

ValiDEPlan – Validity-Driven Software Engineering Experiments Planning Tool

Liliane Fonseca¹, Eudis Teixeira¹, Sergio Soares^{1,2}, Allan Lima¹

¹Centro de Informática – Universidade Federal de Pernambuco (UFPE)
Caixa Postal 7851 – 50.732-970 – Recife – PE – Brasil

²Instituto Senai de Inovação para Tecnologias da Informação e Comunicação (ISI-
TICs)
50.100-000 – Recife – PE – Brasil
{lss4, eot, scbs, aml3}@cin.ufpe.br

***Abstract.** One of the major problems mentioned by researchers in experimental software engineering is dealing with several threats to the validity of their experiments. It is known that the sooner the threat is identified the greater chance of minimizing these potential threats. This paper presents an experiment planning tool that addresses threats to validity to identify and prioritize threats, as well as it suggests control actions to mitigate identified problems. This tool is an evolution of Reviewer EP, a collaborative web platform for reviewing the completeness of experimental plans. The tool can be useful for both experienced and inexperienced researchers.*

***Resumo.** Um dos principais problemas mencionados por pesquisadores na engenharia de software experimental é lidar com as diversas ameaças à validade de seus experimentos. Sabe-se que quanto mais cedo a ameaça for identificada e tratada maior a chance de minimizar essas possíveis ameaças. Este artigo apresenta uma ferramenta de planejamento de experimento orientada a ameaças à validade para identificar e priorizar ameaças, além de sugerir ações de controle para mitigar os problemas identificados. Esta ferramenta é uma evolução da Reviewer EP, uma plataforma web colaborativa para revisão da completude de planos experimentais, podendo ser útil tanto para pesquisadores experientes quanto inexperientes.*

<<https://youtu.be/L-cCINcdoJQ>>

1. Introdução

Uma das principais dificuldades no planejamento de experimento é antecipar possíveis problemas na execução que possam causar vieses no resultado trazendo em muitas vezes resultados irreais ou inconclusivos [Siegmond et al. 2015]. Uma das soluções para aumentar a confiabilidade nos resultados dos experimentos é identificar antecipadamente os principais riscos que possam surgir e se antecipar a eles através do planejamento de ações de controle. No entanto, as atividades de identificação, análise e priorização a ameaças à validade não são atividades simples. Uma das razões se deve ao fato do alto número de possibilidades de problemas que podem acontecer no planejamento, execução e análise de uma avaliação experimental, além de que o tratamento de uma ameaça pode implicar no surgimento de uma outra ameaça a validade não identificada inicialmente [Teixeira et al. 2018]. Embora existam na Engenharia de Software Experimental (ESE) alguns estudos que tratam da temática das

ameaças à validade [Wohlin et al. 2012], [Juristo e Moreno 2010], [Kitchenham et al. 2002] nenhum deles apresenta uma ferramenta que dê suporte de como essas ameaças podem ser identificadas, priorizadas e tratadas, de acordo com o contexto de cada estudo. A partir disto, houve a necessidade de evoluir a ferramenta *Reviewer EP* [Fonseca et al. 2017] a fim de auxiliar pesquisadores no planejamento de seus experimentos com foco na identificação e priorização das ameaças à validade. A *ValiDEPlan* continua dando o suporte à completude dos principais elementos que devem estar contidos no planejamento de um experimento controlado. Além disso, ela antecipa de forma direta riscos em potencial que possam acontecer na execução do estudo e permite a classificação de possíveis ameaças à validade, indicando quais devem ter prioridade ao serem tratadas e quais são as ações de controle mais eficazes para mitigá-las.

O restante deste artigo está organizado da seguinte forma: Seção 2 apresenta a *ValiDEPlan*, com sua arquitetura, principais funcionalidades e potenciais usuários da ferramenta. A Seção 3 apresenta um cenário de uso da ferramenta. A Seção 4 apresenta os trabalhos relacionados, e, por fim, a Seção 5 apresenta as considerações finais deste trabalho.

2. ValiDEPlan - *Validity-Driven Software Engineering Experiments Planning Tool*

A *ValiDEPlan*¹ é uma ferramenta web capaz de auxiliar pesquisadores iniciantes e experientes a planejar experimentos controlados da área de Engenharia de Software (ES), sendo orientada à validade de seus resultados. A ferramenta está na língua inglesa para facilitar o acesso a pesquisadores de diversos países. Esta sessão contempla a arquitetura da ferramenta, as principais funcionalidades da plataforma, incluindo os usuários em potencial da *ValiDEPlan*.

