DTU Library

Online Conformance Checking to Support Human Behavior Study

Di Federico, Gemma

Published in:
CEUR Workshop Proceedings

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

Online Conformance Checking to Support Human Behavior Study (Extended Abstract)

Gemma Di Federico
Department of Applied Mathematics and Computer Science
Technical University of Denmark
Kgs. Lyngby, Denmark
gdfe@dtu.dk

***Abstract*—Human behavior could be represented in form of a process model. Internet of Things systems allow the collection of the behavior in the reality, and the data are then processed by Process Mining algorithms. To achieve this integration, several challenges must be faced. The objective of this PhD project is to adapt online conformance checking to be able to deal with human-related processes.**

I. Problem Context and Motivation

The project aims to contribute into the integration between Internet of Things (IoT) and Business Process Management, combining the potentiality of sensor data toward the implementation of dynamic and enriched models with the objective to support conformance checking techniques. In particular, the project focuses on the modeling of human behavior [8] by means of process mining techniques, obtaining data from IoT devices. There is a desire to dwell on aspects relating to the use of conformance checking techniques in the context of behavioral study. Observing human behavior for a long time, can lead to the collection of various information that can be used in a plenty of application fields [10]. The use case of my PhD project is the dementia disease, under the healthcare field. Patients effected by dementia usually fit into a very tight routine. The repetition of routines helps people to control their world in a predictable manner, adding a sense of order to the days. Repetition of activities is a comfort zone. Establishing a routine contributes to a sense of security and peace, which helps to reduce agitation and troublesome behavior [6]. In the middle to later stages of most types of dementia, a person may start to behave differently. Given that, the breaking of a routine could be symptom of a worsening of the disease [12]. The idea is to observe the behavior of people living with dementia to identify significant variations in the behavior. To do so, an habits model is derived for each patient and it is continuously compared with the actual behavior to find discrepancies. IoT systems are used to capture the behavior in the reality, while online conformance checking techniques to compare the actual behavior with the modeled (and expected) one.

The work in this PhD project also aims to be generalized, so that it can be reused in other contexts. My plan is to extend Process Mining techniques to be applicable in IoT, with the aim of contribute into verifying the compliance of the actual IoT system execution with the expected process schema. In order to succeed in such goal, in the following section I elaborate the research questions.

II. Research Questions

The overall process could be divided into three main topics which are the data collection, the process modeling and the mining. Firstly, there is a simulated environment, equipped with IoT sensors, that captures the behavior of a patient in form of sensors measurements. Secondly, data collected must be organized in an event log that is suitable for discovery algorithms which are used to derive a reference model. This model represents the typical behavior of the patient. The last is the conformance, in which the behavior in the reality is compared with the reference models. The derivation and replication of processes is what process mining is intended to do. Complications emerge as the nature of human processes differs from the structured one that are typically dealt with.

> **RQ1 -** What is the most appropriate way to abstract activities from a stream of sensor measurements? Are semi-supervised approaches the proper manner to organize sensor data into an event log?

Data produced by IoT systems are unstructured and of a different level of granularity compared to the event log expected by Process Mining algorithms. There is a need to bring sensors measurements and model to the same level of abstraction. Actually, there is not a standard technique to organize sensors data into a structured event log. In the literature there are several approaches of event abstraction[15], both supervised or unsupervised. The unsupervised techniques group fine-granular events on the basis of strong re-occurrence of patterns [9]. In the supervised event abstraction techniques, authors propose a set of dimensions based on their understanding of the field and classifying them accordingly [7]. After a first investigation of the available techniques, we would like to focus in the between, i.e., on semi-supervised approaches. A fully supervised abstraction could be influenced by human errors, since it is an expert who must identify every single activity starting from a series of proposed configurations. On the other hand, an unsupervised technique does not allow to verify the way in which the labeling was carried out, or how the sequence of actions for each activity was identified. Therefore, a semi-supervised approach could provide both the right level of automation and allow us to oversee and contribute with knowledge to the activity recognition. Once the data has been structured in an event log, we need to focus on the process modeling language that is able to represent human-related processes.

> **RQ2 -** Are current modeling languages able to represent human behavior and their context? It may be valid to consider stochastic models?

