



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

LIANA MARA CARVALHO DE MENEZES

**UMA ABORDAGEM DE TESTE BASEADA EM GUIAS ESTRUTURADOS PARA
AVALIAR CARACTERÍSTICAS DA INTERNET DAS COISAS**

FORTALEZA

2022

LIANA MARA CARVALHO DE MENEZES

UMA ABORDAGEM DE TESTE BASEADA EM GUIAS ESTRUTURADOS PARA
AVALIAR CARACTERÍSTICAS DA INTERNET DAS COISAS

Dissertação apresentada ao Curso de do
Programa de Pós-Graduação em Ciência
da Computação do Centro de Ciências da
Universidade Federal do Ceará, como requisito
parcial à obtenção do título de mestre em
Ciência da Computação. Área de Concentração:
Engenharia de Software

Orientadora: Prof. Dra. Valéria Lelli
Leitão Dantas

Coorientadora:
Prof. Dra. Rossana Maria de Castro Andrade

FORTALEZA

2022

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

M511a Menezes, Liana Mara Carvalho de.

Uma abordagem de teste baseada em guias estruturados para avaliar características da Internet das Coisas / Liana Mara Carvalho de Menezes. – 2022.
142 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Ciência da Computação, Fortaleza, 2022.

Orientação: Profa. Dra. Valéria Lelli Leitão Dantas.

Coorientação: Profa. Dra. Rossana Maria de Castro Andrade.

1. Internet das Coisas. 2. Teste de Software. 3. Guias. 4. Desempenho. I. Título.

CDD 005

LIANA MARA CARVALHO DE MENEZES

UMA ABORDAGEM DE TESTE BASEADA EM GUIAS ESTRUTURADOS PARA
AVALIAR CARACTERÍSTICAS DA INTERNET DAS COISAS

Dissertação apresentada ao Curso de do
Programa de Pós-Graduação em Ciência
da Computação do Centro de Ciências da
Universidade Federal do Ceará, como requisito
parcial à obtenção do título de mestre em
Ciência da Computação. Área de Concentração:
Engenharia de Software

Aprovada em:

BANCA EXAMINADORA

Prof. Dra. Valéria Lelli Leitão Dantas (Orientadora)
Universidade Federal do Ceará (UFC)

Prof. Dra. Rossana Maria de Castro Andrade (Coorientadora)
Universidade Federal do Ceará (UFC)

Prof. Dr. Rodrigo Pereira dos Santos
Universidade Federal do Estado do Rio de Janeiro (UNIRIO)

Prof. Dra. Atslands Rego da Rocha
Universidade Federal do Ceará (UFC)

Prof. Dr. Ismayle de Sousa Santos
Universidade Federal do Ceará (UFC)

Dedico essa dissertação à minha mãe, Silvia Carvalho, suas orações me deram força e ânimo para persistir.

AGRADECIMENTOS

À Deus por sempre senti-lo perto de mim, não importando o momento.

Aos meus pais e irmãs por tornarem essa jornada mais leve, por sempre fornecerem o suporte e apoio necessário para eu alcançar meus objetivos.

Agradeço especialmente as minhas orientadoras Valéria Lelli e Rossana Andrade, obrigada por acreditarem em mim e me fornecerem todo o suporte necessário para os resultados deste trabalho, muito obrigada.

Aos amigos que me acompanham desde minha graduação, Lavínia Matoso e Alex Felipe, vocês possuem o dom de arrancar sorrisos de mim até nos momentos mais difíceis, vocês são raros.

Agradeço a todos meus amigos do GREat, por fazerem desses anos uma experiência incrível de muito aprendizado, obrigada por me proporcionarem mais que um ambiente de estudo, mas um ambiente de companheirismo e troca de conhecimento.

Agradeço a FUNCAP pelo apoio financeiro a essa pesquisa, do qual foi essencial para sua realização.

