 million compared to the first solution. We can also notice that both solution 1 and solution 2 have a quite similar tailoring risk given that the most required functionalities relating to the first set of modules are well implemented by these two ERP systems.

Conversely, in the second scenario, SEVALERPS suggests choosing solution 1. In fact, in order to cover the functionalities of the overall modules described in the acquisition project's scope, the tailoring risk of solution 2 reaches 65% compared to 28% relating to solution 1. In this situation, solution 1 seems to be more interesting, because it provides the same functional coverage as solution 2 with lower tailoring risk. The difference of the total cost of ownership between these two solutions is insignificant and doesn't justify a difference of tailoring risk estimated at around 37%. We recall that solution 3 was eliminated from further consideration due to its inability to cover some critical business functionalities, even after its tailoring.

The third scenario reconfirms the outcomes of the second one. Actually, the interoperability levels of the first two solutions with the key information systems belonging to other public departments show that solution 1 provides more favorable conditions to deal with integrating the business processes and data of the relevant departments.

4.4. Validation

As mentioned above, despite the tendency of organization "X" to choose solution 2 (For mainly financial considerations), SEVALERPS considers that choosing solution 1 is highly recommended. The results of the in-depth analysis obtained by applying SEVALERPS method were presented to the senior officials and the decision makers of organization "X". They recognized the interest of SEVALERPS method and systematic approach with which this method deals with the most tedious and complex evaluation questions. In spite of its initial higher acquisition cost, they were finally convinced that solution 1 is the most appropriate solution for the case of organization "X" for the long run. The feedback relating to the real implementation of solution 1 has shown that the adoption of this solution was done smoothly and the tailoring risk was controlled. However, there were some gaps between the estimated outputs and the real ones, as show in Table 10.

	Adjustment
Anticipated fitness	-5%
Tailoring Risk	3%
Total cost of ownership	12%

Table 10. Outputs' adjustments

These results lead us to the issue of the estimation accuracy of the inputs and its impact on the outcomes. In fact, these inputs are often uncertain and are based on the experience of the evaluators. For this reason, it is deemed wise to supplement SEVALERPS by a sensitivity analysis to see how outcomes change if the inputs vary within a limited range.

5. Conclusion

The purpose of this paper is to present SEVALERPS, a new *ex-ante* evaluation method proposed to perform ERP system adaptation and selection. This method is mainly developed to handle some of the research questions resulting from investigating shortcomings relating to the previous ERP evaluation methods. The proposed method serves mainly twofold objectives; firstly it introduces a new systematic evaluation process that considers ERP tailoring and mismatch handling as the cornerstone of the evaluation activity. Secondly, it presents a semi-structured evaluation model that helps the selection team to voice out its preferences. The evaluation model relies on many mathematical techniques to handle the various aspects of the evaluation: 0-1 linear programming to determine the anticipated functional coverage of ERP candidates, MACBETH cardinal scales to represent, in an interactive way, the selection team preferences, and the Choquet integral to address interdependencies among evaluation criteria.

The experimentation of SEVALERPS through a large scale project conducted in a public administration shows that our method is practical and improves the ERP selection process. It should be noted that SEVALERPS is also supported by an automation system called AS-SEVALERPS (Automation System of SEVALRPS). The design and main components of this system will be discussed in our further publications.

The first limitation of SEVALERPS lies, like all the previous evaluation methods, in that the accuracy of its results is completely dependent on the one of its inputs. Now, given that these inputs are mainly obtained through human judgments, the reliability question of SEVALERPS results is raised in case these inputs are somehow uncertain. Thus, a research perspective to improve SEVALERPS is to provide organizations with ready to use knowledge databases from which they could validate their evaluation data. This data is mainly relating to the non-adaptation risks, the initial functional coverage levels of the required functionalities, and the anticipated functional coverage levels of these functionalities after tailoring the ERP solutions. The validation of the data contained in these

knowledge databases should be done through the comparison and the review of data obtained through a predictive way with the one obtained from real evaluation cases.

Finally, applying SEVALERPS method requires the determination of the evaluation criteria against which ERP candidate solutions will be judged. These criteria depend on the business requirements of each organization. In this regard, another way to improve SEVALERPS is to define more relevant criteria for each organizational context and validate them through more specific case studies. These mentioned perspectives are the aim of our future research.

6. Bibliographie

- Addo-Tenkorang, R., Helo, P., (2011). *Enterprise resource planning (ERP): a review literature report*. Proceedings of the World Congress on Engineering and Computer Science, 19-21.
- Bana e Costa, C. A., Corte, J. M., Vansnick, J. C., (2011). *MACBETH (Measuring Attractiveness by a Categorical based Evaluation Technique)*. Wiley Encyclopedia of Operations Research and Management Science.
- Brehm, L., Heinzl, A., and Markus, M., (2001). *Tailoring ERP systems: a spectrum of choices and their implications*. Proceedings of the 34th Annual Hawaii International Conference (IEEE), 9–20.
- Bustinza, O. F., Arias-Aranda, D., Barrales-Molina, V., Dominguez, C. H., Arostegui, N. P., (2013). *Integrating ERP benefits, competitive priorities and performance measures*. Proceedings of the 2013 International Conference on Technology Innovation and Industrial Management.
- Chofreh, A. G., Goni, F. A., Shaharoun, A. M., Ismail, S., Klemeš, J. J., (2014). *Sustainable enterprise resource planning: imperatives and research directions*. Journal of Cleaner Production.
- Choquet, G., (1953). *Theory of Capacities*. Annales de l'institut Fourier, 5,131-295.
- Consulting, P., (2013). *Panorama Consulting Solutions' 2013 ERP Report*. <http://panorama-consulting.com/resource-center/2013-erp-report/>.
- Elmir, B. and Bounabat, B., (2012). *Integrated public e-services interoperability assessment*. International Journal of Information Science and Management, 1, 1-12.
- Grabisch, M., (2006). Evaluation subjective. D. Bouyssou and D. Dubois and M. Pirlot and H. Prade. *Concepts et Méthodes pour l'Aide à la Décision*. Hermès, pp.175-232
- ISO9126., (2001). *ISO/IEC 9126-1 Software Engineering - Product Quality - Part I: Quality Model*. ISO International Standards Organisation.
- Khaled, A., Janati-Idrissi, M. A., (2011) (a). *A framework for ERP evaluation and selection using MACBETH technique*. Proceedings of the 13th international conference on enterprise information systems, China, 339-342.
- Khaled, A., Janati-Idrissi, M. A., (2011) (b). *A semi-structured methodology for ERP system selection based on MACBETH and Choquet integral applied to small and medium sized enterprises*. International Journal of E-Entrepreneurship and Innovation, 2/4, 1–19.
- Khaled, A., Janati-Idrissi, M. A., (2012). *A semi-structured tailoring-driven approach for ERP selection*. International Journal of Computer Science Issues, 9/2.
- Katerattanakul, P., Lee, J. J., Hong, S., (2014). *Effect of business characteristics and ERP implementation on business outcomes: an exploratory study of Korean manufacturing firms*. Management Research Review, 37/2, 186-206.
- Martín, M. G., Aguirre, A. O., Vargas, A. M. (2014). *Systematic improvement of IT processes: application of CMMI*. Proceedings of Industrial Engineering 2012, London, 273-280.
- Munkelt, T., Völker, S., (2013). *ERP systems: Aspects of selection, implementation and sustainable operations*. International Journal of Information Systems and Project Management, 1/2, 25-39
- Pérez-Salazar, M. D. R., Rivera, I., Cristóbal-Vázquez, I. M., (2013). *ERP selection: a literature review*. International Journal of Industrial and Systems Engineering, 13/3, 309-324.
- Poba-Nzaou, P., Raymond, L., (2011). *Managing ERP system risk in SMEs: a multiple case study*. Journal of Information Technology, 26/3,170-192.
- Razmi, J., Sangari, M. S., (2013). *A comprehensive decision support system for ERP pre-implementation phase*. European Journal of Industrial Engineering, 7/4, 475-496.
- Saaty, T. L., (2013). *Analytic hierarchy process*. Encyclopedia of Operations Research and Management Science, pages 52-64.
- Sen, C., Baraçlı, H., Sen, S., (2009). *A literature review and classification of enterprise software selection approaches*. International Journal of Information Technology and Decision Making, 8/2, 217–238.
- Yıldız, A., and Yıldız, D., (2014). *Enterprise resource planning (ERP) software selection process with fuzzy TOPSIS method*. Business and Economics Research Journal, 5/1, 87-106.
- Zeng, Y., and Skibniewski, M. J., (2013). *Risk assessment for enterprise resource planning (ERP) system implementations: a fault tree analysis approach*. Enterprise Information Systems, 7/3, 332-353.
- Zoukar, I., Salinesi, C., Rolland, C., (2013). *Évolution du système d'information par l'implantation d'un progiciel de gestion intégrée*. Proceeding of Informatique des Organisations et Systèmes d'Information, 1–17.

Usage et formation

L'art de cultiver les perles rares : changer l'image véhiculée des professionnels en TI

Marie-Christine Roy, Ph. D.

Département des systèmes d'information organisationnels, Université Laval, Québec, Canada

Marie-Christine.Roy@fsa.ulaval.ca

Hager Khechine, Ph. D.

Département des systèmes d'information organisationnels, Université Laval, Québec, Canada

Hager.Khechine@sio.ulaval.ca

Diane Ducharme

Département des systèmes d'information organisationnels, Université Laval, Québec, Canada

Diane.ducharme.1@ulaval.ca

Résumé

Dans plusieurs pays, nous assistons depuis des années à un déclin de ressources humaines spécialisées dans le domaine des technologies de l'information (TI) et des pénuries significatives sont prévues dans le futur. La diminution du nombre des diplômés en TI est une des causes majeures de cette rareté de compétences dans le marché de l'emploi. Cette situation est exacerbée par la place qu'occupent les TI dans cette ère du numérique. Le besoin de compétences spécialisées dans ce domaine ne cesse de grandir. Plusieurs gouvernements et employeurs décrivent haut et fort une situation alarmante de pénurie de ces ressources, situation pour laquelle la rationalité dans la réflexion et la rapidité de l'action revêtent un caractère urgent.

Cet article permet d'expliquer les raisons de la baisse d'intérêt des individus aux programmes de formation et à la carrière en TI. Nos constatations nous amènent à conclure que l'image véhiculée par la formation et la profession TI est l'élément central qui contribue à la pénurie des ressources en TI. Nous proposons des recommandations afin de pallier cette situation préoccupante. Ces recommandations prennent source dans la littérature et dans notre expérience dans le domaine de l'enseignement en TI, ce qui renforce leur faisabilité.

Abstract

Globally, we are witnessing a decline in human resources specialized in Information Technology (IT) and significant shortages are predicted for the future. The reduction in IT graduates, in particular those from information systems programs, is the main cause of the market's lack of these competencies. This situation is further alarming because of the importance of IT and the growing need for these specialists in the numerical era. Many employers and governments are raising their voices to denounce the problem and insist on the urgency of reflection and actions from the IT community.

In this article, we attempt to explain the reasons why students show a lack of interest for IT programs and careers. Our observations lead us to conclude that the social image of the IT profession may be the main factor to address, and suggest different strategies that can correct the situation. The recommendations are based on prior research results and some of our academic experiences.

Mots-clés

Technologies de l'information, formation en TI, image véhiculée de la profession en TI

Keywords

Information technology, IT education, Social image of IT professionals

1. Introduction

Les révolutions en TI ont le potentiel de transformer les économies et les sociétés, et promettent de relever les défis globaux les plus pressants. Dans certains pays, ces révolutions sont accentuées par l'ubiquité d'Internet, la démocratisation des technologies et l'accélération des innovations (Insight report, 2015). Le secteur des TI est décrit comme un générateur de l'activité économique et facilitateur pour la connectivité et la compétitivité dans toute l'industrie.

Les emplois dans ce secteur seront en constante croissance: par exemple aux États-Unis, on prévoit une croissance de 18% de toutes les occupations liées à l'informatique de 2012-2022, comparativement à 11% pour toutes les carrières. Le Conseil des technologies de l'information et des communications (2011) prévoyait un besoin annuel de 17000 travailleurs en TI jusqu'en 2016 au Canada. Au Québec (Habak et Palanca, 2014), ce chiffre s'établit à 6000 emplois en TI à pourvoir chaque année et TechnoCompétences (2015) fait état d'une croissance deux fois plus rapide de l'emploi en TI comparativement à l'emploi total.

Par ailleurs, plusieurs observateurs décrivent une situation alarmante du déclin de ressources humaines compétentes en TI en Amérique (Raymond Chabot Grant Thornton, 2011; Lalonde, 2014; Technocompétences, 2015; ManpowerGroup, 2014). La situation n'est certes pas meilleure en Europe. En France, la demande pour de jeunes diplômés en TI est en forte croissance depuis que l'on s'est remis de l'éclatement de la bulle technologique du début des années 2000. En 2012, une proportion de 40% des grandes entreprises de l'Union européenne a recruté, ou essayé de recruter, du personnel pour des emplois nécessitant des qualifications en TI (Observatoire du numérique, 2012). Des spécialistes en analyse-programmation, des techniciens en maintenance, des architectes de systèmes d'information, des ingénieurs en sécurité, des directeurs de services informatiques et des administrateurs systèmes sont sollicités. Les derniers résultats du baromètre sur les métiers du numérique fait état d'un manque flagrant de candidats pour les postes qui sont offerts (Cap Digital, 2014). Le vice-président de la Commission européenne, Neelie Kroes, lance un appel en 2013 (Commission Européenne communiqué de presse):

"The digital skills gap is growing, like our unemployment queues. We need joint action between governments and companies to bridge that gap. The ICT sector is the new backbone of Europe's economy, and together we can prevent a lost generation and an uncompetitive Europe. So I am expecting concrete pledges by companies, everyone I meet will be getting the same request. The Commission will do its bit but we can't do it alone – companies, social partners and education players – including at national and regional level - have to stand with us."

Des mesures urgentes pour discuter de cette problématique s'organisent partout dans le monde afin que les parties prenantes (collèges, universités, employeurs et organisations gouvernementales) travaillent de concert pour réduire le problème (Lalonde, 2014). Faire l'autruche aura pour conséquence d'exacerber la pénurie des ressources humaines spécialisées et contribuera au déclin du développement et de la recherche, non seulement en TI, mais aussi dans les domaines qui en sont dépendants tels que la santé, l'éducation, l'administration publique, le commerce, les finances, etc. (TechnoCompétences, 2015).

D'un autre côté, contrairement aux croyances populaires, la réalité de la qualité des emplois en technologies de l'information, à l'échelle mondiale, est loin d'être rébarbative: les salaires sont élevés, le travail est varié et les opportunités d'avancement sont excellentes (Hirschheim et Newman, 2010). Un classement des éditeurs de l'US News (2015) portant sur les meilleurs emplois en Amérique donne 3 professions en TI dans les dix premiers pour l'année 2015. Un consultant en TI occupe la septième place, suivi par un gestionnaire de bases de données. Les critères de classification sont la satisfaction, le niveau de stress et la flexibilité. Une étude récente de TechnoCompétences (2015) indique que le taux de chômage dans le secteur des TI est de 2,9% à l'échelle canadienne comparativement à un taux de chômage général de 7,2%. De plus, le salaire moyen d'un professionnel en TI au Québec est 1,5 fois supérieur au salaire moyen, toutes professions confondues (TechnoCompétences, 2015). En dépit de ces données, les Universités nord-américaines et partout dans le monde vivent une réalité accablante que Hirschheim et Newman (2010) qualifient d'«érosion des étudiants» de la discipline des systèmes d'information ou des technologies de l'information. La diminution du nombre de diplômés en TI en Amérique du Nord, comme dans plusieurs autres pays développés, constitue un des trois enjeux clés du secteur d'emploi des TI (Raymond Chabot Grant Thornton, 2011). Ce constat nous amène à réfléchir sur les raisons pour lesquelles l'évolution du nombre d'individus qui embrassent la carrière en TI ne suit pas la même tendance que celle du marché de l'emploi pour des compétences qualifiées en TI.

L'objectif de cet article est de contribuer à notre réflexion sur cette problématique en recensant les données relevées dans des études antérieures pour expliquer le phénomène et de proposer des avenues de solutions fondées sur ces connaissances et sur nos expériences. En particulier, nous examinons les biais de perception qu'ont les jeunes de la carrière en TI et les facteurs qui influencent le choix de poursuivre la formation dans ce domaine. Nous jetons aussi un regard sur des stratégies qui se sont révélées fructueuses pour modifier ces perceptions. Les recommandations proposées sont appuyées par l'expérience et les actions effectuées au département des systèmes d'information organisationnels de l'Université Laval. Ces recommandations sont adressées à toutes les parties prenantes concernées par la formation et la profession TI. Cette discussion pourrait servir comme un point d'ancrage pour des études empiriques futures.

2. Facteurs influençant le choix de se former en TI à travers la TRA

Selon une croyance populaire qui perdure dans les écoles de commerce et des sciences informatiques, les préférences des étudiants quant à la concentration – ou au programme - qu'ils choisissent pour leurs études dépendent de la carrière qu'ils visent (Kuechler *et al.*, 2009). Par contre, les fluctuations dans les inscriptions semblent aussi tributaires non seulement des préférences personnelles des individus ou de leur entourage, mais aussi des occasions d'embauche perçues dans l'industrie (Malgwi *et al.*, 2005). Cette observation a élargi notre compréhension des facteurs pouvant expliquer les motifs de choix des concentrations. Plusieurs chercheurs ont adopté des théories et modèles de prévision du comportement afin de déterminer ces facteurs et aider les universités à mieux cibler les actions pour attirer plus d'étudiants dans les différentes concentrations. Parmi ces théories, la TRA (Théorie de l'action raisonnée ou Theory of reasoned action) d'Azjen et Fishbein (1980) a permis de ressortir 5 facteurs pouvant influencer le choix d'une concentration ou d'une carrière. Le choix de cette théorie comme loupe d'analyse est tributaire du fait qu'elle est la plus inclusive en termes de facteurs explicatifs et la plus fréquemment testée empiriquement pour évaluer le choix des concentrations et carrières. Nous représentons ces facteurs de manière synthétique à la Figure 1.

