

Evoluindo uma aplicação para Cidades Inteligentes através de Avaliação de Métricas de Qualidade e Usabilidade

Improving a smart city application through evaluation of Quality Metrics and Usability

Alexandre Aragão¹, Lucas Machado¹, Nathyane Moreno¹, Davi Viana¹, Francisco Silva¹, Tércio Sousa², Luis Rivero¹, Ariel Teles^{1,2}, Arlindo F. da Conceição³, Inaldo Costa⁴

¹ Universidade Federal do Maranhão (UFMA)
São Luís, Maranhão – Brasil

²Instituto Federal do Maranhão (IFMA)
São Luís, Maranhão – Brasil

³Universidade Federal de São Paulo (UNIFESP)
São José dos Campos, São Paulo – Brasil

⁴Instituto Tecnológico de Aeronáutica (ITA)
São José dos Campos, São Paulo – Brasil

{alexandre.aragao, lucas.machado, davi.viana, fssilva}@lstdi.ufma.br, ariel.teles@ifma.edu.br, arlindo.conceicao@unifesp.br, inaldo@ita.br

Abstract. *Several software applications have been changing people's lifestyles in technological and social aspects. Applications in the context of Smart Cities (SCs) may contain characteristics that are not normally evaluated in traditional applications, since they may involve aspects related to sensor networks, management of large amounts of data and applications that have specific behaviors according to the context of use. This paper presents a study on the evaluation of three software quality characteristics considered relevant for SCs: Context Awareness, Calmness and Mobility. Within the context of the study, two tests were performed to evaluate an application for SCs. Additionally, an heuristic evaluation was carried out, aiming to verify the usability of the application and the impacts of this type of evaluation in applications for SCs. We identified that both the evaluated characteristics and the heuristics are adequate for the evaluation of applications for ICs, but that their results may be influenced by aspects of development and characteristics of the context of use.*

Keywords. *Smart Cities; Quality of Software; Usability.*

Resumo. *Diversas aplicações de software vêm modificando o estilo de vida das pessoas em aspectos tecnológicos e sociais. Aplicações no contexto de Cidades Inteligentes (CIs) podem conter características que normalmente não são*

avaliadas em aplicações tradicionais, uma vez que podem envolver aspectos de redes de sensores, gerenciamento de grande quantidade de dados e aplicações que possuem comportamentos específicos de acordo com o contexto. Este artigo apresenta um estudo sobre a avaliação de três características de qualidade de software consideradas relevantes para CIs, são estas: Sensibilidade ao Contexto, Calmness e Mobilidade. Para isso, realizou-se dois testes de avaliação de uma aplicação para CIs. Adicionalmente, uma avaliação heurística foi realizada visando verificar a usabilidade da aplicação e os impactos deste tipo de avaliação em aplicações para CIs. Identificou-se que tanto as características avaliadas quanto as heurísticas são adequadas para a avaliação das aplicações para CIs, mas que podem ter seus resultados influenciados por aspectos de desenvolvimento e características do contexto.

Palavras-Chave. *Cidades Inteligentes; Qualidade de Software; Usabilidade.*

1. Introdução

O uso de *smartphones* e aplicativos que facilitam a vida cotidiana do ser humano cresce constantemente [AgBrasil 2019]. Essas aplicações buscam auxiliar diversas atividades e otimizar o tempo das pessoas. Concomitantemente, verifica-se a expansão dos centros urbanos que precisam lidar com desafios de planejamento, desenvolvimento e operações das cidades. O uso de aplicações e dispositivos móveis possui um papel fundamental nesta expansão urbana, pois estes auxiliam a gestão das cidades com soluções adequadas e otimizadas [Hernández-Muñoz et al. 2011], como serviços baseados em localização.

O uso de dispositivos móveis e infraestrutura tecnológica na gestão das cidades possibilita a criação das chamadas Cidades Inteligentes (CIs) [Chourabi et al. 2012]. De acordo com [Bakıcı et al. 2013], uma CI é uma “cidade intensiva e avançada de alta tecnologia que conecta pessoas, informações e elementos da cidade usando novas tecnologias para criar uma cidade sustentável e mais verde, comércio competitivo e inovador e um aumento na qualidade de vida”. Do ponto de vista computacional, CIs são um conglomerado de tecnologias para auxiliar na gestão das cidades, como: Internet das Coisas (IoT), computação ubíqua, *big data*, computação distribuída, redes de sensores e atuadores [Santana et al. 2017, Atzori et al. 2010, Jagadish et al. 2014].

Aplicações desenvolvidas para o contexto de IoT e CIs também precisam ser avaliadas para permitir melhorias e evoluções. Todavia, as abordagens tradicionais de avaliação de aplicações de software não consideram especificações de aplicações para estes contextos [Zambonelli 2017, Larrucea et al. 2017] que envolvem múltiplas tecnologias e contextos diferentes [Cavalcante et al. 2017]. Essas aplicações devem ser adequadas para que qualquer pessoa destes centros urbanos se sinta à vontade para utilizá-la. Uma vez que há uma grande diversidade de *stakeholders* [Cavalcante et al. 2017]. Segundo [Carvalho et al. 2018], há um conjunto de características importantes para aplicações no contexto de sistemas ubíquos. Inicialmente neste trabalho, destacaram-se as características de Sensibilidade ao Contexto (do inglês, *Context Awareness*) e *Calmness* [Aragão et al. 2019]. Nessa versão estendida, considerou-se outra característica, a Mobilidade.

A Sensibilidade ao Contexto é definida como a capacidade de perceber o sistema de informações contextuais e adaptar proativamente suas funcionalidades. Ou seja, ela visa tratar da possibilidade das aplicações de identificar as informações do contexto do sistema e dos usuários, para inferirem a situação em que esses encontram-se e, assim, proporcionarem adaptações eficientemente [Dey 2001]. Tal definição é relevante para CIs, pois analisam o grau de adaptação das aplicações ao contexto dos usuários, influenciando a tomada de decisão nas suas rotinas diárias.

A característica de *Calmness* se refere à capacidade do sistema de auxiliar as atividades do usuário na hora e no local correto, fornecendo o melhor serviço possível [Rieki et al. 2004], acarretando portanto, um grande impacto na satisfação do usuário e, finalmente, mas não menos importante, em sua avaliação. Finalmente, a mobilidade é definida como sendo a capacidade do sistema de disponibilizar o acesso contínuo às informações e recursos do sistema, independentemente da localização e dentro dos limites da aplicação [Kourouthanassis et al. 2008]. De acordo com [Ryu et al. 2006], um sistema é dito móvel apenas se suportar interação não estacionária¹.

Este trabalho apresenta uma análise das características de Sensibilidade ao Contexto, *Calmness* e mobilidade. Além disso, descreve-se como essas características podem ser empregadas para a avaliação da qualidade de aplicações no contexto de CIs. Para isto, buscou-se na literatura definições dessas características e realizou-se um estudo de caso em uma aplicação real.

Nesta versão estendida do trabalho de pesquisa, é apresentado um comparativo entre os resultados da avaliação de duas versões de uma mesma aplicação para CIs. A nova versão da aplicação foi criada para que se tornasse mais acessível para os usuários. Como resultado, identificou-se que as características avaliadas são adequadas para a avaliação das aplicações para CIs, incluindo a característica Mobilidade. Por fim, executou-se uma avaliação de usabilidade para aplicações de CIs. O aspecto de usabilidade pode ser definido como sendo a capacidade do sistema de ser aprendido e compreendido adequadamente para seu uso [Carvalho et al. 2018]. Por fim, observou-se a necessidade de avaliação de outras características, como Funcionalidade, Confiabilidade e Eficiência.

Além desta seção introdutória, o restante do artigo está organizado em mais quatro seções. A Seção 2 apresenta o referencial teórico e os trabalhos relacionados sobre avaliação de qualidade para aplicações para IoT e CIs. A Seção 3 apresenta a descrição das métricas definidas na literatura e utilizadas neste trabalho. A Seção 4 apresenta a condução da pesquisa, os resultados e discussão dos resultados obtidos. A Seção 5 apresenta a conclusão do trabalho.

2. Avaliação de Qualidade em IoT e CIs

Uma CI é uma cidade inovadora que utiliza as Tecnologias de Informação e Comunicação (TIC), IoT e outros meios para melhorar a qualidade de vida, eficiência das operações e

¹Um sistema não estacionário é aquele em que suporta interações em movimento, um sistema é dito estacionário se este suportar apenas interações inerciais mesmo que possa ser transferido de local [Ryu et al. 2006].

serviços urbanos². A partir dessa definição, verifica-se que uma CI deve: (a) proporcionar uma alta qualidade de vida aos cidadãos, integrando-os à cidade; (b) monitorar e/ou controlar condições, estados e informações dos diversos sensores e infraestruturas que compõem uma cidade. Com base nessas informações, uma aplicação para CI deve auxiliar seus usuários (e.g., cidadãos em suas atividades diárias) e/ou auxiliar na gestão da cidade. Neste contexto, o paradigma de IoT é fundamental, visto que o uso de tecnologias e soluções desenvolvidas a partir dele é necessário para viabilizar uma CI [Chourabi et al. 2012, Hernández-Muñoz et al. 2011].

