

Testes de Aplicações Móveis: uma análise das pesquisas científicas via revisão sistemática

Ismayle S. Santos^{*}, Valéria L. L. Dantas, Rainara M. Santos, Rossana M. C. Andrade[†]

Grupo de Redes de Computadores, Engenharia de *Software* e Sistemas (GREat)
Departamento de Computação (DC) – Universidade Federal do Ceará (UFC)
Caixa Postal 12.166 – 60.455-760 – Fortaleza – CE – Brazil

{ismaylesantos, valerialelli, rainarasantos, rossana}@great.ufc.br

Resumo. *Com o avanço tecnológico das redes sem fio, os dispositivos móveis (e.g., celulares e tablets) se tornaram mais populares, resultando na demanda de aplicativos mais complexos. Fatores inerentes do ambiente móvel, tais como a mobilidade e as limitações do dispositivo móvel, exigem uma adaptação no processo tradicional de teste de software. Buscando investigar o estado atual das pesquisas na área de testes para aplicações móveis, realizou-se uma revisão sistemática em fontes de pesquisa científica nacionais e internacionais. A estratégia de busca identificou 2167 artigos potencialmente relevantes, dos quais 50 foram analisados de acordo com questões de pesquisa relativas a processo, ferramentas e desafios da área.*

Abstract. *Wireless networks have contributed to the technological advance that popularized the mobile devices usage (e.g., mobile phones), and fostered the development of applications targeted to these devices. Regarding the mobile environment, there are some issues such as mobility and mobile device limitations that demand an adjustment in the traditional process of software testing. Aiming to investigate the current state of research on testing for mobile applications, we carried out a systematic review in sources of national and international scientific research. The search strategy identified 2167 relevant papers from which 50 papers were analyzed according to the research questions concerning the process, tools and challenges in the area.*

1. Introdução

Com o avanço tecnológico das redes sem fio, os dispositivos móveis (e.g., *smartphones*, celulares e *tablets*) se tornaram mais acessíveis aos consumidores. A popularização desses dispositivos resultou na demanda de novos aplicativos que devem interagir com o usuário de uma forma intuitiva, transparente e menos intrusiva.

Nesse cenário, a mobilidade proporcionada pelas tecnologias sem fio e as restrições impostas pelos dispositivos móveis (e.g., pouca memória, baixo poder de processamento e tela pequena) tornam o desenvolvimento de aplicações móveis mais

^{*} Bolsista da CAPES com o número de processo 22001018031P5

[†] Bolsista do CNPq de Produtividade em Desenvolvimento Tecnológico e Extensão Inovadora (DT) 2 com o número de processo 314021/2009-4

complexo. Para que tais aplicações sejam confiáveis, corretas e seguras, processos de melhoria de qualidade intrínsecos ao ambiente móvel são necessários.

Em virtude disso, questões importantes de pesquisa emergem quando se trata de testes para computação móvel: Como testar todos os cenários do ambiente móvel em diferentes tipos de aplicações e dispositivos? Em qual nível de teste há o impacto das limitações desse ambiente? Como o teste de usabilidade pode identificar problemas decorrentes da capacidade do dispositivo (e.g., tamanho da tela)?

Na bibliografia são encontrados trabalhos que documentam processos de desenvolvimento específicos para aplicações móveis [Zeidler *et al.* 2007] e processo de teste específico para determinado tipo de aplicação, por exemplo, jogos [Fernandes e Dias 2006]. Apesar desses trabalhos abordarem temáticas na área de testes para aplicações móveis, não foi encontrado na literatura um esforço sistemático para coletar e sintetizar as iniciativas relacionadas ao teste de aplicações móveis. Dessa forma, este trabalho apresenta uma revisão sistemática sobre testes de aplicações móveis com o objetivo de caracterizar o estado da arte deste tópico de pesquisa.

A revisão sistemática é um meio para identificar, avaliar e interpretar todas as pesquisas disponíveis relevantes para uma questão de pesquisa ou fenômeno de interesse particular [Kitchenham 2004]. Por meio do uso de um processo controlado e formal de pesquisa bibliográfica, acredita-se que os resultados beneficiam pesquisadores e profissionais. Pesquisadores porque os resultados indicam os tópicos mais pesquisados, bem como lacunas e desafios que carecem de mais investigação. Já para os profissionais, os resultados desta pesquisa podem ser úteis como referência para adquirir e/ou desenvolver processos, ferramentas e técnicas de testes de aplicações móveis.

O restante deste artigo é estruturado da seguinte forma: na Seção 2 é descrito o processo da revisão sistemática utilizado neste trabalho; na Seção 3 são apresentados os resultados obtidos e uma análise dos mesmos com base nas questões de pesquisa; na Seção 4 é feita uma discussão sobre a validade dos resultados; por fim, na Seção 5 são sumarizadas as considerações finais bem como as perspectivas de trabalhos futuros.

2. O processo de revisão sistemática

Para execução da revisão, utilizou-se como base o guia para revisão sistemática de Kitchenham (2004), cujas fases e passos são detalhados nas próximas subseções.

2.1. Planejamento da revisão

Os passos envolvidos no planejamento da revisão são descritos a seguir:

Passo 01: Identificação da necessidade da revisão

O GREat³ atua na área de Computação Móvel desde 2001 em projetos de pesquisa e de parceria com a indústria de *software* [Andrade *et al.* 2011]. A partir da experiência profissional e de pesquisas relacionadas a aplicações móveis [Dantas *et al.* 2009, Lima *et al.* 2010] do grupo, percebeu-se a necessidade de uma revisão sistemática para elencar as iniciativas existentes para o teste de aplicações móveis. Na literatura já existem revisões sistemáticas relacionadas ao contexto móvel (e.g., Linha de Produto de

³ GREat - Grupo de Redes de Computadores, Engenharia de *Software* e Sistemas (<http://www.great.ufc.br>)

Software aplicada para o middleware móvel [Bezerra *et al.* 2009]). Contudo, com respeito ao Teste de Aplicações Móveis, não foi encontrado nenhum estudo sistemático com a sintetização das iniciativas voltadas para essa área até a realização deste trabalho.

Passo 02: Desenvolvimento do protocolo da revisão

O protocolo da revisão sistemática define as questões de pesquisa para serem respondidas ao final da revisão, descreve como a literatura será pesquisada e os métodos para sintetizar as evidências objetivando responder as questões de pesquisa [Kitchenham 2004]. Dessa forma, para execução deste trabalho, desenvolveu-se um protocolo⁴, cujos principais tópicos são descritos a seguir.

