

PROPOSITION DE SUJET DE THESE

Campagne 2018

Laboratoire L3i



Sujet de la thèse :

« Modélisation des processus utilisateurs à partir des traces d'exécution, application aux systèmes d'information faiblement structurés. »

Contexte de l'étude:

Le système d'information (SI) représente l'ensemble de ressources et d'outils permettant d'organiser et de donner accès à l'information d'une organisation. Il est vu comme un système socio-technique incluant à la fois la structure organisationnelle, les processus métier ainsi que les données afférentes (Piccoli, 2007). Les processus métier représentent les activités ou les enchaînements d'activités relatives aux affaires de l'organisation concernée.

Ces processus et activités consomment et produisent des données gérées par le SI. Un processus peut correspondre à un assemblage de traitements réalisés par un opérateur humain ou par appel de services extérieurs, tels que les services web, consommés durant le processus. De plus, dans certains contextes, les processus utilisent des données collectées par des sources variées (équipements, humains, capteurs...) et depuis différents supports.

Le travail de recherche proposé dans cette thèse portera essentiellement sur les SI qui ne sont pas structurés autour de processus et dont le modèle des processus n'est pas (complètement) défini. Ces systèmes répondent à deux types de besoins :

- laisser la responsabilité du processus à une entité externe au SI (telle que l'on peut le trouver dans les architectures orientées services);
- permettre une construction agile d'un système d'information complexe (collaboration entre objets connectés).

Tout ou partie des processus impliqués dans la vie d'un SI peuvent être déduits depuis les différentes traces recueillies par l'observation et l'enregistrement des différents événements survenant dans le SI accompagnant ses activités. Ces traces d'exécution, variées mais structurées, peuvent être exploitées pour améliorer le SI (Aalst et al., 2010).

La fouille de processus (Process Mining ou PM) est une discipline de recherche (Aalst et al., 2010) récente qui se situe entre l'intelligence artificielle (BI) et la fouille de données, d'une part, et la modélisation et l'analyse des processus métier d'autre part. L'objectif de la fouille de processus est de découvrir, superviser et améliorer les processus réels (c'est à dire, les processus non définis a priori) par l'extraction de connaissances à partir des traces d'exécution (event logs) facilement disponibles dans les systèmes d'information courants.

Dans le cas de processus faiblement structurés, les techniques traditionnelles de fouille de processus, visant à extraire ou compléter les modèles de processus à partir d'événements de traces d'exécution, aboutissent généralement à des modèles complexes « Butterfly », qui offrent souvent une marge limitée pour l'analyse. Aussi les processus découverts sont rarement exécutés dans le même ordre et sont souvent incomplets, ce qui ne permet pas une modélisation satisfaisante du comportement des acteurs impliqués dans le processus.

Mots clés :

Fouille de processus (process mining, process discovery), modélisation de comportement de processus, (Fouille de comportement : behavioral process mining), systèmes d'information, systèmes faiblement structurés, processus métier (business process).

Résumé du travail proposé :

La fouille de processus est un domaine de recherche assez récent qui s'avère très fertile pour le développement de travaux de modélisation (de découverte et de conformité) orientée événements de traces d'exécution dans les systèmes d'information.

Cette étude s'inscrit dans le domaine de fouille de processus inter-organisationnels où de nouvelles techniques d'analyse doivent être développées. Ces techniques devraient aussi envisager la vie privée et les questions de sécurité.

L'étudiant devra proposer une approche permettant une analyse du comportement des acteurs impliqués dans le processus modélisé et par conséquent, une solution d'intégration de cette dimension dans la modélisation et la correction du processus initial.

Ainsi les verrous à lever dans cette thèse s'articulent autour des thèmes suivants :

1. Identification de processus faiblement structurés à partir des traces d'exécution ;
2. Détection des sous-processus « pertinents » (amenant à des processus concluants : achat, signature, service,...),
3. Création d'un algorithme de modification automatique des modèles initialement créés afin d'adapter le SI à la réalité des utilisateurs.

Prérequis et contraintes particulières :

Master en informatique : des connaissances approfondies en fouille de données, systèmes d'information, processus métiers seraient très appréciées

Financement de la thèse :

Une bourse de l'Université de La Rochelle sera attribuée au candidat retenu.

Références bibliographiques :

(Aalst et al., 2010) W. van der Aalst et al. Process Mining Manifesto, Business Process Management Workshops, LNBIP n° 99, Springer Berlin Heidelberg, 2012, pp. 169-194 (<http://www.win.tue.nl/ieeetfpm/lib/exe/fetch.php?media=shared:pmm-french-v2.pdf>)

(Diamantini et al., 2016) Behavioral process mining for unstructured processes
Claudia Diamantini, Laura Genga... in Journal of Intelligent Information Systems (2016)

(Shalyaeva et al., 2016) 'Events Analysis Based on Internet Information Retrieval and Process Mining Tools'. Conference: 2016 IEEE 10th International - Conference on Application of Information and Communication Technologies (AICT). (October 2016) - Baku, Azerbaijan.

Piccoli G. (2007). Information systems for managers: Texts and cases (1st éd.). Wiley Pub.

Contacts

Email(s) :

ronan.champagnat@univ-lr.fr

jacques.morcos@univ-lr.fr