Metric Proposal for System Testing Models
Verification for Safety Critical Systems

Lukas Spendla, Pavol Tanuska, Lukas Smolarik
Faculty of Materials Science and Technology, Slovak University of Technology, Trnava, Slovakia
lukas.spendla@stuba.sk, pavol.tanuska@stuba.sk, lukas.smolarik@stuba.sk

Abstract—The main goal of this article is the verification of the system testing models for safety critical systems. The proposal is focused on selection of appropriate metric for verification of these models. Proposed metric is based on standard metrics for software systems and is defined according to ISO 9126 standard. Our proposal captures the process of metric definition and demonstrates the usability of the proposed metric on performance testing model.

I. INTRODUCTION

There are number of standards and guidelines for development of standard systems, in which important role plays also testing. From the most important we can mention IEC 12207, IEC TR 15271, IEC 56/575/CD, IEC 14598 and DIN 66272. These standards and guidelines however, deal only with testing of standard systems. Therefor they are not relevant for testing safety critical systems. But the actual testing of these systems is so important, that it is necessary to focus on this task.

Testing of safety critical systems is a very broad area, which cannot be covered as a whole. Therefor we have focused on the specifics of a particular area, namely on the communication subsystem and various issues related to communication in these systems. If we take a typical safety critical system, which instrumentation and control (I&C) system of nuclear power plant certainly is, the system testing model that we need to verify can be used in the area shown on Fig. 1.

From the previous figure it may be not clear, on which communication level is the target area situated. If we should define it more generally, for example on the hierarchy of the industrial control system on Fig. 2, it would be at the control level, including also parts of the production level.

This definition is important due to different requirements for data communication on each level.

A. Safety Critical System

Safety-critical systems are systems where failures can lead to catastrophic results: resulting in loss of life, significant property damage, or damage to the environment. These systems range from ground flight control systems, through medical devices, nuclear power plants control and automotive systems, to military systems. In past years other information systems are also becoming more and more “safety-critical” due to the financial impacts of failures and the fact that human lives depend on them [1].

The constant technology evolution is making these systems more complex and more common. Thus, we need to guarantee maximum dependability and safety properties with better processes and tools [1].

Different industries have various criteria for their information and process control systems. A single organization will usually run combinations of non-critical and mission-critical systems. The type of mission critical systems depends upon the functionality of the system and the importance of that functionality. In regulated industries and in government, these systems are also subject to legal and regulatory criteria [2].
B. System testing

System testing is defined as a testing phase conducted on the complete integrated system to evaluate the system compliance with its specified requirements. It is done after unit, component and integration testing phases [3], as shown on Fig. 3.

A system is a complete set of integrated components that together deliver product functionality and features. A system can also be defined as a set of hardware, software and other parts that together provide product features and solutions. In order to test the entire system, it is necessary to understand the product’s behavior as a whole. System testing helps in uncovering the defects that may not be directly attributable to a module or an interface [3].

System testing brings out issues that are fundamental to design, architecture and code of the whole product [3].

System testing is the only phase of testing which tests the both functional and non-functional aspects of the product. On the functional side, system testing focuses on real life customer usage of the product and solutions [3].

System testing simulates customer deployments. For a general-purpose product, system testing also means testing it for different business verticals and applicable domains such as insurance, banking and asset management [3].

C. Product Quality in Software Engineering

Quality is defined as “the degree to which a set of inherent characteristics fulfills requirement”. In general, description of software quality contains different characteristics which are defined by each model. Our proposal focuses on ISO 9126 model because it also provides the evaluation from user’s perspective [4].

ISO 9126 is a standard for evaluation of software product quality. This model contains two parts. The first part specifies six characteristics (functionality, reliability, usability, efficiency, maintainability and portability) for determination of internal and external software quality [4].

These six characteristics are further subdivided into subcharacteristics. Internal and external quality has same characteristics and subcharacteristics but they are measured differently. Internal quality is measured and evaluated during development process while external quality is measured and evaluated during testing process using simulated environment with simulated data [5].

Fig. 4 shows the relationships between a system and software. An information system consists of computer, communication hardware devices and software products. An External-System consists of information systems, people, machines, buildings and other artifacts [5].

It should be noted, that despite the fact that for testing system as the whole are more appropriate external metrics, for verification of our system testing models are internal metrics more suitable.