(i) Questões de Pesquisa:

Questão Principal (QP)	Qual o estado da arte sobre teste de aplicações móveis?
Secundária 1 (QS1)	Quais problemas são enfrentados durante o teste de aplicações móveis?
Secundária 2 (QS2)	Quais processos e ferramentas existem para o teste de aplicações móveis?
Secundária 3 (QS3)	Em que nível de teste as pesquisas estão mais concentradas?
Secundária 4 (QS4)	Como o teste de usabilidade é feito em aplicações móveis?

(ii) String de Busca: com base na abordagem PICO⁵ e nos termos chave das questões de pesquisa, foi elaborada a *string* de busca.

("mobile application" or "mobile applications") and test) and ("state of the art" or toward or literature or review or survey or problem or difficulty or difficulties or "lessons learned" or complexity or process or approach or method or requirement or technique or tool or framework or pattern or usability or "user design interface" or "user experience")

(iii) Fontes de Pesquisa: neste artigo, apresentam-se os resultados considerando as buscas feitas nas seguintes bases: (a) ACM Digital Library (<http://dl.acm.org/>); (b) Simpósio Brasileiro de Qualidade de Software (SBQS); (c) Simpósio Brasileiro de Engenharia de Software (SBES); e (d) Currículos Lattes (<http://lattes.cnpq.br>) dos principais pesquisadores brasileiros que atuam na área de teste de *software*, os quais foram extraídos a partir do trabalho de Durelli *et al.* (2011).

(iv) Procedimento de busca nas fontes: o procedimento de busca utilizado variou conforme a fonte de pesquisa. Na ACM utilizou-se a *string* de busca para encontrar os trabalhos relacionados com a revisão. Nos congressos nacionais e currículos *Lattes*, por ausência de um mecanismo de pesquisa, a busca foi feita através da análise manual de títulos e resumos (quando disponíveis na *web*), buscando-se pelos termos: "móvel" ou "móveis" ou "*mobile*" e "teste" ou "*test*".

*(v) Critérios de inclusão e exclusão dos estudos primários*⁶: (a) O estudo deve apresentar uma iniciativa relacionada ao teste de aplicação móvel contemplando pelo menos uma das questões de pesquisa; (b) Somente estudos publicados entre 2000 e 2011 serão considerados; (c) O artigo selecionado deve estar disponível na *web*; e (d) O estudo deve estar escrito em Português ou Inglês.

(vi) Procedimento de seleção dos estudos primários

Etapa 1. Um pesquisador aplica a estratégia de busca nas fontes (*item iv*);

⁴ Todas as informações sobre a revisão estão disponíveis em <http://www.great.ufc.br/~ismayle/revisaosistemica>

⁵ PICO - Population, Intervention, Comparison, Outcomes.

⁶ São os estudos com investigação de um determinado tópico de pesquisa em um dado contexto

Etapa 2. Para selecionar um conjunto inicial de estudos, os títulos e resumos de todos os artigos obtidos foram lidos e confrontados contra os critérios do *item v*;

Etapa 3. Por fim, os artigos selecionados na etapa 2 foram lidos por completo e novamente confrontados contra os critérios do *item v(a)*. Os artigos incluídos são documentados e encaminhados para extração dos dados, enquanto que os artigos excluídos foram documentados com uma justificativa de exclusão.

(vii) *Procedimento de extração dos dados:* a extração das informações dos artigos foi feita com base em um formulário com perguntas direcionadas a obter respostas para as questões de pesquisa da revisão.

2.2. Condução da revisão

A condução da revisão é dividida em cinco passos descritos a seguir:

Passo 01: Identificação dos estudos primários

Esta revisão sistemática foi conduzida inicialmente no primeiro semestre de 2011. Neste período, foram realizadas buscas na ACM e nos congressos nacionais. Após a publicação da lista dos 21 principais pesquisadores nacionais relacionados a testes de *software* [Durelli *et al.* 2011] no SBES 2011, foi efetuada a inclusão dos currículos desses pesquisadores como fonte de busca. Além disso, em dezembro de 2011, foi executada uma atualização dos resultados obtidos anteriormente. Como resultado final, as buscas retornaram 2167 artigos potencialmente relevantes à pesquisa, sendo 2154 da ACM, 02 do SBQS e 11 dos *Currículos Lattes* analisados.

Passo 02: Seleção dos estudos primários

A seleção dos estudos primários seguiu o procedimento definido no *item iv* na Seção 2.1. Como resultado da *Etapa 1*, 2167 artigos foram identificados. Após o confronto com os critérios de inclusão e exclusão na *Etapa 2*, foram selecionados 87 artigos, 76 da ACM e 11 dos currículos. Por fim, da *Etapa 3*, ficaram incluídos 46 artigos da ACM e 04 artigos dos *Currículos Lattes*, totalizando 50 artigos.

Passo 03: Avaliação da qualidade dos estudos

A avaliação da qualidade dos artigos dos estudos primários foi feita de forma simplificada verificando apenas a presença ou não de alguma forma de validação. Embora ela não tenha sido utilizada como critério de inclusão/exclusão, essa avaliação foi interessante por revelar que a maioria dos trabalhos apresenta alguma forma de validação dos resultados, conforme será descrito em detalhes na Seção 3.

Passos 04 e 05: Extração e Sintetização dos dados

Uma vez definido o conjunto dos artigos incluídos na revisão, efetuou-se o processo de extração dos dados de acordo com o protocolo (ver resultados na Seção 3).

3. Resultados da Revisão

Esta seção apresenta as contribuições de Pesquisadores Nacionais (PN) e Pesquisadores Internacionais (PI) identificadas durante a revisão sistemática realizada e uma discussão dos resultados com base nas questões de pesquisa investigadas.

3.1. Visão geral dos estudos

Na revisão realizada foram analisados 50 artigos. Destes, somente cinco artigos (10%) eram de pesquisadores brasileiros. Com relação ao ano de publicação, conforme as Figuras 1 e 2, os trabalhos de PN se concentraram mais nos anos de 2006 e 2007. Já com respeito aos trabalhos dos PI, estes foram publicados entre os anos de 2003 e 2011 com um pico no ano de 2007 (10 artigos), sendo o biênio 2007-2008 aquele com a maior quantidade de artigos (19 artigos – 42% do total). Vale ressaltar que o artigo mais antigo encontrado pela revisão foi publicado em 2003 na conferência CHI.

