

Um modelo para avaliação da usabilidade a partir de dados reais de interação dos usuários

Artur Henrique Kronbauer
PMCC – UFBA/UNIFACS/UEFS
Av. Adhemar de Barros s/n sala138
Salvador – BA, Brasil
CEP 40.170-110
arturhk@gmail.com

Celso A. S. Santos
PMCC – UFBA/UNIFACS/UEFS
Av. Adhemar de Barros s/n sala138
Salvador – BA, Brasil
CEP 40.170-110
saibel@dcc.ufba.br

ABSTRACT

This paper presents an ongoing PhD work involves interface evaluation based on real user interaction data, including contextual aspects. The model involves aspects of usability metrics establishment, mapping and instrumentation of the application's source code from experts' interaction, final user interaction capture and analysis of the results obtained. The objective is to automatically gather quantitative data so that can be analyzed in a qualitative and subjective way, making it possible to detect usability problems and measure the quality of use of the interfaces.

Categories and Subject Descriptors

H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems.

General Terms

Design, Performance, Measurement.

Keywords

Usability Evaluation, Mobile, Automatic Evaluation, Model.

1. INTRODUÇÃO

Este artigo apresenta um projeto de doutorado, desenvolvido pelo primeiro autor sob a orientação do segundo, dentro do Programa Multiinstitucional em Ciência da Computação (PMCC), envolvendo a Universidade Federal da Bahia (UFBA), Universidade Salvador (UNIFACS) e Universidade Estadual de Feira de Santana (UEFS). O trabalho se iniciou no primeiro semestre de 2009 e deve ser concluído em meados de 2013.

O foco da pesquisa abrange metodologias e técnicas para avaliações da usabilidade de aplicativos que serão executados em plataformas convergentes, dotadas de recursos tecnológicos como mobilidade, informações de posicionamento, novos paradigmas de interação (gestos, telas sensíveis, voz), entre outros.

A motivação do trabalho está na necessidade de se avaliar o impacto que as mudanças tecnológicas trazem para o cotidiano das pessoas e como elas afetam a realização das suas tarefas diárias. Os recursos tecnológicos disponibilizados pelos mais recentes dispositivos computacionais exigem o uso de novas técnicas, métodos e metodologias de desenvolvimento de interfaces, além de novas formas de avaliação das aplicações geradas dentro desta nova realidade: um número cada vez maior e mais diversificado de usuários utilizando recursos computacionais em diferentes contextos.

Do ponto de vista da Interação Humano Computador (IHC), um dos desafios a ser vencido neste trabalho é a criação de uma metodologia de avaliação que possa contemplar as emoções, crenças, preferências, percepções, reações físicas e psicológicas, comportamentos e realizações do usuário que ocorrem antes, durante e após o uso de um sistema. Esse conjunto de informações é denominado de *Design da Experiência do Usuário (User Experience ou UX)* [12]. Embora este seja um tema abordado amplamente na área de IHC, poucas metodologias, modelos e técnicas conseguem atender à maioria dos requisitos de avaliação propostos pela *UX* [1].

Nesse sentido, o principal objetivo deste trabalho é prover o desenvolvimento de uma solução que possibilite realizar testes de usabilidade contextualizados ao ambiente de uso dos dispositivos e que abranja as diferentes características de seus usuários. Para isso, é proposto um modelo para a obtenção de dados quantitativos a respeito da interação dos usuários com os aplicativos. Posteriormente, é proposto que esses dados sejam interpretados de forma inteligente, provendo informações qualitativas e subjetivas, possibilitando avaliações contextualizadas ao que se propõem a *UX*.

O restante deste artigo está estruturado em cinco seções. A seção 2 aborda trabalhos correlatos. A seção seguinte trata da metodologia usada, seguida da seção 4, que descreve o modelo proposto. A seção 5 mostra o estado atual do trabalho, apresentando os resultados alcançados e as perspectivas e desafios a serem transpostos na sequência da pesquisa.