Por último, aos voluntários que participaram, proativamente, das avaliações deste trabalho.

“Que eu quis, feliz, sonhei, pensei... ser feliz é
bem melhor que ser rei.”

(Emicida)

RESUMO

As aplicações de Internet das Coisas (do inglês *Internet of Things* - IoT) caracterizam-se pela utilização de objetos inteligentes que estão conectados à Internet para fornecer diferentes tipos de serviços. Esses objetos inteligentes normalmente geram dados que precisam ser armazenados e analisados para contribuir para a tomada de decisões (quer sejam imediatas ou não). Neste contexto, tais aplicações podem utilizar diferentes protocolos de comunicação, exigindo um alto desempenho e, em alguns casos, respostas em tempo real. Esses requisitos trazem novos desafios de teste e a necessidade de abordagens específicas, por exemplo, a detecção de falhas de desempenho entre dispositivos IoT heterogêneos que processam uma grande quantidade de dados e, em certas condições, devem ter os seus recursos otimizados. Dessa forma, esta dissertação de mestrado define uma abordagem de teste para avaliar características de aplicações IoT baseada em guias estruturados. Para construir essa abordagem, foi conduzida uma revisão da literatura para identificar guias e normas técnicas, que são utilizados nas avaliações de aplicações IoT. Além disso, dois repositórios de *bugs* foram analisados para compreender as particularidades das atividades de teste de IoT. Com base nessa pesquisa, foi proposto o guia baseado em 11 tópicos: definição da característica, correlação das características, configuração do ambiente, definição de subcaracterísticas, contextualização, casos de testes abstratos, medições, impacto de subcaracterísticas, custo-benefício, sugestões de ferramentas e exemplo de uso. O guia foi instanciado para a característica de Desempenho que é composta por três subcaracterísticas: Comportamento Temporal, Utilização de Recursos e Capacidade. Para automatizar o uso do guia foi construída uma Wiki a fim de apoiar o planejamento do teste de Desempenho. Para avaliação do guia foram realizadas duas avaliações: (i) uma avaliação com os especialistas para validar a estrutura e o conteúdo do guia; e (ii) um experimento controlado para avaliar a utilização do guia em uma aplicação real IoT. Os resultados mostraram que o guia ajuda a avaliar os aspectos de desempenho de aplicações IoT. Além disso, a própria estrutura do guia fornece uma sistematização das atividades de teste e pode ser adaptada para avaliar outras características de IoT. Essa organização gera resultados mais eficazes e uma maior facilidade na detecção de falhas IoT.

Palavras-chave: Internet das Coisas; teste de software; guias; desempenho.

ABSTRACT

Internet of Things (IoT) applications are characterized by the use of smart objects that are connected to the Internet to provide different types of services. These smart objects typically generate data that needs to be stored and analyzed to contribute to decision making (whether immediate or not). In this context, such applications may use different communication protocols, requiring high performance and, in some cases, real-time responses. These requirements bring new testing challenges and the need for specific approaches, for example, the detection of performance gaps between heterogeneous IoT devices that process a large amount of data and, under certain conditions, must have their resources optimized. Thus, this master's thesis defines a testing approach to evaluate IoT applications characteristics based on structured guides. To build this approach, a literature review was conducted to identify guides, guidelines and technical standards, which are used in IoT application evaluations. In addition, two bug repositories were analyzed to understand the particularities of IoT testing activities. Based on this research, the guide was proposed based on 11 topics, which are: characteristic definition, correlation of characteristics, environment configuration, definition of subcharacteristics, contextualization, abstract test cases, measurements, impact of subcharacteristics, cost-benefit, tool suggestions, and usage example. The guide was instantiated for the Performance characteristic, which is composed of three subcharacteristics: Temporal Behavior, Resource Utilization, and Capacity. To automate the use of the guide, we built a Wiki to support the performance test planning. To evaluate the guide, two evaluations were conducted: (i) an evaluation with the experts to validate the structure and content of the guide; and (ii) a controlled experiment to evaluate the use of the guide in a real IoT application. The results showed that the guide is helpful in evaluating performance aspects of IoT applications. Moreover, the structure of the guide itself provides a systematization of testing activities and can be adapted to evaluate other IoT characteristics. This organization generates more effective results and an easier detection of IoT failures.