2.1 Arquitetura da ValiDEPlan

A *ValiDEPlan* é uma aplicação web desenvolvida com tecnologia Angular, responsável por fazer chamadas para um API desenvolvida em Java. A comunicação é efetuada através de REST *web services*. O *setup* de configuração do ambiente de desenvolvimento é simples e dependente da (s) IDE (s) e SGBD que o desenvolvedor preferir utilizar. Nós utilizamos o *IntelliJ* para desenvolver as funcionalidades do *back end* e *front end* a fim de fazer acesso/*queries* ao banco de dados. Para ambientes de produção ou quaisquer outros que porventura venham a ser necessários, a aplicação é executada em contêineres do *Docker*, sendo um contêiner isolando o *front end* e outro isolando a API.

Para as tecnologias utilizadas no *back end* utilizamos Java, Jersey, *Hibernate* e *Jetty*. Para o *Front end* utilizamos as tecnologias angular, *materialize*, *Nginx*. Utilizamos *PostgreeSql* para o banco de dados e para o *deploy shell script* e utilização de containers e *dockers*.

¹ *ValiDEPlan*: Validity-Driven Software Engineering Experiments Planning Tool.
<https://valideplan.cin.ufpe.br>

2.2 Funcionalidades da ValiDEPlan

A ValiDEPlan pode ser utilizada por qualquer pesquisador, seja ele experiente ou iniciante, na construção de planos experimentais orientados à validade de seus resultados. Os perfis de acesso administrador, autor, revisor e colaborador permanecem o mesmo da versão *Reviewer EP* [Fonseca et al. 2017]. As funcionalidades manter planos experimentais, visualizar status da revisão dos planos experimentais, acessar catálogo web externo de mecanismos de suporte, fazer upload e download dos planos experimentais, dentre outras, foram trazidas da versão da *Reviewer EP* para a ValiDEPlan. As funcionalidades produzir e revisar planos experimentais foram adaptadas às modificações do planejamento orientados a ameaças à validade. Novas funcionalidades como identificar características do experimento, classificar ameaças à validade, identificar e priorizar as ameaças e analisar e selecionar ações de controle foram adicionadas na ValiDEPlan. O usuário tem acesso as seguintes funcionalidades:

F1: Produzir um Novo Plano Experimental

Permite ao usuário criar um plano experimental desde o seu início até sua completude, como era feito na ferramenta anterior. No entanto, além disto, ela permite ao pesquisador identificar as características do experimento (F3) ao mesmo tempo que elabora o plano experimental. Ao selecionar a funcionalidade de criar um Novo Plano experimental, o usuário informa nome e descrição do plano e inicia preenchendo o plano experimental de acordo com o instrumento desenvolvido por Fonseca [Fonseca 2016]. Ao longo do processo de criação do plano, o usuário também será exposto a uma lista de verificação² para identificar as características do experimento. Ao final desta funcionalidade o usuário terá o plano experimental juntamente com as características definidas do experimento. Caso deseje, a ValiDEPlan permite ao usuário enviar o plano experimental criado para ser revisado por um pesquisador mais experiente.

F2: Revisar um Plano Experimental

Caso o usuário deseje iniciar por esta funcionalidade, ele pode submeter o plano experimental completo no formato PDF para ser revisado por um pesquisador mais experiente. O usuário deverá incluir revisores disponíveis na base e inserir uma data limite para o retorno da revisão. Os revisores cadastrados como disponíveis na plataforma receberão um e-mail e poderão aceitar ou recusar o pedido de revisão. Esta funcionalidade pode ser acessada também quando o usuário finaliza um plano experimental.

F3: Identificar características do experimento

Esta funcionalidade pode ser acessada diretamente, caso o usuário já tenha elaborado um plano experimental do experimento ou pode ser associada a criação de um novo plano experimental (F1). O objetivo desta função é identificar as características do experimento através de uma lista de verificação². Esta lista é composta por 43 questões dividida em oito categorias. Cada questão é associada a uma escala de respostas: Sim (significa que o experimento possui a característica), Parcialmente (significa que o experimento tem parcialmente a característica), Não (significa que o experimento não tem a característica) e Não Aplicável (significa que o item perguntado não é aplicado ao experimento). A partir da inserção das características do experimento, a ValiDEPlan

² A lista de verificação (em inglês) para a caracterização do experimento está disponível em <https://tinyurl.com/y8qz7cxq>

busca em um banco de dados de ameaças à validade existente e gera uma lista de possíveis ameaças e suas ações de mitigação.

F4: Classificar ameaças à validade sugeridas para o experimento

O usuário deve já ter realizado a identificação das características do experimento através da F1 ou F3. Baseado nas características identificadas, a ValiDEPlan gera uma lista de possíveis ameaças. O usuário deverá classificar cada ameaça sugerida em uma escala de 1 a 3 com base em sua intensidade nos seguintes três critérios: IMPACTO (consequências para o experimento provocada pela ameaça), URGÊNCIA (rapidez com que determinada ameaça deve ser resolvida) e TENDÊNCIA (prognóstico sobre como a ameaça tende a evoluir). A saída dessa atividade é uma lista de ameaças à validade priorizadas.

