

ConTest

Continuous Software Testing: Solving the Oracle Problem in Production

Programm / Ausschreibung	Bridge, Brückenschlagprogramm, Ausschreibungen Bridge 1 (GB 2021)	Status	abgeschlossen
Projektstart	01.10.2021	Projektende	30.09.2024
Zeitraum	2021 - 2024	Projektlaufzeit	36 Monate
Keywords	Continuous Software Engineering; Risikobasiertes Softwaretesten; Testorakel; Testen im Betrieb		

Projektbeschreibung

Moderne Softwareanwendungen wie KI-Systeme, Cloud-Software oder Microservices werden auf der Basis kontinuierlicher Entwicklungsprozesse in kleinen Inkrementen von parallel arbeitenden Teams entwickelt. Im Gegensatz zu früheren Entwicklungsansätzen mit fix getakteten Iterationen wächst die Funktionalität eines Systems stetig durch einen kontinuierlichen Strom von kleinen Änderungen und Erweiterungen.

Die Qualitätssicherung und das Testen werden damit ebenfalls zu kontinuierlichen Prozessen und vor neue noch nicht gelöste Herausforderungen gestellt. Zum einen muss ein kontinuierlichen Durchfluss aller Änderungen und Erweiterungen durch den Test sichergestellt werden. Zum anderen ist jede einzelne Erweiterung dahingehend zu prüfen, ob diese wie vorgesehen funktioniert und keine negativen Seiteneffekte auf andere Teile des Systems hat. Bei einem stetig wachsenden System eine zunehmend unmögliche Aufgabe. Dazu kommt, dass bei modernen Systemen zum Testzeitpunkt häufig noch gar nicht absehbar ist, wie die Benutzer das System im Feld verwenden und mit welchen Daten die KI-Algorithmen konfrontiert sein werden.

Bei kontinuierlichen Entwicklungsprozessen gehen die Entwicklungs- und Betriebsphase zunehmend in einander über. Testen wird dabei zu einem laufenden Prozess, der nicht mit der Freigabe des Systems für die Benutzer abrupt endet. Im Gegenteil, die eigentlich wichtige Phase beginnt, wenn die Benutzer das System verwenden und dadurch Daten über die Nutzung und das Systemverhalten aufgezeichnet werden. Diese Daten dienen zugleich als eine Art Testfall. Um jedoch beurteilen zu können, ob dieser Testfall erfolgreich war oder fehlgeschlagen ist, müssen die aufgezeichneten Daten, interpretiert werden, was nicht durch den Vergleich mit einer Vorversion oder eine Spezifikation möglich ist.

Die Lösung erfordert Forschung zu einem fundamental neuen Testorakelansatz, der von robusten Distanzmessungen basierend auf statistischen Momenten Gebrauch macht und es ermöglicht, unterschiedliche Testläufe zu vergleichen und die graduelle, mehrdimensionale Veränderung der Systemqualität zuverlässig zu messen. Die Ergebnisse der mathematisch formalisierten, statistischen Messung müssen für den Tester im Kontext der konkreten Anwendung auch verstehbar und interpretierbar sein. Dies erfolgt durch die Entwicklung eines Bayes-Risikomodells, mit dem die statistischen Messergebnisse

auf allgemeine und anwendungsspezifische Qualitätsziele abgebildet werden können.

Im Projekt ConTest werden die zur Lösung dieser wissenschaftlichen Fragestellungen notwendigen Kompetenzen der Universität Innsbruck und des Software Competence Center Hagenberg gebündelt. Ziel ist es gemeinsam mit der Firma gepardec, die an einer österreichischen Cloud-Plattform arbeitet, einen neuen Testorakel- und Risikoanalyseansatz für kontinuierliches Testen im Produktionsumfeld zu entwickeln und zu evaluieren.

Abstract

Modern software applications such as AI systems, Cloud software, or Microservices are developed on the basis of continuous development processes in small increments by teams working in parallel. In contrast to earlier development approaches with fixed-cycle iterations, the functionality of a system grows steadily through a continuous stream of small changes and enhancements.

Quality assurance and testing thus also become continuous processes and face new challenges that have not yet been solved. On the one hand, a continuous flow of all changes and extensions through the test must be ensured. On the other hand, each individual extension must be tested to ensure that it functions as intended and does not have any negative side effects on other parts of the system. With a constantly growing system, this is an increasingly impossible task. In addition, with modern systems, it is often not even foreseeable at the time of testing how users will use the system in the field and what data the AI algorithms will be confronted with.

In continuous development processes, the development and operational phases increasingly merge. Testing thereby becomes an ongoing process that does not end abruptly when the system is released to users. On the contrary, the actually important phase begins when the users use the system, and data about usage and system behavior is recorded as a result. This data also serves as a kind of test case. However, in order to judge whether this test case succeeded or failed, the recorded data must be interpreted, which cannot be done by comparing it with a previous version or a specification.

The solution requires research on a fundamentally new test oracle approach that makes use of robust distance measurements based on statistical moments, allowing different test runs to be compared and the gradual, multidimensional change in system quality to be reliably measured. The results of the mathematically formalized statistical measurement must also be understandable and interpretable by the tester in the context of the specific application. This is done by developing a Bayesian risk model that can be used to map the statistical measurement results to general and application-specific quality objectives.

In the ConTest project, the competencies of the University of Innsbruck and the Software Competence Center Hagenberg necessary to solve these scientific questions are bundled. The aim is to develop and evaluate a new test oracle and risk analysis approach for continuous testing in production together with the company gepardec, which is working on an Austrian Cloud platform.

Projektkoordinator

- Universität Innsbruck

Projektpartner

- Geparddec IT Services GmbH
- Software Competence Center Hagenberg GmbH