Keywords: Internet of Things; software testing; guides; performance.

LISTA DE FIGURAS

| | |
|---|----|
| Figura 1 – Fluxo de dados das aplicações IoT. Fonte: O autor. Baseado em (PATEL <i>et al.</i> , 2016) | 22 |
| Figura 2 – Exigências de dispositivos para funcionamento ideal. Adaptada (INTERNATIONAL, 2014). | 25 |
| Figura 3 – Desempenho e suas subcaracterísticas. Fonte: O autor. | 26 |
| Figura 4 – Procedimento metodológico para a construção e estruturação do guia. Fonte: O autor. | 40 |
| Figura 5 – Possíveis organizações estruturais do guia. Fonte: O autor. | 45 |
| Figura 6 – Procedimento metodológico para instanciación do guia para uma característica. Fonte: o autor. | 47 |
| Figura 7 – Procedimento metodológico concretizado para Desempenho, seguindo o modelo da Figura 6. Fonte: O autor. | 50 |
| Figura 8 – Relações das demais características com desempenho. Fonte: O autor. | 55 |
| Figura 9 – Relação das propriedades de Comportamento Temporal com as demais propriedades. Fonte: O autor. | 59 |
| Figura 10 – Relação das propriedades de Utilização de Recursos com as demais propriedades. Fonte: O autor. | 60 |
| Figura 11 – Relação das propriedades de Capacidade com as demais propriedades. Fonte: O autor. | 60 |
| Figura 12 – Interpretação dos resultados do custo-benefício. Fonte: O autor. | 63 |
| Figura 13 – Resumo da aplicabilidade do guia na empresa do João. Fonte: O autor. | 65 |
| Figura 14 – Tela inicial da WiKi. Fonte: O autor. | 66 |
| Figura 15 – Seleção das definições | 67 |
| Figura 16 – Seleção das características correlacionadas | 67 |
| Figura 17 – Seleção das subcaracterísticas | 68 |
| Figura 18 – Seleção das propriedades | 68 |
| Figura 19 – Seleção dos casos de teste | 69 |
| Figura 20 – Seleção das métricas | 69 |
| Figura 21 – Cálculo de custo-benefício | 69 |
| Figura 22 – Resultado de custo-benefício | 69 |
| Figura 23 – Seleção de ferramentas | 70 |

| | |
|--|----|
| Figura 24 – Plano de teste | 70 |
| Figura 25 – Concordância entre especialistas | 74 |
| Figura 26 – Respostas dos especialistas para as perguntas sobre a avaliação do conteúdo do guia | 77 |
| Figura 27 – Tela de gerenciamento das luzes no Automa GREat. Fonte: O autor | 81 |

LISTA DE TABELAS

| | |
|---|----|
| Tabela 1 – Diferença entre testes de desempenho tradicionais e testes de desempenho para IoT (GURIJALA, 2018) | 27 |
| Tabela 2 – Trabalhos realizados como soluções para teste de aplicações IoT | 34 |
| Tabela 3 – Construção da string de busca | 41 |
| Tabela 4 – Construção inicial dos tópicos da estrutura do guia | 42 |
| Tabela 5 – String de busca - Desempenho | 51 |
| Tabela 6 – Estudos levantados na revisão de desempenho | 51 |
| Tabela 7 – Subcaracterísticas e suas propriedades | 57 |
| Tabela 8 – Exemplo de um caso de teste de “Comportamento temporal” | 57 |
| Tabela 9 – Exemplo de uma métrica para o comportamento temporal | 58 |
| Tabela 10 – Ferramentas para teste de desempenho de IoT | 64 |
| Tabela 11 – Perfil dos especialistas | 73 |
| Tabela 12 – Perfil dos participantes do experimento | 83 |
| Tabela 13 – Visão geral dos dados do experimento por grupo | 86 |
| Tabela 14 – Resultados da hipótese $H_{1,1}$ | 87 |
| Tabela 15 – Resultados da hipótese $H_{1,2}$ | 89 |
| Tabela 16 – Resultados da hipótese $H_{1,3}$ | 90 |
| Tabela 17 – Falhas IoT categorizadas como muito graves | 92 |

