Paper Title: Dynamic Random Testing of Web Services: A Methodology and Evaluation

Manuscript ID: TSC-2019-01-0031.R1

Dear Professor ~~James~~ Joshi,

Thank you for your email on October 7, 2019, regarding our paper "Dynamic Random Testing of Web Services: A Methodology and Evaluation," which was submitted to IEEE Transactions on Services Computing (Manuscript ID: TSC-2019-01-0031.R1).

We are submitting a new version of the paper, in which we have made revisions to address and respond to each comment from the associate editor and reviewers. Below are the detailed responses to the comments.

We look forward to hearing from you.

Yours sincerely,
Chang-ai Sun, Hepeng Dai, Guan Wang, Dave Towey, Tsong Yueh Chen, and Kai-Yuan Cai

Response to comments of associate editor and reviewers

In the following, unless otherwise specified, all comments refer to the revised version of the paper.

Associate editor's comments

E1C1: Reviewers have now commented on the revised version of your manuscript. They are mostly satisfied with the revisions and with the answers to their comments. There is however one concern remaining, about the independent verifiability and repeatability of results. In particular, both the experimental subjects and the tool prototype are not made available. (Incidentally, Figure 2, which seems to show the tool configuration interface, is not referenced in the text.) I believe this aspect has to be justified.

Response: Thank you for insightful summary and comments. We have addressed the concern about the independent verifiability and repeatability of results by making both the experimental subjects and the prototype publicly accessible. Please refer to our responses to R3C2 for the concern on the prototype, R2C2 for the concern on the experimental subjects, respectively.

Sorry for missing the reference in the text to Figure 2.

Action: In the revised manuscript, we have added references to the interfaces in Figure 2 in Section 3.3.

E1C2: As for what concerns the tool, if it cannot be made available, the issue needs to be discussed in the threats to validity section, to fairly examine if and to what extent the results may be affected by the implemented tool, which might be, for instance, unwillingly tailored to the subjects, affecting generalizability.

Response: Thank you for the comment. To make things simple, we decide to make the tool prototype available, which has now been released on Github, and can be accessed through the linkage: https://github.com/phantomDai/DRTester.git. Please also refer to our responses to R3C2.

Action: In the revised manuscript, we have added a linkage for accessing the prototype in the first sentence of Section 3.3.

E1C3: As for the subjects, no motivation is provided for not chosing publicly available web services; there is nothing wrong in this choice, but without an

explanation the authors appear not to have done their best efforts to support independent verification, which is scientifically not sound.

Response: Thank you for the comments. We agree that a simple way is to choose publicly available web services for evaluation. However, it is impossible to gain access to service request implementations. Alternatively, we implemented subject web services ourselves according to the real-life specification, because we have to access to the implementation of the web services used in the experiments in order to create the faulty mutants. The experimental subjects have been made available on a public repository, which can be accessed through the following linkage: https://github.com/phantomDai/subjects4tsc.git. Please also refer to our response to R2C2.

Action: In the revised manuscript, we have added a note to explain why not choosing publicly available web services for evaluation in the beginning of Section 3.3.

E1C4: In the manuscript it is stated that "although we have tried to improve the generalisability of the findings ..., we cannot be certain that similar results would be observed in other types of web services"; this statement might raise into the reader legitimate questions about the specificity of the used subjects and tool, because the assumptions do not state that the proposed DRT technique targets a specific "type" of web services, mirrored by the self-implemented subjects.

Response: Thank you for the comment. Originally, we intended to mean that the evaluation results may be a bit different with varying subject web services. However, the current statement is misunderstanding and prone to incur the doubt on the subject web services.

Action: In the revised manuscript, we have changed the statement to better manifest our intent in Section 4.6.2.

E1C5: In summary, the work has been judged sound and almost mature for publication by reviewers, but the comments by two of them, shared by this associate editor, highlight that the matter of independent verifiability should be treated a bit more seriously in the manuscript: more justifications/considerations are in order in section 3.3 and 4.6, and possible solutions for it should be proposed (e.g., if there exist confidentiality constraints, the subjects and the tool could be made available to other researchers not necessarily publicly but upon request and signing a NDA).

Response: Thank you for the summary. We have followed the suggestions to make the experimental subjects and the tool available for independent verifiability and repeatability.

Action: None.

Reviewer 3's comments

R3C1: ***Recommendation: Minor Revision***

The authors have done a good job in the paper revision and have answered most of my comments.

Response: Thanks for the endorsement.

Action: None.

R3C2: ***There is only one point that needs further clarifications. Section 3.3 now describes the tool prototype with a set of steps, however, the tools itself is not provided so the possibility to experiment with the tool for the sake of replicability is still missing.***

Response: Thank for the comment. We decide to make the tool prototype available in order to make the replicability easy. The prototype has been released on Github, and can be accessed through the linkage: https://github.com/phantomDai/DRTester.git.

Action: In the revised manuscript, we have added a linkage of the prototype in the first sentence of Section 3.3.

Reviewer 2's comments

R2C1: ***Recommendation: Accept***

The paper is well written and organized. The redraw of figures and the addition of tables make the contribution improved in terms of clarity. Moreover, the addition of Adaptive Testing (AT) as term of comparison in the experimental part makes the evaluation more solid and impactful: Dynamic Random Testing improves significantly the performance of Random Testing and Partition testing approaches, and it outperforms AT in many cases.

Response: Thank you for the endorsement and recommendation.

Action: None.

R2C2: ***Although the authors reported an accurate description of configuration procedures of DRT approach, they did not make available on a public repository the experimental subjects developed for the experimentation. The authors could make them available to support independent verification and replication of the results. Otherwise, this lack should be explained in Section 4.6.***

Response: Thank you for the suggestion. We decide to make the experimental subjects available on a public repository, which can be accessed through the following linkage: https://github.com/phantomDai/subjects4tsc.git. The released experimental stuff include source code, mutants evaluated in our experiments, and test suites for all subject programs. Also, an introductory note is provided.

Action: In the revised manuscript, we have followed the suggestion and added a linkage of the experimental subjects in the end of the first paragraph in Section 4.2.

Reviewer 1's comments

R1C1: ***Recommendation: Accept***

The manuscript is in good shape after the comments for the initial draft are addressed.

Response: Thank you for the endorsement and recommendation.

Action: None.