Software Runtime Data and Development Data Integration to Monitor External Quality

Aytaj Aghabayli
Dietmar Pfahl
University of Tartu
Tartu, Estonia

Silverio Martínez-Fernández
UPC-BarcelonaTech
Barcelona, Spain

Adam Trendowicz
Fraunhofer IESE
Kaiserslautern, Germany

ABSTRACT

The last years the usage of software analytics in software development companies increases. However, there is little support for software development companies to obtain integrated insightful and actionable information at the right time. This research aims to study the integration of runtime and development data to analyze to what extent external quality is related to internal quality. We followed the CRISP-DM process in collecting and analyzing data of the real software product. We explored the integration possibilities between runtime and development data and implemented. The number of bugs found during the runtime has a weak positive correlation with code quality measures and a moderate negative correlation with the number of rule violations. Other correlations require more data cleaning and higher quality data for their exploration. During our study, several challenges to exploit data gathered both at runtime and during development were encountered. Lessons learned from the research may be useful for practitioners and researchers alike.

ACM Reference Format:
Aytaj Aghabayli, Dietmar Pfahl, Silverio Martínez-Fernández, and Adam Trendowicz. . Software Runtime Data and Development Data Integration to Monitor External Quality. In *Proceedings of* . ACM, New York, NY, USA, 1 page.

1 SOLUTION FRAMEWORK

This abstract is the shortened version of the submitted at the SQUADE workshop paper [2] and Masters thesis work [3]. This research is conducted to cover the gap identified during the literature review. The research question was constructed as the following "How can we integrate runtime data with development/repositories data (internal software quality) to improve software quality?". The main purpose of the study is to integrate runtime errors and crashes during the usage of the software system with development data from the software repositories and study their dependency.

By answering the RQ, we wanted to find out how useful would be to integrate software runtime data with development data in order to understand and predict external quality. Moreover, we tried to understand in what extend problems during the development of the software (internal quality) can cause problems occurring during the use of the software (external quality). It can positively influence the following three factors. First, to improve quality factors (e.g., stability, usability) based on runtime and development data correlation patterns. Second, to predict external quality from known internal quality. Third, since experts-based software quality models are usually costly [5], to develop data-driven software quality designs.

The study use case is the Fraunhofer IESE internal project, a platform providing several digital services. The chosen methodology was CRISP (Cross-Industry Standard Process), hence it is a widely used standard in data analysis problems. The process provides a well-defined structure for planning data-driven projects. For more details please check the full paper [2].

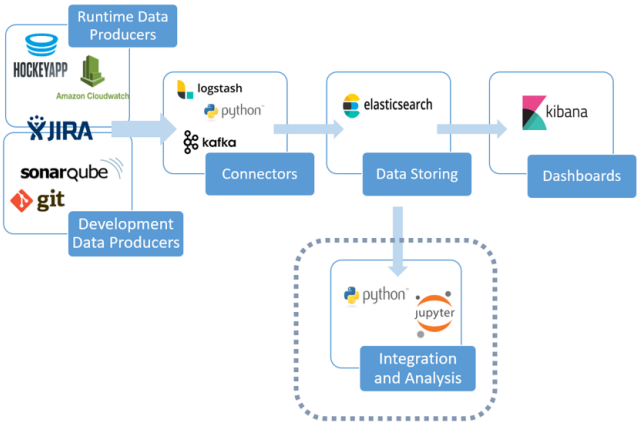

Figure 1: Solution framework (adapted from Q-Rapids).

The data collection part of the solutions is the extended the existing Q-Rapids solution [4] with two connectors to gather real-time data from crashes in HockeyApp and logs in Amazon CloudWatch. We used existing connectors to gather data from Jira, Git, and SonarQube [1]. The solution framework is depicted in Figure 1.

The process flow is the following. First, the automatically generated data are gathered in real-time and stored in elastic. Second, real-time dashboards aim to provide to stakeholders better understand problems in the software and prioritize tasks during the agile planning. Third, the data are integrated and analyzed.

REFERENCES

[1] [n. d.]. Q-Rapids source code. https://github.com/q-rapids/qrapids-connect
[2] Aytaj Aghabayli, Dietmar Pfahl, Silverio Martínez-Fernández, and Adam Trendowicz. 2019. Integrating runtime data with development data to monitor external quality: challenges from practice. 20–26. https://doi.org/10.1145/3340495.3342752
[3] Aytaj Aghabayli. 2019. *Software run time data: visualization and integration of development information - case study.* Master's thesis. University of Tartu. https://comserv.cs.ut.ee/ati{_}thesis/datasheet.php?id=66914{&}year=2019
[4] Lidia López, Silverio Martínez-Fernández, Cristina Gómez, Michał Choraś, Rafał Kozik, Liliana Guzmán, Anna Maria Vollmer, Xavier Franch, and Andreas Jedlitschka. 2018. Q-Rapids Tool Prototype: Supporting Decision-Makers in Managing Quality in Rapid Software Development. Springer, Cham, 200–208. https://doi.org/10.1007/978-3-319-92901-9_17
[5] Silverio Martinez-Fernandez, Andreas Jedlitschka, Liliana Guzman, and Anna Maria Vollmer. 2018. A Quality Model for Actionable Analytics in Rapid Software Development. *2018 44th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)* 732253 (2018), 370–377. https://doi.org/10.1109/SEAA.2018.00067