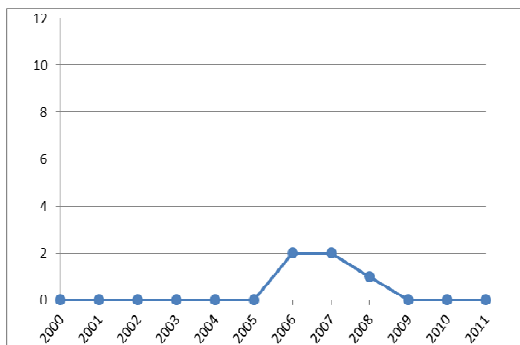


Figura 1. Artigos de PN por Ano

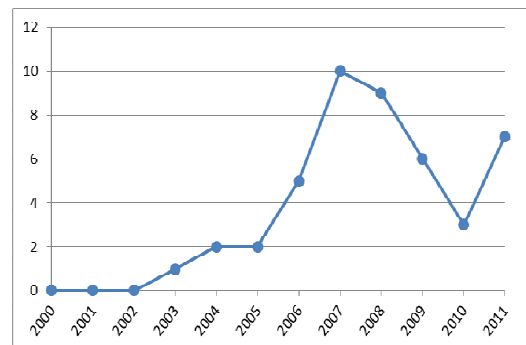


Figura 2. Artigos de PI por Ano

A Tabela 01 apresenta as fontes de publicação dos artigos dos PN. De acordo com a Tabela 01, um artigo (20%) foi publicado em periódico (SEN) enquanto que os outros quatro (80%) foram publicados em congressos ou *workshops*, sendo que destes apenas um era nacional (WTF). Por outro lado, os artigos dos PI se distribuíram em 32 locais de publicação, conforme ilustrado na Tabela 02. Os locais com maior número de artigos foram CHI (5) e Mobility (4). Destaca-se que a maior parte desses artigos foi publicada em conferências ou *workshops* (40 artigos - 89%), enquanto que uma pequena parcela (05 artigos - 11%) foi encontrada em periódicos (JECR, JSS, IPM, JUS e IM).

Tabela 01. Número de artigos dos PN por local de publicação

Local	Sigla	Nº
ACM SIGSOFT <i>Software Engineering Notes</i>	SEN	1
ACM Symposium on Applied computing	SAC	1
International Workshop on Automation of <i>Software Test</i>	AST	1
Workshop de Teste e Tolerância a Falhas	WTF	1
Workshop on Domain Specific Approaches to <i>Software Test Automation</i>	DOSTA	1

Também foram elencadas as universidades com mais publicações na área de testes de aplicações móveis. Nacionalmente, tem-se a Universidade Federal de Campina Grande (UFCG) com a pesquisadora Patrícia D. L. Machado tendo 4 artigos (80%). Já no cenário internacional, as universidades que mais se destacaram, com 02 artigos cada, foram: *Ewha Womans University*, com os pesquisadores Byoungju Choi e Heejin Kim; e *University of Lisbon*, com Luis Carriço e Marco de Sá.

Com respeito à validação dos resultados, dos artigos de PN, apenas 01 artigo (20%) não apresentou nenhuma forma de validação. Dos outros quatro, 01 artigo (20%) descreveu um experimento controlado e os outros 3 (60%) apresentaram um estudo de

caso. Já dentre os artigos de PI, 11 (25%) não apresentaram uma validação, 24 (53%) apresentaram um estudo de caso e 10 (22%) descreveram experimentos controlados.

Tabela 02. Número de artigos dos PI por local de publicação

Local	Sigla	Nº
Conference on Human Factors in Computing Systems	CHI	5
International Conference on Mobile Technology, Applications & Systems	Mobility	4
Conference on Human-Computer Interaction with Mobile Devices and Services	MobileHCI	3
Workshop on Automation of <i>Software Test</i>	AST	3
Conference on Mobile and Ubiquitous Multimedia	MUM	2
Nordic Conference on Human-Computer Interaction	NordiCHI	2
ACM International Conference on Design of Communication	SIGDOC	1
ACM SIGCOMM Workshop on Measurements Up the Stack	W-MUST	1
ACM SIGSOFT <i>Software Engineering Notes</i>	SEN	1
Australasian User Interface Conference	AUIC	1
Australian <i>Software Engineering Conference</i>	ASWEC	1
British HCI Group Annual Conference on People and Computers: HCI	BCS-HCI	1
Conference on Advanced Visual Interfaces	AVI	1
Conference on Secure <i>Software Integration and Reliability Improvement</i>	SSIRI	1
Euromicro Conference on <i>Software Engineering and Advanced Applications</i>	SEAA	1
European Conference on <i>Software Architecture</i>	ECSA	1
International Conference on <i>Software Testing, Verification and Validation Workshops</i>	ICSTW	1
Interactions Magazine	IM	1
International Conference on Business Information Systems	ICBIS	1
International Conference on Computer and Information Science	ICIS	1
International Conference on Convergence and Hybrid Information Technology	ICHIT	1
International Conference on Multimodal Interfaces	ICMI	1
International Conference on Ubiquitous Information Management and Communication	ICUIMC	1
International Symposium on <i>Software Testing and Analysis</i>	ISSTA	1
International Workshop on Advances in Model-based Testing	A-MOST	1
Journal Electronic Commerce Research	JECR	1
Journal of Information Processing and Management	IPM	1
Journal of Systems and <i>Software</i>	JSS	1
Journal of Usability Studies	JUS	1
Participatory Design Conference	PDC	1
Workshop on Dynamic Analysis	WODA	1
Workshop on Mobile Computing Systems and Applications	HotMobile	1

3.2. Estado da arte sobre teste de aplicações móveis (QP)

As contribuições das pesquisas apresentadas nos artigos analisados foram classificadas dentre as seguintes categorias: Abordagem (24%), Ferramenta (24%), *Framework* (14%), Heurística (4%), *Middleware* (2%), Técnica (8%) e Relatos⁷ (24%). As Tabelas 03 a 09 resumem essas contribuições apresentando seus objetivos e referências.