LISTA DE ABREVIATURAS E SIGLAS

| | |
|--------|--|
| AUT | Application under test |
| COAP | Constrained Application Protocol |
| DHCP | Dynamic Host Configuration Protocol |
| FUNCAP | Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico |
| GREat | Grupo de Redes de Computadores, Engenharia de Software e Sistemas |
| GUT | Gravidade, Urgência e Tendência |
| IoT | Internet das Coisas - do inglês <i>Internet of Things</i> |
| LAN | Local Area Network |
| MQTT | Message Queuing Telemetry Transport |
| PDF | Portable Document Format |
| PICO | População, Intervenção, Comparação, Outcome |
| QP | Questão de Pesquisa |
| TCP/IP | Transmission Control Protocol/Internet Protocol |
| WiFi | Wireless Fidelity |

SUMÁRIO

| | | |
|------------|---|-----------|
| 1 | INTRODUÇÃO | 16 |
| 1.1 | Contextualização e Motivação do trabalho | 16 |
| 1.2 | Objetivo e Questões de Pesquisa | 17 |
| 1.3 | Método da pesquisa e Contribuições | 18 |
| 1.4 | Organização da dissertação | 19 |
| 2 | FUNDAMENTAÇÃO TEÓRICA | 21 |
| 2.1 | Internet das Coisas | 21 |
| 2.2 | Teste de software em IoT | 23 |
| 2.3 | Teste de Desempenho de Aplicações IoT | 25 |
| 2.4 | Considerações finais | 27 |
| 3 | DESAFIOS PARA OS TESTES DE APLICAÇÕES IOT | 28 |
| 3.1 | Categorização dos desafios de testes IoT | 28 |
| 3.2 | Lacunas identificadas nos testes IoT | 33 |
| 3.3 | Considerações finais | 33 |
| 4 | TRABALHOS RELACIONADOS | 34 |
| 4.1 | Trabalhos identificados | 34 |
| 4.2 | Discussão | 37 |
| 4.3 | Considerações finais | 37 |
| 5 | ABORDAGEM DE TESTE BASEADA EM GUIAS ESTRUTURADOS | 39 |
| 5.1 | Construção e estruturação do guia | 39 |
| 5.2 | Estrutura do Guia | 43 |
| 5.3 | Considerações finais | 46 |
| 6 | INSTANCIÇÃO DO GUIA DE TESTE | 47 |
| 6.1 | Procedimento metodológico de instanciação do guia | 47 |
| 6.2 | Exemplo de instanciação do Guia - característica de Desempenho | 49 |
| 6.3 | Visão geral do Guia Desempenho | 53 |
| 6.4 | Ferramenta de apoio ao uso do Guia de Desempenho | 66 |
| 6.5 | Considerações finais | 69 |
| 7 | AVALIAÇÃO | 72 |
| 7.1 | Avaliação com especialistas | 72 |

