Ph.D. Thesis Proposal

February 27, 2020

Title: Context-enhanced Process Mining

Advisors: Joyce El Haddad and Daniela Grigori

Contact: joyce.elhaddad@dauphine.fr , daniela.grigori@dauphine.fr

Job location: Université Paris-Dauphine (LAMSADE, PSL), France.

Description:
Process mining is a recent research topic that applies artificial intelligence and data mining techniques to process modelling and analysis [1]. The main idea is to extract knowledge from events recorded in an *events log* in order to discover, monitor and improve processes. Event logs stores *activities* related to process instances as well as additional information such as the *resources* executing the activities, *timestamps*, or *costs*.
Process mining approaches allow the discovery of the process model or its variants (a.k.a. *discovery*), the detection of deviations between the real process and the designed model (a.k.a. *conformance checking*), and the improvement of the process model based on the observed events (a.k.a. *enhancement*). However, most of the existing approaches consider the process to be in steady state and so do not consider the *context* in which the process takes place nor the changes that may affect it while being analyzed [2]. Information about the context could be derived from the process log (resource occupation rate...) or captured from other sources of information that could enrich the log. Dealing with context information is important to detect and analyze changes and is one of the challenges for research described in the Process Mining Manifesto [3].

Objectives:
The aim of this thesis is to consider the context in all the phases of the process improvement life cycle: process discovery, conformance checking, process enhancement. Including the context could improve the precision of the discovered process model and of the analysis enabling better recommendation for process improvement or predictive process monitoring.
It will allow also to address fairness issues (e.g., not blame an overloaded resource for delays) and conduct causality analysis (e.g., which factor or context variable causes delays).
More precisely, the goals of the thesis are:

- Define relevant context attributes and how to model, store and retrieve them
- Propose context-driven process discovery techniques
- Propose context-based conformance checking techniques to check specific constraints
- Use context attributes to propose meaningful improvements
- Study what context attributes to monitor and how to identify when these attributes change
- Propose approaches to detect changes online
- Propose predictive approaches with online learning to make sure that the process model is to up to date

Required skills:
We seek for excellent and highly motivated student with a background in Computer Science. In particular, we need a student with good knowledge of business process management or petri nets, process mining and data mining. Having good programming skills and some notion of machine learning would be an asset.

To apply:
Please send the following material before March 30th, 2020:
- fully detailed CV,
- academic records (Master degree or equivalent),
- cover letter,
- recommendation(s) and supporting letter(s).

References
[1] Van Der Aalst, W. (2016). Data science in action. In Process mining. Springer, Berlin, Heidelberg.

[2] Van Der Aalst, W. M., and Dustdar, S. (2012). Process mining put into context. IEEE Internet Computing, 16(1), 82-86.

[3] Wil M. P. van der Aalst et al., Process Mining Manifesto. Business Process Management Workshops (1) 2011: 169-194

[4] Bose, R. J. C., van der Aalst, W. M., Žliobaitė, I., & Pechenizkiy, M. (2011). Handling concept drift in process mining. In International Conference on Advanced Information Systems Engineering (pp. 391-405). Springer, Berlin, Heidelberg.

[5] Bose, R. J. C., Van Der Aalst, W. M., Žliobaitė, I., & Pechenizkiy, M. (2013). Dealing with concept drifts in process mining. IEEE transactions on neural networks and learning systems, 25(1), 154-171.

[6] Maisenbacher, M., & Weidlich, M. (2017). Handling concept drift in predictive process monitoring. In 2017 IEEE International Conference on Services Computing (SCC) (pp. 1-8). IEEE.