Tabela 03. Heurísticas para o teste de aplicações móveis

Objetivo	Trabalhos
Apoiar avaliações de usabilidade para computação móvel	Bertini <i>et al.</i> (2006)
Apoiar avaliação de jogos móveis	Korhonen e Koivisto (2006)

⁷ Na categoria Relatos foram incluídos trabalhos que reportam resultados de avaliação, por exemplo, avaliações de usabilidade em campo.

Tabela 04. Abordagens/Métodos para o teste de aplicações móveis

Objetivo	Trabalhos
Geração de casos de testes objetivando testar a interação de <i>features</i>	Figueiredo <i>et al.</i> (2006)
Teste de aplicações móveis em campo utilizando o LiLiPUT (<i>A flexible wearable test system</i>)	Reichl <i>et al.</i> (2007)
Avaliação de usabilidade do dispositivo móvel	Lesemann <i>et al.</i> (2007)
Testes de <i>Features</i> de aplicações móveis	Nascimento e Machado (2007)
Geração e seleção de casos de testes de interrupção	Andrade <i>et al.</i> (2007)
Avaliação de aplicações de turistas que combina métodos de recuperação de informação com técnicas de desenvolvimento de IHC	Goker e Myrhaug (2008)
Teste para sistemas móveis orientados a serviços	Bertolino <i>et al.</i> (2008)
Avaliações de usabilidade	Biel <i>et al.</i> (2010)
Simulação “híbrida” em ambientes pervasivos	Leichtenstern <i>et al.</i> (2010)
Suporte a execução, monitoramento e gerenciamentos de testes de aplicações móveis	Ridene e Barbier (2011)
Teste de aplicações móveis usando recursos da Computação nas Nuvens	Baride e Dutta (2011)
Automatizar o teste de aplicações Android, com foco em bugs de GUI	Hu e Neamtiu (2011)

Tabela 05. Ferramentas para o teste de aplicações móveis

Objetivo	Trabalhos	
AMATE (<i>Advanced Mobile Application Test Environment</i>): teste de aplicações móveis distribuídas	Binder e Hanlon (2005)	
Jabuti/ME: Avaliação de cobertura de teste estrutural	Delamaro <i>et al.</i> (2006)	
<i>AppTest</i> : testes de aplicações B3G (<i>Beyond 3G</i>) em múltiplos telefones	Puhakka e Palola (2006)	
<i>MobileTest</i> : teste caixa-preta para dispositivos móveis inteligentes	Bo <i>et al.</i> (2007)	
<i>LTS-BT</i> : ferramenta de seleção e geração de casos de testes que usa como entrada um diagrama de sequencia ou modelos de comportamento LTS	Cartaxo <i>et al.</i> (2008)	
Teste de GUI Móvel	Baseada em cenários	Hwang e Chae (2008)
	Baseado no framework MOGuT	Kwon e Hwang (2008)
Teste de Desempenho	<i>PJunit</i> : baseada no Desenvolvimento Dirigido por Testes	Kim <i>et al.</i> (2009a)
	No nível de teste de unidade	Kim <i>et al.</i> (2009b)
<i>Injector</i> : executa testes em emuladores e dispositivos	Sama e Harty (2009)	
<i>Hermes</i> : suporta o teste de caixa-preta de aplicações feitas em J2ME	She <i>et al.</i> (2009)	
Permitir que avaliadores decidam quais tarefas os usuários devem executar e obtêm dados relacionados com a interação do usuário	Carta <i>et al.</i> (2011)	

Tabela 06. Frameworks para o teste de aplicações móveis

Objetivo	Trabalhos
Especificação e validação de requisitos de desempenho	Lencevicius e Metz (2006)
Avaliações de sistemas e serviços móveis	Jensen (2007)
Projetar e implementar uma estrutura para capturar e reproduzir interações de usuário em aplicativos do ambiente <i>Java 2 Micro Edition</i>	Weiss e Zduniak (2007)
Auxiliar <i>designers</i> no processo de geração e seleção de cenários de Interação Humano-Computador (IHC)	Sá e Carriço (2008)
Teste de usabilidade de aplicações para dispositivos de mão	Au <i>et al.</i> (2008)
Apoiar a construção e avaliação de protótipos para dispositivos móveis	Sá <i>et al.</i> (2008)
Apoiar o teste de usabilidade para o ambiente móvel	Coursaris e kim (2011)

Tabela 07. Middleware para o teste de aplicações móveis

Objetivo	Trabalhos
<i>Pervaho</i> : Desenvolvimento e teste de aplicações móveis e sensíveis ao contexto para dispositivos móveis pequenos	Eugster <i>et al.</i> (2009)

Tabela 08. Relatos relacionados ao teste de aplicações móveis

Objetivo	Trabalhos
Teste de usabilidade em sites XHTML para analisar o estilo de navegação	Kaikkonen e Roto (2003)
Desenvolvimento de uma aplicação de e-mail móvel baseada no XHTML	Nikkanen (2004)
Relatar como foi o projeto, o desenvolvimento e avaliação de uma aplicação móvel multimodal através de estudos de laboratório e campo	Baillie e Schatz (2005)
Apresentar dicas e truques de usabilidade para o teste de aplicações móveis.	Schultz (2006)
Engenharia de usabilidade como um método para o desenvolvimento de mapas para aplicações móveis	Looije <i>et al.</i> (2007)
Apresentar uma avaliação de um protótipo gerador de <i>playlist</i> de celular baseado em conteúdo	Lehtiniemi e Seppänen (2007)
Analisar os resultados de avaliações do usuário final de uma aplicação móvel da plataforma <i>Symbian OS</i>	Koskela <i>et al.</i> (2007)
Avaliar dos resultados do estudo de usabilidade em que a experiência dos usuários de navegação na <i>web</i>	Shrestha (2007)
Apresentar os resultados de um teste em campo de uma aplicação móvel relacionada ao transporte público	Bergvall-Kåreborn e Larsson (2008)
Discutir algumas questões do teste de aplicações móveis, do ponto de vista tecnológico e conceitual.	Nguyen <i>et al.</i> (2008)
Relatar as experiências obtidas com a realização de estudo da experiência do usuário em campo.	Jambon e Meillon (2009)
Apresentar um resultado do estudo de usuários do telefone Android no qual eles avaliaram a qualidade de experiência	Wac <i>et al.</i> (2011)

Tabela 09. Técnicas para o teste de aplicações móveis

Objetivo	Trabalhos
Teste em campo que coleta informações sobre o uso dos dispositivos com as aplicações	Isomursu <i>et al.</i> (2004)
Simular contexto de uso móvel real em um ambiente de laboratório	Crease <i>et al.</i> (2007)
Analisar as técnicas do framework <i>MobileTest</i>	Zhifang <i>et al.</i> (2010)
Teste de regressão e de crash de aplicações Android	Amalfitano <i>et al.</i> (2011)

3.3. Problemas relacionados ao teste de aplicações móveis (QS1)

Com base nos artigos analisados, foram encontrados 13 problemas/desafios relacionados ao teste de aplicações móveis. Esses foram agrupados em uma das cinco categorias: Geral, Tecnologia Móvel, Teste em Laboratório, Teste em Campo e Teste de Usabilidade. Destaca-se que foi criada uma categoria específica para o teste de usabilidade porque uma das questões de pesquisa da revisão era voltada para esse tipo de teste. A Tabela 10 resume os problemas identificados.

