

II Workshop de Dissertações de Mestrado do PPgSI

FICHA DA PESQUISAⁱ

DADOS GERAIS																	
Título do projeto de pesquisa	On the use of control- and data-flow in fault localization																
Orientando	Henrique Lemos Ribeiro																
Orientador(es)	Prof. Dr. Marcos Lordello Chaim																
Momento atual	<input type="checkbox"/> 3º semestre <input type="checkbox"/> 4º semestre <input checked="" type="checkbox"/> 5º semestre <input type="checkbox"/> 6º semestre																
Qualificação	<input checked="" type="checkbox"/> Qualificação já realizada em: 09/02/2015 <input type="checkbox"/> Qualificação planejada para: ____/____/____																
Defesa	Prazo máximo para depósito: 05/07/2016 Depósito planejado para: 30/11/2015																
Linha e Área de pesquisa	<table style="width: 100%; border: none;"> <tr> <td style="width: 50%; border: none;"> <input type="checkbox"/> Gestão e desenvolvimento de Sistemas: <input type="checkbox"/> BD <input type="checkbox"/> Gestão de SI <input checked="" type="checkbox"/> Eng. de Software <input type="checkbox"/> IHC </td> <td style="width: 50%; border: none;"> <input type="checkbox"/> Inteligência de Sistemas: <input type="checkbox"/> IA <input type="checkbox"/> Rec. de Padrões <input type="checkbox"/> Proc. Gráfico </td> </tr> </table>	<input type="checkbox"/> Gestão e desenvolvimento de Sistemas: <input type="checkbox"/> BD <input type="checkbox"/> Gestão de SI <input checked="" type="checkbox"/> Eng. de Software <input type="checkbox"/> IHC	<input type="checkbox"/> Inteligência de Sistemas: <input type="checkbox"/> IA <input type="checkbox"/> Rec. de Padrões <input type="checkbox"/> Proc. Gráfico														
<input type="checkbox"/> Gestão e desenvolvimento de Sistemas: <input type="checkbox"/> BD <input type="checkbox"/> Gestão de SI <input checked="" type="checkbox"/> Eng. de Software <input type="checkbox"/> IHC	<input type="checkbox"/> Inteligência de Sistemas: <input type="checkbox"/> IA <input type="checkbox"/> Rec. de Padrões <input type="checkbox"/> Proc. Gráfico																
Área de aplicação	<table style="width: 100%; border: none;"> <tr> <td style="width: 25%; border: none;"><input type="checkbox"/> Ambientes Corporativos</td> <td style="width: 25%; border: none;"><input type="checkbox"/> Educação</td> <td style="width: 25%; border: none;"><input type="checkbox"/> Linguagem Natural</td> <td style="width: 25%; border: none;"><input type="checkbox"/> Redes Sociais</td> </tr> <tr> <td style="border: none;"><input type="checkbox"/> Bioinformática</td> <td style="border: none;"><input type="checkbox"/> Educação a Distância</td> <td style="border: none;"><input type="checkbox"/> Linguística</td> <td style="border: none;"><input type="checkbox"/> Robótica</td> </tr> <tr> <td style="border: none;"><input type="checkbox"/> Biometria</td> <td style="border: none;"><input type="checkbox"/> Internet</td> <td style="border: none;"><input type="checkbox"/> Processos de Negócio</td> <td style="border: none;"><input type="checkbox"/> Saúde</td> </tr> <tr> <td style="border: none;"><input type="checkbox"/> Economia</td> <td style="border: none;"><input type="checkbox"/> Jogos</td> <td style="border: none;"><input type="checkbox"/> Química</td> <td style="border: none;"><input checked="" type="checkbox"/> Desenv. Software</td> </tr> </table>	<input type="checkbox"/> Ambientes Corporativos	<input type="checkbox"/> Educação	<input type="checkbox"/> Linguagem Natural	<input type="checkbox"/> Redes Sociais	<input type="checkbox"/> Bioinformática	<input type="checkbox"/> Educação a Distância	<input type="checkbox"/> Linguística	<input type="checkbox"/> Robótica	<input type="checkbox"/> Biometria	<input type="checkbox"/> Internet	<input type="checkbox"/> Processos de Negócio	<input type="checkbox"/> Saúde	<input type="checkbox"/> Economia	<input type="checkbox"/> Jogos	<input type="checkbox"/> Química	<input checked="" type="checkbox"/> Desenv. Software
<input type="checkbox"/> Ambientes Corporativos	<input type="checkbox"/> Educação	<input type="checkbox"/> Linguagem Natural	<input type="checkbox"/> Redes Sociais														
<input type="checkbox"/> Bioinformática	<input type="checkbox"/> Educação a Distância	<input type="checkbox"/> Linguística	<input type="checkbox"/> Robótica														
<input type="checkbox"/> Biometria	<input type="checkbox"/> Internet	<input type="checkbox"/> Processos de Negócio	<input type="checkbox"/> Saúde														
<input type="checkbox"/> Economia	<input type="checkbox"/> Jogos	<input type="checkbox"/> Química	<input checked="" type="checkbox"/> Desenv. Software														