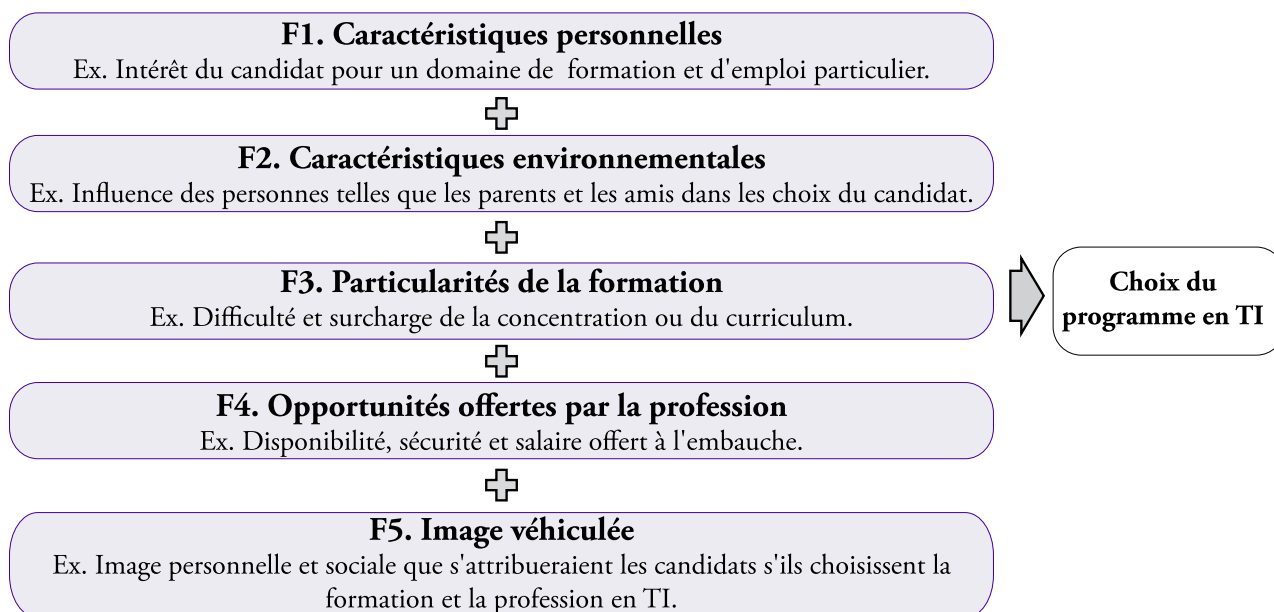


Figure 1. Facteurs influençant le choix d'une formation ou d'une carrière en TI

La TRA a été utilisée par Zhang (2007) pour identifier les déterminants du choix de la concentration «systèmes d'information». Des 114 réponses obtenues d'étudiants inscrits à un cours de baccalauréat en systèmes d'information aux États-Unis d'Amérique, l'auteur a pu déduire que les facteurs qui influencent l'intention de choisir la concentration en systèmes d'information sont l'intérêt pour le domaine (F1), l'avis de la famille et des enseignants sur leur choix (F2), les difficultés perçues dans le curriculum (F3) et la disponibilité de l'emploi (F4).

Le modèle TRA a ensuite été testé par Kuechler *et al.* (2009) pour déterminer les facteurs qui influencent des étudiants au baccalauréat dans le choix d'une concentration de manière générale et de la concentration en systèmes d'information en particulier. Les résultats de cette étude, effectuée auprès de 218 étudiants inscrits à un cours en systèmes d'information organisationnels dans une école de commerce aux États-Unis, ont permis de conclure que l'intérêt à ce domaine (F1), les avis des conseillers et de la famille (F2) et l'image sociale véhiculée par la concentration systèmes d'information (F5) sont des éléments significatifs de l'attitude qui influencent l'intention de choisir la concentration en systèmes d'information.

Plus récemment, Downey *et al.* (2011) ont employé à nouveau le modèle TRA pour déterminer les facteurs qui ont influencé 413 étudiants en sciences de l'administration aux États-Unis dans leur choix de carrière. Les résultats obtenus, toutes concentrations confondues, font ressortir l'intérêt au domaine (F1) et la disponibilité et la sécurité de l'emploi (F4) comme déterminants indirects du choix de la carrière. En scindant l'échantillon en deux groupes – ceux qui suivent la concentration systèmes d'information et ceux qui suivent les autres concentrations (ex. marketing, finances ou comptabilité), les auteurs ont obtenu des résultats différents, mais néanmoins pertinents. En effet, les étudiants qui ont choisi la concentration systèmes d'information étaient influencés, dans leur intention de travailler en systèmes d'information, par leur intérêt personnel au domaine (F1), l'avis des autres étudiants (F2), les difficultés perçues de la concentration et leur aptitude (F3), la sécurité et la disponibilité de l'emploi (F4) et l'image

sociale (F5). Les facteurs qui concernent les étudiants n'ayant pas choisi la concentration en systèmes d'information sont l'intérêt personnel (F1), l'avis des amis et professeurs (F2), la surcharge de travail et l'aptitude (F3) et le salaire ainsi que la disponibilité et la sécurité de l'emploi (F4). Tel que l'illustre la figure 2, l'image véhiculée (F5) n'apparaît pas comme facteur déterminant de l'intention de choix d'une carrière dans d'autres domaines que les systèmes d'information. Ce dernier résultat, que nous avons trouvé intrigant, nous a amenés à penser à la pertinence de mettre l'accent sur l'image véhiculée dans le domaine des TI pour étudier son influence sur le choix de carrière.

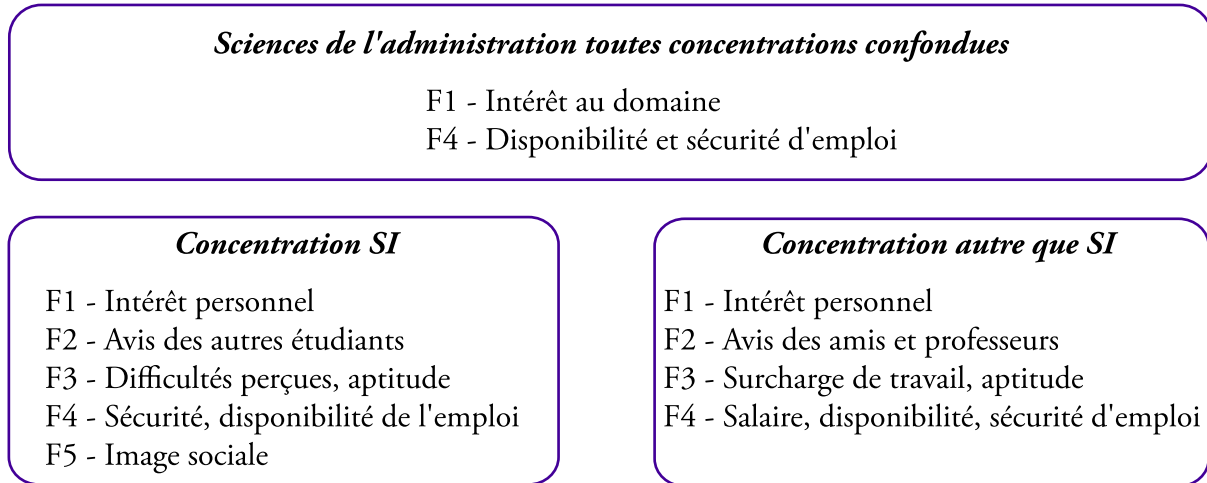


Figure 2. Facteurs influençant le choix de carrière de 413 étudiants en science de l'administration aux États-Unis selon Downey et al. (2011)

À la lumière des résultats des études précédentes, nous confirmons le souci constant des communautés de recherche et de pratique en TI de comprendre les raisons du choix de la spécialisation et de la carrière dans ce domaine. Les caractéristiques personnelles et environnementales, les particularités de la formation et les opportunités offertes par la profession sont certes des facteurs déterminants du choix de toutes les spécialisations. Toutefois, l'image véhiculée sur la carrière en TI sort du lot en se distinguant comme un élément récurrent et un facteur propre à la concentration TI. La section suivante explique les raisons de l'image véhiculée par le domaine des TI.

3. Image véhiculée par le domaine des TI

3.1. Sources de la formation de l'image

L'image véhiculée sur la carrière TI prend son origine de plusieurs sources. Un parcours de la littérature nous a permis de les colliger dans trois groupes. Ces groupes, que nous présentons à la figure 3, sont des éléments qui dissuadent les élèves et étudiants d'opter pour une carrière en TI. Ils se situent sur un continuum entre la perception et la réalité. Le choix du continuum est tributaire du fait que ces groupes «vacillent» entre des biais perceptuels et des faits palpables et que les frontières entre les deux repères sont minces, voire inexistantes.

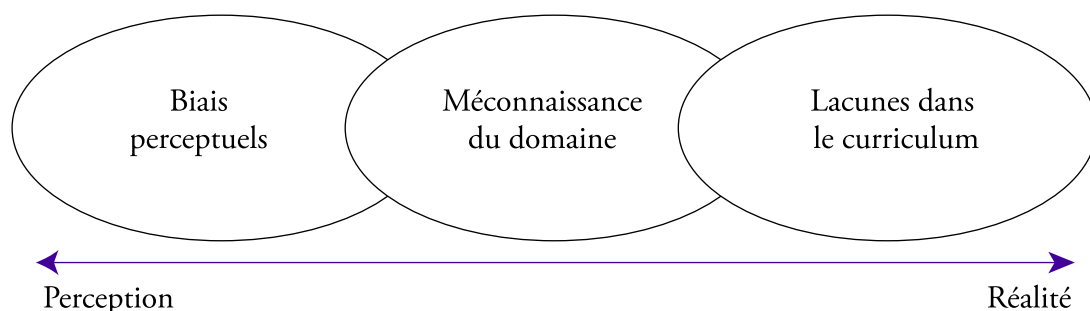


Figure 3. Continuum des sources de l'image véhiculée

Dans le domaine des systèmes d'information, le désintérêt des étudiants pour une carrière dans cette discipline est causé, entre autres, par le stéréotype du «nerd» qui se colle à la spécialisation. L'image de l'informaticien isolé et obsédé par son travail illustre bien le biais perceptuel. La méconnaissance du domaine est une deuxième cause pouvant expliquer le refus des étudiants à choisir la carrière en TI. Cette méconnaissance se trouve à mi-chemin entre la perception (ex. absence d'emplois, salaires bas et monotonie des tâches) et la réalité (ex. disponibilité des emplois, salaires élevés, travail varié et excellentes opportunités d'avancement) puisqu'elle résulte de la conjugaison d'éléments erronés autant perceptuels - ce que l'on pense - que réels - ce que l'on sait. La dernière cause représente une réalité vécue dans plusieurs universités nord-américaines, européennes et australiennes, soit les manquements qu'enregistrent les curriculums des programmes de formation universitaire. Ces manquements se traduisent d'abord par une formation de base des enseignants qui n'est pas centrée sur le développement de compétences en TI. Le manque de temps pour mettre à jour le curriculum est un autre problème souvent rencontré par les enseignants. Enfin, le peu d'adéquation du curriculum avec les besoins du milieu professionnel explique aussi les lacunes dans le curriculum.

3.2. Des biais perceptuels

Un des facteurs déterminants de la rareté des personnes qui choisissent la formation et la carrière en TI est l'image négative diffusée auprès de la population. Un professionnel TI est communément vu comme un «nerd»: une personne solitaire, passionnée, voire obnubilée par des intérêts purement scientifiques, en l'occurrence par l'informatique. L'image de la profession est en partie formée durant les études préuniversitaires. Von Hellens *et al.* (2009) ont mené une étude auprès d'étudiants et d'enseignants au secondaire afin de sonder leurs perceptions de la carrière en TI en Allemagne. La motivation de cette étude prend source du nombre décroissant d'inscriptions dans les formations en TI en Europe de l'Ouest. Les résultats ont révélé que, malgré un usage intensif et des connaissances élevées en TI, la carrière dans ce domaine de compétences n'est pas considérée parmi les premiers choix des étudiants. Les résultats obtenus auprès de 160 répondants montrent que seulement 14% d'entre eux pourraient envisager le reste de leurs études en TI. Selon les auteurs, ceci reflète une tendance générale liée à une opinion erronée que la profession est ennuyeuse et adaptée aux personnes solitaires. Dans une étude exploratoire, Thomas et Allen (2006) ont mis en lumière le fait que les étudiants veulent à tout prix se distinguer du stéréotype du «nerd» véhiculé dans la société et qui peut être attribué à plusieurs facteurs, dont les médias (Thomas et Allen, 2006). La télévision et les films caricaturent les professionnels en technologies comme des personnes qui s'intéressent peu aux relations et caractéristiques humaines. Ce biais perceptuel que les gens ont vis-à-vis de la profession en TI semble lié de manière importante à l'impression que «la programmation» occupe un espace prépondérant dans les fonctions d'un professionnel en technologies. Par conséquent, il y a lieu de distinguer l'image générale de la profession de la méconnaissance du domaine, traitée dans la prochaine section.

3.3. Une méconnaissance du domaine

Certains auteurs ont observé la piètre qualité de l'information sur les carrières en TI diffusée auprès des jeunes. Une étude effectuée auprès de 98 étudiants australiens révèle que la plupart d'entre eux pensent que l'emploi en TI consiste essentiellement à travailler sur un ordinateur et à effectuer des tâches techniques (Thomas et Allen, 2006). Von Hellens *et al.* (2009) ont aussi conclu que les réponses des étudiants qui ont participé à leur étude montrent un manque de compréhension de la nature non technique de la profession et une ignorance de la marge des salaires gagnés à cet emploi. L'étude effectuée par Lasen (2010) abonde dans le même sens dans la mesure où les élèves interrogés avouent n'avoir aucune idée de la profession en TI, si ce n'est de travailler sur des ordinateurs. Les cours suivis à l'école sont la première source d'information sur la formation et la profession en TI. Or, Clayton (2004) souligne que les enseignants et les conseillers en orientation sont eux-mêmes insuffisamment informés sur les formations, l'industrie et les caractéristiques des emplois dans le domaine. Cette lacune est accentuée par le manque de temps des enseignants pour maintenir à jour le matériel pédagogique relié aux TI, ce qui contribue à une transmission d'information tronquée sur cette spécialité.

Alexander *et al.* (2011) ont repris la théorie sociocognitive (SCT) pour étudier les facteurs influençant le choix d'une carrière en TI par les étudiants de deux universités sud-africaines. Un des faits saillants des résultats obtenus dans cette recherche est que les sources d'information pour une profession sont d'une importance capitale dans le choix d'une carrière, spécialement en TI. Toutefois, les 1868 répondants ont affirmé avoir obtenu de l'information sur la formation en TI, mais peu d'information sur la profession correspondante. La recherche qualitative de Clayton *et al.* (2004) abonde dans le même sens en concluant que les étudiants ignorent ce qu'un professionnel en TI fait durant sa journée. Ils recommandent de briser le cycle de la mauvaise perception des TI en offrant aux étudiants de l'information fiable, précise, à jour et de qualité sur la carrière en TI. Néanmoins, malgré ces croyances populaires que les médias exacerbent (Thomas et Allen, 2006), les habiletés requises pour les professionnels en TI

vont bien au-delà de celles purement techniques. L'étude de Gallagher *et al.* (2010) fait valoir que les compétences en gestion de projet et les aptitudes relationnelles sont critiques pour développer des solutions adaptées aux besoins des organisations. Ces compétences non techniques sont souvent méconnues par les jeunes étudiants, à l'instar de ceux qui ont été questionnés dans l'étude de Von Hellens *et al.* (2009).

Être exposé à de l'information de qualité sur la profession en TI pourrait changer les attitudes vis-à-vis de cette profession. Pour le prouver, Walstrom et Schambach (2012) ont effectué deux sondages auprès d'un groupe d'étudiants en administration des affaires. Le premier, au début de l'étude, afin de recueillir leurs attitudes par rapport à une carrière en TI. Le deuxième a eu lieu après leur avoir fait lire l'article de Wieggers (2003) sur la carrière en TI. Les résultats obtenus ont permis de conclure que la perception des étudiants de la carrière TI s'est améliorée grâce à l'information obtenue de l'article lu. Cette perception est en lien avec les caractéristiques du métier telles que l'interaction sociale, le travail en équipe, la disponibilité d'emploi et les tâches effectuées. Cette information concernait plus particulièrement les interactions sociales requises dans le rôle d'un analyste.

3.4. Des lacunes dans le curriculum

Les lacunes dans le curriculum sont observées à trois niveaux :

- *Des enseignants en TI qui n'ont pas l'expertise nécessaire dans le domaine :*

Clayton (2004) a montré que l'enseignement des TI au secondaire était souvent sous la responsabilité d'enseignants spécialisés dans d'autres domaines que les TI. L'étude de Von Hellens *et al.* (2009) a ressorti un profil particulier des enseignants en TI autant en Australie qu'en Allemagne. Ceux-ci ont, en effet, des qualifications d'enseignement en administration des affaires, en sciences ou en mathématiques. Ils ont incarné le rôle d'enseignants en TI pour répondre aux besoins de l'école ou pour un intérêt personnel au domaine.

- *Des enseignants TI qui manquent de temps pour mettre à jour le curriculum :*

Les enseignants se retrouvent face à une difficulté récurrente de maintenir un curriculum d'actualité. Cette difficulté s'explique par le manque de temps au vu de leurs multiples tâches, mais aussi par l'évolution rapide des TI (Clayton, 2004). Un focus group effectué par Lasen (2010) auprès de jeunes au secondaire a permis de conclure que les enseignants ont fait montre d'un manque d'expertise et d'enthousiasme nécessaires pour promouvoir les études en TI dans leurs cours.

- *Les préoccupations du milieu professionnel qui ne sont pas suffisamment considérées lors de la conception et du développement du curriculum :*

Le curriculum est affecté par le manque d'adéquation entre la formation en TI et les besoins du marché. Stevens *et al.* (2011) ont sondé à la fois la perception des enseignants et les praticiens sur les habiletés requises pour un professionnel en TI. Leurs résultats ont montré que les deux groupes sont en désaccord sur l'importance d'au moins 25% de ces habiletés. Ainsi, les enseignants mettent plus d'emphasis sur les capacités techniques et la programmation que les praticiens. Les auteurs ont conclu en la nécessité, pour les programmes en TI, d'être à l'écoute des besoins du marché et de s'y adapter. Le curriculum a besoin d'être modifié afin d'être mieux adapté aux besoins de l'industrie. Les lacunes dans le curriculum généreraient, à notre sens, une insatisfaction de la part des étudiants par rapport à leur expérience éducative avec les TI et diminuerait leur motivation à explorer le domaine.

La revue des facteurs affectant l'image de la profession en TI nous permet d'avoir un aperçu plus clair et global des éléments qui dissuadent les jeunes à opter pour une carrière dans le domaine. Cette information pourra servir comme base pour articuler des plans d'action destinés à accroître le nombre de candidats dans les programmes de formation liés aux TI. C'est ce dont nous traiterons dans les prochaines sections.

4. Recommandations et actions

Le recensement de la littérature nous a permis d'identifier différentes initiatives qui, quand elles sont considérées dans leur ensemble, permettraient d'atteindre un objectif unique: augmenter l'intérêt général des jeunes, des universitaires et de la population envers la profession en TI dans le but de changer l'état alarmant de la situation actuelle. La Figure 4 regroupe les recommandations apportées pour améliorer l'image véhiculée de la carrière en TI. Il est à noter que ces recommandations ont été proposées par plusieurs auteurs dans le domaine. Nous les appuyons par des exemples d'initiatives pratiques que le département des systèmes d'information organisationnels de l'Université Laval à Québec a mis en œuvre pour montrer leur faisabilité. Enfin, aucune recommandation ne revêt une importance supérieure à une autre. Les décideurs ont le choix de poser leurs actions à différents niveaux dépendamment de la source de l'image véhiculée qu'ils veulent combler et des moyens financiers et humains dont ils disposent.