As avaliações de qualidade ocorrem geralmente após a etapa de implementação, mas a concepção do processo de qualidade se inicia concomitantemente a primeira etapa de engenharia das aplicações. Além disso, as avaliações de aplicações para CIs podem ser semelhantes àquelas para IoT, uma vez que compartilham características e necessidades comuns. Portanto, algumas métricas a serem usadas para a avaliação de aplicações de IoT podem ser utilizadas para avaliar uma aplicação para CIs. Métricas como tempo de conexão com a Internet, tempo de alteração de rede, e grau de operação de rede não devem ser utilizadas na avaliação, visto que estão mais relacionadas ao dispositivo em que a aplicação está executando. Por outro lado, o tempo que esta aplicação leva para obter algum dado do banco de dados deve ser medido e avaliado.

De acordo com [Thomas et al. 2016], não existem métricas definidas para a avaliação de usabilidade no contexto de IoT, apesar de sua crescente importância. Adicionalmente, o trabalho desenvolvido em [Yáñez Gómez et al. 2014] verificou resultados positivos quanto à característica de usabilidade. Os autores realizaram uma avaliação heurística adaptada para o contexto de dispositivos IoT. Desta forma, observa-se que há uma lacuna de pesquisa e prática sobre avaliação de usabilidade no contexto de IoT e CIs.

Pesquisas referentes à avaliação de qualidade de aspectos específicos de sistemas complexos, como aplicações de IoT, CIs e Ubíquas têm sido desenvolvidas [Arasteh et al. 2016]. Todavia, por ser uma vertente tecnológica relativamente recente, ainda não há um conjunto sólido de abordagens em Engenharia de Software e Qualidade de Software [Zambonelli 2017, Larrucea et al. 2017]. Alguns trabalhos buscam analisar aplicações de IoT e sistemas ubíquos que podem ser utilizadas no contexto de CIs.

O trabalho realizado em [Darin et al. 2016] apresenta a aplicação de diversas características, entre elas Sensibilidade ao Contexto e *Calmness*, para a avaliação de uma aplicação móvel de IoT. A aplicação é baseada em proximidade para automatizar a criação de grupos virtuais e facilitar o compartilhamento de arquivos e criação de lista de presença durante reuniões [Darin et al. 2016], facilitando interações entre os participantes. Os autores executaram um estudo para validação das métricas reunidas. Este estudo foi conduzido em um ambiente controlado.

A pesquisa realizada em [Carvalho et al. 2018] reuniu métricas para soluções de software relacionadas a sistemas ubíquos encontrados na literatura, partindo-se do trabalho anterior [Maia et al. 2016]. Nesta pesquisa, são reunidas as características, tais como

²Definição de cidades inteligentes: <https://www.itu.int/en/ITU-T/focusgroups/ssc/Pages/default.aspx>

Sensibilidade de Contexto, *Calmness* e Mobilidade, e suas respectivas métricas.

3. Métricas para Avaliação de Sensibilidade ao Contexto, *Calmness* e Mobilidade

As características e métricas selecionadas para o estudo de caso foram obtidas a partir de [Carvalho et al. 2018]. Adicionalmente, algumas métricas para Mobilidade foram obtidas de [Maia et al. 2016] e [Ryu et al. 2006]. Essas características estão relacionadas com aspectos humanos de interação com as aplicações, uma vez que mudanças repentinas nas aplicações podem influenciar no comportamento de seus usuários.

As Tabelas de 1 a 3 apresentam o nome, propósito, método, fórmula e interpretação das métricas selecionadas para este estudo. A coluna “Coleta” refere-se às maneiras que os dados necessários para o cálculo do valor da métrica podem ser coletados. O método de observação consiste na presença de um observador para coletar informações como a quantidade de vezes que o usuário realizou determinada ação (como desviar o foco, abrir determinada tela e entre outros). A Tabela 1 apresenta as métricas para Sensibilidade ao Contexto.

Tabela 1. Métricas selecionadas para avaliação de Sensibilidade ao Contexto.

Nome	Descrição	Coleta	Fórmula	Interpretação
Exatidão de Informação de Contexto	Quantidade de informações de contexto que foram coletadas corretamente.	Log de interação do usuário e observação	$X = \frac{\sum \left(\frac{A}{B}\right)}{N} \times 100\%$ <p>N = Número de tipos de informações de contexto A = Número de informações de contexto corretamente medidas. B = Número de informações e contexto medidas</p>	Quanto mais perto de 100%, melhor.
Granularidade	Nível de detalhamento das informações de contexto.	Questionário	$X = \{1, 2, 3\}$	(1) Baixo: Informação de baixo nível (2) Médio: Situação de contexto (3) Alto: Informação de alto nível

Grau de Adaptação à Mudança de Contexto	Nível de adaptação de funcionalidades referentes às mudanças de contexto.	Log de interação do usuário e observação	$X = \frac{\sum \left(\frac{A}{B}\right)}{N} \times 100\%$ <p>N = Número de tipos de adaptações; A = Número de vezes que os sistema se adaptou; B = Número de vezes que uma adaptação foi acionada.</p>	Quanto mais perto de 100%, melhor.
Exatidão de Adaptação	Medida que intenta verificar se as adaptações foram realizadas de forma eficiente.	Log de interação do usuário e observação	$X = \frac{\sum \left(\frac{A}{B}\right)}{N} \times 100\%$ <p>N = Número de tipos de adaptações; A = Número de adaptações corretamente realizadas; B = Número de adaptações realizadas.</p>	Quanto mais perto de 100%, melhor.
Frequência de Mudança de Contexto	Frequência com que as mudanças de contexto ocorrem.	Log de interação do usuário	$X = \frac{\sum (A_i)}{N} \times 100\%$ <p>A_i = Tempo de uma mudança a outra; N = N intervalos de medida.</p>	Baixa: minutos Média: segundos Alta: milisegundos
Tempo de Adaptação	Tempo necessário para que uma adaptação ocorra desde uma mudança de contexto.	Log de interação do usuário	$X = T$ <p>T = Tempo que o sistema leva para se adaptar desde uma mudança de contexto.</p>	Baixo: Milisegundos Médio: Segundos Alto: Minutos

A Tabela 2 apresenta as métricas para *Calmness*. É possível notar que *Calmness* apresenta algumas das métricas de Sensibilidade ao Contexto, isso ocorre pois algumas métricas são capazes de avaliar mais de uma característica, não se limitando a este exemplo.

Tabela 2. Métricas selecionadas para avaliação de *Calmness*.

Nome	Descrição	Coleta	Fórmula	Interpretação
------	-----------	--------	---------	---------------

Exatidão de Informação de Contexto	Quantidade de informações de contexto que foram coletadas corretamente.	Log de interação do usuário e observação	$X = \frac{\sum \left(\frac{A}{B}\right)}{N} \times 100\%$ <p>N = Número de tipos de informações de contexto; A = Número de informações de contexto corretamente medidas; B = Número de informações e contexto medidas.</p>	Quanto mais perto de 100%, melhor.
Grau de Adaptação à Mudança de Contexto	Nível de adaptação de funcionalidades referentes às mudanças de contexto.	Log de interação do usuário e observação	$X = \frac{\sum \left(\frac{A}{B}\right)}{N} \times 100\%$ <p>N = Número de tipos de adaptações; A = Número de vezes que os sistema se adaptou; B = Número de vezes que uma adaptação foi acionada.</p>	Quanto mais perto de 100%, melhor.
Exatidão de Adaptação	Análise das adaptações e sua realização de forma eficiente.	Log de interação do usuário e observação	$X = \frac{\sum \left(\frac{A}{B}\right)}{N} \times 100\%$ <p>N = Número de tipos de adaptações; A = Número de adaptações corretamente realizadas; B = Número de adaptações realizadas.</p>	Quanto mais perto de 100%, melhor.
Frequência de Mudança de Contexto	Frequência com que as mudanças de contexto ocorrem.	Log de interação do usuário	$X = \frac{\sum (A_i)}{N} \times 100\%$ <p>A_i = Tempo de uma mudança a outra; N = N intervalos de medida.</p>	Baixa: Minutos Média: Segundos Alta: Milisegundos
Tempo de Adaptação	Tempo necessário para que uma adaptação ocorra desde uma mudança de contexto.	Log de interação do usuário	$X = T$ <p>T = Tempo que o sistema leva para se adaptar desde uma mudança de contexto.</p>	Baixo: Milisegundos Médio: Segundos Alto: Minutos