2. TRABALHOS CORRELATOS

O ponto de partida para este estudo teve origem no trabalho publicado por Moldovan e Tarta [10], onde os autores investigam a utilização de Programação Orientada a Aspectos (POA) para implementar métricas de usabilidade em aplicativos para plataformas *WIMP*, com o objetivo de avaliar a sua usabilidade.

Apesar do trabalho se restringir apenas a computadores pessoais, nos quais as interações ocorrem, geralmente, com usuários em uma única posição e com pouca influência externa, os autores comprovam a eficiência da POA para a instrumentação de métricas que permitem a rastreabilidade das ações dos usuários.

Outro trabalho que serve de referência para este artigo, foi a ferramenta MyExperience¹ que apresenta uma arquitetura baseada em sensores possibilitando a captura de interações dos usuários com seu dispositivo móvel. A ferramenta permite a captura de dados quantitativos e qualitativos no cenário natural de uso. Embora não tenha sido proposta para avaliações diretas de interfaces, permite medir a utilização de aplicativos e recursos disponibilizados nos equipamentos. Com ela, é possível relacionar a frequência de uso de aplicativos com horários, datas e localização geográfica. Para isso, a ferramenta monitora o uso do GPS, câmera de vídeo, calendário, chamadas telefônicas, etc.

No *International Workshop on Meaningful Measures: Valid Useful User Experience Measurement (VUUM)*², realizado em julho de 2008, foram apresentados e discutidos os principais trabalhos relacionados a avaliações de *Design* da Experiência do Usuário (UX). Entre eles, podemos destacar o trabalho de Ardito et. al. [1] que apresenta uma avaliação metodológica que combina diferentes técnicas para avaliar o UX, propiciando a coleta de dados quantitativos e qualitativos. A metodologia de avaliação foi testada em um jogo educacional executado em dispositivos móveis. A diferença com relação ao nosso trabalho é que são utilizadas diversas metodologias de avaliação, enquanto propomos uma investigação automática e “transparente” à percepção do usuário final.

Outro assunto que nos ajuda a fundamentar esse trabalho, também discutido no *workshop* citado anteriormente, refere-se a divergência entre avaliações realizadas em laboratório e avaliações realizadas em campo. Avouris [2] apresenta os pontos fortes e fracos de cada metodologia e as contextualiza a avaliações para dispositivos móveis.

3. METODOLOGIA

Conforme apresentado, o trabalho de Tese busca estudar formas de avaliar a usabilidade de aplicações executadas em plataformas móveis e multimodais sob a ótica da UX.

O trabalho se iniciou com o levantamento dos trabalhos seminais [4] [5] [11] e das publicações mais recentes [3] [10] para se estabelecer o estado da arte das áreas relacionadas ao tema do trabalho. A partir deste levantamento, foram examinadas as metodologias e técnicas de avaliação existentes e utilizadas com a finalidade de buscar subsídios que pudessem dar suporte técnico para atingir os objetivos esperados na pesquisa.

A técnica abordada na literatura e escolhida por ser apropriada para o tipo de avaliação pretendida é chamada de Captura Automática Diretamente da Aplicação [5]. Esta é uma técnica de avaliação de usabilidade que consiste na monitoração e coleta automática de dados referentes ao uso dos aplicativos. Os dados a serem obtidos estão associados a métricas de usabilidade que quantificam as interações dos usuários sobre as interfaces das aplicações. Esses dados permitem identificar o padrão de

comportamento dos usuários, tempo gasto para concluir uma tarefa, taxas de erros, frequência de uso de mecanismos de ajuda, entre outros [4].

A partir da avaliação dos dados capturados podem ser feitos ajustes de procedimentos operacionais, aperfeiçoamento de tarefas e personalização dos aplicativos para se adequarem à utilização de grupos de usuários de acordo com características similares de uso. Além disso, é possível coletar dados relativos a um grande número de usuários, atuando em diferentes circunstâncias, de um modo discreto e transparente [11].