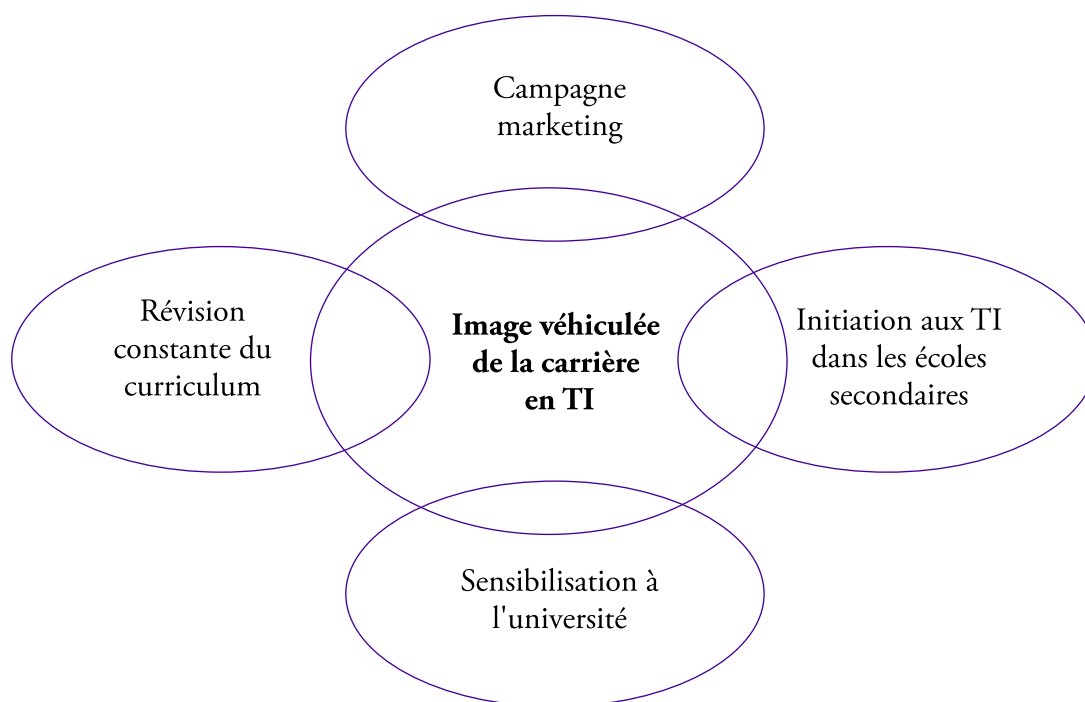


Figure 4. Recommandations d'actions pour améliorer l'image véhiculée sur la carrière TI

4.1. Une campagne Marketing

Tel que le stipulent les écrits précédents (Clayton, 2004; Lasen, 2010; Thomas et Allen, 2006; Von Hellens *et al.*, 2009), une des barrières au choix de la carrière en TI est le manque d'information fiable et précise sur la profession. À cet effet, Ballard *et al.* (2006) suggèrent de faire le marketing des attributs méconnus de la carrière en TI. Ferratt *et al.* (2010) proposent de mettre l'emphase sur la promotion de la réalité du travail des gradués en TI, soit «d'aider les gens d'affaires à réussir en développant des solutions créatives à des problèmes importants et en liant les technologies aux affaires». Dans ces campagnes marketing, les intervenants de l'industrie peuvent travailler en étroite collaboration avec les universités. Ces campagnes viseraient à diffuser des messages courts sur les caractéristiques des travailleurs en TI et sur les biais de perception les plus communs. Le résultat recherché est de briser l'image du «nerd» et de promouvoir les aspects créatifs, relationnels et variés du métier. Si les spécialistes en TI n'agissent pas pour se faire connaître, la profession restera méconnue et la formation en TI demeurera peu attractive pour les générations futures, au risque de disparaître. La faculté des sciences de l'administration de l'Université Laval à Québec a emprunté la voie de la campagne Marketing en effectuant des voyages de sensibilisation à l'étranger, en particulier en Afrique du Nord, afin d'expliquer aux étudiants dans les écoles de gestion les principes de base de la formation et de la profession TI. L'implantation de kiosques dans les salons annuels de l'emploi en Europe – plus particulièrement en France - est une pratique adoptée pour publiciser la spécialisation TI. Des initiatives pour avoir pignon sur rue au Maroc sont aussi en cours d'élaboration afin de familiariser la population nord-africaine à la spécialisation TI.

4.2. L'initiation aux TI dans les écoles secondaires

La sensibilisation des individus à une profession commence dès leur jeune âge. Il serait important de fournir aux jeunes du secondaire de l'information de qualité, fiable, à jour et exacte sur la carrière en TI. La fiabilité de cette information leur permettra de prendre des décisions éclairées sur leur carrière (Clayton, 2004) et éventuellement, d'augmenter les admissions universitaires en TI. Des conférenciers praticiens pourraient rencontrer les jeunes pour leur parler des particularités de leur travail et du rôle important qu'ils jouent pour résoudre des problèmes d'affaires avec des solutions créatives liées aux technologies de l'information. Le département des systèmes d'information organisationnels de l'Université Laval a organisé des interventions dans des classes au secondaire afin d'expliquer aux élèves la différence entre un spécialiste en TI et un informaticien. Un professeur universitaire en TI, qui détient aussi un fort bagage en consultation, s'est chargé de cette mission. L'intérêt des auditeurs au discours était élevé et le message de démythification de la profession a été bien perçu.

4.3. La sensibilisation à l'Université

Selon Roach *et al.* (2011), les perceptions des étudiants de la concentration en TI sont susceptibles de changer durant les premiers cours universitaires dans cette spécialité. Un défi majeur pour les enseignants de ces cours est de donner une bonne impression du domaine afin de susciter l'intérêt des étudiants pour une carrière en TI. De notre expérience personnelle avec les cours de base en technologies d'information, cet intérêt pour le domaine s'est concrétisé grâce à l'approfondissement du lien entre les avancées technologiques et l'utilisation quotidienne de ces technologies. Walstrom et Schambach (2012) suggèrent de remplacer les exposés longs et compliqués de la matière par des éléments courts d'information sur les caractéristiques de la carrière. Ferratt *et al.* (2010) proposent des stratégies d'intervention plus précises et personnelles auprès des étudiants. Ces stratégies, qui nécessitent des changements sur les plans affectif et cognitif des étudiants, varient selon leurs intérêts pour le domaine des TI :

1. Pour les étudiants dont les préférences concordent avec le domaine des TI et qui disposent d'information fiable sur la concentration, il est conseillé de renforcer leur intérêt. Puisque cette catégorie d'étudiants est la plus facile à cibler, il est recommandé de communiquer directement avec les personnes influentes pour ces étudiants ou même d'ajuster le curriculum pour minimiser les coûts liés au changement de la concentration. Dans le cadre des cours en systèmes d'information à l'Université Laval, différents accès gratuits sont offerts aux étudiants pour des activités payantes organisées par la communauté de pratique en TI de la région. Ces activités prennent la forme de dîner-conférence avec les professionnels dont le but est de consolider le réseautage.
2. Pour les étudiants dont les préférences concordent avec le domaine des TI, mais qui ne croient pas que l'aboutissement d'une formation en TI les amènerait à atteindre ces préférences, il serait pertinent de corriger cette situation en ajustant leurs perceptions erronées. Ce qui est communément recommandé est de pallier le manque de connaissances qu'ils ont de cette profession avec de l'information fiable et valide. Nous mettons en pratique cette recommandation dans nos cours de systèmes d'information en accueillant des invités pour partager leurs expériences dans le domaine. Ces invités sont généralement des consultants en TI ou des spécialistes de l'appareil gouvernemental qui acceptent d'intervenir dans les cours sans compensation financière. Les résultats de l'étude de Becerra-Fernandez *et al.* (2010) montrent que la baisse des inscriptions et l'abandon en cours de programme ont diminué grâce à l'invitation de conférenciers praticiens et aux visites sur des sites. Au département des systèmes d'information de l'Université Laval, la place des associations étudiantes en TI revêt une importance particulière. Elles sont constamment encouragées, autant financièrement que par l'appui des compétences humaines, à tisser des liens plus étroits entre les étudiants et les professionnels TI. Nous croyons que l'information peut être plus facilement acceptée et assimilée par les étudiants si elle est obtenue par le biais de leurs pairs.
3. Pour les étudiants dont les préférences ne s'alignent pas avec le domaine des TI, mais qui disposent d'information adéquate sur les caractéristiques de la concentration, une tentative de reconversion peut donner des résultats concluants, même si cette catégorie d'étudiants est la plus difficile à convertir. Des visites dans des entreprises et des discussions avec des professionnels en TI sont des stratégies parmi d'autres pour changer les perceptions et préférences de ces étudiants. Les projets d'intervention en entreprise, qui s'apparentent à un stage pour un projet de fin d'études est une action pratique que le département des systèmes d'information de l'Université Laval a adoptée afin d'implanter cette recommandation.
4. Pour les étudiants dont les préférences ne s'alignent pas avec le domaine des TI et qui ne disposent pas d'information fiable sur ce domaine, il est essentiel d'effectuer des changements tant sur le plan affectif que cognitif. Or, à l'instar de la catégorie précédente d'étudiants, les efforts peuvent être coûteux et les résultats incertains. Les moyens actuels dont dispose la Faculté des sciences de l'administration de l'Université Laval nous permettent d'organiser une activité appelée «Midi des concentrations». Les étudiants sont invités à plusieurs kiosques au sein de leur propre faculté afin de poser des questions sur les différentes spécialisations, dont le programme en TI. Le rôle des responsables des différents programmes présents à ces kiosques est de susciter l'intérêt et de démystifier la spécialisation en TI.

4.4. La révision constante du curriculum

Afin de contrer le manque de spécialisation des enseignants en TI et de temps pour mettre à jour leurs cours, il serait important de les encourager à renouveler leurs connaissances en accommodant leurs charges de travail pour leur permettre de suivre des formations et faciliter la refonte des contenus des cours. Des formations sont souvent offertes par les Universités afin de renforcer les connaissances des enseignants dans leurs domaines respectifs. La participation à des conférences, colloques et congrès doit être encouragée par du financement de partenaires privés ou publics.

Nous sommes certainement conscients que la surcharge des professeurs et enseignants universitaires en TI par l'ensemble de leurs tâches est un frein à ces initiatives. Bien que ces recommandations puissent paraître utopiques, nous croyons que des petites actions ciblées et raisonnables en termes de charge de travail peuvent apporter leurs fruits par un effet boule de neige. À titre d'exemple, le département des systèmes d'information de l'Université Laval a mis sur pied un comité pour analyser les plans de cours du programme de baccalauréat en gestion des technologies d'affaires en vue de l'accréditation BTM (Business Technology Management). Ce comité, chargé de revoir le curriculum du baccalauréat en gestion des technologies de l'information, est composé de cinq professeurs qui ont invité chacun de leurs collègues à remanier ses cours afin que les objectifs de la formation s'inscrivent dans la lignée des expériences pratiques et de l'implication de l'industrie, tel que stipulé par cette accréditation. La répartition du travail a permis de ventiler la charge reliée à cette action de révision du curriculum.

Concernant le lien que le curriculum doit faire avec le milieu professionnel, Stevens *et al.* (2011) proposent d'ajuster constamment le curriculum des programmes afin qu'ils répondent le plus possible aux besoins changeants de l'industrie. Ils suggèrent de mettre moins l'emphase sur les aspects techniques qui semblent préoccuper davantage les enseignants. Comme nous l'avons souligné dans la section précédente, les départements des TI dans les universités peuvent appuyer ces initiatives avec des comités consultatifs faisant appel à des acteurs clés de l'industrie. Le rôle de ces comités est de planifier des rencontres régulières afin de modifier les contenus des cours et de réfléchir à des moyens efficaces pour faire valoir les caractéristiques humaines et relationnelles de la profession. Le rôle de l'industrie dans le changement des modèles mentaux biaisés des jeunes vis-à-vis de la profession TI est essentiel. Il est de l'intérêt des universités d'établir des partenariats avec des entreprises qui œuvrent en TI à travers le mentorat, les stages et les interventions individuelles ou en groupe (Agosto *et al.*, 2008).

5. Conclusion

Depuis plusieurs années, les universités et l'industrie déplorent une situation alarmante en ce qui a trait aux compétences en TI. Les premiers se plaignent de la baisse constante d'inscriptions dans les concentrations en TI et les seconds dénoncent la rareté des ressources dans ce domaine. Plusieurs chercheurs ont tenté d'expliquer cette situation et d'y apporter des solutions par le biais des théories et des modèles explicatifs. Nous avons constaté, à la lumière de ces études, que ce n'est pas tant l'aspect rébarbatif de la profession qui rebute les jeunes, mais plutôt l'image véhiculée de la carrière en TI. En effet, l'image véhiculée par le domaine TI est un facteur prépondérant de l'ensemble des éléments influençant le choix de la concentration TI de par sa récurrence dans les études empiriques. L'image véhiculée du domaine TI a été malmenée avec l'évolution rapide des technologies et la variété des rôles en entreprise en lien avec les TI. Cette raison n'est certainement pas la seule à affecter l'intérêt des individus pour une concentration, mais sa pertinence relève du fait qu'elle est probablement la seule qui est spécifique au domaine des TI.

Cet article n'a nulle intention de proposer une solution miracle au manque de compétences en TI. Il présente par contre les résultats d'un exercice d'agrégation des connaissances apportées par les études théoriques et empiriques précédentes, exercice dont l'objectif ultime est de mettre le doigt sur les raisons de cette problématique et d'y apporter des recommandations. La plupart des recommandations s'articulent autour d'une collaboration étroite à établir avec les différents acteurs, dont le milieu professionnel. Ce dernier joue un rôle important pour fournir l'information fiable et précise sur la profession. Il revient ensuite aux gouvernements et aux établissements scolaires secondaires et universitaires de divulguer cette information pour faire de la carrière en TI un objet de convoitise. Des études empiriques sont certes nécessaires pour tester l'efficacité de ces recommandations.

6. Références

- Agosto, D. E., Gasson, S., Atwood, M. (2008). *Changing mental models of the IT professions: A theoretical framework*. Journal of Information Technology Education, Volume 7, 205-221.
- Alexander, P. M., Holmner, M., Lotriet, H. H., Matthee, M. C., Pieterse, H. V., Naidoo, S., *et al.* (2011). *Factors affecting career choice: Comparison between students from computer and other disciplines*. Journal of Science Education and Technology, Volume 20, Numéro 3, 300-315.
- Azjen, I., Fishbein, M. (1980). *Understanding Attitudes and Predicting Social Behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Ballard, J., Scales, K., Edwards, M. A. (2006). *Perceptions of information technology careers among women in career development transition*. Information Technology Learning and Performance Journal, Volume 24, Numéro 2, 1-9.

- Becerra-Fernandez, I., Elam, J., Clemmons, S. (2010). *Reversing the landslide in computer-related degree programs*. Communications of the ACM, Volume 5, Numéro 2, 127-133.
- Cap Digital (2014). *Baromètre des métiers numériques*. Cap Digital et Multiposting. http://www.capdigital.com/wp-content/uploads/2015/01/barometre_7pages.pdf.
- Clayton, K. (2004). *Information quality: The relationship to recruitment in pre-tertiary IT education*. Issues in Information Science and Information Technology, Volume 1, 65-77.
- Commission européenne (2013). Communiqué de presse. http://europa.eu/rapid/press-release_IP-13-52_en.htm
- Conseil des technologies de l'information et des communications (2011). *Outlook 2011-2016: It's time for action*. http://www.ictc-ctic.ca/wp-content/uploads/2012/06/ICTC_Outlook2011Summary_EN_11-11.pdf.
- Downey, J. P., Mcgaughey, R., & Roach, D. (2011). *Attitudes and Influences toward Choosing a Business Major: The Case of Information Systems*. Journal of Information Technology Education, Vol.10, 231-251.
- Ferratt, T. W., Hall, S. R., Prasad, J., Wynn, D. Jr. (2010). *Choosing management information systems as a major: Understanding the smiFactors for MIS*. Communications of the Association for Information Systems, Volume 27, Numéro 1, 265-284.
- Gallagher, K. P., Kaiser, K. M., Simon, J. C., Beath, C. M., Goles, T. (2010). *The requisite variety of skills for IT professionals*. Communications of the ACM, Volume 53, Numéro 6, 144-148.
- Habak, D., Palanca, A. (2014). *À la conquête du marché québécois de l'emploi en TI: Guide du professionnel formée à l'étranger*. ABACUS.
- Hirschheim, R., Newman, M. (2010). *Houston, we've had a problem ... offshoring, IS employment and the IS discipline: perception is not reality*. Journal of Information Technology, Volume 25, Numéro 4, 358-372.
- Insight Report (2015). *The global Information Technology Report*. http://www3.weforum.org/docs/WEF_Global_IT_Report_2015.pdf.
- Kuechler, W. L., McLeaod, A., Simkin, M. G. (2009). *Why don't more students major in IS?* Decision Sciences Journal of Innovative Education, Volume 7, Numéro 2, 463-488.
- Lalonde, D. (2014). *TI: La pénurie de la main d'œuvre touche l'Amérique du Nord*. Journal Les Affaires.
- Lasen, M. (2010). *Education and career pathways in information communication technology: What are schoolgirls saying*. Computers and Education, Volume 54, Numéro 4, 1117-1126.
- Malgwi, C., Howe, M., Barnaby, P. (2005). *Influences on student choice of college major*. Journal of Education for Business, Volume 80, Numéro 5, 275-283.
- ManpowerGroup, (2014). *Ninth annual 2014 Talent shortage survey*. http://www.manpowergroup.com/talent-shortage-explorer/#.VVOKoPl_Oko.
- Observatoire du numérique (2012). *Compétence*. <http://www.observatoire-du-numerique.fr/usages-2/professionnels/competences>.
- Raymond Chabot Grant Thornton (2011). *Diagnostic sectoriel de la main-d'œuvre du secteur des technologies de l'information et des communications 2011*. TechnoCompétences, https://www.quebecinternational.ca/media/481352/technocompetences_diagnosticsectoriel_final_vr.pdf.
- Roach, D., Downey, J. P., McGaughey, R. (2011). *Attitudes and influences toward choosing a business major: The case of information systems*. Journal of Information Technology Education, Volume 10, 231-251.
- Stevens, D., Totaro, M., Zhiwei, Z. (2011). *Assessing IT critical skills and revising the MIS curriculum*. Journal of Computer Information Systems, Volume 51, Numéro 3, 85-95.
- TechnoCompétences (2015). *Diagnostic Sectoriel de la main-d'œuvre dans le secteur des technologies de l'information et des communications (TIC) au Québec*.
- Thomas, T., Allen, A. (2006). *Gender Differences in Students' Perceptions of Information Technology as a Career*. Journal of Information Technology Education, Volume 5, Numéro 1, 165-178.
- US News (2015). *The 100 best jobs*. <http://money.usnews.com/careers/best-jobs/rankings/the-100-best-jobs>
- Von Hellens, L., Clayton, K., Beekhuyzen, G., Nielson, S. (2009). *Perceptions of ICT careers in German schools: an exploratory study*. Journal of Information Technology Education, Volume 8, 211-228.
- Walstrom, K. A., Schambach, T. P. (2012). *Impacting Student Perceptions about Careers in Information Systems*. Journal of Information Technology Education: Research, Volume 11, Numéro 1, 235-248.
- Wieggers, K. E. (2003). *So You Want To Be a Requirements Analyst?* Software Development. http://www.processimpact.com/articles/be_analyst.pdf.
- Zhang, W. (2007). *Why IS: Understanding undergraduate students' intentions to choose an information systems major*. Journal of Information Systems Education, Volume 18, Numéro 4, 447-458.

Dossier spécial ArabWIC'2015

Trust framework for a secured routing in wireless sensor network

Ouassila Hoceini,

Laboratoire LARI, Université Mouloud Mammeri de Tizi Ouzou, Algerie

ouassila.hoceini@gmail.com

Said Talbi,

Laboratoire LMCS, Ecole nationale supérieure d'informatique (ESI), Alger, Algérie

Said.talbi@yahoo.fr

Rachida Aoudjit

Laboratoire LARI, Université Mouloud Mammeri de Tizi Ouzou, Algerie

rachida_aoudjit@yahoo.com

Résumé

Les techniques traditionnelles développées pour éliminer les attaques internes dans les réseaux filaires et sans fils ne sont pas adéquates pour les réseaux de capteurs, vu les contraintes de ressources. Afin de protéger les réseaux de capteurs sans fils (RCSF) contre les fait malicieux et égoïstes, quelques systèmes basés confiance sont récemment modélisés. L'efficacité et la fiabilité des ressources d'un système de confiance sont les besoins les plus fondamentales pour n'importe quel réseau de capteurs.