The second part is the quality in use. This part shows the quality of software in user’s view and specifies four characteristics [5]:

- Effectiveness is the capability of the software product to enable specified users to achieve specified goals with effectiveness, productivity, safety and satisfaction in specified contexts of use.
- Productivity is the capability of the software product to enable users to expend appropriate amounts of resources in relation to the effectiveness achieved in a specified context of use.
- Safety is the capability of the software product to achieve acceptable levels of risk of harm to people, business, software, property or the environment in a specified context of use.
- Satisfaction is the capability of the software product to satisfy users in a specified context of use.

In general, process quality contributes the improving product quality, and product quality contributes to improving quality in use [5].

II. PROPOSAL OF METRIC

To confirm the correctness of system testing model for testing safety critical systems, it is necessary to carry out the verification and/or validation. Verification should confirm whether the system testing model is created correctly according to its specification. Validation should demonstrate that the system testing model is designed with regards to realistic customer expectations and requirements. Since the verification of the system testing model requires large of realistic data, our metric proposal is focused on the verification of the system testing models.

Metric that we propose is based on standard software metrics. We have analyzed various software quality attributes to identify the appropriate metrics group. From this group we selected a standard metric, which served as a basis for our metric proposal.

Proposed metric however incorporates only requirements related to testing newly developed safety critical systems. This is due to the fact, that analyzed standards and guidelines do not deal with existing or modified safety critical systems.
A. Selection of Appropriate Standard Metric

In the field of information technology, among which we can also include safety critical computer system, are several standards that define software quality attributes and their corresponding metrics. The selection of appropriate metrics is based on the ISO 9126 standard which provides internal and external metrics for software quality assurance.

By analysis of the various software quality attributes we have identified group that met our requirements, namely functional characteristics. Functionality can be defined as the ability of information system to include features that ensure expected or specified user requirements when using system under specific conditions. Therefore functionality characteristics can be used to measure how well the product meets the requirements defined in its specification [6]. Thus this group of metrics is suitable as the basis for system testing models verification.

With regard to validation process definition, we have focused mainly on internal metrics. These metrics are used to measure the product characteristics that cannot be derived from his behavior [6]. This requirement excludes individual quality attributes for which internal metrics do not exists (e.g. conformity of the functionality).

From the individual functional characteristics, adequacy was identified as the most appropriate. Adequacy is defined as the ability to provide functions that ensure specified user tasks and needs. But it is considering only whether the functions are available or not. The basic internal metric of adequacy is completeness of implemented functions, which is defined as [6]:

\[
X = 1 - \frac{A}{B} \quad (1)
\]

Where X is the value of the metric, A is number of functions found during the control according to specification and B is number of functions in the specification.

This metric is modified to work appropriately with weights and serves as the basis for the metric proposed by us.

B. Definition of Metric Attributes

To verify the system testing models, we have defined the following metric. Value of this metric determines whether the system testing model meets or does not meets the specified requirements.

1) Name and identification

Metric determines the Completeness of implemented requirements in model. Raises the question of how complete is the implementation of individual requirements. This metric belongs to the internal metrics for evaluation of adequacy, which is one of the functional characteristics.

2) Definition

Metric to determine completeness of implemented requirements is defined by following function:

\[
X(A_1, A_2, A_3) = 1 - \frac{A_1 W_1}{B_1} + \frac{A_2 W_2}{B_2} + \frac{A_3 W_3}{B_3} \quad (2)
\]

Where X is a function of metric, A1 is the number of implemented general requirements for testing safety critical systems, A2 is the number of implemented requirements on test specification, A3 is the number of implemented requirements on test execution, B1, B2 and B3 is total number of specified requirements for each attribute and W1, W2, W3 are their weights for relative importance.

3) Weight determination for attributes

To simplify the verification process, we have divided the requirements into three categories. For each category we have established level of its importance. Since we could not clearly define which requirements for system testing of safety critical systems are more important, we wanted to use constant weights. This solution (although it is commonly used) may cause inaccuracy in the calculation when using rounded weight values. Therefore we have increased the emphasis on general requirements for testing, which are also the most extensive group of attributes. For this reason, the general requirements for testing have the normalized value set to 4 (from range <1, 5>), requirements for test specification to 3 and requirements for test execution to 3. Based on these normalized values, the weight for each attribute was determined using [7]:

\[
W_i = \frac{E_i}{\sum_{j=1}^{n} E_j} \quad (3)
\]

Where Wi is determined weight and Ei is normalized value of the attribute from interval <1, 5>.

The values of weights for individual attributes are calculated using (3), and their normalized values are listed in Table I.

4) Dimension

Defined metric have all values in the interval <0, 1>. Lower metric values mean a higher quality of system testing model, in term of specified requirements for testing safety critical systems. This range represents the standard range for metrics defined in ISO 9126 standard.