| | | |
|-----|--|-----|
| 7.2 | Experimento controlado | 78 |
| 7.3 | Ameaças à validade | 92 |
| 7.4 | Considerações Finais | 93 |
| 8 | DISCUSSÃO | 94 |
| 8.1 | QP1. Como deve ser organizado o processo de teste da IoT? | 94 |
| 8.2 | QP2. Como avaliar as características de aplicações de IoT? | 95 |
| 8.3 | QP3. Como facilitar a detecção de falhas IoT nas aplicações? | 96 |
| 8.4 | Considerações Finais | 97 |
| 9 | CONCLUSÕES E TRABALHOS FUTUROS | 98 |
| 9.1 | Resultados Alcançados | 98 |
| 9.2 | Trabalhos Futuros | 99 |
| | REFERÊNCIAS | 100 |
| | APÊNDICE A – PROTOCOLO PARA REVISÃO DA CONSTRUÇÃO DA ESTRUTURA DO GUIA | 108 |
| | APÊNDICE B – RESULTADO DA REVISÃO PARA CONSTRUÇÃO DA ESTRUTURA DO GUIA | 110 |
| | APÊNDICE C – FORMULÁRIO DE EXTRAÇÃO DE DADOS | 115 |
| | APÊNDICE D – PROTOCOLO PARA REVISÃO DE DESEMPENHO | 116 |
| | APÊNDICE E – RESULTADOS DA REVISÃO DE DESEMPENHO | 118 |
| | APÊNDICE F – FORMULÁRIO DE EXTRAÇÃO DE DADOS PARA DESEMPENHO | 128 |
| | APÊNDICE G – PLANILHA DE RELAÇÃO DAS CARACTERÍSTICAS DE DESEMPENHO | 129 |
| | APÊNDICE H – RELAÇÃO ENTRE PROPRIEDADES, CASOS DE TESTES ABSTRATOS E MÉTRICAS PARA O GUIA DE DESEMPENHO | 130 |
| | APÊNDICE I – RELAÇÃO ENTRE MÉTRICAS, CASOS DE TES- TES ABSTRATOS E FERRAMENTAS PARA O GUIA DE DESEMPENHO | 131 |
| | APÊNDICE J – CHECKLIST DA AVALIAÇÃO COM ESPECIALIS- TAS | 132 |
| | APÊNDICE K – DESIGN DO EXPERIMENTO | 135 |
| | APÊNDICE L – DEFINIÇÃO DAS CARACTERÍSTICAS | 138 |

| | |
|---|------------|
| APÊNDICE M – CHECKLIST DO EXPERIMENTO | 139 |
| APÊNDICE N – PERGUNTAS ADICIONADAS AO CHECKLIST DO | |
| GRUPO G1 | 142 |

1 INTRODUÇÃO

Essa dissertação define uma abordagem de teste para avaliar características de aplicações IoT baseada em guias estruturados. Esse capítulo introduz na Seção 1.1 o contexto no qual esse trabalho está inserido e são apresentadas as lacunas identificadas nos testes IoT, que motivaram este trabalho. Na Seção 1.2 são apresentados o objetivo dessa pesquisa e as questões de pesquisa que foram investigadas. Na Seção 1.3 é detalhada a metodologia utilizada para o desenvolvimento da dissertação e as contribuições deste trabalho. Por fim, na Seção 1.4 é apresentado a estrutura organizacional desta dissertação.

1.1 Contextualização e Motivação do trabalho

A Internet das Coisas (do inglês *Internet of Things* - IoT) é caracterizada pelo uso de objetos inteligentes, que se comunicam por meio da rede com ou sem fio com intuito de fornecer serviços (ANDRADE *et al.*, 2017). Com o aumento no número de dispositivos conectados à Internet, a IoT está se difundindo cada vez mais (RAZZAQUE *et al.*, 2015). Esse crescimento introduz novos desafios para diversas áreas que envolvem a construção de um software.