Dans ce papier, nous avons proposé une architecture de confiance pour un routage sécurisé dans un réseau de capteurs sans fils, qui emploie une topologie hiérarchique. Cette approche peut réduire considérablement le coût d'évaluation de la confiance et garantit une meilleure sélection de chemins sécurisés qui mènent vers la station de base. La théorie et les résultats de simulation montrent que notre schéma utilise moins de ressources et d'énergie comparé aux systèmes de confiance actuels dans les réseaux de capteurs sans fils. De plus, il détecte les noeuds malicieux et défectifs et nous protège des attaques internes.

Abstract

Traditional techniques to eliminate insider attacks developed for wired and wireless ad hoc networks are not well suited for wireless sensors networks due to their resource constraints nature. In order to protect WSNs against malicious and selfish behavior, some trust-based systems have recently been modeled. The resource efficiency and dependability of a trust system are the most fundamental requirements for any wireless sensor network (WSN).

In this paper, we propose a Trust Framework for a Secured Routing in Wireless Sensor Network (TSR) scheme, which works with clustered networks. This approach can effectively reduce the cost of trust evaluation and guarantee a better selection of safest paths that lead to the base station.

Theoretical as well as simulation results show that our scheme requires less communication overheads and consumes less energy as compared to the current typical trust systems for WSNs. Moreover, it detects selfish and defective nodes and prevents us of insider attacks.

Mots-clés

Réputation, Management de confiance, Evaluation de confiance, Energie, Sécurité, Réseau de Capteurs sans fils

Keywords

Reputation, Trust management, Trust evaluation, Energy, Security, Wireless Sensor Network

1. Introduction

Traditional crypto-schemes developed for wired and wireless networks may not prevent sensor networks of malicious attacks. In other words, they may not be suitable for networks with small sensor nodes due to limited bandwidth and stringent node constraints in terms of power and memory. Therefore, it is important to develop trust management schemes and protocols that take into account the intrinsic features of wireless sensor networks. In sensor network security, trust is used as a measure of node's competence in providing required service (Riaz Ahmed Shaikh, *and al.* 2009), (A.A. Pirzada, 2004), (Y.L. Sun *and al.* 2006), (R.A. Shaikh *and al.*, 2006), (M. Momani, *and al.* 2007). It is the level of assurance about a key's authenticity that would be provided by some centralized trusted body to the sensor node (E. Shi and A. Perrig, 2004), (H.S. Ng *and al.* 2006), (Riaz Ahmed Shaikh, *and al.* 2009), (R.A. Shaikh *and al.* 2006). In case of multihop clustering, it helps to select trusted route through which a node can send data to the cluster head. During inter-cluster communication, trust management helps to select trusted route gateway or other trusted cluster heads through which the sender node will forward data to the base station (R.A. Shaikh, *and al.* 2006). In this work we focus on the inter-cluster communication, we give a way to select the safest path that lead to the base station on reducing the cost of trust evaluation. Our proposed scheme focus on the following features:

- TSR does not evaluate trust values of individual nodes. It builds the safest paths by evaluating nodes composed them.
- Our scheme works in a clustered topology.
- Traditional trust management schemes consume a lot of energy during recommendation phase. Our scheme allows recommendations exchange between cluster-heads with applying a light mechanism to reduce communication overheads.
- TSR detects insider attacks, defectives and selfish nodes and eliminates them from network communication.
- It eliminates unsecured paths to guarantee successful communication in a clustered network.

The remainder of the paper is organized as follows: Section 2 presents some related work to trust systems in wireless sensor networks. The proposed trust scheme is described in Section 3. We provide in section 4 a set of tests and evaluate our reputation system with respect to overall network performance; energy consumption and resistance against presence of selfish and malicious nodes. Finally section 5 concludes the paper and suggests some recommendations for further research.

2. Related Work

Recently, some trust management schemes have been proposed such as GMTS (R.A. Shaikh, *and al.* 2006), PLUS (Z. Yao *and al.* 2006), RFSN (S. Ganeriwal and M.B. Srivastava, 2004), LDTS (Li, X.; *and al.* 2013). Trust-Based Security for Wireless Ad Hoc and Sensor Networks (A. Boukerche, *and al.* 2007), TSRF (A Trust-Aware Secure Routing Framework in Wireless Sensor Networks, Junqi Duan *and al.* 2014), 2-ACKT (Two-Way Acknowledgment-Based Trust Framework for Wireless Sensor Networks, X. Anita, *and al.* 2013), A Secure Trust Establishment Scheme for Wireless Sensor Networks (Farruh Ishmanov, *and al.* 2014). In the following, we describe briefly some of them. In (Z. Yao *and al.* 2006), Z. Yao *and al.* have proposed PLUS protocol for wireless sensors networks. The authors adopt a localized distributed approach. Trust evaluation is based on either direct or indirect observations. S. Ganeriwal *and al.* have proposed RFSN (S. Ganeriwal *and al.* 2008), (S. Ganeriwal and M.B. Srivastava, 2004) scheme for sensor networks, where each sensor node maintains the reputation for neighboring nodes only. Trust values are calculated on the basis of that reputation and they use Bayesian formulation for representing reputation of a node. RFSN assumes that the node has enough interactions with the neighbors so that the reputation can reach a stationary state.

However, if the rate of node mobility is higher, reputation information will not stabilize. In RFSN, no node is allowed to disseminate bad reputation information. If it is assumed that bad reputation is implicitly included by not giving out good reputation, then in that case, the scheme will not be able to cope with uncertain situations (H. Chen *and al.*, 2007).

Shaikh R.A. *and al.* have proposed lightweight Group-based Trust Management Scheme (GTMS) (R.A. Shaikh *and al.*, 2006) for wireless sensor networks. Within a cluster, each sensor node calculates individual trust values for all other nodes based on the direct or indirect observations. Based on the trust value, nodes are classified into three categories: 1) trusted, 2) un-trusted or 3) un-certain. In the same way, each cluster maintains the trust value of other clusters (Riaz Ahmed Shaikh *and al.*, 2010).

3. Trust framework for a secured routing in wireless sensor networks

Trust Framework for a Secured Routing in Wireless Sensor Network (TSR) is a light trust scheme that works in a clustered topology and evaluates the paths reliability in a sensor network basing on:

- Sink's acknowledgements.
- Neighbors recommendations.

Our protocol guarantee a secured routing with employing a best path searching technique. The new of this protocol is the global trust measurement for a group of nodes forming a path. In the trust schemes presented until now, if sensor node forwards received packets to the next hop, we qualify it as trusted. Without taking into account what is the end receiver of these packets. In our protocol, if a trusted node has selfish nodes as neighbors, the trustworthiness of path to sink via this node will be reduced. The trusted next hop selection is not sufficient to be sure that the packet will be received by destination. The reliable path construction is the goal of our scheme. Consider as a node "S" send packet to one reliable neighbor "V", if node "V" relays the received packet from "S" to one selfish neighbor "A", the reliability of the path $S \rightarrow V \rightarrow A$ will be reduced. So the trust score of "V" will be decremented in the node memory of "S". We assume in the following that the sink has a signal power that can reach the entire network.

For evaluating the trust degree, our scheme requires three phases:

1. Successful interactions calculation
2. Exchange recommendations
3. Aggregation at CH level.

3.1. Successful interactions calculation

Each *CH* (Cluster-head) calculates successful interactions using the acknowledgements received from the sink. Thus, a sequential number *Seq* of 2 bytes is assigned for each data packet created to identify it. Once a *CH_i* makes decision about the direct destination (first hop) of packet forwarding, it save its identity *IDI* (first hop) of 2 bytes. When a *CH* send packet to sink, it send it with its sequential number *Seq*. At the beginning of communication, that one is considered successful. When the sink receives a data packet, it reads its cache and compares the sequential number of current packet with the one of the last packet coming from the same *CH_i*. If the current sequential number *Seq_i* is not the following of the last one *Seq_j* (*i* is not equal to *j+1*), the sink will send the value of *Seq_j* to *CH_i*. This represents also an acknowledgement of previous packets. Periodically, an acknowledgement is received by all the *CHs* from the sink, even if there is no packet loss. This acknowledgement informs *CHs* of lost packets and knowledge the previous packets. At the reception of an acknowledgement from the sink, in case of packet loss, the *CH* search in its cache the identity of direct destination (*IDI*) of the lost packet. Once finding, it decrements the number of successful interactions with it. With this manner a sensor node has not need to use promiscuous mode to check if its neighbor's forward packets, so communication overheads is reduced.

3.2. Exchange recommendations

Each round is devised into sub-periods, after each sub-period, *CHs* broadcast the vector of recommendations of $n \times 2$ bytes that contains the identities of neighbors judged not reliable.

- *n* is the number of nodes judged not reliable.

At the reception of these recommendation vectors, the number of good recommendations is incremented for nodes that do not include within this vector. The process of recommendations exchange is executed periodically, without calling of recommendations requests, and then we conserve considerably node battery. This exchange is less costly compared with others schemes presented in literature. Taking for example the PLUS protocol, for evaluating one neighbor, a pair of packets must be exchanged, a request *EReq* of 2 bytes for asking a recommendation and a response *Erep* of 6 bytes for responder to this request.

3.3. Aggregation at CH Level

The trust value of a neighbor will be calculated as shown in formula (1):

$$T_{ij} = \frac{\alpha}{2} \left(\frac{GI_{ij}}{NI_{ij}} + \frac{GR_{ij}}{NR_i} \right) \quad (1)$$

When:

- T_{ij} Trust value of a link $i \rightarrow j$.

- $GI_{i,j}$ The number of successful interactions at the link $i \rightarrow j$.
- $NI_{i,j}$ The number of interactions carried out by the link $i \rightarrow j$.
- $GR_{i,j}$ The number of good recommendations that node i received about the node j .
- NR_i The total number of recommendations received at node i (the number of times a node receives the trust vector).
- α Each node has a value of this coefficient, it equal to 1 when no neighbor declares this node as "attack". For each reception to recommendation "attack", the α coefficient is decremented so as:

$$\alpha = \frac{K_{ij}}{NbNeighbor_i} \quad (2)$$

- K_{ij} is the number *nodei's* neighbors that do not declare the *node j* as "attack".
- $NbNeighbor_i$ is the total number of *nodei's* neighbors.
- This division out of 2 because a trust value is considered to be a numerical quantity lying between 0 and 1 (inclusive) as suggested in (Y.L. Sun *and al.*, 2006), (G. Theodorakopoulos and J.S. Baras, 2006), and (H. Jameel *and al.* , 2005).

3.4 Representation of Trust levels

One of TSR feature is once the trust values are updated the decision making became a simple process. Each CH separates its neighbors in three groups according to these following demands:

$$\text{Level}(CH_j) = \begin{cases} \text{GOOD} & \text{if } T_{ij} \geq \gamma \\ \text{MIDDLE} & \text{if } \gamma < T_{ij} < \beta \\ \text{ATTACK} & \text{if } T_{ij} < \beta \end{cases}$$

- β is the threshold to judge a neighbor as attack and γ is the value from what a node is judged GOOD.
- With the reception of a message declared one node as attack, α will be decremented, as shown in formula (2).

A selfish node can improve its score and join the group of reliable nodes if its trust degree overtakes the threshold fixed beforehand (β). Else, if its trust degree is less than the threshold, node is considered as attack or defective node and it cannot improve its score. Table.1 shows the different types of TSR packets.

Type	Packet	Size
DataPacket (CM \rightarrow CH, CH \rightarrow Sink)	Data (variable), SeqNum (2 bytes), the source (2 bytes)	4 bytes + size of data
RecPacket (CH \rightarrow CH)	N (number of nodes judged no reliable), ID of node (2 bytes)	2 bytes x N (size of vector)
Ack (Sink \rightarrow CH)	SeqNum (2 bytes), ID of destination (2 bytes)	4 bytes

Table 1. TSR packets

4. Tests And Evaluations

For evaluating the performances of our protocol TSR, a comparative study between our protocol TSR and two other ones PLUS and RFSN will be presented in this section using OMNET ++ simulator. For energy consumption analysis, we assume first order radio model that is widely used by the researchers as in (H. sook Kim and K. jun Han, 2005) and (Y. Massad *and al.*, 2008), in which the energy expanded to transfer a k -bit packet to a distance d and to receive that packet, as suggested by H.O. Tan and I. Korpeoglu in (H. O. Tan and I. Korpeoglu, 2003) is:

$$\begin{aligned} ET_x(k, d) &= k * E_{elec} + k * d^2 * E_{amp} \\ ER_x(k) &= k * E_{elec} \end{aligned} \quad (3)$$

Here, E_{elec} is the energy dissipation of the radio in order to run the transmitter and receiver circuitry and is equal to $50nJ/bit$. The E_{amp} is the transmit amplifier that is equal to $100pJ/bit/m^2$. The E_{elec} and E_{amp} are the device specific parameters. The values that we use here for the theoretical analysis are the assumed values, which are commonly used in the literature (H. sook Kim and K. jun Han, 2005), (Y. Massad *and al.* 2008).

Scenario 1: Recommendations between cluster-heads

In the case of PLUS and RFSN, when a sensor node has need a recommendation about other nodes, it send a request to its neighbors. In the case of our scheme TSR, a CH receives periodically from its neighbors a recommendation vector that specifies the selfish nodes. The request for asking recommendation does not exist in our scheme. In the case of TSR the consumed energy by the transmitter of recommendation packet is:

$$E = ET_x (16*m, d) \tag{4}$$

The consumed energy by the receiver of recommendation packet:

$$E = ER_x (16*m) \times Nb \tag{5}$$

- m is the size of recommendation vector.
- Nb represents the number of neighbors.
- 16 represents the size of field reserved to node identity.

In the case of RFSN, the consumed energy by the transmitter of recommendation packet is:

$$E = n \times [ET_x (16, d) + ER_x (48)] \tag{6}$$

When:

- n represents the number of trusted nodes into a cluster.
- 16 and 48 present respectively the size of request packet and the size of recommendation packet of RFSN scheme.

Also, in RFSN the consumed energy by the transmitter of response packet is:

$$E = ET_x (48, d) + ER_x (16) \\ E = 16 * E_{elec} + 48 (E_{elec} + d^2 * E_{amp}) \tag{7}$$

In case of PLUS protocol, the consumed energy by the transmitter of recommendation packet is:

$$E = ET_x (16, d) + (n-2)ER_x (48) \\ E = 16(E_{elec} + d^2E_{amp}) + (48 E_{elec}) \tag{8}$$

When:

- n is the number of cluster-heads.

For consumed energy by the transmitter of response packet is:

$$E = ET_x (48, d) + ER_x (16) \\ E = 48(E_{elec} + d^2 * E_{amp}) + (16 E_{elec}) \tag{9}$$

- 16 and 48 present respectively the size on bits of request and response packet of PLUS scheme.

The summary of energy consumption during recommendations between cluster-heads is presented in table 2. When: n is the total number of neighbors in PLUS and RFSN and m is the size of recommendation vector in TSR protocol.

	TSR	RFSN	PLUS
Number of sanded packets	0	$t \leq n-2$ For evaluating a one neighbor	1 For evaluating a one neighbor
Number of recommandations received	1 For evaluating all neighbors	$t \leq n-2$ For evaluating a one neighbor	$n-2$ For evaluating a one neighbor
Size of request	/	16 bits	16 bits
Size of response	$m*16$ bits	48 bits	48 bits
The consumed energy on request	$E_{Tx} (m*16, d)$	$n \times [E_{Tx} (16, d) + [E_{Rx} (48)]]$	$E_{Tx} (16, d) + (n - 2) E_{Rx} (48)$
The consumed energy on response	$E_{Rx} (m*16)$	$E_{Tx} (48, d) + E_{Rx} (16)$	$E_{Tx} (48, d) + E_{Rx} (16)$

Table 2. Pairs recommendations between cluster-head

In order to compare energy consumption during recommendation scenarios between cluster-heads. We assume to have the following simulation parameters:

Parameter	Value
Surface	100*100*50
Localization of base station	(0,0,0)
Number of nodes	100, 200, 300
Number of base stations	1
Period of Round	10 s
Simulation Time	200 s
Number of selfish CHs	3, 5, 7

Table 3. Simulation parameters of scenario 1

The figure 1 shows clearly that TSR consumes less energy compared with RFSN and PLUS schemes. In TSR protocol, nodes do not send requests for asking recommendations from their neighbors. The sending of recommendation packets is carried out periodically. While, in RFSN and PLUS, cluster-heads send recommendation requests every time a CH has need recommendation about one neighbor. This figure shows also that the PLUS scheme consumes less energy than RFSN, because that in PLUS scheme the request packet is broadcasted of all its neighbors, when needed recommendation. While, in the RFSN protocol, the request packet is sanded on unicast to all its trusted neighbors.

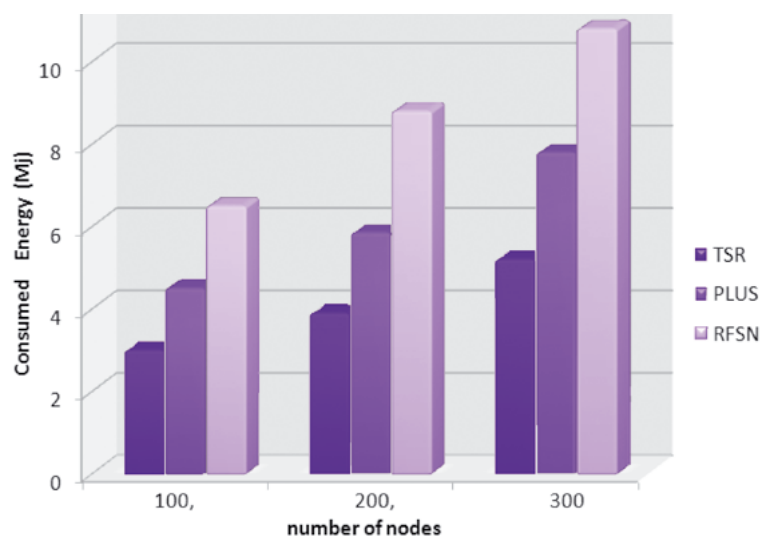


Figure 1. Consumed energy in each scheme TSR, RFSN and PLUS for recommendations scenarios

Scenario 2: Loss packet rate in the presence of selfish nodes

In a scenario of 100 nodes, using these parameters we are shown the following results:

Parameter	Value
Surface	100*100*50
Localization of base station	(0,0,0)
Number of nodes	100
Number of clusters	8
Number of base stations	1
Period of Round	10 s
Simulation Time	200 s
Maximum rate of selfish CHs	37,5

Table 4. Simulation parameters of scenario 2

Figure 2 compares the packet loss rate with applying the TSR protocol, and without applying it. Values of γ and β are fixed after several simulations at respectively 0,50 and 0,23.

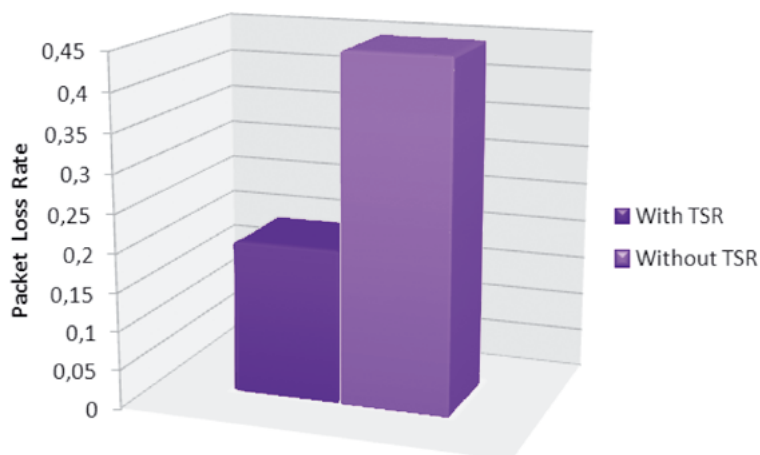


Figure 2. The packet loss rate with TSR and without TSR

Results presented in this figure are obtained, for a network size of 100 nodes. The trust degree attributed to nodes reduces interactions with selfish nodes, so we reduce considerably the packet loss rate.