De forma geral, relacionados ao teste de aplicações móveis, foram encontrados os seguintes problemas: (i) dificuldade de detecção dos defeitos relacionados à GUI devido ao alto grau de acoplamento entre componentes GUI e MIDlets; (ii) falta de formalismo para suportar atividades de V&V; (iii) altos custos, por necessitar de usuários reais, tarefas reais e tecnologia de ponta; e (iv) poucas iniciativas na área.

Vale ressaltar que segundo as análises, os testes de aplicações móveis em campo são relevantes por serem conduzidos em um ambiente real, no qual o fator mobilidade pode ser testado. No entanto, este tipo de teste é oneroso, difícil de configurar, conduzir e capturar dados. Por outro lado, a execução de testes em laboratório é uma alternativa com menor custo, mas ainda com avaliações pouco realistas.

Tabela 10. Problemas reportados no teste de aplicações móveis

Categories	Problemas/Desafios	Referências
Geral	Alto grau de acoplamento	Sama e Harty (2009);
	Custos elevados	Bertini <i>et al.</i> (2006)
	Falta de formalismo	Nguyen <i>et al.</i> (2008);
	Poucas iniciativas na área	Delamaro <i>et al.</i> (2006)
Tecnologia Móvel	Dinamicidade e Heterogeneidade	Isomursu <i>et al.</i> (2004); Bo <i>et al.</i> (2007); Sá e Carriço (2008); Zhifang <i>et al.</i> (2010)
	Mobilidade	Looije <i>et al.</i> (2007); Goker e Myrhaug (2008)
	Restrições dos Dispositivos (e.g., pouca memória, tamanho de tela e limitação de bateria)	Baillie e Schatz (2005); Puhakka e Palola (2006); Bo <i>et al.</i> (2007); Looije <i>et al.</i> (2007); Au <i>et al.</i> (2008); Kim <i>et al.</i> (2009a); Kim <i>et al.</i> (2009b)
Teste em Laboratório	Pouco realistas	Isomursu <i>et al.</i> (2004); Goker e Myrhaug (2008); Leichtenstern <i>et al.</i> (2010);
Teste em Campo	Alto custo	Leichtenstern <i>et al.</i> (2010)
	Captura dos dados	Crease <i>et al.</i> (2007); Jambon e Meillon (2009)
	Configuração	Leichtenstern <i>et al.</i> (2010); Jambon e Meillon (2009)
	Escassez de suporte	Sá <i>et al.</i> (2008)
Teste de Usabilidade	Ausência de material	Kaikkonen e Roto (2003); Schultz (2006)
	Dificuldade na coleta de dados	Jambon e Meillon (2009)

3.4. Processos e ferramentas para o teste de aplicações móveis (QS2)

Pela revisão sistemática realizada não foi encontrado nenhum processo de teste para aplicações móveis. Contudo, por meio de buscas *ad hoc* foi encontrado na literatura um processo de teste específico para jogos móveis [Fernandes e Dias 2006]. Por ele ser específico para um tipo de aplicação móvel e não ter sido capturado nas buscas desta revisão optou-se por não incluí-lo nas análises. As ferramentas encontradas por meio desta revisão são apresentadas na Tabela 05 da Seção 3.2.

3.5. Nível de teste com mais pesquisas (QS3)

Adotando a definição da IEEE (2004) para níveis de teste, 48 artigos (96%) analisados nesta revisão focam no Teste de Sistema enquanto que apenas 02 (4%) são relacionados ao Teste de Unidade. Vale ressaltar que a classificação dos artigos em relação aos níveis de testes não foi uma tarefa trivial, pois essa informação nem sempre era explícita no artigo e, dessa forma, exigiu uma análise mais detalhada.

3.6. Teste de usabilidade de aplicações móveis (QS4)

Nesta revisão foram identificados 19 artigos relacionados ao teste de usabilidade para aplicações móveis. Destes, existem trabalhos que relatam resultados de testes de usabilidade em sites XHTML com telefones móveis [Kaikkonen e Roto, 2003], em protótipos de geradores de *playlist* de celular [Lehtiniemi e Seppänen, 2007], em aplicações *web* móvel [Shrestha, 2007], em aplicações móveis relacionadas à saúde e bem-estar [Koskela *et al.* 2007], em aplicações de *e-mail* móvel [Nikkanen, 2004] e também em aplicações móveis voltadas ao esporte [Jambon e Meillon, 2009].

Para auxiliar o teste de usabilidade de aplicações móveis foram encontrados trabalhos que apresentam dicas de usabilidade [Schultz, 2006], heurísticas de usabilidade [Bertini *et al.*, 2006] e heurísticas de jogabilidade para avaliação de jogos

móveis [Korhonen e Koivisto, 2006]. A revisão também identificou trabalhos voltados para a engenharia de usabilidade de aplicações móveis [Looije et al., 2007], *frameworks* [Sá e Carriço 2008, Au et al. 2008, Jensen 2007, Coursaris e Kim 2011], uma ferramenta para avaliação de usabilidade [Carta et al. 2011] e uma abordagem para o teste de usabilidade de aplicações móveis [Lesemann et al., 2007].

Além disso, foram encontrados três trabalhos com propostas diferenciadas. Biel et al. (2010), apresentam como a análise de arquitetura de *software* e avaliações de usabilidade podem ser combinados e como os resultados se relacionam entre si. Reichl et al. (2007) descrevem o LiLiPUT (*a flexible wearable test system*) e mostram como utilizá-lo para realizar vários tipos de testes de aplicações móveis em campo. Leichtenstern et al. (2010) apresentam uma abordagem para avaliação de estágios iniciais de aplicações móveis onde telefones móveis reais são usados como dispositivos de interação para uma simulação virtualizada de um ambiente *pervasivo*.