DESCRIÇÃO DO PROJETO DE PESQUISA

Contextualização / motivação	Depuração de programas é uma das atividades mais custosas da etapa de desenvolvimento de software. Recentemente, técnicas que utilizam cobertura de código gerado durante o teste de software têm sido propostas para indicar trechos mais suspeitos de conter defeitos. A maioria dessas técnicas utilizam coberturas de teste baseadas em fluxo de controle: cobertura de comandos e de condições.
Problema de pesquisa	Essa pesquisa visa comparar o uso de coberturas baseada em fluxo de controle (e.g., cobertura de comandos) e o uso de coberturas baseadas em fluxo de dados (e.g., duas – definition-use associations) na localização de defeitos.
Objetivo geral da pesquisa	Analisar e comparar a utilização de informações de teste de fluxo de controle e dados na localização de defeitos, realizando experimentos com programas industriais e usados na literatura, além de experimentos com usuários que deverão utilizar o ambiente desenvolvido para localizar os defeitos.
Trabalhos relacionados	Santelices et al. (2009) compara o uso de fluxo de dados e fluxo de controle, mostrando vantagem para o uso de fluxo de dados. Mao (2014), Alves (2011) e Ju (2014) apresentam soluções que utilizam combinações de slices (um tipo de cobertura de fluxo de dados) para a localização de defeitos. Entretanto, a coleta de slices é custosa a ponto de inviabilizar seu uso industrial e a coleta de duas utilizada por Santelices et al. (2009) é ineficiente.
Justificativa e relevância	Embora as informações de fluxo de dados tenham mostrado um desempenho melhor do que de fluxo de controle para localizar o defeito, a alta sobrecarga para coletar essa informação tem impedido a sua utilização em código de nível industrial. Esta pesquisa irá utilizar uma ferramenta desenvolvida recentemente (BA-DUA) que apresentou 38% de sobrecarga de tempo de execução para programas grandes e realizar experimentos com usuários para avaliar a sua eficácia e eficiência na prática.
Proposta para Solução	Desenvolvimento de um ambiente que possibilitará 1) coletar informações de fluxo de dados e de fluxo de controle em programas industriais e usados na literatura 2) analisar a eficácia da aplicação da técnica de localização de defeitos automatizada quando utilizada pelos desenvolvedores.
Dados	Número de comandos necessários para localizar o defeito utilizando cobertura de fluxo de controle e de fluxo de dados. Eficácia (número de acertos) e eficiência (tempo necessário) de usuários utilizando cobertura de fluxo de controle e de dados.
Forma de validação	Serão utilizados programas de código aberto com características industriais (processo de desenvolvimento e tamanho) com defeitos reais e semeados. As coberturas serão utilizadas para identificar trechos de código mais suspeitos. O número de comandos que deverão ser investigados no pior caso até o sítio do defeito serão comparados. No experimento com pessoas, serão avaliadas o número de acertos e o tempo de localização dos defeitos pelos usuários da ferramenta que será desenvolvida utilizando tanto fluxo de controle como de dados.
Limitações	Os experimentos com pessoas não apresentarão uma população suficiente para a confirmação ou refutação da hipótese de pesquisa – fluxo de dados é mais eficaz e eficiente que fluxo de controle para a localização de