Um dos desafios desta técnica é romper a barreira da análise direta das ações que podem ser monitoradas, permitindo análises inteligentes para identificar como as pessoas se sentem quando interagem com um sistema. Nesse sentido, os esforços desse trabalho direcionam-se para a obtenção de dados quantitativos que possam ser interpretados, provendo informações qualitativas e subjetivas a respeito da interação do usuário final, dessa forma, indo de encontro à proposta da UX.

A técnica de Captura Automática de Dados impõe outro desafio: a integração do código associado às métricas de usabilidade com os sistemas submetidos à avaliação, uma vez que esses códigos se tornam intrusivos no contexto das aplicações [10].

Diante desse fato, buscou-se uma metodologia na área de Engenharia de Software que proporcionasse a solução desse problema. Após investigações exploratórias, constatou-se que a utilização da Programação Orientada a Aspectos (POA) [5] permite a instrumentação das aplicações a serem investigadas com métricas de avaliação, sem interferência direta no código original das aplicações [10]. A POA permite a separação dos interesses pela criação de uma nova dimensão chamada de Aspecto. Dessa forma, os interesses funcionais (relacionados ao domínio da aplicação) são codificados nas Classes e os interesses sistêmicos (os recursos de avaliação da usabilidade) nos Aspectos.

O próximo passo do trabalho foi idealizar um modelo que pudesse contemplar os requisitos necessários para a avaliação proposta. Para a validação do modelo e verificação de suas potencialidades, são previstas investigações exploratórias pela instrumentação do código de aplicações piloto que executem em dispositivos móveis, como *smartphones* e *tablets*. Para contemplar avaliações sob a ótica da UX, as aplicações pilotos devem ser submetidas a interações reais com usuários em seu ambiente natural, possibilitando avaliações quantitativas, qualitativas e subjetivas com dados coletados em campo.

4. O MODELO PROPOSTO

O modelo proposto tem como base três fases distintas, que abrangem todos os passos envolvidos no processo de avaliação idealizado neste trabalho:

1. O mapeamento da interação do Engenheiro de Testes de uma aplicação, com relação à execução das tarefas que a aplicação provê a seus usuários. Esse mapeamento deve ser realizado em vários contextos permitindo, assim, diferentes investigações.
2. A instrumentação do código da aplicação com métricas de usabilidade para possibilitar a coleta automática de dados relacionados à interação do usuário final com a aplicação. As métricas adotadas são escolhidas de acordo com o contexto de investigação baseando-se, por exemplo, nas características do

¹ <http://myexperience.sourceforge.net>

² <http://141.115.28.2/cost294/upload/523.pdf>

aparelho usado (com GPS, 3G, etc), cenário de uso (em movimento, interagindo com outras pessoas, etc) e características do usuário (idade, escolaridade, etc).

3. A avaliação dos resultados obtidos nas duas fases anteriores tem a finalidade de obter informações contextualizadas às particularidades da aplicação, cenário de uso e tipo de dispositivo usado, com o objetivo de mensurar a qualidade de uso da interface disponibilizada.

