Process Analysis in Issue Tracking Environment

Iryna Oleksandrivna Shvydyuk – 70352
Supervisor: Professor Miguel Leitão Bingolas Mira da Silva
Co-Supervisor: Professor Diogo Manuel Ribeiro Ferreira
Departamento de Engenharia Informática
Instituto Superior Técnico

Abstract. Issue tracking systems are well-known for the recording, tracking, solving and archiving of issues. Due to its nature, they serve as a link between the stakeholders in the supported process. Traditionally, the researchers have been focusing on the system’s operational level, i.e. on the understanding of how they are being used in practice. Thus, the information intrinsic in the process was forgotten. This paper proposes the uncovering of the stored processes in the issue tracking environment with process mining techniques to address the lack of awareness of the stored information from process perspective. Design Science Research Methodology was used to conduct the present work. In terms of proposed solution, three applications were performed on one of the most used issue trackers and these were evaluated according to a holistic view of criteria and standard methods. It was composed by logical reasoning and analysis and unstructured interviews.

Keywords: issue tracking systems, process mining, agile, methodology, process analysis, awareness

1 Introduction

Among a wide range of tools, ITS are the ones that are well-known for its organizational ability. These systems were proved to be a main repository for recording, tracking, solving and archiving of issues [1]–[3]. Moreover, it also serves as a link between the stakeholders involved in the supported process [2]. Traditionally, an issue was a bug, a feature or an inquiry. The designation varies according to the person who reported it.

Throughout the time, the project tracking functionality started to be supported by ITS and the system has evolved. In the software projects, they are referred as bug tracking systems and in other context as project management tools. Consequently, the concept of an issue started to mean a main task.

The research community performed several studies on ITS, however, their main focus was on its operational level, i.e. on understanding of how these systems were used in practices [3]–[5]. Moreover, the reasoning was made through a qualitative study that, despite allowing a better understanding of the problem, are time consuming.

Process mining discipline has appeared recently and offers the opportunity to understand actual behavior of business processes through event log analysis, i.e. diagnose its weaknesses based on omnipresent information produced by the
information systems [6]. Moreover, it allows to improve the existing processes. For the last few years, this area is becoming more and more popular among organizations. To address the problem of lack of awareness of the stored information from process perspective, the proposal relies on exploring the process mining techniques and uncover the stored processes in the ITS environment.

The present research was conducted in compliance with DSRM’s specifications and incorporated the application of the proposed solution in three contexts of the one of the most used ITS. The evaluation process was based on the holistic view of criteria and standard methods. The results of this research were analyzed according to previously selected criteria, such as efficacy, validity and accuracy.

The presented document is structured as follows: section 2 develops the theoretical background related to the present work, section 3 provides a description of the proposal, section 4 aims to demonstrate the applicability of the proposed solution, section 5 comprises the evaluation of the results of each application and, finally, section 6 concludes the present report with a summary of the performed work and the overall findings.

2. Related work

2.1 Process mining

Process mining is a recent discipline that aims to extract knowledge from event logs recorded by an information system in order to discover, monitor and improve business processes [6].

Usually an event log contains information about events that occurred in an information system and mention a process instance, also referred as case, and an activity, also known as task. To apply process mining, one assumes that beyond process instance and activity, it is possible to record the person that is performing an activity (performer) and all events have a timestamp associated to it and are ordered.

There are three types of process mining, such as discovery, conformance checking and enhancement. Discovery stands for a construction of a model based on the information contained in an event log and does not requires a-priori model. Conformance checking allows to verify the accordance between the predefined model and the discovered one. Finally, the enhancement aims to enrich the existing model with data from event log.

Additionally, process mining can be applied to analyze three different perspectives, such as process perspective (how?), the case perspective (what?) and organizational perspective (who?) [7].

Process mining applicability: real-life cases

Process mining’s application has been conducted in several real-life cases incorporated in different areas [8]–[11].
Aast et. al. [8], aimed to show the applicability of process mining techniques. Their goal was to improve the main process responsible for efficiency loss in an organization. The process perspective was conducted in an iterative way over which several meetings with organization people were made to validate the obtained process and focus on the core process. In case of an organizational perspective, three different categories of personnel which were not involved in the process as expected and the personnel’s passivity were discovered. By analyzing the case perspective, it was possible to discover another cause of efficiency problem related to the human component. Moreover, the results of previous perspective analysis were used to clearly understand its impact in the process (personnel’s passivity).

Process mining techniques were also successfully applied in the health-care area. In Mans et. al., the subject of interest was finding a compromise between high-quality services and low costs [9], [10]. Similarly to Aalst et. al., during the analysis phase various meetings with experts were made to validate the information obtained.

Process mining was also proved to be valuable when incorporated in a software development project based on Scrum. Rubin et. al. experienced blocker performance issues and problems with system usage [11]. It turned out to be an asset that allowed to understand the overall behavior (system and users).

As a final point, it became obvious that the applicability of process mining in real-life cases has taken a very successful path. The offered mining perspectives allow us to get inside the organizational processes and understand which occurs and is stored in information systems. Moreover, by streamlining their processes, the companies can assure its competitive advantage.