F5: Identificar e priorizar as ameaças à validade

Esta funcionalidade unifica as funcionalidades F3 e F4, ou seja, o usuário tem a opção de, a partir de um plano experimental construído, identificar as características e classificar por ordem de prioridade as ameaças à validade sugeridas juntamente com suas respectivas ações de controle.

F6: Analisar e Selecionar ações de controle

Nesta funcionalidade, o experimentador classifica um conjunto de ações de mitigação para cada ameaça à validade. Depois de classificar as ameaças sugeridas pela ValiDEPlan para o experimento, ela define uma escala de magnitude para cada ameaça de acordo com sua prioridade (Muito alta, Alta, Moderada, Baixa e Muito Baixa). O usuário pode acessar através de um filtro, qual prioridade interessa. Em seguida, o usuário classifica cada ação de controle utilizando uma escala de importância/efetividade, considerando os seguintes valores possíveis: (i) não importante; (ii) ligeiramente importante; (iii) moderadamente importante; (iv) importante; e (v) muito importante. O resultado desta atividade é uma lista de ações de controle para cada ameaça identificada.

3. Exemplo de uso

A ValiDEPlan pode ser usada tanto pelo pesquisador autor que vai produzir o plano experimental quanto por um colaborador ou revisor do plano. Para exemplificar o uso da ValiDEPlan, apresentamos o fluxo de identificação, priorização e mitigação das ameaças à validade. Para realizar esta atividade, o usuário já deve ter acessado a ValiDEPlan <<https://valideplan.cin.ufpe.br>> com login e senha e ter elaborado um plano experimental. O pesquisador pode salvar o seu progresso a qualquer momento.

3.1 Identificando e priorizando Ameaças

O pesquisador deve acessar a área principal da ValiDEPlan (*Dashboard*) e clicar no item <*Identify and Prioritize Threats*>, preencher os campos obrigatórios: Nome (*Draft Name*) e Descrição de seu plano experimental (*Experiment Description*) e clicar na opção SAVE. O usuário deve utilizar como suporte o plano experimental do seu experimento. Criada a representação do plano experimental, o usuário será redirecionado automaticamente para a tela de caracterização (*Experiment Characteristics*). O pesquisador deve preencher integralmente ao questionário de acordo com as informações do plano experimental já realizado. Se por algum motivo, o usuário sair da tela de preenchimento do *checklist*, ele poderá acessá-lo na opção *Experimental Drafts* do menu esquerdo da ValiDEPlan. Ao concluir o preenchimento do questionário, o usuário deve clicar em SAVE. O usuário será questionado se deseja classificar as

ameaças à validade sugeridas para o seu experimento. Caso, sim, o usuário deverá clicar em YES e será redirecionado para a tela de classificação de ameaças. Caso não, o usuário poderá optar por terminar o preenchimento em outro momento.

Na tela de classificação de ameaças, todas as ameaças sugeridas para o plano experimental serão listadas. Em seguida, o pesquisador deve classificar em uma escala de 1 a 3 os critérios *impacto*, *urgência* e *tendência*. As informações sobre o entendimento da ameaça apresentada ou dos critérios a serem classificados podem ser obtidas através do sinal de interrogação em cada um dos itens.

Após classificar todas as ameaças e clicar em *SAVE*, o usuário da ValiDEPlan será questionado se deseja definir e classificar as ações de controle sugeridas para mitigar as ameaças do seu plano experimental.

3.2 Classificando as Ações de Controle

Após ler as instruções da tela de classificação de ações de controle, o pesquisador deve clicar em *Define Actions*. A tela de classificação de ações de controle informa que as ameaças classificadas anteriormente podem ser filtradas/agrupadas de acordo com uma escala de magnitude/prioridade (*muito baixa*, *baixa*, *moderada*, *alta* e *muito alta*). Nesta tela, serão exibidas, individualmente, cada ameaça, junto às suas ações de controle. O usuário poderá selecionar as ações de controle desejadas para o seu plano experimental e classificá-las utilizando uma escala de importância que vai de “*Não importante*” até “*Extremamente importante*”. As ações classificadas com maior nível de importância serão as que terão maior efetividade ao mitigar as ameaças sugeridas ao plano experimental.