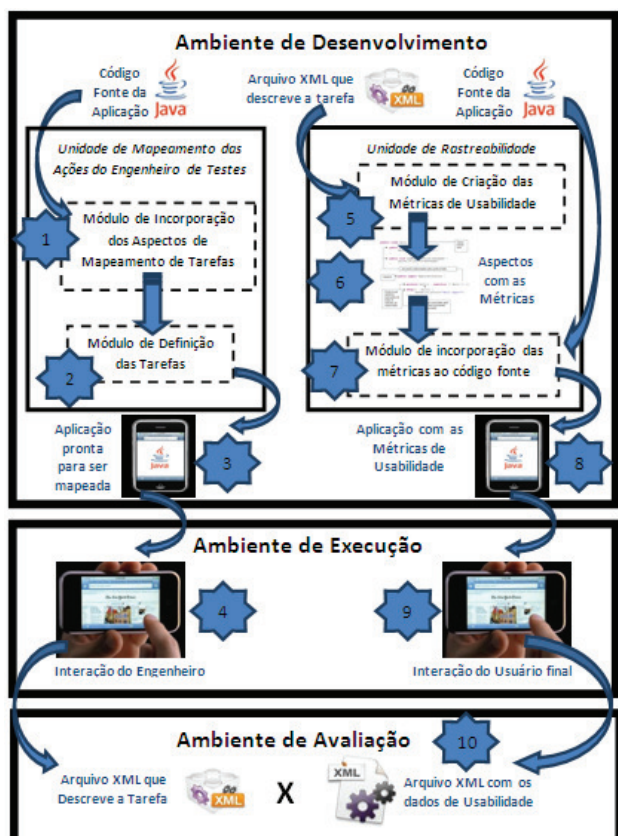


Figura 1. Modelo proposto.

Como pode ser observado na Figura 1, o modelo está estruturado em três ambientes distintos: ambiente de desenvolvimento, ambiente de execução e ambiente de avaliação.

3.1 Ambiente de Desenvolvimento

Nesse ambiente são tratados todos os aspectos de instrumentação do código fonte das aplicações que serão avaliadas. Ele está dividido em duas unidades: (i) Unidade de Mapeamento das Ações do Engenheiro de Teste (etapas 1 e 2) e (ii) Unidade de Rastreabilidade da interação do usuário final (etapas 5, 6 e 7).

Na Unidade de Rastreabilidade das Ações do Engenheiro de Teste, são criados aspectos que permitam guardar os passos executados pelo Engenheiro no momento em que define como realiza uma tarefa. Para isso, é criado um *Aspecto* que permite identificar todos os métodos executados no caminho da tarefa. Com a adição de um filtro para guardar apenas os métodos que tratam eventos do usuário, é possível identificar as interações necessárias com o aplicativo para concluir a tarefa. A lógica para

realizar essa atividade é definida no Módulo de Incorporação dos Aspectos de Mapeamento de Tarefas (etapa 1 do modelo).

O Módulo de Definição de Tarefas (etapa 2) prevê que o aplicativo seja compilado com AspectJ, tornando a aplicação apta para ser mapeada (etapa 3).

Na Unidade de Rastreabilidade, o Módulo de Criação das Métricas de Usabilidade (etapa 5) recebe como entrada o arquivo XML (criado na etapa 4) com os métodos que descrevem os passos de uma tarefa. Isso permite associar métricas aos métodos de acordo com as necessidades de investigação. O resultado desse módulo é a geração automática de um Aspecto com as métricas de usabilidade relacionado a tarefa que será rastreada (etapa 6).

No Módulo de Incorporação das Métricas ao Código Fonte (etapa 7), é compilado o aplicativo com o Aspecto gerado anteriormente, resultando na instrumentação da aplicação para ser rastreada durante a interação do usuário final (etapa 8).

3.2 Ambiente de Execução

Nesse ambiente são previstas as interações do Engenheiro de Teste e do Usuário Final. O Engenheiro de Teste interage com o aplicativo para mapear as tarefas. O resultado dessa interação é o arquivo XML com os métodos referentes a execução da tarefa e alguns dados de usabilidade sob a ótica do Engenheiro. Esse arquivo será usado para dar continuidade a instrumentação do código fonte da aplicação (etapa 5), bem como será uma das entradas de dados para o Ambiente de Avaliação (etapa 10).

Por outro lado, o Ambiente de Execução também prevê a interação do usuário final. É nesse momento que as interações do usuário serão rastreadas, gerando os dados quantitativos que poderão ser interpretados posteriormente no Ambiente de Avaliação.

3.3 Ambiente de Avaliação

Nesse ambiente os arquivos XML gerados na interação do Engenheiro e do usuário final serão interpretados, permitindo comparações entre as interações. A partir desses arquivos, é possível realizar análises inteligentes sobre os dados capturados com a finalidade de obter informações a respeito de como as pessoas se sentem ao interagir com o aplicativo, levando em conta aspectos dinâmicos, circunstâncias e mudanças de comportamento.