**Process mining methodology**

A methodology for conducting process mining analysis was proposed by Bozkaya et. al. one [12] (Figure 1). The authors stated that process mining can be a “repeatable service” and does not require any previous knowledge about the context of the problem, i.e. only the organization should analyze the outcome. As well, the methodology should provide whole information needed in “short time-period”. Following these objectives, a six steps methodology was defined in which the data is pre-processed and the event log is built; the familiarization with process through the statistic collections about the log are performed; the process is discovered, hence in the case of having the predefined one, the conformance checking is performed; the interactions inside the organization are understood and, finally, the results are transferred and validated. In terms of technology, the methodology relies on ProM’s functionalities.

![Figure 1 Methodology for Process Mining Analysis][12]
2.2 Agile

The business environment continues to change at a dramatically increasing pace and the question of handling these changes throughout the project’s life-cycle was stressed [13]. As a result, the agile approach appeared as a combination of already existing practices and relies on the four core values, such as individuals and interactions over processes and tools, working software over comprehensive documentation, customer collaboration over contract negotiation and responding to change over following the plan [14].

Agile does not provide a set of rigid practices but it requires certain management practices to be implemented. Thus, it suggests to perform frequent releases, dynamic prioritizations of the requirements and emphasizes the importance of constant feedback and the importance of interactions among people and the teamwork.

Scrum, the one of the most used methodologies, focuses on the way that software development team members should function. It starts with initial planning and functions with iterations, i.e. sprints. After each sprint, the results are presented and discussed with the business individuals. The feedback obtained is integrated in form of requirements in the product backlog. Moreover, the team self-evaluates its own performance after each sprint. With this iterative process Scrum intends to implement the agile philosophy in the organizations.

3 Research proposal

Process mining discipline is a relatively new area that offers the opportunity to understand the actual behavior of business processes through event log analysis. With the application of its techniques, the organizations obtain additional insights about its processes and improve the existing processes.

To make the analysis more systematic, a general methodology was provided as a repeatable service and focused on a specific perspectives and features. The aspects that were considered important in each application of process mining in business environment were not considered.

With these considerations in mind and, to solve the presented problem, we propose to uncover the processes stored in ITS by using process mining techniques with a development of a method. It will combine process mining analysis with the agile approach to impose continuous feedback and involve the customer in the analysis process.

2.2 Methodology

The proposed methodology is based on agile and Scrum development processes, heavily incorporating their iterative nature. It is composed by three major phases.

**Preparation phase.** The definition of the aspects of the process that the customer wants to explore. Based on the current knowledge, requirements are gathered and
added to product backlog. These also will be prioritized according to information available. Furthermore, the familiarization with the business process and technology is performed.

**Development phase.** It can be divided in five different components, such as:

- **Data extraction** - the data is extracted from ITS. The team defines a rapid and simple way of extracting the information.
- **Log familiarization** - familiarization with the information stored in the event log will be performed;
- **Control flow analysis** - obtaining the central process stored in the log by applying different discovery algorithms;
- **Performance analysis**;
- **Transferring the results** - when the sprint is over, the results are transferred and the product backlog is updated. The feedback provided by experts will serve as input for the next sprint.

**Transferring phase.** The conclusions about the overall process are drawn and the documentation is produced.

### 3 Demonstration

The following section corresponds to the demonstration step of DSRM. Three applications of the proposed solution took place on one of the most used issue trackers: JIRA. Since all projects in JIRA’s context are represented as collections of issues, they became the focus of the analysis.

#### Issue management

In order to provide a better understanding of the organizations processes, an overall analysis took place regarding four projects, namely Xray, Xray-Support, XPORTER and XPORTER-Support.

**Preparation phase.** The organization A was interested in exploring the issue’s behavior in different contexts and understand where the main loses were.

**Development phase.** Throughout the first week, the familiarization with the process and technology was performed. The four-week period that followed was dedicated to data extraction and its storage according to the format assumed by process mining.

The approach taken to perform the data extraction, was divided in three main components, such as obtaining the list of projects, extracting the issues per project and transforming the XML to CSV document. During the log inspection phase, the projects were treated in separate ways, thus the constructed logs presented different number of cases and events. The data gathered was from one year and half period. Moreover, through the inspection it became obvious that the projects of support were considerably smaller.
During the control flow and performance analysis, the discovery of a process and its performance was possible. However, during the transferring the results, the experts stated that there were some additional activities that did not belong to the process. Thus, a second three-week long iteration that aimed to filter those activities took place. The final process is represented in Figure 2. The workflow at the right represents its performance. Also, the observation about the project’s size was taken into account and the main focus was performed on the Xray and Xray-Support. In terms of roles, the presented results reported the centrality of the same performer in each project.

![Workflow Diagram](image)

**Figure 2** Issue Management Workflow (Xray project)

**Transferring phase.** During one week, the documentation was produced in order to explain the performed work. The software developed among with process mining analysis from each iteration were documented and transferred to the organization.

**Idea management**

Organization B explores JIRA’s ability of administration of projects in the idea management context. Here, the goal is to support employee’s creativity and stimulate ideas’ production.

