Statistical Testing within the Scope of the System Validation of a RIS/PACS System

S. Holpp\textsuperscript{1}, H. Goetz\textsuperscript{2}, W. Dulz\textsuperscript{3}

\textsuperscript{1}IKM ES SV, Siemens Healthcare
Hartmannstr. 16, 91052 Erlangen
Stefan.Holpp.ext@siemens.com

\textsuperscript{2}IKM IQ SQ, Siemens Healthcare
Hartmannstr. 16, 91052 Erlangen
Helmut.H.Goetz@siemens.com

\textsuperscript{3}Department of Computer Science 7
Martensstr. 3, 91058 Erlangen
Dulz@informatik.uni-erlangen.de

Abstract: The division IKM of Siemens Healthcare develops and tests the software platform \textit{syngo} Suite, an integrated RIS/PACS solution for radiology departments. Within the scope of a diploma thesis, the practicality of the Statistical Testing approach was proven in comparison to present deterministic testing.

1 Introduction

The core objective of statistical testing is not to improve the systems quality by finding more software defects, but to estimate the achieved quality. Statistical testing is based on Markov chain usage models which represent all possible uses of a software system at some level of abstraction. From these models, test cases can be derived which represent a statistical sample of possible uses (see figure 1). Based on this sample, statistics like a reliability estimator can be obtained ([BD03]). A usage model describes the expected use of a software system in terms of a Markov chain by adaption of the transition probabilities. This provides the possibility to create customer profiles which reflect the expected system usage of specific customers or customer types, based on a common software system.
2 Generation of Markov Chain Usage Profiles

In the department of System Validation, UML activity diagrams are used to model use cases from which test cases are derived. To handle growing model complexity, these activity diagrams follow a hierarchical\(^1\) approach. One main objective was to find a feasible way to map those models to Markov chains. We have shown that it is possible to define a mapping between the hierarchical structure of present UML activity diagrams and a semantical equivalent (flat) Markov chain. This mapping has further been implemented so that the conversion can be performed automatically.

![Hierarchical UML](image1)

![Resulting Markov Chain](image2)

Figure 2: Generation of Customer Usage Profiles from hierarchical UML Diagrams

In order to define customer usage profiles, appropriate transition probabilities had to be obtained. In [JHPW00], the authors propose a constraint based approach to define customer usage behavior. This approach is based on statements like "I do not know the probability of \(p_1\), but I know that it is at least twice as big as \(p_2\)". and therefore provides an intuitive way to describe customer behavior.

These statements can be expressed by (in)equations. In this case the resulting inequation would be \(p_1 > 2 \cdot p_2\). A new mathematical approach has been developed [Hol08] to obtain a suitable probability distribution from a set of such constraints (see figure 3).

For the creation of customer profiles, the long years of customer experience within IKM has been utilized. Customer experts and application trainers have been interviewed and it turned out that the constraint based approach is a very suitable way to convert informal customer knowledge into mathematical constraint definition.

---

\(^1\) Actions within an activity diagram may reference other activities. This ends up in a tree-like composition of multiple activity diagrams.
3 Metrics

Markov chains provide the possibility to derive a multitude of metrics out of the created customer profiles. These metrics can be obtained analytically so that no further effort is necessary. For example, two customer profiles can be compared in order to obtain a quantitative measure of behavior deviation. Of course, this information may have a huge influence regarding test planning. Imagining two customers with strongly deviating system usage. Intuitively, both customers are likely to demand for different testing in order to assure high quality results.

Further, the expected fraction of times a user performs a specific action, observed over a long period of time, can be obtained. While the specific use case is continuously repeated (for example several days or weeks), the user will take different paths through the underlying model. The average of this set of paths can be taken to calculate the long run probability of being in each state while executing this use case. This information can be used to analyze the usage of specific features of the system. Highly used features then may ask for improved usability or performance to satisfy customer needs.

4 Test cases and Reliability

Statistical testing demands the choosing of test cases in a statistical correct manner. To achieve an unbiased set of test cases, sampling must be done randomly but based on the underlying usage profile. Consequently, paths with high probability are more likely to occur in a test set than low probable paths.

In comparison to present deterministic test case selection based on path prioritization or path coverage algorithms, the statistical approach provides the possibility of a customer-based quality assessment by performing a reliability estimation. Therefore, the Miller Reliability Model ([PP04]) has been chosen since it offers many advantages like the possibility to consider a-priori knowledge about the system. After a set of test cases has been executed, test cases, test results and a usage profile are taken to calculate the Miller reliability estimator. This can be also done during test execution so that the reliability progress can be observed over the whole testing period. Further, usage profiles can be exchanged in order to obtain customer specific reliability estimators.
From this, several possibilities to define a test end criterion emerge. For example, testing can be stopped if the reliability for a specific customer type or an average over available usage profiles has reached a certain stability level which can be verified by the reliability variance. Another possibility is to stop testing if the tested paths do correspond closely to a chosen usage profile. The amount of correspondence between tested paths and expected usage can thereby be obtained analytically.

Suggestions on how to apply the statistical testing approach over a full release cycle of the syngo Suite product are presented in [Hol08] as well as results based on present test data. In addition to that, a future pilot project is expected to verify the described approaches for their applicability in the domain of IKM System Validation.

5 Conclusions

The statistical testing approach is a promising extension to present deterministic testing. We have shown that the creation of Markov chains based on present UML models can be done with a minimum of manual effort. The constraint based approach represents a suitable technique to transform present customer knowledge into Markov chain usage profiles. From this profiles, a multitude of metrics can be obtained analytically to improve test planning and support management decisions. A possibility to achieve a customer based reliability estimator at every stage of a test execution phase has been shown as well as the definition of possible test stopping criteria. By being able to obtain quantifiable customer based statements during the complete test cycle, decision processes for quality assurance can be significantly supported.

References