Aplicação Proativa	Analisa o quanto a aplicação pode automatizar ações do usuário.	Questionário e Observação	$X = \frac{A_i}{A_i + N_i} \cdot 100\%$ <p>A_i = Número de informações de contexto automatizadas pelo sistema; N_i = Número de requisitos.</p>	Quanto mais perto de 100%, melhor.
Grau de Disponibilidade	Indicação da disponibilidade da aplicação em qualquer lugar e a qualquer momento.	Questionário	$X = \{1, 2, 3, 4\}$	(1) Alta (2) Média (3) Baixa (4) Muito baixa
Grau de Sincronização de Sensibilidade de Contexto	Impressão do usuário em relação às interações com a aplicação.	Questionário	$X = \{1, 2, 3, 4\}$	(1) Alto (2) Médio (3) Baixo (4) Muito baixo
Grau de Cortesia de Interação	A aplicação requisiu a atenção do usuário somente quando necessário.	Questionário	$X = \{1, 2, 3, 4\}$	(1) Alto (2) Médio (3) Baixo (4) Muito baixo
Grau de Relevância de Interação	A aplicação requisitou e mostrou informações úteis da percepção do usuário.	Questionário	$X = \{1, 2, 3, 4\}$	(1) Alto (2) Médio (3) Baixo (4) Muito baixo
Quantidade de Mudança de Foco	Quantidade de ações que um usuário toma para que a aplicação funcione corretamente e que mudam o foco de sua atividade principal.	Observação	$X = A$ <p>A = Número de ações que mudam o foco do usuário durante o uso da aplicação</p>	Quanto mais longe de zero, melhor.
Tempo Médio Entre Falhas	Tempo entre falhas ocorridas durante o uso da aplicação.	Log de interação do usuário e observação	$X = \frac{Y}{Z}$ <p>Y = Tempo de operação; Z = Número de falhas detectadas.</p>	Quanto maior, melhor.

A Tabela 3 apresenta as métricas para Mobilidade. As métricas Taxa de Informações Entregues pelo Dispositivo, Grau de Cobertura, Taxa de Adaptação e Taxa de Acesso Adaptável foram obtidas de [Maia et al. 2016] e [Ryu et al. 2006]. Essas métricas se mostraram relevantes para a avaliação da característica e optou-se por adaptá-las e utilizá-las neste trabalho, com exceção da métrica Taxa de Informações Entregues pelo Dispositivo. A nomenclatura das métricas citadas é sugestão dos autores, visto que não foram nomeadas no trabalho original.

Tabela 3. Métricas selecionadas para avaliação de Mobilidade.

Nome	Descrição	Coleta	Fórmula	Interpretação
Taxa de Informações Entregues pelo Dispositivo	Taxa de transações bem-sucedidas relativas às transações mal sucedidas.	Log de interação do usuário	$X = \frac{A}{B}$ <p>A = Número de transações bem-sucedidas; B = Número de transações malsucedidas</p>	Quanto maior melhor.
Indicador de Mobilidade Transparente	Grau de interação dinâmica da aplicação, isto é, verifica a capacidade de um usuário em movimento de utilizar a aplicação.	Observação	$X = \{0, 1, 2, 3\}$	(0) Interação estacionária (1) Baixa (2) Média (3) Alta
Grau de Operações em Rede	Quantidade de falhas de conexão em rede o sistema apresentou.	Log de interação do usuário	$X = N$ <p>N = Quantidade total de falhas.</p>	Quanto menor o número de ocorrências melhor.
Delay de Transferência	Tempo de troca de informações em rede.	Log de interação do usuário	$X = T$ <p>T = Tempo requerido para execução de transação ou troca de informação.</p>	Quanto menor melhor.
Suporte para Acesso ao Sistema	Quantidade de tecnologias em acesso à rede que o sistema utiliza.	Observação	$X = A$ <p>A = Número de tecnologias de acesso a rede fornecidas pelo sistema.</p>	Quanto maior melhor.
Grau de Cobertura	Análise da amplitude de cobertura do sistema.	Log de interação do usuário e observação	$X = A$ <p>A = Número de casos encontrados pelo usuário com desconexão além do permitido.</p>	Quanto menor melhor.
Taxa de Adaptação	Taxa de tempo do sistema para localização de um componente perante à uma realocação.	Log de interação do usuário e observação	$X = A/B$ <p>A = Tempo para o sistema localizar o componente realocado; B = Tempo de localização antes de se mover.</p>	Quanto mais próximo de 1 melhor.

Taxa de Acesso Adaptável	Taxa de tempo de acesso do sistema para um componente perante a uma realocação.	Log de interação do usuário e observação	$X = \frac{A}{B}$ <p>A = Tempo para o sistema acessar o componente realocado; B = Tempo de acesso antes da realocação.</p>	Quanto mais próximo de 1 melhor.
--------------------------	---	--	---	----------------------------------

Ao todo, 21 métricas foram selecionadas, sendo seis para Sensibilidade ao Contexto, 12 para *Calmness* e oito para Mobilidade. Observa-se que cinco métricas são aplicadas tanto para Sensibilidade ao Contexto quanto para *Calmness*. Apesar de uma métrica cobrir mais de uma característica, ela não pode ser analisada isoladamente para uma determinada característica. Desta forma, para se obter conclusões sobre uma determinada característica, é necessário analisar um conjunto de métricas.

4. Utilização das Métricas na Avaliação de Aplicação de CIs

Para a avaliação de aplicações para CIs fora desenvolvida uma aplicação para cidades inteligentes no Laboratório de Sistemas Distribuídos Inteligentes (LSDi) da Universidade Federal do Maranhão (UFMA). Com a aplicação, foram testes cuja objetivos eram analisar a utilização das características e suas métricas na análise da qualidade de uma aplicação para CIs. Desta forma, pode-se verificar dois resultados dos testes: a adequabilidade das características e métricas para avaliação deste tipo específico de aplicação; e aspectos de melhoria da própria aplicação para cidades inteligentes. A seguir, apresenta-se a aplicação para cidades inteligentes e o planejamento dos testes.

4.1. Aplicação Para Cidades Inteligentes

A aplicação desenvolvida é voltada para o monitoramento de frotas de veículos utilizando *smartphones*. Dentre as principais funcionalidades da aplicação estão: monitoramento da velocidade de cada veículo da frota; detecção da localização dos veículos exibida em um mapa em tempo real; acompanhamento de um histórico de paradas do veículo (quando o veículo atingir a velocidade zero); e alertas sobre solicitações de pedidos de ajuda feitos pelo motorista.

A aplicação divide-se em dois módulos: o primeiro, denominado *Mobile Vehicle Monitoring System* (MVMS), que executa em dispositivos móveis com a plataforma *Android*, os quais ficam embarcados nos veículos da frota. Este módulo captura os dados de contexto a partir dos sensores do *smartphone* (acelerômetro, o sensor que detecta o nível de bateria e o GPS). A partir do GPS é obtida a coordenada geográfica (latitude e longitude) do veículo que será utilizada para o monitoramento da sua localização. A Figura 1 apresenta telas da aplicação.