Considerando as fases comuns de construção de software e seus desafios, alguns estudos em IoT apresentam problemas enfrentados na etapa de desenvolvimento, por exemplo, as dificuldades para desenvolver rapidamente a comunicação entre sistemas que possuem linguagens de programação diferentes (MAIA *et al.*, 2014), e na arquitetura - conectar diversas arquiteturas e direcionar a um mesmo objetivo (RAZZAQUE *et al.*, 2015). Outros estudos tratam dos desafios na qualidade de software, como definir a qualidade de maneira unificada para diferentes sistemas conectados (MATEEN *et al.*, 2015) e há estudos que elencam de maneira geral as problemáticas enfrentadas na área, por exemplo, quais adaptações são necessárias para que as aplicações se comuniquem em uma rede IoT? (REETZ *et al.*, 2013)(ROSENKRANZ *et al.*, 2015).

Além das etapas mencionadas, alguns estudos destacam o teste como sendo um desafio para a IoT (CARVALHO, 2018)(BURES *et al.*, 2019). Comparando o teste de sistemas convencionais e o teste de sistemas IoT, um dos desafios IoT identificados nesses estudos é a dificuldade de fornecer padronização a uma área que envolve a comunicação de aplicações de diversos domínios com objetos inteligentes que possuem hardware, protocolos e capacidade de armazenamento distintas. Além disso, a garantia de qualidade de aplicações IoT requer dos engenheiros de software um conhecimento detalhado das características IoT (*e.g.*, interopera-

bilidade, desempenho, interconectividade) que serão validadas, das limitações dos recursos, e do comportamento da aplicação no contexto em que ela será inserida. Por exemplo, o teste de *Desempenho* traz desafios principalmente devido à quantidade de protocolos existentes (*e.g.*, TCP/IP, DHCP, entre outros) e à quantidade de dados que precisam ser processados (KIM; ZIEGLER, 2017), (DAHMANN *et al.*, 2010).

Em um trabalho anterior a esta dissertação da autora Carvalho (2018), foram identificados, por meio de um mapeamento sistemático, os desafios no teste de aplicações IoT. Nesse trabalho, os desafios foram relacionados com as características IoT. As características de *Segurança, Interoperabilidade e Desempenho*, nesta ordem, foram as que obtiveram o maior número de desafios associados (CARVALHO, 2018). Além disso, no trabalho a área de teste é apontada como pouco padronizada, e não foram identificadas abordagens de teste relacionadas às características IoT. Ainda há outros estudos na literatura que apontam o pouco aprofundamento na condução testes de IoT como uma problemática (CORTÉS *et al.*, 2019)(MURAD *et al.*, 2018).

Por meio dos desafios elencados no trabalho de Carvalho (2018) foi possível identificar as lacunas que se tratadas poderiam vir a auxiliar na solução desses desafios de teste IoT, tais lacunas são:

- Ausência de soluções estruturadas para os testes IoT;
- Carência de abordagens flexíveis, sendo possível validar diferentes características IoT;
- Ausência de ferramentas e medidas que auxiliem nos testes IoT;
- Falta de abordagens de teste IoT que considerem a correlação das características.

1.2 Objetivo e Questões de Pesquisa

Sabendo dos desafios e lacunas identificadas este trabalho tem por objetivo definir uma abordagem baseada em guias estruturados para auxiliar no teste de aplicações IoT. Para isso são investigadas as seguintes questões de pesquisa (QP):

QP1. Como o processo de teste para aplicações IoT deve ser organizado? Com esta questão de pesquisa, busca-se compreender como a execução dos testes em aplicações IoT pode ser melhor estruturada, com o objetivo de desenvolver uma solução que promova uma abordagem mais sistemática e padronizada para os testes nesse contexto.

QP2. Como avaliar as características das aplicações IoT? Visto que muitos dos desafios de teste são relacionados às características de IoT (*e.g.*, Desempenho, Segurança e Interoperabilidade), com esta questão de pesquisa busca-se investigar os métodos que são

utilizados para avaliá-las e, assim, propor uma solução que se baseia nessas características para testar aplicações IoT.