4. Discussão sobre a validade dos resultados

Como ameaças à validade dos resultados da revisão sistemática realizada tem-se: (i) tendência na identificação dos pesquisados brasileiros mais atuantes na área de testes de *software*; (ii) currículos *Lattes* dos pesquisadores brasileiros desatualizados; (iii) vieses na seleção dos estudos analisados; e (iv) imprecisão na extração dos dados.

No que diz respeito à lista dos pesquisadores brasileiros, ela se refere aos pesquisadores com mais publicações no SBES na área de testes. Essa lista foi publicada por Durelli et al. (2011) no SBES, um dos eventos nacionais mais relevantes para a comunidade brasileira de Engenharia de *Software*. Além disso, os currículos extraídos da plataforma *Lattes* foram analisados de acordo com as informações disponibilizadas pelos próprios pesquisadores. Vale ressaltar que dos 21 pesquisadores, apenas 3 estavam com seus currículos *Lattes* desatualizados (com mais de 1 ano sem atualização).

A seleção dos estudos analisados seguiu o procedimento de seleção dos estudos primários (*item vi* do Passo 2 da Seção 2.1). No caso da biblioteca digital ACM, visando minimizar a perda na extração de dados, foi utilizado o modo de busca avançado sem restringir o local da busca, o que justifica a grande quantidade de artigos retornados. No caso das buscas dos trabalhos publicados nos congressos nacionais e os elencados nos currículos *Lattes*, alguns trabalhos que não utilizaram nos seus títulos e/ou resumos as palavras chave da *string* de busca podem não ter sido selecionados.

O SBQS e o SBES foram usados como fonte de busca por serem os principais simpósios nacionais de Engenharia de *Software*. Logo, com as buscas nestes simpósios e nos currículos dos principais pesquisadores brasileiros de teste de *software*, acredita-se ter alcançado uma parte significativa das pesquisas nacionais potencialmente relevantes para esta revisão. É importante destacar que durante a revisão percebeu-se que boa parte das publicações nacionais não estão disponíveis na *web*. Contudo, a maioria dos artigos nacionais foi eliminada por não estar relacionada a alguma questão de pesquisa.

Por fim, para minimizar vieses na extração dos dados, esta foi feita em pares e no caso de impasse, discussões eram feitas com um terceiro revisor até chegar a um consenso.

5. Considerações Finais

Este artigo apresentou os resultados de uma revisão sistemática no contexto de teste de aplicações móveis. A abordagem de pesquisa utilizada foi interessante, pois permitiu, dentre outras coisas, a descrição do estado atual da área.

Na revisão sistemática realizada, 50 artigos foram analisados. Os resultados trazem indícios de que não existe um processo específico para testes de aplicações móveis, embora já existam várias ferramentas disponíveis. As pesquisas têm se concentrados mais no nível de teste de sistema e as dificuldades mais recorrentes ao testar uma aplicação móvel são advindos das restrições dos dispositivos móveis (e.g, tela pequena e pouca memória). Por fim, os 19 trabalhos (38% do total) encontrados com foco no teste de usabilidade fornecem indícios de que as restrições dos dispositivos móveis e a mobilidade do usuário enquanto usa a aplicação móvel, são uma preocupação tanto da área de Teste como de Interação Humano-Computador (IHC).

Para a continuação e aprimoramento deste trabalho, pretende-se como trabalhos futuros a extensão da revisão sistemática em outras fontes de pesquisa, por exemplo, a IEEE, *Scopus*, *Science Direct* e DBLP e também a extensão para o contexto da Computação Ubíqua e Pervasiva.

Referências

- Andrade, R.M.C.; Maia, P.H.M.; Castro, R.N.S.; Dantas, V.L.L.; de C. Andrade, T., (2011). Industry is from mars, academia is from venus. *In Proceedings of the 25th Brazilian Symposium on Software Engineering*, pp. 108–113.
- Andrade, W. L. ; Neto, F. G. O.; Machado, P. D. L. (2007). Geração de Casos de Teste de Interrupção para Aplicações de Celulares. *In VIII Workshop de Teste e Tolerância a Falhas (WTF 2007)*, Belém. Porto Alegre: SBC, 2007, vol. 1, pp. 129-142.
- Amalfitano, D.; Fasolino, R. A.; Tramontana, P. (2011). A GUI Crawling-Based Technique for Android Mobile Application Testing. *In Proceedings of the 2011 IEEE Fourth International Conference on Software Testing, Verification and Validation Workshops*. IEEE Computer Society, Washington, USA, pp. 252-261.
- Au, F. T. W.; Baker, S.; Warren, I.; Dobbie, G. (2008) Automated usability testing framework. *In Proceedings of the ninth Conference on Australasian User Interface*, vol. 76, Australian Computer Society, Inc., Darlinghurst, Australia, pp. 55-64
- Baillie, L.; Schatz, R. (2005) Exploring multimodality in the laboratory and the field. *In Proceedings of the 7th International Conference on Multimodal Interfaces*, pp.100-107.
- Baride, S.; Dutta, K. (2011). A cloud based *software* testing paradigm for mobile applications. *In SIGSOFT Software. Eng. Notes*, vol. 36, n. 3 (May 2011), pp. 1-4.
- Bergvall-Kåreborn, B.; Larsson, S. (2008) A case study of real-world testing. *In Proceedings of the 7th International Conference on Mobile and Ubiquitous Multimedia (MUM '08)*. ACM, New York, NY, USA, pp. 113-116.