	defeitos.
Resultados esperados	Contribuições científicas: Contribuições científicas: os resultados dos experimentos poderão indicar se o uso de informações do fluxo de dados devem ser utilizados amplamente para localização de defeitos automatizados.
	Contribuições tecnológicas: Contribuições tecnológicas: será desenvolvido um Plug-in para o IDE Eclipse em que será possível aplicar a técnica em qualquer projeto de software Java, esta ferramenta será disponibilizada abertamente (GitHub).

MÉTODO DE PESQUISA			
Gênero	<input type="checkbox"/> Pesquisa teórica	<input checked="" type="checkbox"/> Pesquisa prática	<input checked="" type="checkbox"/> Pesquisa empírica <input type="checkbox"/> Pesquisa metodológica
Natureza	<input type="checkbox"/> Pesquisa básica/pura	<input checked="" type="checkbox"/> Pesquisa aplicada	
Objetivo	<input type="checkbox"/> Pesquisa descritiva	<input type="checkbox"/> Pesquisa exploratória	<input checked="" type="checkbox"/> Pesquisa explicativa
Abordagem	<input checked="" type="checkbox"/> Pesquisa quantitativa	<input type="checkbox"/> Pesquisa qualitativa	<input type="checkbox"/> Pesquisa mista (quali-quant)
Procedimento(s) técnico(s)	<input checked="" type="checkbox"/> Pesquisa experimental <input type="checkbox"/> Pesquisa bibliográfica <input type="checkbox"/> Pesquisa documental <input type="checkbox"/> Pesquisa <i>ex-post-facto</i> <input type="checkbox"/> Pesquisa de levantamento	<input type="checkbox"/> Pesquisa com <i>survey</i> <input checked="" type="checkbox"/> Estudo de caso <input type="checkbox"/> Pesquisa participante <input type="checkbox"/> Pesquisa-ação <input checked="" type="checkbox"/> Pesquisa etnográfica	<input type="checkbox"/> Pesquisa netnográfica <input type="checkbox"/> Teoria fundamentada em dados (<i>grounded theory</i>) <input type="checkbox"/> Ciência do projeto (<i>Design science research</i>)
Fonte(s) de dados	<input type="checkbox"/> pesquisa de laboratório	<input type="checkbox"/> pesquisa de campo	<input checked="" type="checkbox"/> pesquisa bibliográfica
Técnica(s) / Instrumento(s) de coleta de dados	<input checked="" type="checkbox"/> medição <input checked="" type="checkbox"/> questionário <input type="checkbox"/> entrevista <input type="checkbox"/> grupos focais	<input type="checkbox"/> formulário <input checked="" type="checkbox"/> <i>benchmark</i>	<input checked="" type="checkbox"/> observação (direta / participante) <input type="checkbox"/> diário de campo / notas de campo <input type="checkbox"/> análise documental (ou de artefatos)
Técnica(s) de análise de dados	<input checked="" type="checkbox"/> Análise quantitativa: <input checked="" type="checkbox"/> Estatística descritiva <input type="checkbox"/> Estatística inferencial		<input type="checkbox"/> Análise qualitativa: <input type="checkbox"/> Análise de conteúdo <input type="checkbox"/> Análise do discurso

CRONOGRAMA																								
	2013					2014					2015													
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
Estudo da literatura																								
Revisão sistemática																								
Escrita de artigo Rev. Sis.																								
Definição da proposta																								
Preparo da qualificação																								
Exame de qualificação																								
Execução da proposta																								
Coleta/obtenção de dados																								
Análise dos dados																								
Validação da proposta																								
Escrita da dissertação																								
Depósito da dissertação																								
Escrita artigos Resultados																								

¹ Esta ficha é uma adaptação da usada no “VIII Workshop de Teses e Dissertações em Sistemas de Informação (WTDSI 2015)” realizado como parte do “XI Simpósio Brasileiro de Sistemas de Informação (SBSI 2015)”