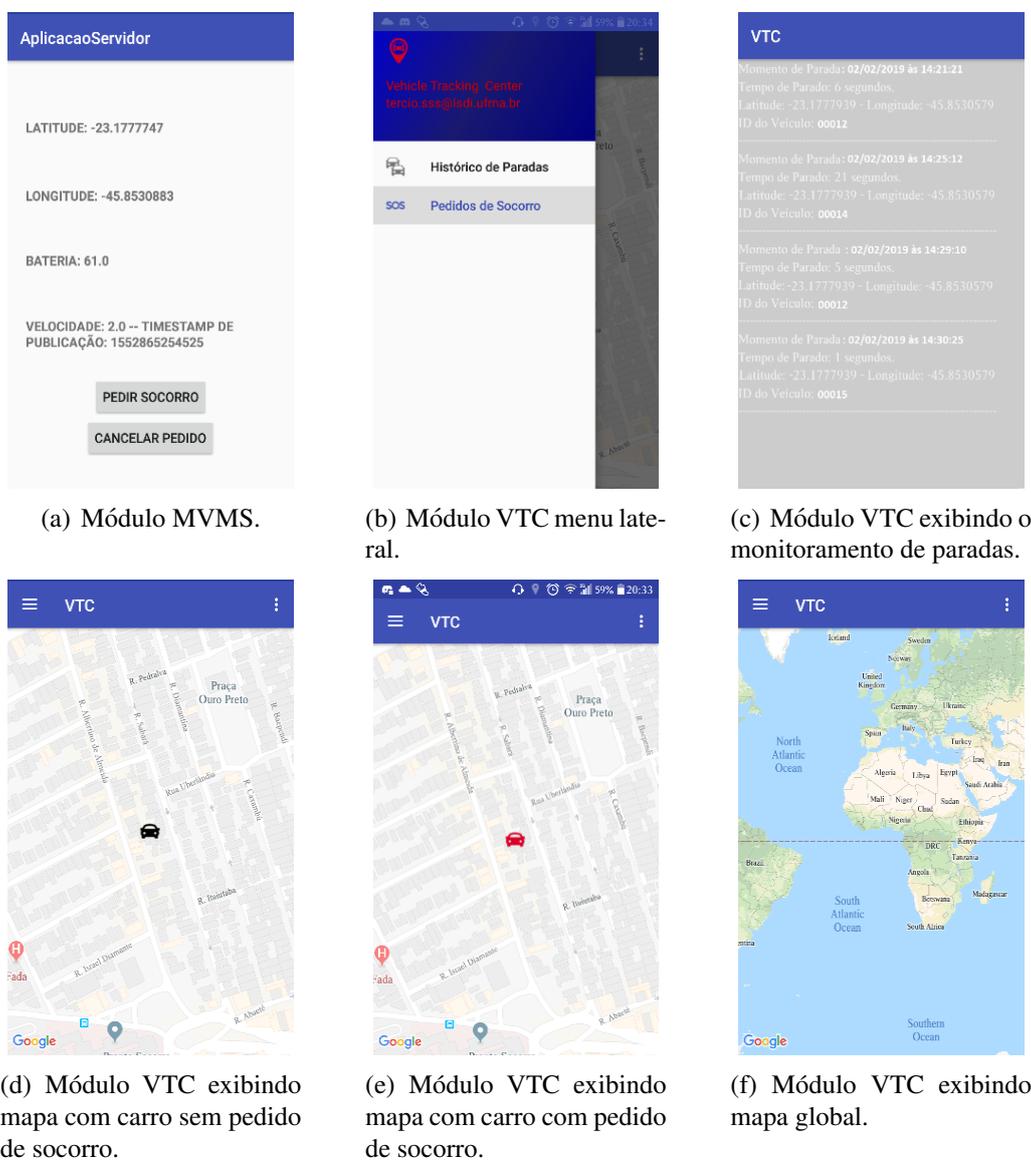


Figura 1. Capturas de tela dos módulos MVMS e VTC.

Para detectar a velocidade com que o veículo está se deslocando e identificar que realizou-se uma parada, é utilizada uma API do Google Android (*Location do Google Play Services*). O nível de bateria do *smartphone* é um outro dado de contexto capturado e utilizado para ajustar a frequência que as demais informações são enviadas através de uma conexão com a Internet para o outro componente da arquitetura. O objetivo deste controle é economizar a bateria do dispositivo móvel à medida em que seu nível diminui. A Figura 1(a) mostra a interface gráfica do módulo MVMS e exibe os dados de contexto capturados do dispositivo móvel.

O outro módulo da aplicação, denominado *Vehicle Tracking Center (VTC)*, tem por objetivo manter uma central de monitoramento de todos veículos da frota. As Figuras 1(b) e 1(c) mostram as interfaces gráfica do módulo VTC. O módulo permanece obtendo

os dados de contexto enviados pelo MVMS, e assim que os recebe, atualiza a interface gráfica de acordo com as informações de contexto recebidas. Por exemplo, na Figura 1(c) é exibido um histórico de paradas, atualizado cada vez que é detectado que a velocidade de algum dos veículos é igual a zero.

4.2. Planejamento das avaliações

Os testes foram executados utilizando os mesmos cenários, questionário e aplicação. Inicialmente, elaborou-se um cenário de utilização da aplicação. Neste cenário, definiu-se que dois usuários executariam a aplicação. Desta forma, os testes foram realizados em pares. Um usuário percorre uma rota previamente escolhida usando um veículo ou a pé, enquanto o outro usuário realizava o monitoramento através do módulo VTC. Posteriormente, foi feita a inversão dos papéis dos usuários, para que eles pudessem utilizar os dois módulos e avaliar a aplicação completa.

Por fim, um questionário³ foi entregue para os usuários e para o desenvolvedor da aplicação. Esse questionário captura dados relacionados as métricas analisadas neste artigo, como a métrica Granularidade, da característica Sensibilidade ao Contexto.

Para o segundo teste, inseriu-se um observador para o usuário do módulo MVMS. O papel do observador foi confirmar e contar a quantidade de pedidos de socorro efetuadas, não sendo oferecido nenhum auxílio para o usuário durante o teste.

Neste segundo teste, algumas métricas foram coletadas diretamente a partir de dados da aplicação e dispositivo móvel. A aplicação foi ajustada para prover dados necessários para os cálculos das métricas. O desenvolvedor da aplicação também apoiou a execução do estudo e coleta das métricas.

4.3. Execução e Resultados do Primeiro Teste

Este teste foi realizado com quatro pares de participantes. Cada par de usuários demorou entre três e seis minutos para executar o cenário proposto. O desenvolvedor da aplicação também forneceu dados para uma métrica (granularidade de contexto) através de um formulário. Em seguida, os dados foram tabulados pelos pesquisadores. Os resultados são apresentados nas Tabelas 4 e 5.

Em relação à característica de Sensibilidade ao Contexto, verificou-se que a métrica Exatidão de Informação de Contexto obteve um resultado satisfatório. Cabe ressaltar que a conexão intermitente e a fraca conectividade do dispositivo com a rede de telefonia móvel para acessar a Internet pode interferir na recepção e/ou envio dos dados de contexto, influenciando a taxa de atualização ou atualidade da posição no mapa.

Tabela 4. Primeiros resultados para as métricas de Sensibilidade ao Contexto.

Métrica	Dado Coletado	Resultado
Exatidão de Informação de Contexto	Posição no mapa Velocidade Nível da bateria	90%

³<http://goo.gl/nfAj77> - será inserido como apêndice na versão final do artigo

Granularidade	Questão 1 do questionário de desenvolvedor	Alto
Grau de Adaptação à Mudança de Contexto	Localização nos pontos marcados para solicitação de socorro após mudar de local Mostrar o tempo de parada após movimentar-se novamente Veículo muda de dor Nível de bateria (alto/médio/baixo)	88%
Exatidão de Adaptação	Localização de acordo com posição real do usuário Dados coletados de acordo com nível de bateria	55%
Frequência de Mudança de Contexto	—	—
Tempo de Adaptação	—	—

A Granularidade de Contexto foi avaliada como alta pelo desenvolvedor da aplicação, sendo um resultado adequado já que os dados necessários para se obter a localização, nível de bateria e velocidade dependem do *smartphone*, culminando em uma aplicação com menos mudanças bruscas devido ao contexto. Deve-se ressaltar que a granularidade de contexto pode interferir no desempenho da aplicação e no comportamento do usuário.

O Grau de Adaptação à Mudança de Contexto mostra-se aceitável, já que normalmente a aplicação exhibe o dispositivo (que possui o MVMS) em movimento de acordo com a velocidade mostrada na tela do VTC. Quanto à exatidão de adaptação, o valor referente atende à demanda dos dados coletados, porém a adaptação referente à bateria não se mostrou perceptível devido não chegar a haver mudança na frequência durante a avaliação (3s, 5s e 8s para os níveis de baterias alto, médio e baixo, respectivamente). Todavia, verifica-se um valor médio para esta métrica (55%). Não foi possível coletar duas métricas sobre esta característica, são elas Frequência de Mudança de Contexto e Tempo de Adaptação, uma vez que a conexão à Internet para comunicação do MVMS e VTC interferia no resultado. Essas não foram utilizadas neste trabalho, pois a coleta de dados foi inviabilizada pela aplicação móvel.

Em relação à característica *Calmness*, observou-se que as métricas de Grau de Disponibilidade, Grau de Sincronização de Sensibilidade ao Contexto, Grau de Cortesia de Interação, Grau de Relevância de Interação obtiveram um resultado satisfatório. A aplicação forneceu os serviços de localização, pedidos de socorro e verificação de paradas. Esses serviços foram prestados durante a comunicação entre os módulos MVMS e o VTC da aplicação móvel, sem perturbações desnecessárias aos usuários, uma vez que houve conexão com a Internet.

Tabela 5. Primeiros resultados para as métricas de *Calmness*.