Scenario 3: Packet loss rate in the presence of 37, 5% of selfish nodes and of 12, 5 % of Black Hole attacks

In a network of 100 nodes, we show the following results. Simulations have shown an important packet loss rate during the first rounds in the case of application of TSR scheme. That can be explained, with considering all neighbors as trusted at deployment. By disobeying to forward packets, communication with selfish nodes weakens. This reduces in parallel the packet loss rate. The figure 3 presents the number of packet lost in the presence of Black Hole attacks and selfish nodes in a network of 9 clusters.

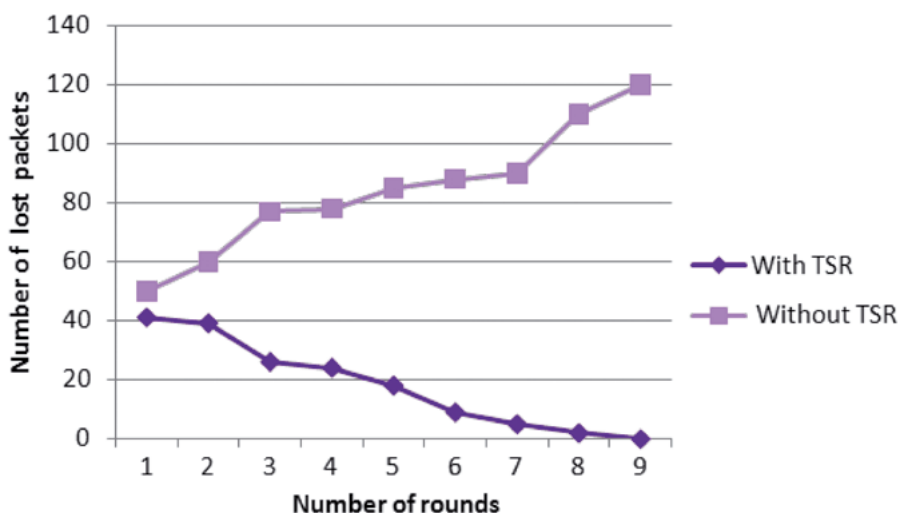


Figure 3. Number of lost packets with 12,5 % of attacks and 37,5% of selfish nodes

5. Conclusion

In this paper, we have proposed a light scheme that can cope with insider attacks in WSNs. TSR allows carrying out successful interactions in sensor networks, with its relevant technique that focus on safest path searching. TSR minimizes the energy consumption in the recommendations exchange phase. Although, the choice of the best path ensure that the destination will receive packets, but these packets can suffer from falsification. So, the challenging problem is to ensure that the data does not been modified. This motivates future work.

6. References

- A.A. Pirzada and C. McDonald. (2004). *Establishing Trust in Pure Ad-Hoc Networks*. Proc. 27th Australasian Computer Science Conf. (ACSC '04), pp. 47-54, Jan. 2004.
- A. Boukerche, X. Li, and al.(2007). *Trust-Based Security for Wireless Ad Hoc and Sensor Networks*. Computer Comm., vol. 30, pp. 2413-2427, Sept. 2007.
- A. Perrig, R. Szewczyk, and al.(2002). *SPINS: Security Protocols for Sensor Networks*. Wireless Networks Journal (WINET), 8(5):521–534, September 2002.
- E. Shi and A. Perrig. (2004). *Designing Secure Sensor Networks*. IEEE Wireless Comm., vol. 11, no. 6, pp. 38-43, 2004.
- Farruh Ishmanov, Sung Won Kim and al. (2014). *A Secure Trust Establishment Scheme for Wireless Sensor Networks*. Sensors 2014, 14, 1877-1897; doi:10.3390/s140101877.
- G. Theodorakopoulos and J.S. Baras.(2006). *On Trust Models and Trust Evaluation Metrics for Ad Hoc Networks*. IEEE J. Selected Areas in Comm., vol. 24, no. 2, pp. 318-328, Feb. 2006.
- H. Chen, H. Wu, and al.(2007). *Reputation-Based Trust in Wireless Sensor Networks*. Proc. Int'l Conf. Multimedia and Ubiquitous Eng. (MUE '07), pp. 603-607, Apr. 2007.
- H. Jameel, L.X. Hung, and al.(2005). *A Trust Model for Ubiquitous Systems Based on Vectors of Trust Values*. Proc. Third IEEE Int'l Security in Storage Workshop (SISW '05), pp. 674-679, Dec. 2005.
- H. sook Kim and K. jun Han. (2005). *A power efficient routing protocol based on balanced tree in wireless sensor networks*. in Proc. of the 1st Int. Conference on Distributed Frameworks for Multimedia Applications (DFMA '05), Feb. 2005, pp. 138–143.
- H.S. Ng, M.L. Sim, and al.(2006). *Security Issues of Wireless Sensor Networks in Healthcare Applications*. BT Technology J., vol. 24, no. 2, pp. 138-144, Apr. 2006.
- H. O. Tan and I. Korpeoglu.(2003). *Power efficient data gathering and aggregation in wireless sensor networks*. ACM SIGMOD Record, vol. 32, no. 4, pp. 66–71, Dec. 2003.
- Junqi Duan, Dong Yang, and al. (2014). *TSRF: A Trust-Aware Secure Routing Framework in Wireless Sensor Networks*. International Journal of Distributed Sensor Networks Volume 2014 , Article ID 209436, 14 pages.
- Li, X.; Zhou, F, and al.(2013). *LDTS: A Lightweight and Dependable Trust System for Clustered Wireless Sensor Networks*. IEEE Trans. Inf. Forensics Security 2013, 8, 924–935.
- M. Momani, S. Challa, and al.(2007). *Modelling Trust in Wireless Sensor Networks from the Sensor Reliability Prospective*. Innovative Algorithms and Techniques in Automation, Industrial Electronics and Telecomm., T.S. et al., ed., pp. 317-321, Springer, 2007.
- R.A. Shaikh, H. Jameel, and al.(2006). *Trust Management Problem in Distributed Wireless Sensor Networks*. Proc. 12th IEEE Int'l Conf. Embedded Real-Time Computing Systems and Applications (RTCSA '06), pp. 411-414, Aug. 2006.
- Riaz Ahmed Shaikh, Hassan Jameel, and al.(2009). *Group-Based Trust Management Scheme for Clustered Wireless Sensor Networks*. IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 20, NO. 11, NOVEMBER 2009.
- Riaz Ahmed Shaikh, Young-Koo Lee, and al.(2010). *An Extended Energy Consumption Analysis of Reputation-based Trust Management Schemes of Wireless Sensor Networks*. JOURNAL OF NETWORKS, VOL. 5, NO. 3, MARCH 2010.
- S. Ganeriwala and M.B. Srivastava. (2004). *Reputation-Based Framework for High Integrity Sensor Networks*. Proc. ACM Workshop Security of Ad Hoc and Sensor Networks (SASN '04), pp. 66-67, Oct. 2004.
- S. Ganeriwala, L. K. Balzano, and al. (2008). *Reputation-based framework for high integrity sensor networks*. ACM Trans. Sen. Netw., vol. 4, no. 3, pp. 1–37, 2008.
- X. Anita, J. Martin Leo Manickam, and al.(2013). *Two-Way Acknowledgment-Based Trust Framework for Wireless Sensor Networks*. International Journal of Distributed Sensor Networks Volume 2013, Article ID 952905, 14 pages, 4 April 2013.
- Y.L. Sun, W. Yu and al.(2006). *Information Theoretic Framework of Trust Modeling and Evaluation for Ad Hoc Networks*. IEEE J. Selected Areas in Comm., vol. 24, no. 2, pp. 305-317, Feb. 2006.
- Y. Massad, M. Goyeneche, and al. (2008). *Data aggregation in wireless sensor networks*. in Proc. of the 3rd Int. Conference on Information and Communication Technologies: From Theory to Applications (ICTTA 2008), April 2008, pp. 1–6.
- Z. Yao, D. Kim, and al.(2006). *PLUS: Parameterized and localized trust management scheme for sensor networks security*. in Proc. of the 3rd IEEE Int. Conf. on Mobile Adhoc and Sensor Systems, Vancouver, Canada, Oct. 2006, pp. 437–446.

Translating Discrete Time SIMULINK to SIGNAL

Safa Messaoud

University of Illinois at Urbana-Champaign, Department of Electrical and Computer Engineering
messaou2@illinois.edu

Neda Saeedloei

University of Minnesota Duluth, Department of Computer Science
nsaeedlo@d.umn.edu

Sandeep Shukla

Virginia Polytechnic and State University, Department of Electrical and Computer Engineering
shukla@vt.edu

Résumé

Les systèmes cyber physiques (CPS) devenant de plus en plus complexes et critiques, la conception orientée modèle (*Model Based Design*) qui consiste à élaborer les modèles formels d'un système afin qu'ils soient utilisés pour la vérification et la génération de code correct-par-construction, est devenue une méthode prometteuse pour le développement du logiciel embarqué de tels systèmes. Ce paradigme de conception réduit considérablement le coût et le temps de développement, tout en garantissant une meilleure robustesse et exactitude par rapport aux spécifications d'origine, en comparaison avec les méthodes traditionnelles de conception *ad hoc*. SIMULINK a été l'outil le plus réputé pour la conception de contrôle embarqué, aussi bien dans la recherche que dans l'industrie, durant les dernières décennies. Comme SIMULINK n'est pas doté d'une sémantique formelle, l'application de la méthodologie et des outils de conception orientée modèle à ses modèles est très limitée. Dans cet article, nous présentons un traducteur sémantique qui transforme les modèles SIMULINK de temps discret en programmes SIGNAL. Le choix de SIGNAL est motivé par son formalisme polychrone qui améliore la programmation synchrone par du parallélisme asynchrone ainsi que par la capacité de son compilateur de générer du code déterministe *multi-thread*. Notre traduction comporte trois grandes étapes : l'inférence de l'horloge, l'inférence de type et la traduction ascendante hiérarchique. Nous validons notre prototype en le testant avec différents modèles SIMULINK.

Abstract

As Cyber Physical Systems (CPS) are getting more complex and safety critical, Model Based Design (MBD), which consists of building formal models of a system in order to be used in verification and correct-by-construction code generation, is becoming a promising methodology for the development of the embedded software of such systems. This design paradigm significantly reduces the development cost and time while guaranteeing better robustness and correctness with respect to the original specifications, when compared with the traditional *ad-hoc* design methods. SIMULINK has been the most popular tool for embedded control design in research as well as in industry, for the last decades. As SIMULINK does not have formal semantics, the application of the model based design methodology and tools to its models is very limited. In this paper, we present a semantic translator that transforms discrete time SIMULINK models into SIGNAL programs. The choice of SIGNAL is motivated by its polychronous formalism that enhances synchronous programming with asynchronous concurrency, as well as, by the ability of its compiler of generating deterministic multi thread code. Our translation involves three major steps: clock inference, type inference and hierarchical top-down translation. We validate our prototype tool by testing it on different SIMULINK models.

Mots-clés

Système cyber physique, SIMULINK, SIGNAL, méthodes formelles, génération de code.

Keywords

Cyber Physical Systems, SIMULINK, SIGNAL, Formal Methods, Code Generation

1. Introduction

Cyber Physical Systems (CPS) are engineering systems consisting of the integration of computational control and physical components with continuous dynamics. As these systems are becoming more complex and their reliability and safety requirements are becoming more and more crucial, and harder to guarantee by the traditional design tools and methodologies, new design paradigms are emerging. Model Based Design is a much discussed approach for developing such systems. It consists of building mathematical models that capture the specifications as well as the critical design decisions for the system in the different stages of the development life cycle. Different tools have been developed to generate correct by construction code from these models, as well as for the verification of the system behavior in early design phases. Despite the intensive research in the Model Based Design, the Mathwork's graphical environment SIMULINK[5] is still the most widely used tool for the design of embedded software. Although it is very convenient to use, SIMULINK does not have published and authentic formal semantics. Hence, its models can not be used with the Model Based Design framework. Its generated diagrams are verified through numerical simulations and its behavior is strongly correlated with the simulation configuration parameters. Although simulation based analysis is a well accepted technique in industrial practice, it becomes impossible to exhaustively simulate the system for verification purposes, once it gets very complex. The preservation of semantic is another issue, since the behavior equivalence between the simulated model and the generated code is unclear. Formal models, on the other hand, are less applied as they are less intuitive to use and harder to learn. In order to close the gap between formal methods and industrial practices, researcher have attempted to either give formal semantics[8] to SIMULINK or translate it into formal models of computation[15][6][16]. In this paper, we present a prototype tool, SIM2SIG, that translates the discrete time blocks of SIMULINK to SIGNAL. SIGNAL is a data flow synchronous programming language, which was developed by IRISA[1]. Each variable (signal) within the SIGNAL program has its own clock, giving us the multi-rate (polychronous) formalism of SIGNAL. This timing model allows for streams to be computed asynchronously, which fits very easily to a multi-thread environment. This increases the embedded software reactivity and capabilities. Moreover, a number of formal verification tools such as the model checker SIGNALI[3] and the graphical developing interface SME[2] exist for Signal. These characteristics make SIGNAL an interesting model of computation for embedded software design. We follow the same translation methodology proposed in [16], for translating LUSTRE to SIMULINK, namely type and clock inference, and hierarchical block by block translation. The novelty in this work consists of bridging the gap between the 'almost' synchronous model of computation of SIMULINK and the polychronous model of computation of SIGNAL. In the past work by [16], the translation was straight forward due to the fact that the target language is synchronous and a global clock driven, whereas in SIGNAL language there is no global clock per se. A global clock may be calculated using the clock calculus if the translated SIMULINK model has the endochrony property. If a single global clock driver does not emerge, a polychronous model leading to multi-threaded behavior emerges. The other addition in this work is the use of affine clock relations between SIGNAL sub-processes, when multiple SIMULINK blocks have sampled inputs with varying sampling rates. The rest of the paper is organized as follows: Section 2 is a survey of the translation of SIMULINK to different models of computation. Section 3 is an overview of the SIGNAL formalism. In Sections 4 and 5, we compare SIMULINK and SIGNAL formalisms and present the translation goals and assumptions. Sections 6, 7 and 8 represent the three steps of the translation. Our prototype tool Sim2Sig is described in Section 9. In Section 10, Sim2Sig is tested on a SIMULINK model of a discretized DC-motor closed loop controller. We close this paper with some concluding remarks and suggested future work

2. Related work

A hand full of research efforts in the past have tried to give formal semantics to SIMULINK either by converting its models into a synchronous language [16] [9], hybrid automata[12], or I/O extended Finite Automata[6] or into a system of mathematical equations[8]. The main motivation for translating SIMULINK models to a formal language program lies in gaining access to the analysis and verification tools of the target language. In [16], discrete time SIMULINK was translated to LUSTRE[11] following three steps: clock inference, type inference and hierarchical bot-tom up block by block translation. Basic blocks like Addition or Multiplication are translated to primitive LUSTRE operators. Complex nodes like Subsystems are translated to Lustre nodes, which are carefully named in order to keep track of the original SIMULINK hierarchy. In [6], a semantic translator from SIMULINK to Hybrid System Interchange Format (HSIF) was introduced. HSIF is a network of hybrid automata, which can interact with each other using signals and shared variables. The translation was limited to continuous SIMULINK blocks and STATEFLOW Diagrams. The translation from SIMULINK/STATEFLOW is based on graph transformation. HA can model both continuous and discrete systems. However, they have not been formally standardized. Chapoutot et. al. [8] proposed to assign formal semantics to SIMULINK's simulation engine, solver and a subset of blocks that span discrete and continuous operations. The dynamical SIMULINK system is represented as a state space with continuous time, as well as discrete time state functions to represent the fact that SIMULINK models are hybrid systems. The simulation goal consists of finding the solution for the set of state space equations. This approach was validated by comparing the outputs of the SIMULINK simulator and the equation-based one for different case studies. Although the system of equations approach is able to cover both discrete and continuous blocks, this approach lacks tools for the equation grammar verification and simulation. In order to take advantage of the high computing power resulting from multi-core architectures, multi threading is very desired. Due to its multi-rate formalism, the polychronous language SIGNAL, leads naturally to multi-threaded code synthesis. This justifies our motivation for choosing SIGNAL as a target language.

3. The polychronous language signal

SIGNAL is a declarative multi-rate synchronous language. It satisfies the synchrony hypothesis, which assumes that the computation and communication time are instantaneous. While the synchronous languages have a totally ordered model of logical time, SIGNAL's model of logical time is partially ordered. The semantics of the language does not assume an a priori existence of a reference clock. Each variable (signal) is characterized by its own clock. In the following section, we introduce some preliminaries notions related to SIGNAL.

3.1. Preliminaries

The basic entity in a polychronous language is an event.

Definition 3.1. (Event). An event is an occurrence of a new value. We denote the set of all events in a system by Ξ . The relative occurrences of events can then be represented using the following binary relations over Ξ :

Definition 3.2. (Precedence, Preorder, Equivalence). Let $<$ be a precedence relation between events in Ξ . It is defined such that $\forall a, b \in \Xi, a < b$ if and only if a occurs before b . The relation $<$ defines a partial order on Ξ such that $\forall a, b \in \Xi, a < b$ if and only if a occurs before b or a, b occur logically simultaneously, or their order does not matter. Finally the equivalence relation \sim , is defined on Ξ such that $a \sim b = a < b \wedge b < a$, meaning that a and b are equivalent only if they occur simultaneously or their order does not matter. Thus \sim represents synchronicity of events.

An instant can also be seen as a maximal set of events that occur in reaction to any one or more events. Formally:

Definition 3.3. (Logical Instant). The set of all instants is denoted by Y . Each instant in Y can be seen as an equivalence partition obtained by taking the quotient of Ξ with respect to \sim such that $Y = \Xi / \sim$. For each set $S \in Y$, all events in S will have the property $\forall a, b \in S, a \sim b$, and $\forall a, b, (a \in S1 \wedge b \in S2 \wedge S1 \neq S2 \wedge S1, S2 \in Y \rightarrow a \not\sim b$. Each instant contains events on signals. If a signal has no event in an instant then it is considered absent. We denote a specific value of a signal x by function $x(t)$ where $t \in \mathbb{N}$ and t represents the t th instant in the totally ordered set of instants where signal x is different from \perp .

Definition 3.4. (Epoch, Clock). The epoch, $\sigma(x)$, of a signal x is the maximum set of instants in Y where for each instant in $\sigma(x)$, x takes a value from T . The clock of the signal x is a characteristic function that tells whether or not an event in x is absent or is in the set T . Clock is a function of type $Y \rightarrow [\text{true}, \text{false}]$ such that for a signal x it returns another signal \hat{x} defined by $\hat{x}(t) = \text{true}$ if $x(t) \in T$.

Note that not all inputs and outputs are present or computed during every instant in Y which means that not all signals have the same epoch or clock. This gives the multi-clocked or polychronous behavior. Using the above definitions and characteristics, three possible relationships can be drawn between any two clocks x and y: equivalent, sub-clocked, or unrelated. If the clocks of x and y are true for the exactly the same set of instants, $\hat{x} = \hat{y}$, then it is said that these two clocks are equivalent, and the corresponding signals are also synchronous. If the clock of a signal x is true for a subset of instants where the clock of y is true then it is said that x is a sub-clock of y. If the clocks of x and y are not equivalent or subset or superset of the other then the clocks are said to be unrelated [14].