A análise dos arquivos leva em consideração o contexto onde as tarefas foram mapeadas pelo Engenheiro e o cenário de uso da aplicação onde ocorreram as interações do usuário final. Essa relação é possível ser inferida devido ao conjunto de métricas escolhido para a análise. Por exemplo, poderiam ser incorporadas métricas específicas relacionadas ao GPS do dispositivo móvel, permitindo correlacionar as interações de um usuário com informações relativas ao seu posicionamento, velocidade e percurso realizado. Nesse caso, a análise da usabilidade poderia levar em conta não somente como o usuário realiza uma tarefa, mas também como o ambiente influencia suas interações.

O modelo prevê adequações ao contexto no qual o aplicativo será avaliado. Isto inclui características do cenário e dos recursos tecnológicos do dispositivo. Por exemplo, a detecção das formas de interação do usuário final (voz, gestos, toques) pode servir de subsídio para determinar a utilização de determinadas métricas ou a supressão das mesmas. Para atender a este requisito, foi previsto no Módulo de Incorporação das Métricas (etapa 7) a recuperação

de informações sobre os recursos do dispositivo e a configuração do cenário de interesse de investigação. Com isso, é possível associar diferentes conjuntos de métricas aos diferentes contextos de uso.

5. ESTÁGIO ATUAL E PERSPECTIVAS

Para dar suporte ao mapeamento das tarefas, o modelo inicialmente foi concebido para que fosse realizada a captura das interações do projetista da aplicação [8]. Para isso, foi criada uma ferramenta baseada em um formalismo para a descrição de tarefas. Entretanto, após os primeiros estudos de caso, foram constatados os seguintes problemas: (i) a dependência do projetista da aplicação para a realização do mapeamento das tarefas; (ii) a possível interferência do projetista nos resultados obtidos, já que o principal responsável pela aplicação fazia um dos papéis mais importantes do processo e (iii) a dificuldade de mapear as tarefas com o formalismo proposto.

Diante das constatações apresentadas acima, foram feitos ajustes significativos no modelo para permitir que as tarefas fossem mapeadas de forma automática [9]. Para isso, o Engenheiro de Testes (ou especialista) interage com o aplicativo, enquanto Aspectos de mapeamento guardam os passos realizados para completar as tarefas. Com isso, o processo de mapeamento se torna automático, além de eliminar a necessidade da participação do projetista (e, portanto, do conhecimento do código da aplicação), já que o mesmo está indiretamente sendo avaliado.

Essa alteração acarretou a necessidade da concepção de uma nova ferramenta chamada de *Automatic Task Description*, a qual está em estágio avançado de desenvolvimento.

Para dar suporte a Unidade de Rastreabilidade, foi construída a ferramenta *Creation Usability Aspect* que possibilita a incorporação de métricas de usabilidade de forma automática. Para isso, a ferramenta utiliza como entrada de dados o arquivo XML criado na Unidade de Mapeamento das Ações do Engenheiro de Teste com o objetivo de capturar os métodos relacionados à execução da tarefa. De posse dessas informações, as estruturas desenvolvidas na concepção da ferramenta são utilizadas para gerar os aspectos ligados à usabilidade.

As contribuições do trabalho, até o presente momento, referem-se a proposta do modelo, a construção das ferramentas que dão suporte ao mapeamento automático das tarefas e da geração automática das métricas de usabilidade. Além de alguns estudos de caso preliminares realizados para testar as funcionalidades do modelo e avaliar as ferramentas que dão suporte à sua implementação [7] [8] [9].

Os desafios e perspectivas futuras são:

(i) Desenvolver novas métricas contextualizadas a determinados aplicativos e que investiguem a utilização de recursos tecnológicos disponibilizados pelos aparelhos. O objetivo é avaliar as restrições do modelo quanto a incorporação de métricas mais específicas a um determinado contexto.