Métrica	Dado Coletado	Resultado
Exatidão de Informação de Contexto	Posição no mapa Velocidade Nível da bateria ID do dispositivo	90%
Grau de Adaptação à Mudança de Contexto	Localização nos pontos marcados para solicitação de socorro após mudar de local Mostrar o tempo de parada após movimentar-se novamente Veículo muda de dor Nível de bateria (alto/médio/baixo)	88%
Exatidão de Adaptação	Localização de acordo com posição real do usuário Dados coletados de acordo com nível de bateria	55%
Frequência de Mudança de Contexto	—	—
Tempo de Adaptação	—	—
Aplicação Proativa	Posição no mapa Velocidade Nível da bateria	47%
Grau de Disponibilidade	Disponibilidade da aplicação para o usuário	Alto
Grau de Sincronização de Sensibilidade de Contexto	Interação no momento correto para o usuário	Alto
Grau de Cortesia de Interação	Nível de atenção de interação para o usuário	Médio
Grau de Relevância de Interação	Nível de importância de informações apresentadas ou requisitadas para o usuário	Médio
Quantidade de Mudança de Foco	Ações de configurações de GPS e Internet, reiniciar a aplicação.	Dois ações do usuário
Tempo médio entre falhas	Falhas durante uso da aplicação	Dois minutos

Quanto ao Tempo Médio Entre Falhas, observou-se que a aplicação parava de funcionar devido à rotação da tela. Por fim, em relação à métrica Quantidade de Mudança de Foco, verificou-se que a ação de ativar o GPS e a conexão com a Internet (3g/4g), para que o serviço funcionasse, fazia com que o usuário precisasse reiniciar a aplicação para que suas funcionalidades fossem executadas adequadamente.

O resultado para Grau de Adaptação à Mudança de Contexto e Exatidão de Informação de Contexto são os mesmos coletados para Sensibilidade ao Contexto. A métrica de Aplicação Proativa não foi coletada nesta pesquisa, pois as informações sobre sua interpretação não possuíam detalhes suficientes para explicá-la, principalmente sobre as ações do usuário que devem ser consideradas na avaliação.

4.4. Execução e Resultados do Segundo Teste

O teste contou com dez voluntários, totalizando 5 pares para aplicação e execução do cenário planejado. O tempo do teste variou dependendo da rota escolhida pelo usuário do módulo MVMS, ao final dos testes notou-se que todos os testes duraram entre três e seis

minutos. O teste foi realizado em apenas uma tarde dentro de um mesmo bairro e com todos os dispositivos totalmente carregados. O usuário do módulo MVMS era acompanhado por um observador, para confirmar se o pedido de parada foi feito corretamente.

Posteriormente, o desenvolvedor da aplicação respondeu ao seu respectivo formulário. O resultado deste formulário forneceu dados para algumas métricas, como, por exemplo, a métrica Granularidade, da característica Sensibilidade ao Contexto. Ao final de todos os testes, foi possível tabular os resultados obtidos. Para classificação dos resultados percentuais utilizou-se a Tabela 6.

Tabela 6. Classificação dos resultados percentuais.

Medida	Classificação
Superior ou igual a 0% e inferior a 10%	Muito Ruim
Igual ou superior a 10% e inferior a 50%	Ruim
Igual ou superior a 50% e inferior a 70%	Regular
Igual ou superior a 70% e inferior a 95%	Bom
Igual ou superior a 95% e inferior ou igual a 100%	Excelente

Os resultados do teste incluem, *logs* em formato .txt da aplicação VTC (Figura 2(a)), *logs* e publicações no banco de dados (Figura 2(b)) comum a ambos módulos, resultados do questionário do desenvolvedor e resultados do questionário do usuário.



Marcador created at: Thu Oct 31
17:20:56 GMT-03:00 2019
Pedido created at: Wed Oct 31 2019
17:18:08 GMT+0000 (UTC)



(a) Captura de tela do *log* do VTC.



(b) Captura de tela do banco de dados.

Figura 2. Exemplos de *logs* utilizados no teste.

O texto em amarelo na Figura 2(b) indica que aquele valor sofreu uma atualização recente. Os resultados das métricas selecionadas são apresentados nas Tabelas de 7 a

9. A Tabela 7 apresenta os resultados das métricas para a característica Sensibilidade ao Contexto.

Tabela 7. Resultados para as métricas de Sensibilidade ao Contexto.

Métrica	Dado Coletado	Resultado
Exatidão de Informação de Contexto	Posição no mapa Velocidade Nível da bateria ID do dispositivo	80%
Granularidade	Questão 1 do questionário de desenvolvedor	Alto
Grau de Adaptação à Mudança de Contexto	Localização nos pontos marcados para solicitação de socorro após mudar de local Mostrar o tempo de parada após movimentar-se novamente Veículo muda de dor Nível de bateria (alto/médio/baixo)	88%
Exatidão de Adaptação	Localização de acordo com posição real do usuário Dados coletados de acordo com nível de bateria	68%
Frequência de Mudança de Contexto	Quantidade de pedidos efetuados Intervalo entre pedidos	Médio (54.69s)
Tempo de Adaptação	Momento de efetuação do pedido Veículo muda de cor	Médio

Para a característica Sensibilidade ao Contexto, verificou-se que a métrica Exatidão de Informação de Contexto obteve o resultado igual a 80%, considerado como bom. Cabe ressaltar que a conexão intermitente e a fraca conectividade do dispositivo com a rede de telefonia móvel para acessar a Internet pode interferir na recepção e/ou envio dos dados de contexto, influenciando a taxa de atualização ou atualidade da posição no mapa.

A métrica Granularidade é avaliada pelo desenvolvedor, este a classificou como Alta. O resultado é adequado, já que os dados necessários para se obter a localização, nível de bateria e velocidade dependem do *smartphone*, culminando em uma aplicação com menos mudanças bruscas. Deve-se ressaltar que a granularidade de contexto pode interferir no desempenho da aplicação e no comportamento do usuário.

O resultado da métrica Grau de Adaptação à Mudança de Contexto foi igual a 88%, classificado como bom. Desta forma é possível afirmar que maior parte das adaptações foram realizadas corretamente. Os erros de adaptação foram oriundos de falhas de conexão, como: conexão intermitente e fraca conectividade do dispositivo com a rede de telefonia móvel.

Sobre a Exatidão de Adaptação, entende-se que o valor observado ainda é aceitável para o propósito da aplicação. Acredita-se que a nova implementação da aplicação tenha afetado o resultado desta métrica, o resultado, que foi 68%, é classificado como regular e se difere bastante do resultado do primeiro teste, que foi 94%.

A métrica Frequência de Mudança de Contexto apenas indicou que o tempo médio entre os pedidos estava no intervalo de segundos, por isso seu resultado foi classificado como médio. O Tempo de Adaptação apresentou o resultado médio, isto é, a aplicação

leva alguns segundos até se adaptar à uma mudança de contexto.

Em relação à característica *Calmness*, o resultado das métricas Exatidão de Informação de Contexto, Grau de Adaptação à Mudança de Contexto, Exatidão de Adaptação, Frequência de Mudança de Contexto e Tempo de Adaptação já foram discutidos. Desta forma, somente as métricas Aplicação Proativa, Grau de Disponibilidade, Grau de Sincronização de Sensibilidade ao Contexto, Grau de Cortesia de Interação, Grau de Relevância de Interação, Quantidade de Mudança de Foco e Tempo Médio entre Falhas serão discutidas. A Tabela 8 apresenta os resultados das métricas para a característica *Calmness*.

Tabela 8. Resultados para as métricas de *Calmness*.

Métrica	Dado Coletado	Resultado
Exatidão de Informação de Contexto	Posição no mapa Velocidade Nível da bateria ID do dispositivo	80%
Grau de Adaptação à Mudança de Contexto	Localização nos pontos marcados para solicitação de socorro após mudar de local Mostrar o tempo de parada após movimentar-se novamente Veículo muda de dor Nível de bateria (alto/médio/baixo)	100%
Exatidão de Adaptação	Localização de acordo com posição real do usuário Dados coletados de acordo com nível de bateria	68%
Frequência de Mudança de Contexto	Quantidade de pedidos efetuados Intervalo entre pedidos	Médio (54.69s)
Tempo de Adaptação	Momento de efetuação do pedido Veículo muda de cor	Médio
Aplicação Proativa	Posição no mapa Velocidade Nível da bateria	44.44%
Grau de Disponibilidade	Disponibilidade da aplicação para o usuário	Alto
Grau de Sincronização de Sensibilidade de Contexto	Interação no momento correto para o usuário	Alto
Grau de Cortesia de Interação	Nível de atenção de interação para o usuário	Médio
Grau de Relevância de Interação	Nível de importância de informações apresentadas ou requisitadas para o usuário	Médio
Quantidade de Mudança de Foco	Ações de configurações de GPS e Internet, reiniciar a aplicação.	2 ações do usuário
Tempo Médio Entre Falhas	Falhas durante uso da aplicação	25min 11s

A métrica Aplicação Proativa obteve o resultado igual a 44%, um valor considerado ruim segundo a classificação estabelecida. O resultado, todavia, era esperado, uma vez que a aplicação não tem como função automatizar ações do usuário, apenas fornecer uma ferramenta de controle dentro do contexto de CIs.