5) Initial and target value

Given the fact that the system testing models for the safety critical systems are very specific, it is not easy to establish the target value we seek. Our metric was mainly designed for verification of newly designed system testing models for safety critical systems, therefor no previous (or initial values) were available to facilitate the estimation of metric’s desired target value.

Since our testing models focus on testing of safety critical systems, the metric’s achieved value should be as

<table>
<thead>
<tr>
<th>Index</th>
<th>Attributes</th>
<th>Normalized attribute value</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Implemented general requirements for testing</td>
<td>4</td>
<td>0.4</td>
</tr>
<tr>
<td>2</td>
<td>Implemented requirements for test specification</td>
<td>3</td>
<td>0.3</td>
</tr>
<tr>
<td>3</td>
<td>Implemented requirements for test execution</td>
<td>3</td>
<td>0.3</td>
</tr>
</tbody>
</table>
The second source of data is the appropriate system testing model for safety critical systems that we want to verify. Since this model should implement requirements for system testing of safety critical systems, its various activities, events and states must reflect and incorporate these requirements. However it may not be directly specific actions or events, but these requirements can be incorporated in various elements in the proposed system testing model.

7) Confirmation (verification) procedures

To determine the value of metric it is necessary to define all requirements which shall the model implement. In the requirements specification for the system testing of safety critical systems we identified individual requirements which must be implemented in the corresponding system testing model. These individual requirements need to be allocated into one group of attributes that represent them. This step determines the number of specified requirements for each group of attributes.

Subsequently it is necessary to analyze all steps, actions and states of the appropriate system testing model and assign them to the specified requirements for system testing of safety critical systems. Although some of the requirements can be implemented only verbally in text, they also must be assigned to the relevant requirement in specification. Based on these assignments we can determine the number of implemented requirements in the appropriate model.

After determination of these parameters we can calculate the value of the proposed metric using (2).

III. VERIFICATION OF THE SYSTEM TESTING MODEL

One of the system testing models, which were verified by the proposed metric, was the performance testing model for testing communication subsystem of safety critical systems. This type of test (together with load testing) is one of the most used tests for testing communication subsystem as part of the whole system. There are several variants of performance tests (e.g. network sensitivity test) that focus on the specific attributes of data communication.

This performance testing model deals with testing of safety critical systems in general, and was outlined using appropriate UML diagrams. These diagrams capture the

<table>
<thead>
<tr>
<th>Requirement groups</th>
<th>Specified requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test plan</td>
<td>Documentation of all untested items</td>
</tr>
<tr>
<td>Created test</td>
<td>Determination of test input requirements</td>
</tr>
<tr>
<td>Test scenario</td>
<td>Unambiguous identification and name of the test</td>
</tr>
<tr>
<td>Test scenario</td>
<td>Definition of guidelines for test execution</td>
</tr>
</tbody>
</table>

The data needed to establish the metric value are based on two main sources. The first source are specified requirements for safety critical systems testing that are used to design the system testing model.

These requirements are based on the generalized testing requirements obtained by analysis of standards and guidelines for safety critical systems. These requirements represent different groups of attributes which completeness of implementation we want to determine.

It should be noted, that these attributes are valid only for newly developed safety critical systems. This condition is mainly due to the fact that analyzed standards and guidelines on which these attributes relay, mainly focus on newly developed safety critical systems.

The first group of attributes represents general requirements for testing. These requirements are mainly based on the general requirements for testing of safety critical systems, according to international standard IEC 61508. Various requirements that fall under this category are shown in Table II.

Some of the requirements from the standard IEC 61508, which also fall into other groups, are included in one of the following groups of attributes.

The second group of attributes represents requirements on the test specification. This group of requirements is mostly based on the GAMP guidelines, since other standards for safety critical systems do not address testing in sufficient detail. These requirements are supplemented by general requirements that fall into this category of attributes. Various requirements that fall under this category are shown in Table III.

The last group of attributes represents requirements on test execution. These requirements, as well as the previous group of attributes, are mostly based on the GAMP guidelines. They are supplemented by general requirements that fall into this category of attributes. Various requirements groups that fall under this category are shown in Table IV.