- Bertini, E; Gabrielli, S.; Kimani, S (2006) Appropriating and assessing heuristics for mobile computing. *In Proceedings of the Working Conference on Advanced Visual Interfaces (AVI '06)*. ACM, New York, NY, USA, pp. 119-126.
- Bertolino, A.; Angelis, G.; Lonetti, F.; Sabetta, A. (2008) Let The Puppets Move! Automated Testbed Generation for Service-oriented Mobile Applications. *In Proceedings of the 2008 34th Euromicro Conference Software Engineering and Advanced Applications*. IEEE Computer Society, DC, USA, pp. 321-328.
- Bezerra, Y. M., Pereira, T. A. B.; Silveira, G. E. (2009) A Systematic Review of Software Product Lines Applied to Mobile Middleware, *In Proceedings of the 2009 Sixth International Conference on Information Technology*, DC, USA, pp. 1024-1029.
- Biel, B.; Grill, T.; Gruhn, V. (2010) Exploring the benefits of the combination of a software architecture analysis and a usability evaluation of a mobile application. *In Journal of Systems and Software*, vol. 83, n. 11, pp. 2031-2044.
- Binder, R. V.; Hanlon, J. E. (2005) The advanced mobile application testing environment. *In Proceedings of the 1st International Workshop on Advances in Model-Based Testing (A-MOST '05)*. ACM, New York, NY, USA, pp. 1-1.
- Bo, J.; Xiang, L.; Xiaopeng, G. (2007) MobileTest: A Tool Supporting Automatic Black Box Test for Software on Smart Mobile Devices. *In Proceedings of the 2nd International Workshop on Automation of Software Test*. IEEE, Washington, USA.
- Carta, T.; Paternò, F.; Santana, V. (2011). Support for remote usability evaluation of web mobile applications. *In Proceedings of the 29th ACM international conference on Design of communication (SIGDOC '11)*. ACM, NY, USA, pp. 129-136.
- Cartaxo, E. G.; Andrade, W. L.; Neto, F. G. O.; Machado, P. D. L. (2008). LTS-BT: a tool to generate and select functional test cases for embedded systems. *In Proceedings of the 2008 ACM Symposium on Applied Computing*, pp.1540-1544.
- Coursaris, K. C.; Kim, J. D. (2011). A Meta-Analytical Review of Empirical Mobile Usability Studies. *Journal of Usability Studies* vol. 6, n. 3, Article 11, 55 pages.
- Crease, M.; Lumsden, J.; Longworth, B. (2007) A technique for incorporating dynamic paths in lab-based mobile evaluations. *In Proceedings of the 21st British HCI Group Annual Conference on People and Computers*, vol. 1, Swinton, UK, pp. 99-108.
- Dantas, V. L.L.; Marinho, F. G.; Costa, A. L.; Andrade, R. M. C. (2009). Testing Requirements for Mobile Applications. *In: Proceedings of the 24th International Symposium on Computer and Information Sciences - ISCIS 2009*. Institute of Electrical and Electronics Engineers (IEEE) Xplore Digital Library, pp. 555-560.
- Delamaro, M. E.; Vincenzi, A. M. R.; Maldonado, J. C. (2006) A strategy to perform coverage testing of mobile applications. *In Proceedings of the 2006 international workshop on Automation of software test*. ACM, New York, USA, pp. 118-124.
- Durelli, V. H. S.; Araujo, R. F.; Silva, M. A. G.; Oliveira, R. A. P.; Maldonado, J. C.; Delamaro, M. E. (2011) What a Long, Strange Trip It's Been: Past, Present and Future Perspectives on Software Testing Research. *In Proceedings of the Brazilian Symposium on Software Engineering*.

- Eugster, P.; Garbinato, B.; Holzer, A. (2009) Pervaho: A specialized middleware for mobile context-aware applications. *In Journal Electronic Commerce Research*, vol. 9, n. 4 (December 2009), pp. 245-268.
- Fernandes, T.; Dias, A. (2006). Processo de Testes para Desenvolvimento de Jogos Celulares. *In: Proceedings of the V Brazilian Symposium on Computer Games and Digital Entertainment (SBGames Computing)*. Short Paper. 2006.
- Figueiredo, A. L. L.; Andrade, W. L.; Machado, P.D.L. (2006). Generating interaction test cases for mobile phone systems from use case specifications. *In SIGSOFT Software Engineering*. Notes vol. 31, nº 6 (November 2006), pp. 1-10.
- Goker, A.; Myrhaug, H. (2008). Evaluation of a mobile information system in context. *In Journal Information Processing and Management: an International Journal*, vol. 44, n. 1 (January 2008), pp. 39-65.
- Hwang, S.; Chae, H. (2008) Design & Implementation of Mobile GUI Testing Tool. *In Proceedings of the 2008 International Conference on Convergence and Hybrid Information Technology*. IEEE Computer Society, DC, USA, pp. 704-707.
- Hu, C.; Neamtiu, I. (2011). Automating GUI testing for Android applications. *In Proceedings of the 6th International Workshop on Automation of Software Test (AST '11)*. ACM, New York, NY, USA, pp. 77-83.
- IEEE (2004). Guide to the *Software Engineering Body of Knowledge*. IEEE Society.
- Isomursu, M.; Kuutti, K.; Vainamo, S. (2004) Experience clip: method for user participation and evaluation of mobile concepts. *In Proceedings of the eighth Conference on Participatory Design: Artful integration: interweaving media, materials and practices - Volume 1 (PDC 04)*, Vol. 1. ACM, New York, NY, USA, pp. 83-92.
- Jambon, F.; Meillon, B. (2009) User experience evaluation in the wild. *In Proceedings of the 27th International Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '09)*. ACM, New York, NY, USA, pp. 4069-4074.
- Jensen, K. L. (2007) New methods for conducting experiments with mobile systems and services in the field. *In Proceedings of the 9th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '07)*. ACM, New York, NY, USA, pp. 167-170.
- Kaikkonen, A.; Roto, V. (2003). Navigating in a mobile XHTML application. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, pp. 329-336.
- Kim, H.; Choi, B.; Yoon, S. (2009a) Performance testing based on test-driven development for mobile applications. *In Proceedings of the 3rd International Conference on Ubiquitous Information Management and Communication (ICUIMC '09)*. ACM, New York, NY, USA, pp. 612-617.
- Kim, H.; Choi, B.; Wong, W. E. (2009b) Performance Testing of Mobile Applications at the Unit Test Level. *In Proceedings of the 2009 Third IEEE International Conference on Secure Software Integration and Reliability Improvement (SSIRI '09)*. IEEE Computer Society, Washington, DC, USA, pp. 171-180.