3.3 Visualização de Reports

O usuário será redirecionado para o espaço de trabalho de seu plano experimental (*Workspace*). Neste espaço, o usuário deverá visualizar os relatórios (*Reports*) de ameaças gerados para o seu plano experimental. A saber: relatório de ameaças sugeridas, relatório de ameaças classificadas e relatório de ameaças e ações de controle. Após esta etapa, o fluxo é encerrado, podendo o usuário criar um novo experimento ou editar cada uma das etapas anteriores.

4. Trabalhos Relacionados

Pesquisadores da comunidade de ESE tem proposto guias, *checklists* e resumos [Feldt, and A. Magazinius, 2010]. No entanto, a estratégia mais utilizada é o *checklist* onde o experimentador verifica se as ameaças presentes na lista podem ocorrer em seu experimento [Teixeira, 2018]. Em 2013, Neto & Conte propuseram um modelo conceitual que contém as relações entre ameaças à validade e suas respectivas ações de controle e desenvolveram uma ferramenta chamada *Threats to Validity Assistant* (TVA), cujo objetivo foi identificar e mitigar ameaças à validade durante o planejamento de um experimento da área de ES. A ValiDEPlan é capaz de auxiliar pesquisadores iniciantes e experientes a planejar experimentos controlados da área de ES, sendo orientada à validade de seus resultados. Dessa forma, seu principal diferencial relacionado à TVA é o suporte à classificação/priorização dos riscos identificados, possibilitando ao experimentador definir ações de controle mais precisas ao mitigar as ameaças à validade de seus estudos.

5. Conclusão

Neste artigo, apresentamos a ferramenta ValiDEPlan, evolução da ferramenta *Reviewer EP* para planejamento de experimentos com o diferencial de auxiliar os pesquisadores a planejarem seus experimentos orientados a ameaças à validade. A ValiDEPlan dá suporte a identificação, priorização e controle das ameaças à validade já na fase de planejamento, a fim de reduzir os potenciais riscos e problemas que podem ocorrer desde o planejamento, execução e reporte do experimento. A ferramenta está sendo avaliada por especialistas e pesquisadores inexperientes, sendo estes últimos o principal público alvo da ferramenta. Em breve os resultados serão publicados.

6. Agradecimentos

Este trabalho é parcialmente financiado pelo INES (Instituto Nacional de Engenharia de Software), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico do Governo Brasileiro), do subsídio 465614 / 2014-0, e do FACEPE (Fundação de Ciência e Tecnologia de Pernambuco). APQ-0399-1.03 / 17 e PRONEX APQ / 0388-1.03 / 14. Liliane Fonseca, Eudis Teixeira e Sergio Soares são parcialmente apoiados respectivamente pelo CNPq grant 88887.351815/2019-00, 141705/2015-9 e 304499/2016-1.

Referências

- Fonseca, L., Lucena, E., Soares, S., Seaman, C. (2017) “Reviewer EP– Uma Plataforma Web Colaborativa para Revisão da Completude de Planos Experimentais” (2017) In: Tools Session in VIII Brazilian Conference on Software: Theory and Practice (CBSOFT), Brazil, pp. 97–104.
- Feldt, R. and Magazinius, A., 2010, July. Validity Threats in Empirical Software Engineering Research-An Initial Survey. In SEKE (pp. 374-379).
- Fonseca, L. (2016) “An instrument for reviewing the completeness of experimental plans for controlled experiments using human subjects in software engineering”, Ph.D. thesis, Federal University of Pernambuco, Recife, Brazil.
- Juristo N. and Moreno A. M. (2010) “Basics of Software Engineering Experimentation”. Springer Publishing Company, Incorporated, 1st edition.
- Kitchenham B. A., Pfleeger S. L., Pickard L. M., Jones P. W., Hoaglin D. C., Emam K. E., and Rosenberg J. (2002) “Preliminary guidelines for empirical research in software engineering”. IEEE Trans. Softw. Eng., 28(8):721–734.
- Neto, A.A. and Conte, T., 2013, April. A conceptual model to address threats to validity in controlled experiments. In Proceedings of the 17th International Conference on Evaluation and Assessment in Software Engineering (pp. 82-85). ACM.
- Siegmund, J. Siegmund, N., Apel, S. (2015) “Views on internal and external validity in empirical software engineering”, In: Proceedings of the 37th International Conference on Software Engineering-Volume 1, IEEE Press, pp. 9–19.
- Teixeira, E., Fonseca, F., Soares, S. (2018) “Threats to validity in controlled experiments in software engineering: what the experts say and why this is relevant” In: Proceedings of the XXXII Brazilian Symposium on Software Engineering, ACM, pp. 52–61.
- Wohlin, C., Runeson, P., Host, M., Ohlsson, M. C., Regnell, B., Wesslen A. (2012) “Experimentation in software engineering”, Springer Science & Business Media.