<table>
<thead>
<tr>
<th>Requirement group</th>
<th>Specified requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test case</td>
<td>Specification of test data</td>
</tr>
<tr>
<td>Test case</td>
<td>Determination of the expected test results</td>
</tr>
<tr>
<td>Test environment</td>
<td>Definition of support software for testing</td>
</tr>
<tr>
<td>Test environment</td>
<td>Describe test system configuration</td>
</tr>
<tr>
<td>Documentation</td>
<td>Documentation of test results</td>
</tr>
<tr>
<td>Configuration data</td>
<td>Input configuration data in allowed range</td>
</tr>
<tr>
<td>Data communication</td>
<td>Ensure the accuracy of the data in terms of data transmission</td>
</tr>
<tr>
<td>Support tools</td>
<td>Standardized inputs of the testing process</td>
</tr>
<tr>
<td>General</td>
<td>Determination of the test criteria, after which the test process is ended</td>
</tr>
</tbody>
</table>
process of performance testing of safety critical systems from multiple perspectives. Part of the performance testing model is captured as sequence UML diagram on Fig. 5. This diagram presents the basic steps of performance testing in terms of time.

It should be noted, that this diagram is only one of the views on the proposed model of performance testing for safety critical systems. In [8] are presented further details regarding this performance testing model.

The individual testing steps are based on standard testing phases of the IEEE 829 testing standard. These steps capture the performance testing steps for standard systems. However the testing of the safety critical systems varies from the testing of standard systems and has specific requirements for testing. These requirements in our testing models are mainly based on analysis of multiple standards and guidelines, which can be divided into following groups:

- Main standard for safety critical systems (IEC 61508)
- Major regional and product standards for safety critical systems (e.g. BS EN 50128, IEC 61511, IEC 62061, MISRA guidelines …)
- Major standards for nuclear power plants (mainly IEC 60880, IEC 61513, IEC 62138 …)
- GAMP guidelines

Despite the fact that the analyzed standards and guidelines focus also on the specific safety critical systems, proposed model focuses on testing safety critical systems in general. Requirements obtained by analysis of these standards and guidelines were generalized after their identification.

For the verification of this performance testing model we used the proposed metric. Individual requirements were assigned into one of the following groups of attributes:

- General requirements for testing
- Requirements for test specification
- Requirements for test execution

During this step we have excluded requirements that could not be fulfilled due to the nature of the performance test. Subsequently we have identified number of implemented requirements of the relevant attributes.

Examples of this data are shown in Table V.

![Example of requirements implementation](image)

<table>
<thead>
<tr>
<th>Specification of test data</th>
<th>Way of implementation in performance testing model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple steps in TestCaseSpecification phase that define required test data</td>
<td>Action defineTestData in ProblemDomainSpecification state</td>
</tr>
<tr>
<td>Determination of prerequisite for the testing</td>
<td>Included in testObjectives step in TestCaseSpecification phase, Action defineTestData in ProblemDomainSpecification state</td>
</tr>
</tbody>
</table>

Therefor we can consider the given performance testing model to be verified according to the requirements for testing the safety critical systems. The proposed metric meets our expectations also at verification of other system testing models. Therefore, the proposed metric can be considered as suitable for verification of system testing models for safety critical systems.

IV. CONCLUSION

In this paper, we have proposed a metric for verification of the system testing models for safety critical systems. The emphasis was put on verification of testing models for testing newly developed safety critical systems. Our proposed metric was based the internal metric of adequacy defined in the ISO 9126. Internal metric was chosen with regards to the verification process specifications. For this proposed metric we have defined all attributes that are required by the ISO 9126 standard.

We have demonstrated the suitability of the proposed metric on the performance testing model for testing the communication subsystem of the safety critical system. This performance testing model was outlined using appropriate UML diagrams. The requirements for testing the safety critical systems are based on the analysis of multiple standards and guidelines for various safety critical systems. These requirements served as the basis for the verification process of given performance testing model using the proposed metric.

After the analysis of metric’s confirmation procedure and resulting value for given performance testing model and other system testing models, we can consider the proposed metric to be suitable for the verification of the system testing models of newly developed safety critical systems.

The main application are for proposed metric and system testing models was testing of industrial control systems in nuclear power field. Given the requirements on which our proposal is made, we can consider this application area as possible. To be fully applicable in this area, the metric and also systems testing models would need to incorporate partial sub requirements on lowest level possible. Due to the nature of these systems, some requirements can be exactly and fully specified only for specific test case of particular safety critical system.

ACKNOWLEDGMENT

This publication is the result of implementation of the project: “Increase of Power Safety of the Slovak Republic” (ITMS: 26220220077) supported by the Research & Development Operational Programme funded by the ERDF.
REFERENCES


---

Figure 5. Verified performance testing model for safety critical systems