



PROPOSITION DE STAGE

Année 2018



Laboratoire L3i

Sujet de stage :

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

Résumé du travail proposé :

Un processus peut correspondre à un enchaînement 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. Tout ou partie, des processus impliqués dans la vie d'un système d'information (SI), peuvent être déduits depuis les différentes traces recueillies dans le SI. Les travaux du stagiaire se focaliseront sur des SI « faiblement structurés », comme ceux des systèmes de e-commerce, pour lesquels des processus principaux d'interaction ont été définis mais que les utilisateurs utilisent et abandonnent ou suivent d'autres processus pour atteindre leur objectif personnel. Dans de tels SI, les techniques de fouille de processus traditionnelles, visant à extraire ou compléter les modèles de processus à partir des traces d'exécution, aboutissent généralement à des modèles complexes « spaghetti », qui offrent souvent une marge limitée pour l'analyse. L'objectif du stage est d'explorer l'état de l'art dans le domaine de l'analyse de comportement à partir d'événements de traces d'exécution dans les systèmes faiblement structurés et notamment les techniques « sequence mining » et « behavioral mining ». Le stagiaire effectuera une comparaison des techniques existantes et en expérimentera l'apport sur des données issues de cas de navigations réelles sur des sites de tourisme ou de e-commerce.

Mots clés :

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

CPER NUMERIC:

- **Thème : e-Education/e-Formation**
- **Actions concernées : POLARIS**

Informations complémentaires :

Encadrant(s) : Jacques Morcos, Mourad Rabah, Ronan Champagnat

Équipe :

- Images et Contenus
- Dynamique des systèmes et adaptativité
- Modèle et Connaissance

Domaine d'application stratégique :

E-éducation

Environnement et développement durable

E-culture

Valorisation de contenus numériques

Cadre de coopération : Les travaux proposés dans le cadre de ce stage constituent une phase préliminaire à un travail de recherche proposé pour une thèse de Doctorat qui débutera en septembre 2018. Ils pourront déboucher sur des applications dans différents domaines : e-éducation pour la personnalisation des parcours des apprenants ; découverte des processus menant à des actes d'achats dans des sites web ; sécurité, découverte de processus anormaux.

Ces applications pourront faire l'objet de projets de partenariat industriels et institutionnels.

Date de début du stage : mars 2018

Durée du stage : 6 mois

Financement : CPER

LOCALISATION DU STAGIAIRE (dans quel bureau) : L3i, bureau 001

BESOIN MATERIEL (indiquer qui fournit ce(s) matériel(s)) : une machine

Contexte de l'étude:

La fouille de processus (*Process Mining* ou PM) est une discipline de recherche [1] récente qui se trouve 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 événements des traces d'exécution (*event logs*) facilement disponibles dans les systèmes d'information courants.

Nous concentrons nos travaux sur des SI « faiblement structurés », comme par exemple des systèmes de e-commerce, pour lesquels des processus principaux d'interaction avec les visiteurs ont été définis et que les utilisateurs utilisent et abandonnent ou suivent d'autres processus pour développer un travail en coopération.

Dans les 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 « spaghetti », 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.

Description du sujet :

Dans ce contexte le stage porte sur plusieurs aspects :

- études des algorithmes de fouille de processus (process mining)
- études des algorithmes de sequence mining
- études des algorithmes de behavioral mining
- sélection, test et illustration du fonctionnement des algorithmes les plus pertinents sur des jeux de données issues de traces de navigations sur des sites de e-commerce

Prérequis et contraintes particulières :

Formation générale en informatique de niveau M1 avec des connaissances en génie logiciel, en UML, Modélisation de Processus (réseaux de Petri ou BPMN), en formalisation mathématique et pratique des outils et méthodes de développement classiques : IDE, Java, GIT. Des notions et en fouille de données seraient un plus.

Le stage pourrait déboucher sur une thèse de Doctorat dans la continuité des travaux menés.

Références bibliographiques :

- 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>)
- W.M.P. van der Aalst. Business Process Simulation Survival Guide. In J. vom Brocke and M. Rosemann, editors, Handbook on Business Process Management 1, International Handbooks on Information Systems, pages 337-370. Springer-Verlag, Berlin, 2015.
- Behavioral process mining for unstructured processes
Claudia Diamantini, Laura Genga... in Journal of Intelligent Information Systems (2016)

Contacts – liens :

Email : {jacques.morcos.ronan.champagnat, mourad.rabah}@univ-lr.fr

Lien vers le fichier de description : (PDF) (si nécessaire)