As métricas Grau de Disponibilidade, Grau de Sincronização de Sensibilidade ao Contexto, Grau de Cortesia de Interação e Grau de Relevância de Interação foram obtidas

a partir da avaliação do usuário. Mais especificamente, as métricas Grau de Cortesia de Interação e Grau de Relevância de Interação foram obtidas a partir da questão 3 do formulário e as outras questões avaliaram as demais métricas. Vale ressaltar que o resultado alto para a métrica Grau de Disponibilidade é essencial para aplicações de CIs.

A presença do observador foi fundamental para avaliar a Quantidade de Mudança de Foco, pois, como o nome indica, é necessário estar atento às ações do usuário. A escolha de acompanhar o usuário do módulo MVMS deu-se pela razão deste estar ao volante, assim sendo escolhido por motivos de segurança.

Quanto ao Tempo Médio Entre Falhas, entende-se que o valor obtido não reflete a realidade. A nova versão da aplicação apresenta falsos positivos, isto é, indica um cenário aparente correto para os usuários de ambos módulos. Ao verificar o console do Firebase⁴ 3 Cloud Functions (responsável pela transição de dados) é possível notar que muitas das requisições foram mal-sucedidas (Figura 3). Isso ocorre pela limitação do pacote free do serviço, que limita a quantidade de requisições em um período de tempo.

Hora ↓	Nível	Função	Mensagem de evento
31 de out. de 2019			
5:40:00.334 ...	▲	getData	▶ Error: quota exceeded (Quota exceeded for quota group 'DNSResolutionsNonbillable' and limi...
5:39:59.359 ...	▲	getData	▶ Error: quota exceeded (Quota exceeded for quota group 'DNSResolutionsNonbillable' and limi...
5:39:58.320 ...	▲	getData	▶ Error: quota exceeded (Quota exceeded for quota group 'DNSResolutionsNonbillable' and limi...
5:39:57.346 ...	▲	getData	▶ Error: quota exceeded (Quota exceeded for quota group 'DNSResolutionsNonbillable' and limi...
5:39:56.208 ...	▲	getData	▶ Error: quota exceeded (Quota exceeded for quota group 'DNSResolutionsNonbillable' and limi...
5:39:55.191 ...	▲	getData	▶ Error: quota exceeded (Quota exceeded for quota group 'DNSResolutionsNonbillable' and limi...
5:39:54.198 ...	▲	getData	▶ Error: quota exceeded (Quota exceeded for quota group 'DNSResolutionsNonbillable' and limi...
5:39:53.166 ...	▲	getData	▶ Error: quota exceeded (Quota exceeded for quota group 'DNSResolutionsNonbillable' and limi...

Figura 3. Captura de tela do log do Firebase Cloud Functions.

As características Sensibilidade ao Contexto e *Calmness* já foram avaliadas no trabalho anterior [Aragão et al. 2019], na mesma aplicação descrita neste trabalho. Este novo teste visou principalmente avaliar a Mobilidade, a nova característica selecionada para estudo. A Tabela 9 apresenta os resultados das métricas para a característica Mobilidade.

⁴Site do Firebase <https://firebase.google.com/>

Tabela 9. Resultados para as métricas de Mobilidade.

Métrica	Dado Coletado	Resultado
Taxa de Informações Entregues Pelo Dispositivo	Quantidade de requisições feita ao servidor Quantidade de requisições bem-sucedidas	42%
Grau de Cobertura	Quantidade de requisições mal-sucedidas Tempo de execução	1.16 users/min
Taxa de Adaptação	—	—
Taxa de Acesso Adaptável	—	—
Indicador de Mobilidade Transparente	—	—
Grau de Operações em Rede	Quantidade de requisições mal-sucedidas	58%
<i>Delay</i> de Transferência	Tempo de realização do pedido Tempo de publicação Tempo de exibição do pedido	8s
Suporte para Acesso ao Sistema	Wifi e Redes móveis (2G, 3G e 4G)	2

A Taxa de Informações Entregues pelo Dispositivo acabou se mostrando bem baixa, classificada como ruim. Esta nova indica os erros de troca de informações entre os módulos. Em outras palavras, apenas 42% das informações entregues pelo dispositivo foram bem realizadas. A métrica Grau de Cobertura está diretamente ligada a duas outras métricas, Taxa de Informações Entregues pelo Dispositivo e Grau de Operações em Rede. Ela indica que, a cada minuto, 1.16 usuários recebem um erro, tendo em vista que o teste era realizado em pares, 58% dos usuários recebem erros no período de um minuto.

O *Delay* de Transferência mensurou o tempo que a aplicação VTC levou para receber o pedido da aplicação MVMS, o valor de 8s pode ser considerado alto, visto que a aplicação intenta ser de tempo real. As métricas Taxa de Adaptação, Taxa de Acesso Adaptável e Indicador de Mobilidade Transparente não foram consideradas adequadas para a avaliação dessa aplicação em específico.

4.5. Comparativo entre os Testes

Com o novo teste, é possível comparar os resultados obtidos para as métricas e as características, evoluindo, portanto, o modelo de teste para aplicações para cidades inteligentes exposto ao longo deste artigo. A Tabela 10 apresenta a comparação entre os resultados dos dois testes.

Tabela 10. Comparação entre os resultados das métricas entre os dois testes.

Métrica	Resultado Anterior	Resultado Atual
Exatidão de Informação de Contexto	90%	80%
Granularidade	Alto	Alto
Grau de Adaptação de Mudança de Contexto	88%	88%
Exatidão de Adaptação	94%	68%
Frequência de Mudança de Contexto	—	Médio (54.69s)
Tempo de Adaptação	—	Médio
Aplicação Proativa	47%	44.44%
Grau de Disponibilidade	Alto	Alto
Grau de Sincronização de Sensibilidade de Contexto	Alto	Alto
Grau de Cortesia de Interação	Alto	Alto
Grau de Relevância de Interação	Médio	Médio
Quantidade de Mudança de Foco	2 ações do usuário	2 ações do usuário
Tempo Médio Entre Falhas	2 minutos	25min 11s
Taxa de Informações Entregues Pelo Dispositivo	NÃO FOI AVALIADA	42%
Grau de Cobertura	NÃO FOI AVALIADA	1.16 users/min
Taxa de Adaptação	NÃO FOI AVALIADA	—
Taxa de Acesso Adaptável	NÃO FOI AVALIADA	—
Indicador de Mobilidade Transparente	NÃO FOI AVALIADA	—
Grau de Operações em Rede	NÃO FOI AVALIADA	58
Delay de Transferência	NÃO FOI AVALIADA	8s
Suporte para Acesso ao Sistema	NÃO FOI AVALIADA	2

Apesar de em ambos testes a mesma aplicação estar sendo avaliada, a comunicação entre os módulos VTC e MVMS se diferem. No primeiro teste a aplicação utiliza um serviço próprio do laboratório em que foi desenvolvida, o CDDL do LSDi⁵. No segundo teste, a comunicação entre as duas se deu utilizando o serviço de *Cloud Functions* do plano *free* do Firebase.

A principal diferença notada foi a velocidade de comunicação entre as aplicações, pois a primeira versão apresentava melhor desempenho que a segunda. A métrica *Delay de Transferência* é um dos indicadores, apesar de não ter sido avaliada no primeiro teste, os usuários que participaram em ambos testes puderam notar a diferença.

A métrica *Exatidão de Adaptação* também corrobora com este argumento, devido ao limite de requisições e ao grande tempo de resposta das *Cloud Functions* implementadas, a localização exibida pelo módulo VTC não correspondia à real localização do usuário do módulo MVMS. Neste sentido, fica claro que para aplicações para CIs, as tecnologias utilizadas para seu desenvolvimento devem ser adequadas e fornecer o maior suporte possível para estas aplicações.

⁵<http://www.lsd.ufma.br/projetos/cddl/doku.php?id=doc:doc-api>

4.6. Avaliação Heurística

A interação humano-computador (IHC) é um campo de estudo multidisciplinar focado em compreender a interação entre humanos e computadores, as características aqui abordadas e avaliadas podem ser utilizadas para mensuração de IHC, assim como feito e classificado para sistemas ubíquos em [Maia et al. 2016]. A avaliação heurística possui a vantagem de possuir baixo custo, ser intuitiva, motivadora e não requer planejamento complexo [Nielsen and Molich 1990]. A seguir, apresenta-se o planejamento, execução e resultados obtidos com esta avaliação.

4.6.1. Planejamento da Avaliação Heurística

Neste trabalho, aplicou-se as 10 heurísticas de Nielsen, definidas em [Nielsen 1994] e apresentadas na Tabela 11. Definiu-se a avaliação apenas na aplicação MVMS, pois é o módulo utilizado pelo usuário final. Desta forma, ele necessita de uma avaliação mais criteriosa da sua interface.