(ii) Escolher ou trocar automaticamente o conjunto de métricas que será utilizado de acordo com a identificação do contexto de uso do aplicativo. Por exemplo, no caso de um usuário interagir muito rápido e com poucos erros, pode ser inferido que ele é experiente, dessa forma, será desnecessário continuar medindo dados como a taxa de erros e o tempo para completar uma tarefa.

Entretanto, medidas recursos utilizados pode ser de extrema importância para a personalização do aplicativo.

(iii) Construir uma ferramenta para auxiliar nas avaliações dos dados capturados, embasada nos conceitos introduzidos pela metodologia UX. Para isso, será necessário extrair dados que relacione o momento da interação do usuário e cenário de uso, com os dados qualitativos obtidos pelo rastreamento das suas ações.

(iv) Realizar novos experimentos com um número suficientemente grande de usuários, aplicativos e cenários, possibilitando validar o modelo e avaliar suas reais contribuições.

6. REFERÊNCIAS

- [1] Ardito, C.; Buono, P.; Costabile, M. F.; Angeli, A.; Lanzilotti, R. Combining Quantitative and Qualitative Data for Measuring User Experience of an Educational Game. Proc. Int. Workshop on Meaningful Measures: VUUM, Reykjavik, Iceland, 2008, pp. 27-31.
- [2] Avouris, N.; Fiotakis, G.; Raptis, D. On Measuring Usability of Mobile Applications. Proc. Int. Workshop on Meaningful Measures: VUUM, Reykjavik, Iceland, 2008, pp. 33-44.
- [3] Betiol, A. H.; Cybis, W. A. Usability Testing of Mobile Devices: A Comparison of Three Approaches. INTERACT, LCNS, 2005, pp. 470-481.
- [4] Downtown, A. C. Evaluation techniques for human – computer systems design. London: McGraw-Hill, Maidenhead, 1991, pp. 325-355.
- [5] Ivory, M. Y.; Hearst, M. A. The state of the art in automating usability evaluation of user interfaces. ACM Computing Surveys, volume 33, number 4, pp. 470-516.
- [6] Kiczales, G. *et al.* Aspect-oriented programming. In: ECOOP'97 — European Conf. on Object-Oriented Programming, Springer-Verlag, Finland, 1997, pp. 220-242.
- [7] Kronbauer, A. H.; Santos, C. A. S. Um Modelo para Avaliação de Usabilidade de Aplicações Imperativas em GingaJ. InfoBrasil, Fortaleza, Brasil, 2011. 6p.
- [8] Kronbauer, A. H. ; Santos , C. A. S. Um Modelo de Avaliação da Usabilidade Baseado na Captura automática de dados de Interação do Usuário em Ambientes Reais. In: XVII WebMedia, Florianópolis, Brasil, 2011. 8p. (*to appear*)
- [9] Kronbauer, A. H.; Santos , C. A. S. Um modelo de avaliação da usabilidade baseado na captura automática de dados de interação do usuário em ambientes reais. In: X IHC e V CLIHC, Porto de Galinhas, Brasil, 2011. 10p.
- [10] Moldovan, G. S.; Tarta, A. Automatic Usability Evaluation using AOP. IEEE-TTTC Int. Conf. on Automation, Quality and Testing, Robotics, Cluj-Napoca, 2006, pp. 84-89.
- [11] Nielsen, J.; Mack, R. L. Usability Inspection Methods. John Wiley & Sons, New York, 1994.
- [12] Väänänen-Vainio-Mattila, K., Roto, V., Hassenzahl, M. Towards Practical User Experience Evaluation Methods. Proc. Int. Workshop on Meaningful Measures: VUUM, Reykjavik, Iceland, 2008, pp. 19-22.
- [13] Waterson, S.; Landay, J.A.; Matthews, T. In the lab and out in the wild: remote web usability testing for mobile devices. Extended Abstracts, CHI 2002, pp. 796-797.