- Kitchenham, B.(2004) Procedures for Performing Systematic Reviews, Joint Technical Report *Software Engineering Group*, Department of Computer Science Keele University, United King and Empirical *Software Engineering*,National ICT Australia.
- Korhonen, H.; Koivisto, Elina M. I. (2006) Playability heuristics for mobile games. *In Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '06)*. ACM, New York, NY, USA, pp. 9-16.
- Koskela, T.; Kassinen, O.; Ala-Kurikka, J.; Ylianttila, M. (2007) Empirical evaluation of mobile peer-to-peer wellness group communication application for Symbian OS. *In Proceedings of the 6th International Conference on Mobile and Ubiquitous Multimedia (MUM '07)*. ACM, New York, NY, USA, pp. 98-102.
- Kwon, O.; Hwang, S. (2008) Mobile GUI Testing Tool based on Image Flow. *In Proceedings of the Seventh IEEE/ACIS International Conference on Computer and Information Science*. IEEE Computer Society, Washington, DC, USA, pp. 508-512.
- Lencevicius, R.; Metz, E. (2006) Performance assertions for mobile devices. *In Proceedings of the International Symposium on Software Testing and Analysis*. ACM, pp. 225-232.
- Lehtiniemi, A.; Seppänen, J. (2007) Evaluation of automatic mobile playlist generator. *In Proceedings of the 4th International Conference on Mobile Technology, Applications, and Systems*. ACM, New York, NY, USA, pp. 452-459.
- Leichtenstern, K.; André, E.; Rehm, M. (2010) Using the hybrid simulation for early user evaluations of pervasive interactions. *In Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, pp. 315-324.
- Lesemann, E.; Woletz, N.; Koerber, S. (2007) Combining methods to evaluate mobile usability. *In Proceedings of the 9th International Conference on Human Computer Interaction with Mobile Devices and Services*, ACM, NY, USA, pp. 444-447.
- Lima, F. F. P.; Nogueira, T. P.; Braga, A. D.; Neto, J. S. R; Filho, J. B. F.; Lima, C. R. F.; Leite, S. J. C.; Viana, W.; Dantas, V. L. L.; Andrade, R. M. C. (2010) Manutenção Adaptativa de *Software* Embarcado para Telefones Celulares Apoiado por Ferramentas de Automação, In *IX Simpósio Brasileiro de Qualidade de Software*
- Looije, R.; Brake, G. M.; Neerincx, M. A. (2007) Usability engineering for mobile maps. *In Proceedings of the 4th International Conference on Mobile Technology, Applications, and Systems*. ACM, New York, NY, USA, pp. 532-539.
- Nascimento, L. H. O. and Machado, P. D. L. (2007). An experimental evaluation of approaches to feature testing in the mobile phone applications domain. *In Workshop on Domain Specific Approaches to Software Test Automation: in conjunction with the 6th ESEC/FSE joint meeting (DOSTA '07)*. ACM, NY, USA, pp. 27-33.
- Nguyen, M. D.; Waeselynck, H.; Riviere, N. (2008) Testing mobile computing applications: toward a scenario language and tools. *In Proceedings of the 2008 International Workshop on Dynamic Analysis*. ACM, NY, USA, pp. 29-35.
- Nikkanen, M. (2004) User-centered development of a browser-agnostic mobile e-mail application. *In Proceedings of the third Nordic Conference on Human-Computer Interaction (NordiCHI '04)*. ACM, New York, NY, USA, pp. 53-56.

- Puhakka, T.; Palola, M. (2006) Towards automating testing of communicational B3G applications. *In Proceedings of the 3rd International Conference on Mobile Technology, Applications & Systems (Mobility '06)*. ACM, New York, NY, USA.
- Reichl, P.; Froehlich, P.; Baillie, L.; Schatz, R.; Dantcheva, A. (2007) The LiLiPUT prototype: a wearable lab environment for user tests of mobile telecommunication applications. *In CHI '07 Extended Abstracts on Human Factors in Computing Systems (CHI EA '07)*. ACM, New York, NY, USA, pp. 1833-1838.
- Ridene, Y. and Barbier, F. (2011). A model-driven approach for automating mobile applications testing. *In Proceedings of the 5th European Conference on Software Architecture: Companion Volume (ECSA '11)*. ACM, New York, NY, USA, 7 pages.
- Sá, M.; Carriço, L. (2008) Defining scenarios for mobile design and evaluation. *In CHI '08 extended abstracts on Human factors in Computing Systems (CHI EA '08)*. ACM, New York, NY, USA, pp. 2847-2852.
- Sá, M.; Carriço, L.; Duarte, L.; Reis, T. (2008) A framework for mobile evaluation. *In CHI '08 Extended Abstracts on Human Factors in Computing Systems (CHI EA '08)*. ACM, New York, NY, USA, pp. 2673-2678.
- Sama, M.; Harty, J. (2009) Using code instrumentation to enhance testing on J2ME: a lesson learned with JInjector. *In Proceedings of the 10th Workshop on Mobile Computing Systems and Applications (HotMobile '09)*. ACM, New York, USA.
- Schultz, D. (2006) 10 usability tips & tricks for testing mobile applications. *Interactions* vol. 13, n. 6 (November 2006), pp.14-15.
- She, S; Sivapalan, S.; Warren, I. (2009) Hermes: A Tool for Testing Mobile Device Applications. *In Proceedings of the 2009 Australian Software Engineering Conference (ASWEC '09)*. IEEE Computer Society, DC, USA, pp. 121-130.
- Shrestha, S. (2007) Mobile web browsing: usability study. *In Proceedings of the 4th International Conference on Mobile Technology, Applications, and Systems*. ACM, NY, USA, pp. 187-194.
- Zeidler, C.; Kittl, C. e Petrovic, O. (2007). An Integrated Product Development Process for Mobile Software. *In: Proceedings of the IX International Conference on the Management of Mobile Business (ICBM), 2007*, p. 23-31.
- Zhifang, L.; Bin, L.; Xiaopeng, G. (2010). Test automation on mobile device. *In Proceedings of the 5th Workshop on Automation of Software Test*, pp. 1-7.
- Weiss, D.; Zduniak, M. (2007) Automated integration tests for mobile applications in java 2 micro edition. *In Proceedings of the 10th International Conference on Business Information Systems (BIS'07)*, Springer-Verlag, Berlin, Heidelberg, pp. 478-487.
- Wac, K.; Ickin, S.; Hong, J; Janowski, L.; Fiedler, M.; Dey, K. A. (2011). Studying the experience of mobile applications used in different contexts of daily life. *In Proceedings of the 1st ACM SIGCOMM Workshop on Measurements Up the Stack*. ACM, pp. 7-12.