Tabela 11. Heurísticas de Nielsen.

Heurística	Descrição
Diálogo simples e natural	Informações devem ser apresentadas de forma natural, lógica e adequada ao contexto.
Falar a linguagem do usuário	Informações devem ser apresentadas na linguagem do usuário e não orientadas ao sistema.
Minimizar a carga de memória do usuário	Usuário não deve necessitar lembrar de informações específicas do sistema para a sua utilização.
Consistência	Diferentes informações, ações e situações não podem possuir conflito. Usuário não deve necessitar se preocupar quanto a significados ou resultados diferentes para uma mesma informação, ação ou situação.
<i>Feedback</i>	Usuário deve ter conhecimento se sua ação foi recebida e realizada. Sistema deve informar ao usuário continuamente sobre as suas ações.
Saídas evidentes	Usuário deve ser capaz de desfazer ou refazer ações no sistema.
Atalhos	Sistema deve ser flexível para a utilização de usuários avançados, eficiência de uso.
Mensagens de erros	Informações de erro devem ser precisas e apresentadas na linguagem do usuário.
Prevenção de erros	Situações de erro devem ser previstas e prevenidas pelo sistema.
Documentação e ajuda	Sistema deve disponibilizar ajuda. Qualquer informação deve possuir fácil acesso e não ser extensa.

Durante o planejamento, definiu-se que toda a aplicação MVMS seria avaliada, uma vez que é uma aplicação relativamente simples. Adicionalmente, buscou-se definir a quantidade de participantes. Foram selecionados cinco voluntários para participar desta avaliação.

4.6.2. Execução e Resultados da Avaliação Heurística

Inicialmente, um dos pesquisadores envolvidos neste trabalho realizou a apresentação da aplicação MVMS e o objetivo de avaliação heurística. Em seguida, cada participante realizou um percurso pré-definido em automóvel, acompanhado do observador de forma que tivesse uma experiência interativa com a aplicação. Ao final, todos os participantes receberam um questionário avaliativo das 10 heurísticas de Nielsen para a aplicação MVMS. Cada violação das heurísticas foi avaliada pelo grau de severidade, de acordo com o apresentado na Tabela 12. Cada participante teve cinco minutos para realizar a avaliação.

Tabela 12. Graus de severidade.

Grau de Severidade	Tipo	Descrição
0	Sem importância	Não afeta a operação da interface para os usuários, pode não ser um problema de usabilidade.
1	Cosmético	Problema cosmético, corrigido se houver tempo disponível.
2	Simple	Pequeno problema de usabilidade, baixa prioridade de correção.
3	Grave	Grande problema de usabilidade, alta prioridade de correção.
4	Catastrófico	Problema de usabilidade catastrófico, necessária correção urgente, imperativa.

Os dados foram sintetizados de forma com que o grau de severidade seja representado para cada heurística em função do número de ocorrências. A Tabela 13 corresponde ao resultado da avaliação heurística.

Tabela 13. Resultados para as heurísticas de Nielsen.

Heurística Avaliada	Quantidade de Ocorrências Relatadas	Sem importância	Cosmético	Simple	Grave	Catastrófico
Diálogo simples e natural	2	-	2	-	-	-
Falar a linguagem do usuário	1	-	1	-	-	-
Minimizar a carga de memória do usuário	1	-	1	-	-	-
Consistência	4	-	-	3	1	-
Feedback	5	-	-	-	2	3
Saídas evidentes	1	-	-	-	1	-
Atalhos	2	-	-	2	-	-
Mensagens de erros	2	-	-	1	1	-
Prevenção de erros	4	-	1	2	1	-
Documentação e ajuda	5	-	-	-	1	4

As heurísticas de Nielsen se mostraram adequadas na avaliação da aplicação pois todas as heurísticas foram violadas, ao menos uma vez. Apenas as heurísticas de Documentação e ajuda e *Feedback* obtiveram violação classificada como catastrófica, apresentando respectivamente 80% e 60% das ocorrências classificadas como tal. A heurística de *Feedback* obteve 100% das ocorrências redundantes, as quais relataram o mesmo problema, apesar de ter sido classificada em graus de severidade diferentes, indicando iminência. As heurísticas Falar a linguagem do usuário e Minimizar a carga de memória do usuário, compreenderam mínimas violações, classificadas ainda com baixo grau de severidade.

Apesar da simplicidade da aplicação, através da heurística de “documentação e ajuda”, foi possível identificar a ausência de documentação, ausência de auto apresentação da aplicação, além de que o objetivo da aplicação não se mostrou claro e de fácil compreensão. Na heurística de *Feedback* foi observada uma ausência de informação referente à sucesso ou falha da funcionalidade de pedir socorro e cancelar pedido. Esse aspecto foi classificado como uma falha crítica para a proposta da aplicação. Na atual versão da aplicação, o usuário desconhece se o pedido de socorro ou cancelamento de pedido foi efetuado.

Na heurística de consistência, 50% dos relatos foram classificados como inconsistência na informação de velocidade e os demais como inconsistência geral da aplicação perante a rotação de tela. Na heurística de atalhos, foi relatado ausência de ações otimizadas durante o uso da aplicação. Na heurística de Prevenção de erros, os relatos podem ser classificados como problemas já então mencionados e mau uso de cores em ações críticas. A heurística de diálogo simples e natural, os participantes relataram que eram necessárias mais informações além da velocidade, latitude e longitude. Adicionalmente, também foi observada a ausência de formatação legível para o ser humano no campo *timestamp de publicação* da aplicação, tornando a informação incompreensível para o usuário.

4.7. Discussão sobre as avaliações realizadas

A realização de avaliações de qualidade que busquem analisar se determinadas características são adequadas para um contexto específico de aplicações é relevante, pois criam-se indícios e evidências sobre essas características. Neste trabalho, inicialmente, buscou-se realizar duas avaliações de características e suas respectivas métricas. Essas características são avaliadas em sistemas ubíquos. Todavia, ainda não se tinha indícios ou evidências dessas características em sistemas específicos para CIs.

Após o primeiro teste, observou-se que outra característica é importante para o contexto de CIs, a Mobilidade. Além disso, as aplicações, MVMS e VTC, foram reimplementadas. Desta forma, fazia-se necessário repetir o teste nas características já avaliadas. Apesar das mesmas aplicações terem sido avaliadas, as mudanças realizadas no desenvolvimento da nova versão da aplicação tiveram impacto nos resultados das características e métricas. Vale ressaltar a importância da seleção de ferramentas adequadas para construção de aplicações para CIs, sendo necessários novos estudos que as identifiquem de maneira adequada.

Como limitações desses testes, observa-se que coletar as métricas apenas no uso

de uma aplicação não fornece indícios suficientes para analisar de maneira abrangente todas métricas estudadas neste trabalho. Além disso, há limitação da aplicação em si em relação aos dados necessários de coleta. Os desenvolvedores podem desconhecer determinadas métricas e isso dificulta a geração e coleta dos dados para a análise. Adicionalmente, observa-se que aspectos relacionados ao *hardware* e infraestrutura de comunicação podem impactar nas métricas analisadas. Todavia, neste trabalho, não foi possível considerar tais aspectos. Outra limitação identificada é o aspecto de generalização dos resultados. Nesta pesquisa, utilizou-se uma aplicação para cidades inteligentes que pode não representar a variedade de aspectos desses sistemas.

Adicionalmente, observou-se que se fez necessário uma avaliação de usabilidade da aplicação a ser utilizada pelo usuário final. Este tipo de avaliação é relevante para mostrar a adequabilidade da aplicação para seu o objetivo final. Neste trabalho, empregou-se a avaliação heurística. As heurísticas de Nielsen são genéricas e facilitam a avaliação de usabilidade a diversos tipos de aplicações. Todavia, observa-se que a usabilidade de aplicações para CIs sofrem influência direta de características inerentes deste cenário, como mudança de contexto e mobilidade. Desta forma, é necessário um estudo mais aprofundado sobre a adequação destas heurísticas para avaliar a usabilidade deste tipo de aplicação. Este é um passo inicial na análise e definição de um conjunto de abordagens para a avaliação da qualidade nas aplicações para CIs.

5. Conclusão

A avaliação de qualidade de aplicações para cidades inteligentes ainda é um problema em aberto de pesquisa, uma vez ainda não há um conjunto de abordagens consolidadas para o desenvolvimento de sistemas e sua engenharia neste contexto específico. Neste trabalho, o objetivo foi analisar características de qualidade de software para verificar a sua utilidade em avaliação de aplicações para cidades inteligentes. Para isso, buscou-se na literatura características e métricas relevantes para aplicações deste contexto.