QP3. Como facilitar a detecção de falhas IoT nas aplicações? Diante dos desafios nos testes de aplicações IoT, esta questão de pesquisa busca investigar estratégias de teste que viabilizem e facilitem a detecção de falhas, com o objetivo de desenvolver uma solução que as incorpore de forma eficaz

1.3 Método da pesquisa e Contribuições

Inicialmente, nesta pesquisa foram identificadas as lacunas inerentes aos desafios de testes IoT (CARVALHO, 2018), por exemplo, soluções estruturadas, abordagens flexíveis, catalogação de ferramentas e medidas, e correlação das características. O preenchimento dessas lacunas contribui para minimizar os desafios na área de teste. Na pesquisa conduzida não foi identificado nenhum trabalho que preenchesse as lacunas apresentadas na subseção 1.2. Dessa forma, esta dissertação de mestrado define uma abordagem de teste baseada em guias estruturados. Para atender às lacunas identificadas, foi conduzida uma revisão inicial da literatura com o objetivo de elencar os principais tópicos que comporiam o guia. Como resultado, o guia foi estruturado em 11 tópicos, entre eles: "Definição da Característica", "Casos de Teste Abstratos", "Correlação das Características", "Sugestões de Ferramentas", "Medições", dentre outros.

Após a construção da estrutura do guia, este foi instanciado para a característica de *Desempenho*, tendo em vista a relevância da mesma e a familiaridade dos envolvidos nessa pesquisa com essa característica. Para a instanciação, foi realizada uma segunda revisão da literatura, com foco propriamente dito na característica de *Desempenho*. Com a análise dos estudos selecionados na revisão, foram preenchidos os tópicos do guia. Em seguida, o guia foi refinado baseado nas análises obtidas de duas aplicações reais. Além disso, para automatizar o guia, foi desenvolvida uma Wiki a fim de apoiar o seu uso no planejamento dos testes de desempenho.

Para avaliar o guia, foram conduzidas duas avaliações: (i) uma avaliação com os especialistas para validar a estrutura e o conteúdo do guia; e (ii) um experimento controlado do guia de desempenho na utilização de uma aplicação real de IoT. Para a primeira avaliação foram reunidos 7 especialistas das áreas de Teste e/ou IoT. Para essa avaliação foi fornecido um *checklist* de perguntas sobre a estrutura do guia e outras sobre o conteúdo do guia. A segunda

avaliação envolveu 12 participantes e seguiu o modelo de experimento controlado, sendo utilizada uma aplicação real IoT. Os resultados obtidos nas avaliações mostraram que o guia auxiliou na condução dos testes de desempenho, na detecção de falhas IoT de forma mais eficiente e a sua estrutura ajudou na sistematização do processo de teste.

A principal contribuição desta dissertação é definir uma abordagem de teste para aplicações IoT. O objetivo dessa abordagem é prover uma estrutura na forma de um guia que possa ser reutilizada para avaliar outras características IoT. Dessa forma, as contribuições deste trabalho são sumarizadas a seguir:

1. o conjunto de desafios de teste relacionados com as características de qualidade relevantes para IoT;
2. o modelo de um guia estruturado em tópicos relevantes para o contexto de aplicações IoT;
3. a metodologia de instanciação do guia para uma característica IoT;
4. o guia de teste de desempenho para avaliar aplicações IoT; e
5. a WiKi como ferramenta de apoio para o uso do guia de desempenho.

1.4 Organização da dissertação

Essa dissertação está organizada da seguinte forma:

No Capítulo 2 de “Fundamentação Teórica” são apresentados os termos e assuntos principais para compreensão deste trabalho. Nesse capítulo é detalhado o funcionamento da IoT e as comunicações presentes nesse contexto. No capítulo também é apresentado o teste dessas aplicações IoT, as dificuldades e peculiaridades e, por fim, o teste de desempenho IoT, uma vez que a instanciação do guia para essa pesquisa ocorreu para essa característica.