3.2. The SIGNAL Formalism

The primitive SIGNAL operators are Function, Delay, Under-sampling and Priority Merging (see Figure 1).

1. Function Operator: performs user defined operations on a set of signals x_1, \dots, x_n that must be present simultaneously and produces an output y at the same instant.

$$\begin{aligned} \text{Operation: } & y := f(x_1, x_2, \dots, x_n) \\ \text{Clock Relation: } & \hat{y} = \hat{x}_1 = \hat{x}_2 = \dots = \hat{x}_n \end{aligned} \tag{1}$$

2. Delay Operator: sends a previous value of the input to the output with an initial value k as the first output. The original and delayed signals are synchronous.

$$\begin{aligned} \text{Operation: } & y := x \$ \text{ init } k \\ \text{Clock Relation: } & \hat{y} = \hat{x} \end{aligned} \tag{2}$$

3. Under-Sampling Operator: down-samples an input signal x based on the true occurrence of another input signal z. The output signal clock is thus equal to the intersection of the clocks of x and z=true, noted [z].

$$\begin{aligned} \text{Operation: } & y := x \text{ when } z \\ \text{Clock Relation: } & \hat{y} = \hat{x} * [z] \end{aligned} \tag{3}$$

4. Priority Merging Operator: merges two signals x and z into one signal y. At any logical instant, if x is present, then y will have the value present on x, else y will have the value present on z. If neither x nor z are present, y is absent.

$$\begin{aligned} \text{Operation: } & y := x \text{ default } z \\ \text{Clock Relation: } & \hat{y} = \hat{x} + \hat{z} \end{aligned} \tag{4}$$

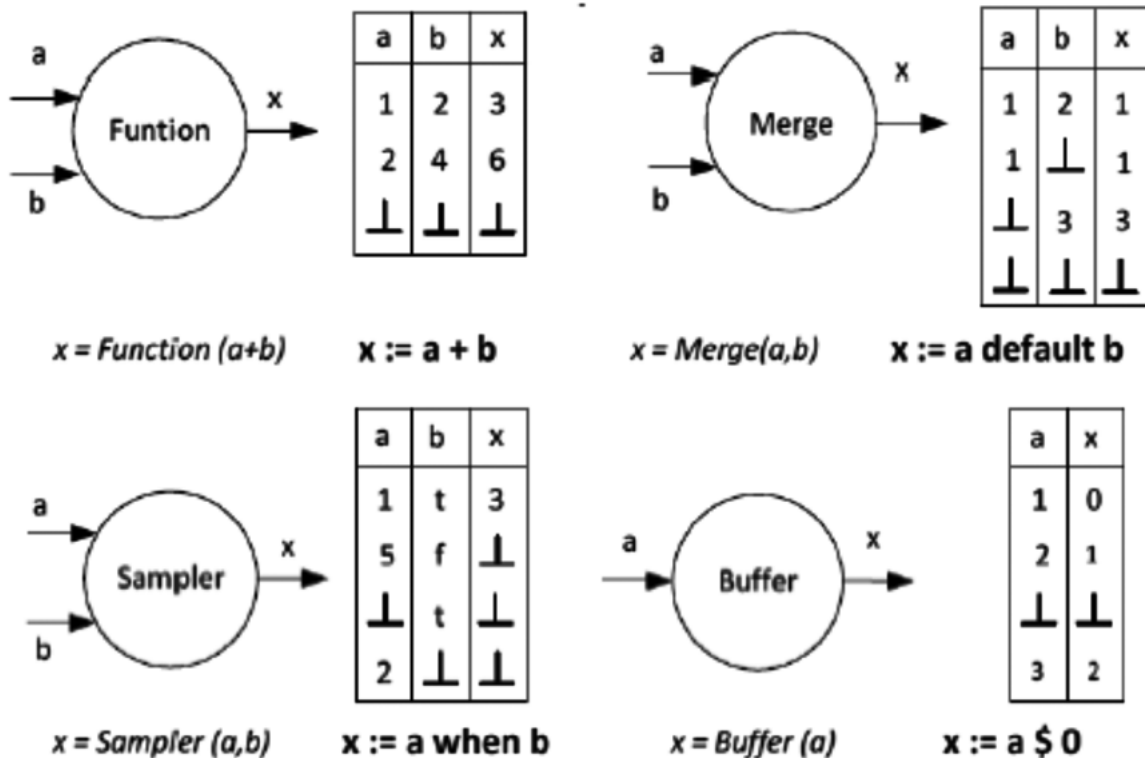


Figure 1. Primitive SIGNAL operations [7]

3.3. Advanced SIGNAL Constructs

Clock relations are not only inferred from the SIGNAL statement, they can be given explicitly [10]: The equation, $clk := when\ b$, implies that clk represents the set of instants at which b holds true. The equation, $clk^{\wedge} = s$, implies that clk is the clock of s . The equation, $s1^{\wedge} = s2$, specifies that the signals $s1$ and $s2$ are synchronous. Another useful construct is *Cell*: $y := x\ cell\ z\ init\ k$. In this case, the output signal contains the values of the first input signal x for all its instants and retains the previous value of x during the true instances of the second boolean input k . The clock of y is the union of the clocks of x and z . An example of the Cell operator is shown below:

$$\begin{aligned} x: & \perp\ v2\ v3\ \perp \\ z: & \text{t}\ \text{t}\ \perp\ \text{f} \\ y: & k\ v2\ v3\ v3 \end{aligned}$$

3.4. SIGNAL Processes

A SIGNAL program is a process. The parallel composition of two processes P and Q, noted P|Q is the union of equation systems defined by both processes. P and Q communicate via their common signals. The template of a SIGNAL process is:

```
Process MODEL = { %parameters%
    (? %inputs%; ! %outputs%; )
    ( | %body of the process% | )
where
    %local declarations%
end ;
```

The input-output ports are declared using the symbol ? and ! respectively. Each input or output is associated with its type (event, integer, boolean, real). Each SIGNAL statement consists of the four primitive operators.

4. Comparing simulink and signal

Both SIMULINK and SIGNAL are data-flow languages. They both manipulate signals. In SIMULINK, signals are the wires that connect the blocks in a model. In SIGNAL, a signal is the program variable corresponding to a stream. A system performs a specified operation on an input signal and produces an output signal. The systems in SIMULINK are library blocks that could be simple (e.g., Adder, Product) or composed (subsystems). In SIGNAL, systems are built-in operators (e.g., when, default), as well as user defined ones, called processes. Another similarity consists of the hierarchical composition of systems. In SIMULINK, the subsystems are drawn graphically within their parent system, to form a tree structure. In SIGNAL, as well, a parent process can contain multiple subprocesses. Despite of these similarities, SIMULINK and SIGNAL are different in several major ways: First, SIGNAL has a well defined formal semantic, whereas SIMULINK's behavior strongly depends on the choice of the simulation parameters. For example, some models are accepted if we allow to handle rate change automatically, others are rejected if the automatic rate change opting is unchecked. Second, SIGNAL has a discrete time semantics, whereas SIMULINK has a continuous one. Even the blocks belonging to the discrete library produce piecewise constant continuous-time signals. Third, SIGNAL is a strongly typed language that explicitly specifies the type of each flow. However, SIMULINK does not require the type specification for each block. This can be done, using, for instance, a Data Type Converter Block (see Section 6). Finally, SIGNAL is a multi-rate language, which means that two variables can be of different rates and can remain unrelated throughout the program. However, SIMULINK, both in sample-driven and event-driven cases, has a global clock, namely the simulation clock, that is synchronous with every clock in the model (see Section 7).

5. Translation goals and assumptions

The problem of semantic translation can be formulated as follows: Given a SIMULINK model of a dynamic system, compute a flow equivalent dynamic system model in SIGNAL which produces the same execution traces as the simulation output in SIMULINK. Our tool rejects the models with typing or timing errors flagged by SIMULINK. We limit our translation to the discrete time part of SIMULINK. This is justified by the fact that only the controller in safety critical systems is implemented on the computer, hence it must be designed in discrete time. The list of supported SIMULINK blocks is shown in Figure 2. As SIMULINK semantics depend on the simulation method, we limit our translation to one method. We chose the solver to be discrete and fixed step, the simulation

mode to be auto and to automatically handle rate transition for data transfer deterministically. We also assume that the boolean logic signals flag is on. We developed our translation method using MATLAB 7.12.01 (R2011a) and SIMULINK block Library V7.7.

6. Type inference

6.1. Types in SIMULINK

Unlike in SIGNAL, variable types are not explicitly declared in SIMULINK. However, implicitly, SIMULINK has some typing rules. The simulation engine rejects some models because of typing errors. The basic types for SIMULINK are: boolean, double, single, int8, uint8, int16, uint16, int32, uint32. The main SIMULINK typing rules are:

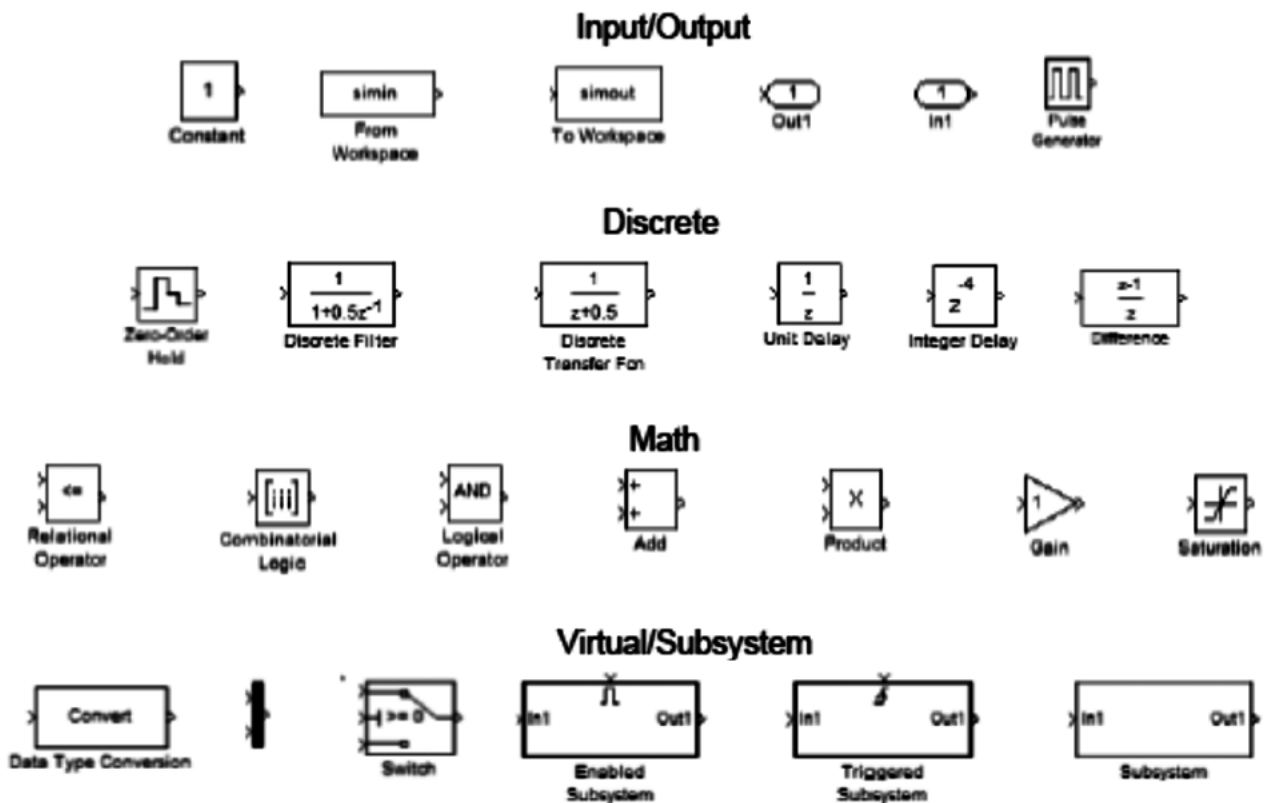


Figure 2. Supported SIMULINK blocks

- By default, all signals are of type double, except when a block requires a defined type. For example, the inputs of Logical Operator blocks must be of type boolean.
- The user can explicitly set the type of a signal to another type (e.g., by a Data Type Converter Block)
- An error type occurs when incompatible types are fed in one block, for example, when a boolean and an integer are fed to the same Adder block. The typing rules for each block are given in Table 1. We define $T_{Num} = \{\text{double, single, int8, uint8, int16, uint16, int32, uint32}\}$, and $T_{Bool} = \{\text{boolean}\}$. Let $\{\alpha, \phi\} \in T_{Num}$, $\theta \in T_{Bool}$ and $\{\gamma, \beta\} \in \{T_{Bool}, T_{Num}\}$

SIMULINK Block	Typing Rule
Constant	α
Adder	$\alpha \cdot \dots \cdot \alpha \rightarrow \alpha$
Gain	$\alpha \rightarrow \alpha$
Relational Operator	$\alpha \cdot \alpha \rightarrow \theta$
Logical Operator	$\theta \cdot \dots \cdot \theta \rightarrow \theta$
Discrete Transfer Function	$\alpha \rightarrow \alpha$
Unit Delay, Inport, Outport	$\gamma \rightarrow \gamma$
Data Type Converter	$\gamma \rightarrow \beta$
Switch	$\alpha \cdot \phi \cdot \alpha \rightarrow \alpha$

Table 1. Typing Rules for Some SIMULINK blocks

6.2. Types In Signal

SIGNAL is a strongly typed language: variables have a declared type and operations have precise type signatures. The basic types for SIGNAL are integer, real and boolean. Type casting can be performed as in C. For example, an integer x is converted to a real y as follows: $y = \text{integer}(x)$. The array type allows grouping synchronous elements of the same type. An array of size N with elements of type $element_type$ is defined as follows: $[\text{inp}_1, \dots, \text{inp}_N]$ $element_type$.

6.3. Type inference

The goal of this step consists of inferring the type of each signal in SIMULINK, so that its corresponding type in SIGNAL can be used in the translation. For the type inference, we use a fix-point algorithm on the lattice shown in Figure 3. \perp means undefined type and error means typing error. We call $x^T \in T_{Sim}$ the type variable corresponding to the variable x , with $T_{Sim} = \{T_{Num}, T_{Bool}\}$. We define a monotonic function $\text{sup}: (T_{Sim})^n \rightarrow T_{Sim}$ in the type lattice, where n is the number of blocks in the SIMULINK model. $\text{Sup}(x^T, y^T) = z^T$, denotes that z^T is a least common upper bound of x^T and y^T . The fixed point is calculated on the set of equations shown in Table 2.

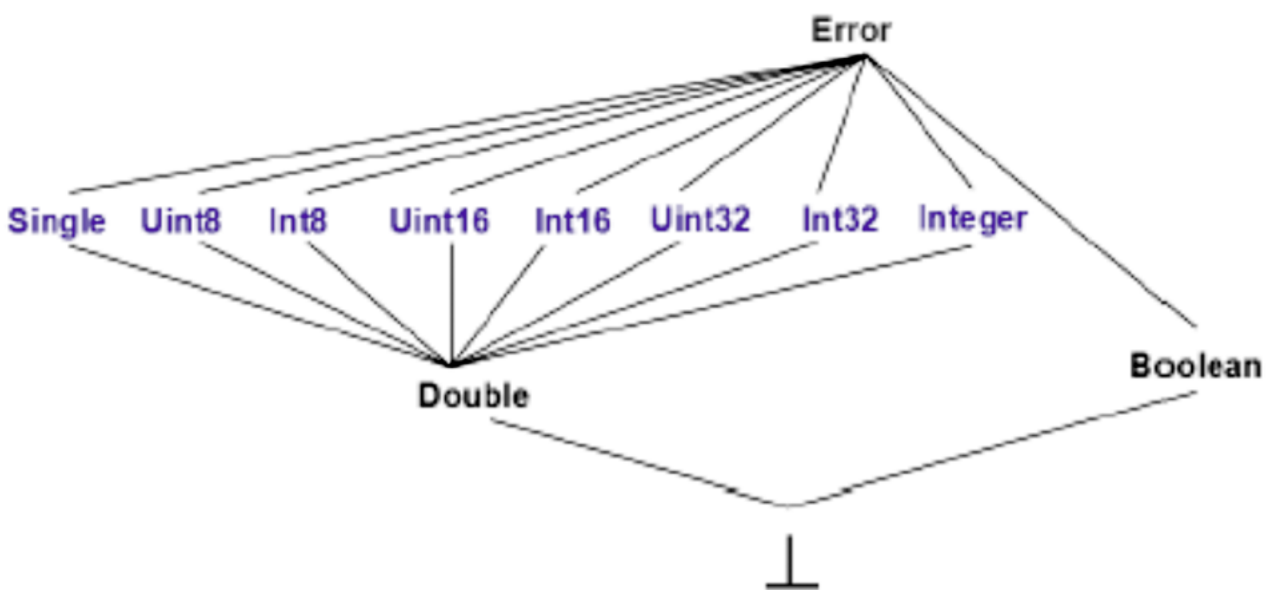


Figure 3. The Type Lattice

SIMULINK Equation	Type Equation
$y = \text{Adder}(x_1, \dots, x_k)$	$y^T = x_1^T = \dots = x_k^T = \text{Sup}(\text{double}, y^T, x_1^T, \dots, x_k^T)$
$y = \text{Constant}_\alpha$	$y^T = \text{if } y^T \leq \alpha \text{ then } \alpha \text{ else error}$
$y = \text{DataTypeConv}_\alpha(x)$	$y^T = \text{if } y^T \leq \alpha \text{ then } \alpha \text{ else error}$
$y = \text{UnitDelay}(x)$	$x^T = y^T$
$y = \text{RelOp}(x_1, x_2)$	$x_1^T = x_2^T = \text{Sup}(\text{double}, x_1^T, x_2^T), y^T = \text{bool}$
$y = \text{LogOp}(x_1, \dots, x_k)$	$x_1^T = x_2^T = \dots = x_k^T = y^T = \text{bool}$
$y = \text{Switch}(x_1, x_2, x_3)$	$x_1^T = x_2^T = x_3^T = y^T = \text{Sup}(x_1^T, x_2^T, x_3^T)$

Table 2. Type Inference Equations [16]

7. CLOCK INFERENCE

7.1. Time in SIMULINK

SIMULINK has two different timing mechanisms, namely samples and triggers.

7.1.1. Sample Time

The discrete time SIMULINK signals are piecewise-constant continuous-time signals. Blocks in SIMULINK can be assigned sample times, as configuration parameters. A sample time equal to 2, means that the block should be evaluated every two ticks of the global simulation clock. The sample time corresponds then to the period π of the block output signals. Some blocks (e.g, Pulse Generator) can also be characterized by initial phase θ , which is propagated to the neighboring blocks. Hence, in general, every block is characterized by a period π and a phase θ . It is evaluated every $k\pi + \theta$ ($k=0, \dots, n$). By default, blocks have their sample time set to -1, which corresponds to an inherited (from the inputs or the parent subsystem) value. We assume that the configuration option, *Automatically handle rate transition* for deterministic data transfer, is chosen. SIMULINK has some timing rules, if violated, the model is rejected:

- The inputs of a simple block B must have sample times that are multiplier or divisor of the block sample time: $(\pi_{\text{Inp}1\dots n} = k \pi_B)$ OR $\pi_{\text{Inp}1\dots n} = 1/k * \pi_B$, with $k = 0, \dots, n$
- Enabled and Triggered Subsystems' inputs should have the same sample times: $\pi_{\text{Inp}1} = \dots = \pi_{\text{Inp}n}$.

