Taxonomy Based Testing using SW91, a Medical Device Software Defect Taxonomy

Hamsini Ketheswarasarma Rajaram
Dundalk Institute of
Technology
Co Louth, Ireland
hamsini.Ketheswarasarma@dkit.ie

John Loane
Dundalk Institute of
Technology
Co Louth, Ireland
john.loane@dkit.ie

Silvana Togneri MacMahon
Dundalk Institute of
Technology
Co Louth, Ireland
silvana.macmahon@dkit.ie

Fergal Mc Caffery
Dundalk Institute of
Technology
Co Louth, Ireland
fergal.mccaffery@dkit.ie

Abstract

This paper presents a summary of research undertaken to investigate and assess the use of a taxonomy based testing approach to improve medical device software (MDS) quality.

Keywords— Defect taxonomy, Taxonomy based testing, Empirical data, Root cause analysis, Defect minimisation, and Medical device software

I.   RESEARCH PROBLEM

Poor quality software in medical devices has caused deaths and serious injuries to patients [1],[2]. The US Food and Drug Administration (FDA) oversees defective medical device recalls, which could be potentially harmful to the public [3]. From 2005 to 2011, 19.4% of medical device recalls were related to software [4]. Between 2010 and 2015, software related recalls increased by 111% [5]. There are many software quality assurance activities (SQA) such as testing, inspections, reviews, walkthroughs, and analyses which can be used to improve software quality [6]. However, MDS development companies continue to face challenges in improving software quality [7].

II.   RESEARCH CONTRIBUTION

Software defect taxonomies have been successfully used to improve software quality in other software domains such as safety critical, business and telecommunications domains [8]–[11]. A defect taxonomy is a system of hierarchical categories designed to be a useful aid for reproducibly classifying defects [12]. There are many defect taxonomies such as the orthogonal defect classification scheme and the IEEE standard classification for software anomalies. Classification of Defects in Health Software - SW91 is a new defect taxonomy being developed by the Association for the Advancement of Medical Instrumentation (AAMI). AAMI will publish SW91 as a standard by early this year. SW91 includes multi-level defect categories such as parent level and child level from the planning to the maintenance phase of software development. Each defect category has a unique defect code which is part of a hierarchical structure. The annexes included in SW91 demonstrate how the defect categories from SW91 can be used in software development. This research is focused on improving MDS quality by using SW91. The next section details the research methodology and the research questions formulated by considering poor quality software in medical devices and how SW91 might be utilised in addressing this problem.

III.   RESEARCH METHODOLOGY AND RESEARCH TO DATE

The chosen research methodology is action design research (ADR). It has four stages. The remainder of this section explains the four stages and the research to date.

A.   Stage 1: Problem Formulation

SW91 was identified as an artefact to contribute the identified research problem. An approach called taxonomy based testing was proposed by considering the benefits of defect taxonomies in software testing and SW91. The following main research question was developed: *"To what degree can taxonomy based testing be used to improve medical device software quality?"*
The following three research sub questions were formulated to achieve the main research question:
1. How were software defect taxonomies used in other software domains?
2. How can a taxonomy based testing approach be used to improve medical device software quality?
3. What are the potential benefits of applying taxonomy based testing in a medical device software development company?

B.   Stage 2: Building, Intervention, and Evaluation (BIE)

A research collaboration was formed between this research and the work been undertaken by the SW91 development team from AAMI. The research collaboration enabled this research to participate in the development activities of SW91 via mapping vulnerabilities from the Common Weakness Enumeration (CWE) into SW91 defect categories. The next section explains this mapping.

1. Mapping defects from existing databases into SW91 defect categories.

As a part of this research, a mapping of the vulnerabilities from the Common Weakness Enumeration (CWE) (http://cwe.mitre.org/) into SW91 defect categories was conducted. The CWE is an open source list containing common software weaknesses and their vulnerabilities. This mapping work considered the defect coverage of SW91 and changes were made to SW91 by considering the CWE mapping. A subset of the mapping was published as annex D in SW91. This mapping work lead to me becoming a member of the SW91 development team. This mapping work was explained in a conference paper "Benefits of Defect

Taxonomies and Validation of a new Defect Classification for Health Software" [13].

C. Stage 3: Reflection and Learning

This stage moves from building an artefact into applying the artefact into a broader class of problems. This stage of ADR will involve applying the taxonomy based testing approach into several medical device software development companies. Analysis of the taxonomy based testing results will be used to assess the implications of taxonomy based testing and to validate SW91. This stage of ADR will answer RSQ 2. The results from the taxonomy based testing will be used to answer RSQ 3. SW91 validation results will be input to the initial design of the artefact, SW91. The following section explains what taxonomy based testing is and how it was retrospectively investigated using data from a MDS development company.