A utilização de métricas para avaliar as características de Sensibilidade ao Contexto, *Calmness* e Mobilidade no contexto de CIs mostrou-se relevante nesta pesquisa. Algumas métricas foram coletadas a partir de *logs* de funcionalidades da aplicação. Devido à aplicação ter sido desenvolvida no mesmo laboratório desta pesquisa, conseguiu-se solicitar ao desenvolvedor que fossem fornecidos dados para calcular as métricas. Em caso de aplicações prontas e desenvolvidas por terceiros, seria difícil conseguir analisar determinados eventos e interações com a aplicação.

Com os dados coletados, foi possível analisar como a aplicação poderia impactar positiva ou negativamente no comportamento do usuário. As características de Sensibilidade ao Contexto, *Calmness* e Mobilidade fornecem indicativos de como a aplicação pode influenciar o comportamento dos usuários. Este tipo de avaliação de qualidade é relevante para que as aplicações sejam as mais adequadas possíveis do ponto de vista do usuário.

A avaliação heurística mostrou-se relevante para este estudo, pois foi possível avaliar a utilização real da aplicação por potenciais usuários. Dada a limitação dos testes, não é possível afirmar que é válida para qualquer aplicação para CIs, pois depende do tipo

de interface e interações da aplicação.

É possível ainda destacar a importância da avaliação da característica Mobilidade, pois os resultados complementam a métrica Grau de Disponibilidade da característica *Calmness*. Apesar da aplicação ter um alto grau de disponibilidade, a métrica Grau de Operações em Rede mostra que 58% das requisições foram mal-sucedidas. Dessa forma, apensar do serviço estar disponível, as informações exibidas estão desatualizadas.

Assim, identificou-se que as características avaliadas são adequadas para a avaliação das aplicações para CIs. Adicionalmente, executou-se uma avaliação de interação humano-computador para aplicações de CIs. Por fim, observou-se a necessidade de avaliação de outras características, como Funcionalidade, Confiabilidade e Eficiência [Koscianski and Soares 2007]. Aspectos de qualidade de aplicações para cidades inteligentes precisam ser levadas em consideração na especificação e projeto de aplicações para este domínio. Assim como qualquer aplicação, isso é importante para aceitação e ampla utilização pelos usuários.

Como contribuições para a pesquisa, o trabalho busca apresentar indícios da aplicabilidade da avaliação de qualidade de características específicas para o contexto de cidades inteligentes. Além disso, apresenta-se como os dados das métricas relacionadas as características foram coletados e utilizados. Por fim, a avaliação heurística evidencia como uma técnica tradicional pode ser utilizada neste contexto emergente. Desta forma, abrem-se novas oportunidades de pesquisa como: (i) adaptações da avaliação heurística; (ii) definição de novas métricas e possibilidades de operacionaliza-las durante a execução de avaliações de qualidade. Para a prática, observou-se a necessidade de definir padrões ou abordagens que busquem uma avaliação mais completa deste tipo de aplicação. Além disso, este conjunto de características e métricas podem guiar futuros desenvolvedores a pensarem na qualidade de software durante o desenvolvimento deste tipo de aplicação.

Em trabalhos futuros, os autores esperam realizar novos testes em outras aplicações para CIs, bem como encontrar melhores maneiras de avaliar as características selecionadas para avaliação. Além disso, é necessário levar em conta aspectos da infraestrutura física e tecnologias de comunicação utilizadas como fatores de influência das métricas.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001 através do PROCAD-Amazônia. O quarto autor agradece à FAPEMA (Edital Universal). O quinto autor agradece ao apoio do INCT de Internet do Futuro para Cidades Inteligentes (CNPq 465446/2014-0, CAPES 88887.136422/2017-00, FAPESP 14/50937-1 e 15/24485-9), CNPq (311608/2017-5, 420907/2016-5 e 312324/2015-4). O trabalho teve o apoio de bolsas de Iniciação Científica da FAPEMA, UFMA e CNPq.

Referências

AgBrasil (2019). Brasil é 5º país em ranking de uso diário de celulares no mundo. Acessado em: 25 de fevereiro de 2019. Disponível em:

- <http://agenciabrasil.ebc.com.br/geral/noticia/2019-01/brasil-foi-5o-pais-em-ranking-de-uso-diario-de-celulares-no-mundo>.
- Aragão, A., Moreno, N., Viana, D., Silva, F., Sousa, T., Rivero, L., Teles, A., da Conceição, A., and Costa, I. (2019). Utilizando métricas de qualidade na avaliação de uma aplicação para cidades inteligentes. In *Anais do IV Workshop sobre Aspectos Sociais, Humanos e Econômicos de Software*, pages 51–60. SBC.
- Arasteh, H., Hosseinezhad, V., Loia, V., Tommasetti, A., Troisi, O., Shafie-Khah, M., and Siano, P. (2016). Iot-based smart cities: a survey. In *2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC)*, pages 1–6. IEEE.
- Atzori, L., Iera, A., and Morabito, G. (2010). The Internet of Things: A survey. *Computer Networks*, 54(15):2787–2805.
- Bakıcı, T., Almirall, E., and Wareham, J. (2013). A smart city initiative: the case of barcelona. *Journal of the Knowledge Economy*, 4(2):135–148.
- Carvalho, R. M., de Castro Andrade, R. M., and de Oliveira, K. M. (2018). Aquarium - a suite of software measures for hci quality evaluation of ubiquitous mobile applications. *Journal of Systems and Software*, 136:101 – 136.
- Cavalcante, E., Cacho, N., Lopes, F., and Batista, T. (2017). Challenges to the development of smart city systems: A system-of-systems view. In *Proceedings of the 31st Brazilian Symposium on Software Engineering*, pages 244–249.
- Chourabi, H., Nam, T., Walker, S., Gil-Garcia, J. R., Mellouli, S., Nahon, K., Pardo, T. A., and Scholl, H. J. (2012). Understanding smart cities: An integrative framework. In *2012 45th Hawaii international conference on system sciences*, pages 2289–2297. IEEE.
- Darin, T., Barbosa, J., Rodrigues, B., and Andrade, R. (2016). Greatroom: Uma aplicação android baseada em proximidade para a criação de salas virtuais inteligentes. In *Workshop de Ferramentas e Aplicações (WFA). WebMedia 2019*.
- Dey, A. K. (2001). Understanding and using context. *Personal Ubi. Computing*, 5(1):4–7.
- Hernández-Muñoz, J. M., Vercher, J. B., Muñoz, L., Galache, J. A., Presser, M., Gómez, L. A. H., and Pettersson, J. (2011). Smart cities at the forefront of the future internet. In *The future internet assembly*, pages 447–462. Springer.
- Jagadish, H., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., and Shahabi, C. (2014). Big data and its technical challenges. *Communications of the ACM*, 57(7):86–94.
- Koscianski, A. and Soares, M. d. S. (2007). Qualidade de software: aprenda as metodologias e técnicas mais modernas para o desenvolvimento de software.[s.l]: Novatec, 2007. *Citado na*, page 26.
- Kourouthanassis, P. E., Giaglis, G. M., and Karaiskos, D. C. (2008). Delineating the degree of 'pervasiveness' in pervasive information systems: An assessment framework and design implications. In *2008 Panhellenic Conference on Informatics*, pages 251–255.

- Larrucea, X., Combelles, A., Favaro, J., and Taneja, K. (2017). Software engineering for the internet of things. *IEEE Software*, 34(1):24–28.
- Maia, R., Andrade, R., Oliveira, K., Santos, I. D. S., and Bezerra, C. I. M. (2016). Quality characteristics and measures for human–computer interaction evaluation in ubiquitous systems. *Software Quality Journal*.
- Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 152–158. ACM.
- Nielsen, J. and Molich, R. (1990). Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 249–256. ACM.
- Riekkki, J., Isomursu, P., and Isomursu, M. (2004). Evaluating the calmness of ubiquitous applications. volume 3009, pages 105–119.
- Ryu, H., Hong, G., and James, H. (2006). Quality assessment technique for ubiquitous software and middleware. *Res. Lett. Inf. Math. Sci*, 9:13–87.
- Santana, E. F. Z., Chaves, A. P., Gerosa, M. A., Kon, F., and Milojcic, D. S. (2017). Software platforms for smart cities: Concepts, requirements, challenges, and a unified reference architecture. *ACM Computing Surveys (Csur)*, 50(6):1–37.
- Thomas, M. O., Onyimbo, B. A., and Logeswaran, R. (2016). Usability evaluation criteria for internet of things. *Int J Inf Technol Comput Sci*, 8:10–18.
- Yáñez Gómez, R., Cascado Caballero, D., and Sevillano, J.-L. (2014). Heuristic evaluation on mobile interfaces: A new checklist. *The Scientific World Journal*, 2014.
- Zambonelli, F. (2017). Key abstractions for iot-oriented software engineering. *IEEE Software*, (1):38–45.