7.1.2. Triggered Subsystem

The second timing mechanism of SIMULINK is the *triggers*. Only subsystems can be triggered by a signal Trig. The triggered subsystem is evaluated if Trig has a *rising* or *falling* transition. The sample time of the blocks inside a triggered subsystem are all equal to the period T of the trigger signal. The example below shows the execution of a triggered subsystem (rising trigger). We assume that the triggered subsystem has no sub-blocks inside.

Trigger:	0	1	1	1	0	1	0	1	1
Input:	0	1	2	3	4	5	6	7	8
Output:		1				5		7	

7.1.3. Enabled Subsystem

The timing mechanism of the enabled subsystem is ambiguous [16]. It cannot be understood from a set of experiments. For the sake of the translation, we assume that the enabled subsystems have the same timing mechanisms as the triggered ones. The only difference lies in evaluating the block, if the *Enable* signal is equal to 1. The example below shows the execution flow of an enabled subsystem. We assume that the subsystem contains no sub-blocks.

Enable:	0	1	1	1	0	1	0	1	1
Input:	0	1	2	3	4	5	6	7	8
Output:		1	2	3		5		7	8

7.2. Time in Signal

SIGNAL has a partially ordered logical time. This means that the duration is abstracted to a point, namely the logical instant, and the time instants are partially ordered. Similar to the synchronous languages, SIGNAL also assumes the synchrony hypothesis. However, SIGNAL does not have a global clock, as a reference for sampling all the signals at each tick. Each SIGNAL flow x is characterized by a boolean flow b_x , called the clock of x . If x is present at instant i , $b_x(i)$ is equal to true, otherwise it is equal to false. The signal clocks can be independent until the end of the program. In case of synchronization requirements, extra timing constraints can be added. Epoch analysis is performed, in order to determine whether a sequential program can be synthesized from the SIGNAL specifications. In other words, it determines whether a *Master Trigger* can be found. If not, exogenous constraints are required from the user to form a *Master Trigger* [13]. We refer the reader to [4] for more detailed discussion on SIGNAL timing model.

7.3. Clock Inference

The blocks inside a triggered or enabled subsystem must have a sampling period and phase equal to the ones of the enclosing Triggered/Enabled subsystem. Otherwise, we consider two cases. In a first case, the sample time of a given block b_i is defined ($\text{Periodes}[i] \neq -1$). If it is a multiplier or divisor of the input signals' periods, it is kept. If, however, the sample time of the block is undefined ($\text{Periodes}[i] == -1$), it is inferred as the greatest common divisor of the input signals' periods [16] (See Formulas 5, 6 and 7).

$$(\pi_B, \theta_B) = GCD_{rule}((\pi_i, \theta_i)_{i=1..n}) \quad (5)$$

$$\pi_B = \begin{cases} \gcd(\pi_1, \dots, \pi_n) & \text{if } \theta_1 = \dots = \theta_n \\ \gcd(\pi_1, \dots, \pi_n, \theta_1, \dots, \theta_n) & \text{otherwise} \end{cases} \quad (6)$$

$$\theta_B = \begin{cases} \theta_1 \bmod \pi & \text{if } \theta_1 = \dots = \theta_n \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

For Example:

$GCD_{rule}((12,4), (12,0)) = (4,0)$, $GCD_{rule}((12,4), (4,4)) = (4,4)$,

$GCD_{rule}((12,4), (12,3)) = (1,0)$, $GCD_{rule}((4,0), (3,0)) = (1,0)$.

8. Translation

8.1. Type Translation

Once the type Inference step is completed, the obtained SIMULINK types are mapped to their corresponding SIGNAL ones, as it is shown in Table 3.

SIMULINK Type	SIGNAL Type
bool	Boolean
int8, uint8, int16, uint16, int32 or uint32	Integer
⊥, double, single	Real

Table 3. Type Translation

8.2. Clock Translation

Once all the sample times are inferred, we use these information to reproduce the same traces of the SIMULINK model simulation. The blocks inside a triggered or enabled subsystems are assigned the periods and phases of the trigger/enable signal. For the rest of the blocks, according to the ratio between the periods of the every two directly connected blocks A and B, we distinguish three cases:

- Case 1 : $\alpha = \text{Periode}_A / \text{Periode}_B = 1$

The output signal clock of A is equal to the output signal clock of B:

$$\text{output}_A^{\wedge} = \text{input}_B^{\wedge} = \text{output}_B$$

- Case 2 : $\alpha = \text{Periode}_A / \text{Periode}_B > 1$

Undersampling should be performed in B. The clock relation between the input and output signals of B is: $\text{output}_B^{\wedge} = \alpha \cdot \text{input}_B^{\wedge} + \beta$. β here is the phase difference between output_B^{\wedge} and input_B^{\wedge} : $\beta = \text{Phase}_B - \text{Phase}_A$. In order to implement the discussed affine clock relation, we consider a counter variable cnt, that has the same clock as the input signal outputA. Hence, starting from the initial phase, cnt is incremented every time a new input is read. When cnt reaches a multiplier of α , the function f performed by the block is evaluated and the output is produced. The following SIGNAL code illustrates the above explained algorithm for a Unit Delay block:

```
| cnt := (cnt + 1) $ init (PHASE_B - PHASE_A);
| cnt2 := cnt modulo (PERIODE_A / PERIODE_B);
| cnt^ = input_B;
| tmp := input_B $ init 1;
| output_B := tmp when (cnt2=0);
```

The flow of the output signal outputB for the Unit Delay block, with $\alpha=2$ and $\beta=0$ and initial value v_0 , is:

Output _A	v_1	v_2	v_3	v_4	v_5	v_6
Cnt	0	1	2	3	4	5
Cnt2	0	1	0	1	0	1
tmp	v_0	v_1	v_2	v_3	v_4	v_5
Output _B	v_0	.	v_2	.	v_4	.

- Case 3 : $1/\alpha = \text{Periode}_A / \text{Periode}_B < 1$

Oversampling should be performed in B. Similar to the Undersampling case, the clock relation between the input and output signals of B is: $\text{output}_B^{\wedge} = \alpha \cdot \text{input}_B^{\wedge} + \beta$. β here is the phase difference between output_B^{\wedge} and input_B^{\wedge} :

$\beta = \text{Phase}_B - \text{Phase}_A$. In order to implement the discussed affine clock relation, we consider a counter variable cnt, that has the same clock as the output signal outputB. Hence, starting from the initial phase, cnt is incremented. When cnt reaches a multiplier of α , a new input is read, the function f performed by the block is evaluated and a new output is produced. Otherwise, the old output is emitted. The following SIGNAL code illustrates the oversampling algorithm in case of a Unit Delay block:

```
| cnt := (cnt+1) $ init (PHASE_B - PHASE_A);
| cnt2 := cnt modulo (PERIODE_B/PERIODE_A);
| cnt^ = output_B;
| input_B^ = when (cnt2=0);
| tmp := input $ 1 init 1;
| output_B := tmp cell^ outputB;
```

The flow of the output signal Output_B for the Unit Delay block, with $\alpha = 2$ and $\beta = 0$ is shown below:

Output _A	v_1	.	v_2	.	v_3	.	v_4	.	v_5
Cnt	0	1	2	3	4	5	6	7	8
Cnt2	0	1	0	1	0	1	0	1	0
tmp	v_0	.	v_1	.	v_2	.	v_3	.	v_4
Output _B	v_0	v_0	v_1	v_1	v_2	v_2	v_3	v_3	v_4

8.3. Basic SIMULINK Blocks translation

In this section, we illustrate how the main basic blocks are translated. The remaining blocks are translated similarly.

- Sum Block: performs addition or subtraction on its inputs. The operation of the block is specified by the list of signs parameters ((+) and (-)), indicating the operations to be performed on the inputs:


```
| out := inp1 + inp2 -inp3
```

- Gain Block: performs a multiplication of the input with a constant:
| $out := inp * GAIN$
- Logical Operator Block: performs a boolean operation $Op \in \{AND, OR, NOT\}$ on its inputs:
| $out := inp_1 Op inp_2$
- Unit Delay and Integer Delay Block: The output of the Unit Delay and Integer Delay blocks is a delayed version of the input by NB DELAY instants. NB DELAY is equal to 1, in case of a Unit Delay Block. INIT VALUE is the initial value of the output:
| $out := inp \$ NB DELAY init INIT VALUE$
- Data Type Conversion Block: is translated into a Type Casting operation. The following code translates a real input into an integer output:
| $out := (integer) inp$
- Zero-Order Hold Block: If the sample time of the Zero-Order Hold Block is set to -1, it is equivalent to the identity function:
| $out := inp$. Otherwise, the clock translation, as explained in Section 8.2 is performed.
- Constant Block: The Constant Block value is added in SIGNAL to the list of the parent process parameters.
- Saturation Block: truncates its inputs according to an upper limit (LIM_{UP}) and a lower limit (LIM_{LOW}) bounds given by the user:
| $out := (LIM_{UP} \text{ when } (inp > LIM_{UP})) \text{ default } (LIM_{LOW} \text{ when } (inp < LIM_{LOW})) \text{ default } inp$
- Switch Block: The Switch Block has three inputs. It compares its middle input inp_2 to a threshold value. If it is greater than the THRES, the first input is passed to the output, otherwise the third input is emitted as an output:
| $out := (inp_1 \text{ when } (inp_2 > THRES)) \text{ default } (inp_3 \text{ when } (inp_2 < THRES))$
- Pulse Generator: The Pulse Generator with a period= 4, a phase= 2, an amplitude= AMP and a pulse width=2 is translated into the following SIGNAL code:

```

process PulseGenerator = { real AMP }
(? ! real out;)
(| dpg1:= dpg2 $1 init AMP
 | dpg2 := dpg3 $1 init AMP
   | dpg3 := dpg4 $1 init real(0)
 | dpg4 := dpg5 $1 init real(0)
 | pha1 := dpg1 $1 init real(0)
 | pha2 := pha1 $1 init real(0)
 | out := pha2 |)

```

where
real dpg1, dpg2, dpg3, dpg4, pha1, pha2;
end;
- Discrete Filter and Discrete Transfer Function: The Discrete Filter/Discrete Transfer Function's parameters are the nominator coefficients number values COEFFN, the denominator coefficients values COEFFD and the initial state value INIT VAL. The transfer function $1/(1+0.5z^{-1})$ is translated into the following SIGNAL code:

```

| output := (input * COEFFN[0] + tmp0)/COEFFD[0]
| tmp0 := (- COEFFD[1] * output)$1 init INIT_VAL

```
- Mux: The Mux block combines its inputs into a single vector output. It is generally used to merge the output of different blocks. The SIGNAL code for a Multiplexer with three inputs $inp1$, $inp2$ and $inp3$ is:
| $out := [inp1, inp2, inp3]$
- Combinatorial Logic: implements a truth table. It reads a boolean number, and outputs the row in the boolean table corresponding to the read input.

```

Process CombinatorialLogic =
{ integer N, M, K; [N] boolean TruthTab; }
(? [M] boolean inp; ! [K] boolean out );
(| array i to (M-1) of
  (| Row:= Row[?]+ ((1 when inp[i])
   default (0 when not inp[i])) |)
 with
  (| Row:=0 |)
 end
 | index0:= Row*K-K..Row*K
 | output := TruthTab[index0] |)

```

- FromWorkspace/ToWorkspace: are translated respectively into an input and an output of the SIGNAL process.
- Trigger: takes a real/integer flow and transforms it into a boolean flow. We distinguish between *Rising Trigger*, *Falling Trigger* or *Either*. The *Rising Trigger* gets the value true when an input transition from a negative number to a positive one happens. The *Falling Trigger* is true when an input transition from a positive to a negative value occurs. The *Either trigger* is true, if either a rising or a falling transition happens. The following example, illustrates the trigger mechanism:

Input	-1	0	1	2	-2	3	1	4	-1
Rising Trigger	f	t	f	f	f	t	f	f	f
Falling Trigger	f	f	f	f	t	f	f	f	t
Either Trigger	f	t	f	f	t	t	f	f	t

The following SIGNAL code generates a Rising Trigger flow. The not_before variable ensures that the trigger is only produced, if no one happened in the previous time step. The Falling Trigger is defined similarly:

```

| RiseTriggerold := Trig $ init false
| Trig := neg_to_nonneg OR (nonpos_to_pos and not_before)
| neg_to_nonneg := ((inp_old < 0) AND (inp >= 0))
| nonpos to pos := ((inp_old <= 0) AND (inp > 0))
| not_before := NOT (RiseTrigger_old)
    
```

- Enable: Similar to the Trigger, the Enable block transforms a real/integer input flow to a boolean flow en. en has the value true, when the input inp is positive:
 | en:=(true when (inp > 0)) default (false when (inp <= 0)).

8.4. Subsystems translation

As the model increases in size, its complexity can be reduced by grouping the functionality related blocks together into subsystems. A subsystem can be executed conditionally or unconditionally. A conditionally executed subsystem may or may not execute depending on a control signal. We distinguish between triggered and enabled subsystems.

8.4.1. Plain Subsystems Translation

A SIMULINK diagram can be constructed in SIGNAL by recursively translating subsystems into processes and the enclosed atomic blocks into subprocesses. The first step in the top down translation is the blocks parameter list generation, then the outputs and inputs definition. After that, the subprocesses in the first hierarchy level are called. Additional equations for the block connections are generated. In the Where part, the local variables are defined. The subprocesses body is also implemented. For every block, type and clock translation are performed as described in Section 8.1 and Section 8.2.

```

Process P = {integer N;}
            (? integer inp; boolean b; ! integer out;)
            | tmp := Q{N}(inp)
            | out := tmp when b|
Where
Integer tmp ;
Process Q = {integer M; }
            (? integer s1; ! integer s2;)
            | s2 := s1 * M |);
    
```

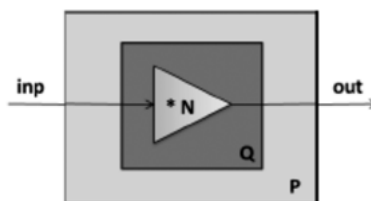


Figure 4. Plain Subsystem Translation

8.4.2. Triggered Subsystems Translation

The Triggered Subsystem is a subsystem with a control input, namely the *trigger* input. The subsystem is executed, each time a trigger event occurs. If no trigger happens, the output is either reset or it holds its old value. Figure 5 shows a Triggered Subsystem enclosing a Unit Delay block.

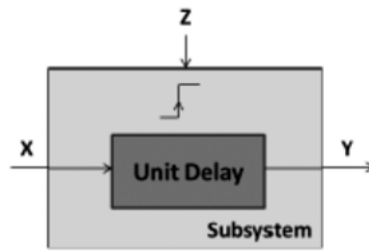


Figure 5. Triggered Subsystems Translation

Below is the corresponding SIGNAL code. The parent process *Sim2Sig* calls the subprocesses *Trigger* and *SubSys*. *Trigger* generates the rising trigger *SubSys_{trig}* from the in-put *Y*. *SubSys* is the Triggered Subsystem. It is only executed when *SubSys_{trig}* is true, otherwise it emits its old output, each time a new input with no trigger event arrives.

```

Process Sim2Sig = {integer NB_D, INIT_V; }
  (? integer X, Y; ! integer Z;)
  (| SubSystrig := Trigger(Y)
  | Z := SubSys {NB_D, INIT_V} (X when SubSystrig)cell ^ X | )
Where
  boolean SubSystrig;
  Process Trigger = (As defined in Section 8.3)
  Process SubSys = { integer NB_D, INIT_V; }
    (? integer in; ! integer out;)
    (| UnitDelayout := UnitDelay {NB_D, INIT_V} (UnitDelayin)
    | UnitDelayin := in
    | out := UnitDelayout |)
  Where
    integer UnitDelayin, UnitDelayout;
    Process UnitDelay = (As defined in Section 8.3)
  End ;
    
```

8.4.3. Enabled Subsystems Translation

The same translation method discussed for the case of a triggered subsystem, applies for the enabled one. The only difference is replacing the Trigger block with an Enable one.

9. Implementation: the tool SIM2SIG

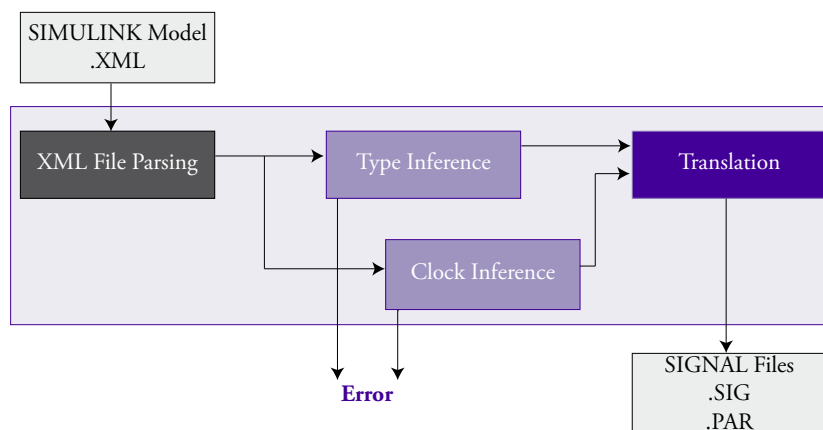


Figure 6. Translation Framework

Our prototype tool SIM2SIG is written in C++. Its architecture is shown in Figure 6. It reads an XML file generated by SIMULINK, parses it and builds a data structure representing the original SIMULINK model. After that, type inference (Section 6), clock inference (Section 7) and the translation (Section 8) steps are performed. The tool outputs a SIG file containing the SIGNAL program and a PAR file with the parameters, as well as error messages.

10. Case study: discretized dc-motor closed loop controller

In this section, we use our tool to translate into SIGNAL a SIMULINK model consisting of a system consisting of an input sampler and a discretized DC-motor in a loop with a PID controller. The electrical and mechanical dynamics of the three-level DC-motor are represented by the following equations:

$$V_{in} - R \cdot i - K_c \cdot \theta[n + 1] = L \cdot i[n + 1] \tag{8}$$

$$K_t \cdot i - b \cdot \theta[n + 1] = J \cdot \theta[n + 2] \tag{9}$$

R stands for the resistance, b for the damping factor L^{-1} is the inductance, J^{-1} is the inertia, i is the current and θ is the angular frequency. Equation 8 is implemented in the subsystem S1. Equation 9 is implemented in the subsystem S2.

The system inputs are sampled by the *Input sampling* block. The *Adder* Block in the PID subsystem down-samples the voltage with a factor of two. The *Rate Adjustment* block has a sampling rate equal to 1. Hence, the DC-Motor output is over-sampled with a factor of two. This case study shows then how the hierarchical translation is performed, as well as how different SIMULINK Timing mechanisms are translated into SIGNAL. The flow equivalence is validated by comparing the traces generated from the SIMULINK model and its corresponding SIGNAL translation.