Human related process differs from business one. Human behavior evolves over the time [5] and it is influenced by the participant and by the context. The process is completely guided by the participants, who decide how and when to perform activities [3]. The concept means that there are a series of variables, external or internal to the process execution, which influence the process. These variables could also change over the time [14]. For example, the factors that influence the behavior could be the seasonality, the weather, the day of the week, but also a dependency between executed activities. Based on each of them, the behavior adapts [13]. Often the cause of change is hidden, not known a priori, making the learning task more complicated. Hidden contexts may be expected to recur, generating cycling phenomena. In this domain, is a key element to keep track of these changes in order to use them to adapt models. Another important fact to consider is the non-determinism of human behaviors. Human behavior is usually characterized by the repetition of activities [1] and, each time an activity is executed it follows a different execution pattern. In addition, human beings could perform multiple actions at the same time, mixing the execution of activities, i.e. concurrency. Since the objective of the use case is to identify discrepancies in the behaviors, also the frequency concept must be taken into account [11]. Human behavior is governed by uncertainty and variability, therefore it is not possible to establish ex ante which is a frequent or a not frequent behavior. As a starting point I identified a list of requirements that the model must fulfill [4], hence now the objective is to position the work toward one direction. Once the modeling language has been identified (or extended) I can focus on the process conformance.

> **RQ3 -** Are the actual mining algorithms able to deal with unstructured scenario? Which one is the most appropriate? Is there the need of extending these algorithms?

In conformance checking, the behavior of a process model and the behavior recorded in an event log are compared to find commonalities and discrepancies. This model is the reference model, on which all the replay work is based, so it is very important that it is accurate and representative of the behavior recorded in the log. The discovery of a process related to human behavior, unstructured by nature, is an aspect not to be underestimated. Considering the above, there is a close dependence between mining algorithms and modeling languages. Most of the current process mining algorithms, assume the derivation of Petri nets and consequently conformance checking techniques have been developed along this line. But there is no certainty that a Petri net has enough expressiveness to represent these processes. For this reason, the conformance algorithm must be able to deal with the process modeling language selected, perhaps, also in this case, the algorithm should be adapted to respond to the process requirements. To conclude, given the large amount of data produced by IoT systems, it is desirable that the conformance is carried out online [2, 16].

References

[1] N. Banovic et al. “Modeling and understanding human routine behavior”. In: *Proc. of CHI*. 2016, pp. 248–260.

[2] A. Burattin and J. Carmona. “A framework for online conformance checking”. In: *BPM*. Springer. 2017, pp. 165–177.

[3] C. Di Ciccio, A. Marrella, and A. Russo. “Knowledge-intensive processes: characteristics, requirements and analysis of contemporary approaches”. In: *J. Data Semant.* 4.1 (2015).

[4] G. Di Federico, A. Burattin, and M. Montali. “Human Behavior as a Process Model: Which Language to Use?” In: *ITBPM* (2021).

[5] M. S. Feldman and B. T. Pentland. “Reconceptualizing organizational routines as a source of flexibility and change”. In: *Adm. Sci. Q.* 48.1 (2003), pp. 94–118.

[6] H. C. Kales, L. N. Gitlin, and C. G. Lyketsos. “Assessment and management of behavioral and psychological symptoms of dementia”. In: *BMJ* (2015).

[7] M. de Leoni and S. Dündar. “Event-log abstraction using batch session identification and clustering”. In: *Proc. SAC20*. 2020, pp. 36–44.

[8] F. Leotta, M. Mecella, and J. Mendling. “Applying process mining to smart spaces: Perspectives and research challenges”. In: *Lect. Notes Bus. Inf. Process.* 2015.

[9] F. Leotta, M. Mecella, and D. Sora. “Visual process maps: A visualization tool for discovering habits in smart homes”. In: *J. Ambient Intell. Humaniz.* (2019), pp. 1–29.

[10] S. Mandal et al. “A classification framework for IoT scenarios”. In: *BPM*. Springer. 2018, pp. 458–469.

[11] A. Rebuge and D. R. Ferreira. “Business process analysis in healthcare environments: A methodology based on process mining”. In: *Inf. Syst.* (2012).

[12] K. SantaCruz and Jr. Daniel L. Swagerty. *Early Diagnosis of Dementia*. Tech. rep. 2001.

[13] A. Stefanini et al. “A process mining methodology for modeling unstructured processes”. In: *Knowl. Process. Manag.* 27.4 (2020), pp. 294–310.

[14] A. Tsymbal. *The problem of concept drift: definitions and related work*. Tech. rep. 2004.

[15] S. J. van Zelst et al. “Event abstraction in process mining: literature review and taxonomy”. In: *Granul.* (2020).

[16] S. J. van Zelst et al. “Online conformance checking: relating event streams to process models using prefix-alignments”. In: *Int. J. Data Sci. Anal.* (2019).