2. Mapping defects from a MDS development company into SW91 defect categories and Taxonomy based testing

From the literature, I identified that defect taxonomies were successfully used in the testing phase of software development to improve software quality [9], [14], [15]. Then, as a contribution to the research problem identified, an approach called taxonomy based testing was proposed. The application of this approach is expected to improve software quality in MDS development companies. Generally, in testing, the test cases are written based on the requirements. These test cases will be executed to check whether the software complies with its requirements or not. In the taxonomy based testing approach, the requirements will be mapped into potential defect categories from the defect taxonomy and the test cases will be written based on the requirements and the mapped defect categories. The test cases will be executed to verify whether the software complies with the relevant requirements and does not contain the mapped defect categories from the defect taxonomy. A retrospective study was conducted using empirical data from company A in Ireland which develops MDS applications. This retrospective study successfully investigated the applicability of taxonomy based testing to the empirical data from a MDS development company and detailed the benefits of taxonomy based testing. This study explained how taxonomy based testing can be used to improve the defect reporting and how it can be used to minimise defects. It also identified how risks can be minimised and how root cause analysis can be conducted by implementing taxonomy based testing. A report has been compiled on the study and it has been submitted to the company A for their consideration. This work partially answered RSQ 2. A conference paper based on the retrospective study will be submitted to a relevant conference.

D. Stage 4: Formalization of Learning

The identified artefact which is SW91 will be published as a standard including annex D. Findings and results from taxonomy based testing using SW91 will be published in conferences. A protocol for taxonomy based testing will be published for future implementation of taxonomy based testing.

IV. FUTURE RESEARCH

Another retrospective analysis will be performed at company B in Ireland to investigate the taxonomy based testing approach further. The company B retrospective study will be conducted with SQA engineers, developers and a SQA manager. A taxonomy based testing protocol has been written for future implementation at MDS development companies. This protocol will enable taxonomy based testing to be conducted without the researcher's involvement. Company C is expected to implement the taxonomy based testing protocol by early next year.

The results from Company B and C will be used to

- Assess the benefits of taxonomy based testing in MDS development companies. The results from the assessments will be used to answer RSQ 3.
- Validate SW91 in terms of the efficiency, the reliability, defect coverage and the useful analyses enabled by SW91.

References

[1]   F. M. Caffery and G. Coleman, 'The Need for a Software Process Improvement Model for the Medical Device Industry', vol. xx, no. January 2006.

[2]   R. K. Iyer, Z. Kalbarczyk, J. Raman, and J. Raman, 'Analysis of safety-critical computer failures in medical devices', *IEEE Security and Privacy Magazine*, vol. 11, no. 4, pp. 14–26, 2013.

[3]   Food and Drug Administration, 'FDA 101: Product Recalls', 2010.

[4]   L. K. Simone, 'Software-related recalls: An analysis of records', *Biomed. Instrum. Technol.*, vol. 47, no. 6, pp. 514–522, 2013.

[5]   J. R. Dix, S. Ramachandran, and D. S. Oppenheimer, 'What's Behind Medtech's Recall Epidemic?', 2016. .

[6]   J. Tian, *Software quality engineering*, vol. 32, no. 1. Dallas, TX: John Wiley & Sons, Inc, 1990.

[7]   N. BLACK, 'Quality Assurance of Medical Care', *J. Public Health (Bangkok).*, 1990.

[8]   N. Bridge and C. Miller, 'Orthogonal Defect Classification Using Defect Data to Improve Software Development', *Softw. Qual.*, vol. 3, no. 1, pp. 1–8, 1998.

[9]   R. R. Lutz and I. C. Mikulski, 'Empirical analysis of safety-critical anomalies during operations', *IEEE Trans. Softw. Eng.*, vol. 30, no. 3, pp. 172–180, 2004.

[10]   N. Mellegård, M. Staron, and F. Törner, 'A light-weight defect classification scheme for embedded automotive software and its initial evaluation', 2012.

[11]   N. Li, Z. Li, and X. Sun, 'Classification of software defect detected by black-box testing: An empirical study', in *WCSE 2010*, 2010, vol. 2, pp. 234–240.

[12]   International Software Testing Qualifications Board, 'Standard glossary of terms used in Software Testing.'

[13]   H. K. Rajaram, J. Loane, S. T. MacMahon, and F. M. Caffery, 'Benefits of Defect Taxonomies and Validation of a new Defect Classification for Health Software', in *EuroAsiaSPI²*, 2017, p. 15.

[14]   R. Black, *Advanced Software Testing - Vol. 2*, 2nd ed., vol. 1. Santa Barbara: Rocky Nook Inc, 2008.

[15]   M. Felderer and A. Beer, 'Using defect taxonomies to improve the maturity of the system test process: Results from an industrial case study', in *SWQD 2013*, 2013, vol. 133, pp. 125–146.