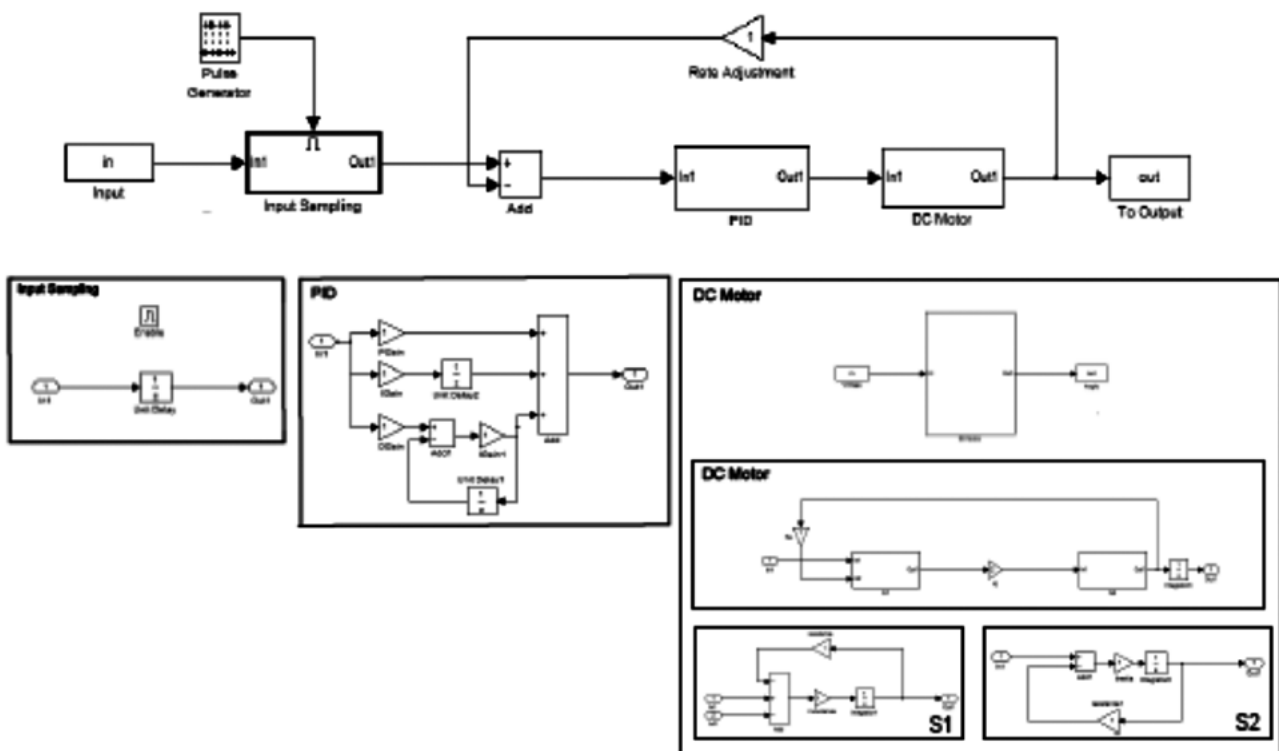


Figure 7. Case Study. Discretized DC-Motor Closed Loop Controller

11. Conclusion

In this paper, we developed a tool that translates a discrete time subset of SIMULINK into the polychronous formal language SIGNAL. The motivation behind this work, lies in the lack of formal semantics of the most popular embedded software design tool, SIMULINK. Hence, the correctness of the generated models can not be completely and efficiently verified. On the other hand, formal languages, are less popular as they are harder to learn. They have, however, clear and precise semantics, that allow the application of powerful design methodologies.

The choice of SIGNAL as a target language for this work, is justified by its multirate nature that allows for signal streams to be computed asynchronously, which fits very easily to a multi-threaded environment. The novelty in this work consists of bridging the gap between the synchronous and polychronous models of computation, through constructing affine clock relations between every block's inputs and outputs. This allows the generation of flow equivalent SIGNAL programs from the SIMULINK models. The translation follows three major steps: type inference, clock inference and hierarchical topdown translation. The SIGNAL program is generated by recursively translating the SIMULINK blocks. Subsystems are translated into SIGNAL processes and their enclosing blocks are translated into subprocesses. Our tool is tested on a discretized DC-Motor controller. Apart from SIGNAL code generation, our tool can be used for checking typing and timing rules of the SIMULINK models. Models that are rejected in SIMULINK are also rejected by our tool. The main drawback of this tool is its dependency on SIMULINK semantics, which keeps changing from one version to another. Besides, this tool is still incomplete, as it does not translate all the SIMULINK blocks. In fact, the behavior of many blocks is ambiguous, despite of the multiple experiments performed to understand it (ex. Enabled SubSystem). In the future, this work can be extended in different ways. One research direction would be to translate STATEFLOW to SIGNAL, since, SIMULINK and STATEFLOW are complementary tools, that are used together in many applications. Another interesting direction, would be to compare the concurrency of the SIGNAL produced C code, with the one generated from LUSTRE and the one provided by the SIMULINK code generator. This would prove the advantage of choosing SIGNAL as a target language instead of other synchronous languages. Formally proving the flow equivalence between SIMULINK and SIGNAL is also within the scope of our future work. Finally, the scalability of the tool can be further tested by applying the translation tool to more complicated SIMULINK models from the industry. In this case, the fault coverage obtained from using SIGNAL verification tools over the SIMULINK ones can be compared.

12. References

- [1] <http://www.irisa.fr/>.
- [2] <http://www.irisa.fr/espresso/polychrony/index.php>.
- [3] <http://www.irisa.fr/vertecs/logiciels/sigali.html>.
- [4] <http://www.isr.umd.edu/sites/default/files/sandeepshukla.pdf>.
- [5] Simulink: User's Guide. The Mathworks.
- [6] Aditya Agrawal, Gyula Simon, and Gabor Karsai. *Semantic translation of simulink/stateflow models to hybrid automata using graph transformations*. Electronic Notes in Theoretical Computer Science, 109:43-56, 2004.
- [7] Julien Ouy Bijoy A. Jose, Abdoulaye Gamatie and Sandeep K. Shukla. Smt based false causal loop detection during code synthesis from polychronous specifications.
- [8] Olivier Bouissou and Alexandre Chapoutot. *An operational semantics for simulink's simulation engine*. ACM SIGPLAN Notices, 47(5):129-138, 2012.
- [9] Marius Dorel Bozga, Vassiliki Sfyrila, and Joseph Sifakis. *Modeling synchronous systems in bip*. In Proceedings of theseventh ACM international conference on Embedded software, pages 77-86. ACM, 2009.
- [10] Abdoulaye Gamati. *Designing Embedded Systems with the SIGNAL Programming Language: Synchronous, Reactive Specification*. Springer Publishing Company, Incorporated, 1st edition, 2009.
- [11] N. Halbwachs, P. Caspi, P. Raymond, and D. Pilaud. *The synchronous data flow programming language lustre*. Proceedings of the IEEE, 79(9):1305-1320, 1991.
- [12] Thomas A Henzinger. *The theory of hybrid automata*. Springer, 2000.
- [13] Bijoy A Jose and Sandeep K Shukla. *An alternative polychronous model and synthesis methodology for model-driven embedded software*. In Proceedings of the 2010 Asia and South Pacific Design Automation Conference, pages 13-18. IEEE Press, 2010.
- [14] Julien Ouy Mahesh Nanjundappa, Matthew Kracht and Sandeep K. Shukla. *A new multi-threaded code synthesis methodology and tool for correct-by-construction synthesis from polychronous specifications*. 2013.
- [15] Vassiliki Sfyrila, Georgios Tsiligiannis, Iris Safaka, Marius Bozga, and Joseph Sifakis. *Compositional translation of simulink models into synchronous bip*. In Industrial Embedded Systems (SIES), 2010 International Symposium on, pages 217-220. IEEE, 2010.
- [16] Stavros Tripakis, Christos Sofronis, Paul Caspi, and Adrian Curic. *Translating discrete-time simulink to lustre*. ACM Transactions on Embedded Computing Systems (TECS), 4(4):779-818, 2005.

Fiche équipe

Laboratoire d'Analyse des Systèmes, Traitement de l'Information et Management Intégré (LASTIMI)

Naoual Berbiche

École Supérieure de Technologie de Salé, Université Mohammed V de Rabat, Maroc
nberbiche@hotmail.com

Jamila El Alami

École Supérieure de Technologie de Salé, Université Mohammed V de Rabat, Maroc
alamijamila1@gmail.com

Résumé

Le Laboratoire d'Analyse des Systèmes, Traitement de l'Information et Management Intégré (LASTIMI) s'est fixé pour objectif la recherche dans différentes disciplines en vue de la compréhension globale des systèmes complexes par l'analyse, le traitement et l'évaluation, ce qui permet d'en déceler les défauts et les limites. Grâce à la détermination et la quantification des critères de performances et d'optimisation, on cherche à améliorer la qualité de ces systèmes. On retrouve les systèmes complexes dans différents domaines tels que l'industrie, l'entreprise, les réseaux et télécommunications, systèmes d'information, etc.

Abstract

Systems Analysis, Information Processing and Integrated Management Laboratory (LASTIMI) has set a goal of research in different disciplines to the global understanding of complex systems through analysis, processing and evaluation, which allows to detect faults and limitations. Thanks to the determination and the quantification of performance and optimization criteria, we try to improve the quality of these systems. We find the complex systems in various fields such as industry, business, networking and telecommunications, information systems, etc.

Mots-clés

Robustesse, traitement de l'information, diagnostic des défauts, optimisation des critères de performance, MDA, sécurité, apprentissage par les jeux, Web sémantique.

Keywords

robustness, information processing, fault diagnosis, optimization of performance criteria, MDA, safety, learning games, semantic Web.

1. Présentation

Le Laboratoire d'Analyse des Systèmes, Traitement de l'Information et Management Intégré (LASTIMI) est une structure de recherche de l'École Supérieure de Technologie de Salé, appartenant à l'Université Mohammed V de Rabat.

Il a été créé en avril 2005, suite à la structuration de la recherche lancée par l'université Mohammed V-Agdal. Depuis mars 2008, il est accrédité par l'université, et dès avril 2009, il est membre du CEDOC «**Sciences et Techniques pour l'Ingénieur**» de l'École Mohammadia des ingénieurs.

Le laboratoire est constitué de 21 enseignants-chercheurs permanents, un enseignant chercheur associé, 3 membres administratifs et environ 70 doctorants.

Le LASTIMI se compose de trois équipes de recherche pluridisciplinaire:

- Automatique et analyse des défauts
- Management intégré
- Mathématiques Appliquées et Traitement de l'Information.

Les différents champs disciplinaires du laboratoire sont décomposés en thèmes de recherche alliant les aspects théoriques, les aspects appliqués et le transfert de technologie. En effet, le laboratoire a pour objectif de développer des recherches de grande qualité dans les domaines de l'Automatique, du Traitement de l'Information, des Mathématiques appliquées et du Management industriel. Ce faisant, le laboratoire aspire à ce que les résultats de ses recherches soient diffusés, reconnus et utilisés par la communauté scientifique et au-delà, par les utilisateurs finaux dans le milieu socioprofessionnel. Par ailleurs, d'autres missions non moins importantes sont menées par le laboratoire à travers l'action de formation universitaire et d'expertise.

En outre, le laboratoire s'est consacré, depuis sa création, à l'organisation d'événements scientifiques menés par ses équipes dans le but de s'inscrire positivement dans le paysage de recherches scientifiques national et international. Le LASTIMI a accueilli ses premiers doctorants lors de l'année universitaire 2009/2010. Grâce aux efforts de ces doctorants et de leurs enseignants, il a pu entreprendre une activité non négligeable de publications et de communications dans des revues indexées, de renommée internationale, dans des conférences et colloques nationaux ou internationaux avec comité de lecture. Par ailleurs, ses membres participent à de nombreux comités de programme de manifestations scientifiques nationales.

Dans le cadre du partenariat, le LASTIMI a une convention de coopération avec le Centre d'excellence en Logistique (CELOG) de L'École Supérieure des Industries du Textile et de l'Habillement (ESITH).

2. Thématiques de recherche

Automatique et analyse des défauts

Modélisation et stabilisation, Robustesse, Commande optimale, Commande adaptative, Commande des systèmes à retard, Contrôle des systèmes électro-énergétiques et des systèmes à sources d'énergie renouvelable, Traitement de signal, Traitement d'image

Management intégré

Approche systémique pour la quantification des critères de performance et prise de décision, Approches classique et méta-heuristique pour l'Optimisation des chaînes logistiques, Développement de méthodes optimisées pour le management industriel.

Mathématiques appliquées et traitement de l'information

Mathématiques appliquées

Optimisation et théorie des systèmes, Optimisation des critères de performance, Jeux différentiel, Analyse Numérique, Problème d'Inégalité Variationnelle et Complémentarité Non-Linéaire, Problème d'Inégalité Variationnelle mixte générale

Traitement de l'information

Web sémantique et Ontologies, eLearning, Transformation de modèles, Reverse engineering, Interopérabilité des systèmes, Bases de données et Indexation, Système d'information décisionnel, L'apprentissage automatique, Sécurité des systèmes d'information, Traitement de l'information, Analyse des données, L'utilisation des technologies d'information et de communication dans l'enseignement et l'éducation : cas de l'intégration des jeux dans l'apprentissage, Sécurité des réseaux informatiques (filaire et sans fil), Sécurité des applications web, Télé-enseignement, Traitement Automatique des langues Naturelles.

Parallèlement aux thématiques principales du laboratoire, certains de ses membres s'activent aussi dans l'axe de la Pédagogie universitaire notamment dans les thèmes suivants:

Amélioration de la qualité dans l'enseignement supérieur, Pratiques pédagogiques dans l'enseignement supérieur et Ingénierie de la formation et qualité du système.

La diversité des champs disciplinaires du laboratoire constitue sa force.

4. Productions scientifiques

La production scientifique du laboratoire LASTIMI est constitué de :

- Environ 110 publications dans des journaux indexés, internationaux et nationaux pour la période allant de 2006 à Mars 2015
- Plus de 28 communications dans des congrès, conférence et colloques pour la période 2010-2014
- Trois thèses de doctorat ont été soutenues.
-
-

5. Manifestations scientifiques

Le LASTIMI a organisé depuis sa création plusieurs manifestations scientifiques, à savoir:

- Séminaire sur «la Sécurité des Systèmes d'Information» le 6 juin 2006
- Formation sur l'Intelligence Économique dans le Management des Risques de l'Entreprise et la Qualité de l'Informatisation : les 17 et 18 Novembre 2007
- Colloque National « Grands chantiers et développement socio-économique au Maroc : le 19 février 2009 en collaboration avec le Forum des économistes Marocains «FEM»
- Organisation de deux journées doctorales sur l'analyse des systèmes traitement de l'information et management intégré - décembre 2013 et juin 2014
- Organisation de deux ateliers de formation pour les doctorants sur L'Art de la Rédaction Scientifique les 28 mars 2015 et 18 avril 2015
- Animation de la recherche interne à la structure.

Le laboratoire prévoit d'organiser le 25 février 2016 la journée scientifique d'analyse des systèmes et traitement de l'information – troisième édition- JASTI'3 (<http://www.lastimi.ma/jasti/>).

J'ai lu

Ingénierie des exigences. Méthodes et bonnes pratiques pour construire et maintenir un référentiel

Stéphane Badreau & Jean-Louis Boulanger

Ounsa Roudies

Univ Mohammed V-Rabat. Ecole Mohammadia d'Ingénieurs. BP 765 Rabat Agdal. Maroc

roudies@emi.ac.ma

Résumé

Cet ouvrage présente les différentes facettes de l'ingénierie des exigences, allant des techniques d'élucidation à la certification des acteurs, en passant par les outils d'assistance.

Abstract

This book enlightens the various facets of requirements engineering such as elicitation technics, engineer's certification or support tools.

Mots-clés

Ingénierie des exigences, SysML, Systèmes complexes, certification IREB, Association Specief.

Keywords

Requirements engineering, SysML, complex systems, IREB certification.

Cet ouvrage publié en 2014 chez Dunod est un guide pratique pour gérer les exigences d'une organisation. Orienté ingénierie, c'est une synthèse concise et très bien documentée d'un domaine très dynamique depuis la fin des années 90. Il aborde aussi bien les activités liées à l'ingénierie des exigences que les standards apparentés et la certification des acteurs. L'ouvrage s'appuie sur l'expérience professionnelle des auteurs et leurs participations aux associations dédiées telles que l'IREB (International Requirements Engineering Board).

L'ouvrage identifie et présente les techniques à mettre en œuvre à chaque étape-clé du cycle de développement des exigences (Elucider, Analyser, Spécifier et Valider), mais aussi du processus de gestion des exigences. Il s'appuie pour cela principalement sur des langages standard de modélisation tels que UML et SysML.

J'ai bien apprécié le chapitre 10 consacré aux standards et normes dédiés. Ce chapitre positionne également l'ingénierie des exigences dans le cadre de référentiels génériques tels que UP ou Togaf.

La clarté de l'ouvrage le rend accessible à toutes les parties intéressées, même s'il me semble particulièrement pertinent pour l'ingénieur des exigences, vu son niveau de détails. Signalons en conclusion qu'un tel ouvrage reste unique dans l'édition francophone alors que de nombreux ouvrages rédigés en anglais et en allemand ont été consacrés à ce sujet.

Appel aux articles

Appel aux articles pour la 9^e édition

en accès libre : www.revue-eti.net

La Revue électronique en Technologies de l'Information sollicite aussi bien les universitaires que les industriels pour présenter leurs résultats de recherche, leurs réflexions et les tendances dans les thématiques liées aux Technologies de l'Information (TI). Son but est de favoriser les échanges des connaissances en TI, entre pays du Nord et du Sud. La revue e-TI est en libre accès. Elle est référencée par plusieurs moteurs de recherches tels que DOAJ, EBSCO, IMIST, revue.org, ResearchGate, et Scholar Google.

Cet appel aux articles est ouvert, ce qui signifie que vos articles soumis seront évalués en continu pour la 9^e édition, et selon le résultat, ils sont insérés dans le site Web de la revue, au fur et à mesure de leur acceptation. Le délai d'évaluation est au plus de deux mois. A la clôture de la 8^e édition, l'ensemble des articles parus sera rassemblé en un volume, en version papier ainsi qu'en version électronique, téléchargeable à partir du site de la revue.

Vos soumissions peuvent s'inscrire dans les rubriques suivantes :

- **Etat de l'art** propose un état de l'art d'un thème dans le domaine des TI ;
- **Recherche** rassemble des articles portant sur la théorie, la conception, la spécification ou l'implémentation d'outils liés aux TI. Les articles traitant d'axes émergents sont particulièrement les bienvenus ;
- **R&D** met en évidence des expériences de recherches et de développement ainsi que leur validation ;
- **Fiche équipe** présente une équipe de recherche afin d'encourager la coopération ;
- **J'ai lu** introduit et critique un ouvrage récent.

La liste suivante suggère (mais ne limite pas) les thèmes de e-TI :

- Systèmes d'information
- Approche à base de composants
- Méthodologie de conception, Ingénierie dirigée par les modèles
- Bases de données, Big Data
- Sémantique, ontologie
- Intelligence artificielle, data mining, base de connaissances, SIAD
- Technologies Internet et Web, cloud computing, web mining
- Systèmes répartis, interopérabilité et intégration, SOA
- Mobilité, pervasivité, MDM
- Qualité, sécurité et aspects non fonctionnels
- Utilisabilité, accessibilité, personnalisation, environnement collaboratif
- Applications e-gov, e-business, SIG.

Soumission

Les auteurs sont invités à soumettre des articles en **français** ou en **anglais** dans les rubriques citées : 22 pages maximum pour les articles **Etat de l'art**, 20 pages maximum pour les articles de **Recherche**, 10 pages maximum pour les **Expériences R&D**, 2 pages maximum pour les **Fiches des équipes** et la rubrique **J'ai lu**. Ces articles doivent respecter les recommandations aux auteurs. Le *template* des articles ainsi que les recommandations sont téléchargeables sur le site de la revue.

La soumission est à effectuer via le site EasyChair eTI8 <<https://www.easychair.org/account/signin.cgi?key=18571941.PGsvQyBmIeXPCLRu>> ou par voie électronique à l'adresse eti@revue-eti.net. La sélection sera réalisée suite à une double lecture anonyme.

Directeurs de publication

Ounsa Roudies (EMI) et Dalila Chiadmi (EMI).

Rédacteurs en chef de la 9^e édition

Ounsa Roudies (EMI) et Mounia Fredj (ENSIAS).