**Preparation phase.** The organization was not familiar with the process mining concept. During the first meeting, an overview of process mining area and its techniques was performed in order to provide some insights and do not create false hopes or mislead the client. The management, experts and Scrum team were present. The focus was made on the business process regarding the idea and initiatives.

**Development phase.** The iterations were considerably smaller. The data extraction process was already implemented from the application in the issue tracking environment. Thus, the first iteration was one-week long and the second one took three weeks.

The data gathered was from two-years period and comprised many events, activities and performers. Thus, the variability of information was not questioned.

In terms of control flow analysis, the discovered process was very similar to the predefined one. Its representation is shown in Figure 3 and the figure on the left represents its performance.
The process presented some deviating behavior that was detected with the help of information gathered during the preparation phase. The results were presented and validated with the experts. Inspired, the experts and management were interested in exploring the relation between the ideas and the employees and in uncovering the most participative units of the organization. However, during the second iteration it was discovered that due to the confidentiality issues, the required analysis was not possible.

**Transferring phase.** The process mining analysis from each iteration was documented and transferred to the organization.

**Request management**

The third application of the proposed solution was performed on the same organization. Here, JIRA was used to assist the request management system, i.e. the request management workflow was supported. In terms of time-period, the first iteration took two weeks and the second one took one month.

**Preparation phase.** The individuals from this area were not familiar with the concept of process mining. Thus, similarly to the previous application of the methodology, an overview of the process mining was performed. The experts were interested in uncovering the process and compare it with the predefined one. Moreover, they provided an overview of the request management process.

**Development phase.** The data was extracted from a four-month period and included many events, activities and performers. When performing the analysis, four additional activities were found. When comparing with the information gathered during the first meeting, there was no evidence of such. Considering the expert’s availability, an in-person meeting took place. He/she was surprised to find those activities and stated that they belong to the tranche component that is intrinsic in a request. Thus, the decision was made to filter the data regarding to the additional activities found.

The obtained process from the second iteration is represented in Figure 4. In terms of control flow analysis, the compliance with the predefined process was verified. When analyzing the performance associated to the process (represented by the process...
from the left in Figure 4), the reduced transition time present at the end of the process raised additional questions.

When this observation was communicated during in-person meeting. The expert pointed out that the transition time associated to the final phase of the process was effectively too fast. Thus, a doubt arose about the influence of the filtered activities on the overall time.

![Figure 4 Request Management Workflow](image)

**Figure 4 Request Management Workflow**

**Transferring phase.** The process mining analysis from each iteration was documented and transferred to the organization. Additionally, it should be noted that the third application was made during the vacation period. Thus, its durability was longer than expected and did not allowed to conclude the process mining analysis.

5 Evaluation

The importance of the evaluation process has been recognized among the research community, however a predefined criteria and methods that could turn the assessment process more systematic were absent. A recent research presented the holistic view of criteria and standard methods [15].

The authors considered an artifact as a system and stated that the criteria are selected by the researchers. They should consider the nature of the developed work and choose the more suitable ones.

In this work, the results were assessed regarding the efficacy, validity and accuracy. Here, the overall evaluation of the three application is presented.

**Efficacy.** From the feedback obtained during the analysis process and the final meetings, it is believed to be achieved. The experts from idea and request managements expressed interest in continuing with the analysis in the future. In the case of request management, due to the time limits and the fact that it comprised the vacation period, there was no possibility to conclude the assessment of this criterion.

**Validity.** The proposed artifact was applied in the isolated environments. The manipulations that were used as input in each iteration were based on the experts’ feedback from process perspective.
Accuracy. The requires processes were discovered and analyzed according to the process, organizational and case perspectives. However, in the case of request management, the performance analysis was left open. The variants of the processes were discovered along with the performers associated to them.

Concerning the question of the customer involvement, with a frequent release it was possible to stimulate ideas for the process mining and solve the issues found.

6 Conclusion

This study aimed to uncover the processes stored in the ITS environment by using process mining techniques. To address this problem, a development of a method was proposed which was enriched with the notions of agile approach and incorporated the iterative nature of Scrum.

The proposed methodology was applied in three projects available in JIRA, one of the most used ITS. In each one of applications, the processes were discovered and in two of them, the compliance with the predefined one was verified. Moreover, the additional knowledge expressed by the insights about the processes was extracted. Each of the performed applications allowed to explore a different component of the methodology.

In the first case, the most interesting component was represented by data gathering phase. With the application of the methodology in the idea management environment, it was possible to explore the requirements variability and prove that involving the customer in the process is a great asset. Finally, the third case is an example of the impact of the external factors in the business environment.

Through the applications, the most valuable lessons learned were:

- The importance of the expert’s involvement from the beginning;
- The organizational interest in the further use of process mining;
- The impact that the frequent iterations have on the validation and, consequently, on the formulation of requirements;
- The bureaucracy associated to the business environment should be considered;
- Regardless to the degree of interest and utility, the terms of confidentiality establish the boundaries for the process mining analysis.

In conclusion, the techniques developed during this work offer a possibility to raise awareness about the stored information from the process perspective.

References