No Capítulo 3 de “Desafios” são apresentados os desafios de testes relacionados às aplicações IoT. Nesse capítulo, os desafios identificados na literatura foram categorizados e detalhados. Além disso, foi feita uma subseção à parte (ver 3.1.1) a fim de detalhar os desafios específicos de desempenho. Ao final do capítulo, são apresentadas as lacunas identificadas a partir dos desafios.

No Capítulo 4 de “Trabalhos Relacionados” são discutidos os trabalhos obtidos na literatura que se relacionam a essa pesquisa. Durante a revisão da literatura, não foram encontrados muitos trabalhos que possuem uma estrutura de guia; portanto, esse capítulo foi dividido em trabalhos que não possuem uma estrutura de guia e demais trabalhos da área, resultando em uma tabela comparativa entre esses trabalhos, da qual é apresentada nesta dissertação.

No Capítulo 5 de “Metodologia” é apresentada a construção e instanciación do guia, bem como a estrutura. Nesse capítulo, é especificado o procedimento metodológico para a construção e a estruturação do guia.

No Capítulo 6 de “Instanciación do guia de teste” é apresentado de forma sistemática, o procedimento metodológico para instanciar a abordagem para uma característica IoT, por exemplo, é feita a instanciación do guia para a característica de Desempenho. Nesse capítulo também é apresentada a WiKi que serve de ferramenta de apoio para a utilização do guia de desempenho.

No Capítulo 7 de “Avaliação” são apresentadas as avaliações do guia que foram conduzidas e os seus resultados. Neste capítulo é detalhada a metodologia utilizada para execução das avaliações bem como a coleta e análise dos dados. Além disso, nesse capítulo são apresentadas as ameaças à validade.

No Capítulo 8 de “Discussão” são relembradas as questões de pesquisa deste trabalho e as mesmas são discutidas.

No Capítulo 9 de “Considerações Finais” são apresentadas as considerações finais e um resumo das contribuições dessa dissertação. Ao final do capítulo são detalhados os trabalhos futuros para essa pesquisa.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta os assuntos principais que fundamentam este trabalho. Inicialmente, será abordado sobre a IoT de maneira geral (Seção 2.1), seguido por Teste de Software em IoT (Seção 2.2) e, então, será apresentada uma visão dos Testes de Desempenho para aplicações IoT (Seção 2.3), visto que a instanciação deste trabalho foi realizada para a característica de Desempenho.

2.1 Internet das Coisas

A Internet das Coisas (IoT) surgiu como um meio revolucionário para o modelo de comunicação. Através de mudanças em contextos é possível tomar decisões, por meio da comunicação com diferentes dispositivos em ambientes diversificados (GIUSTO *et al.*, 2010).

O ambiente IoT é heterogêneo, contendo interações entre diferentes dispositivos, unindo o real e o virtual, por exemplo, com o intuito de promover cidades e carros inteligentes, melhores organizações no trânsito, melhorias na agricultura e soluções diversas (AMALFITANO *et al.*, 2017). As interações presentes em IoT podem ser realizadas com o usuário (interações *human-thing*) ou entre os dispositivos (interações *thing-thing*) para prover algum serviço (ANDRADE *et al.*, 2017).

A IoT é composta por diferentes características que a permeiam. A partir dos resultados do estudo de (CARVALHO *et al.*, 2020) e de outros trabalhos que o próprio autor cita (UCHÔA *et al.*, 2017), é possível identificar ao menos 35 características que permeiam a IoT, como, por exemplo, interoperabilidade, interconectividade, segurança, usabilidade, desempenho, entre outras. Essas características interagem entre si de forma positiva e/ou negativa (CARVALHO *et al.*, 2020). Um exemplo dessas relações é a característica de Desempenho em conjunto com Confiabilidade, uma vez que possa haver regras de criptografias exigidas na aplicação que acabam resultando em um baixo desempenho. Deste modo, é imprescindível considerar o papel ímpar que cada característica pode exercer, sendo, inclusive, fator de possíveis interferências na aplicação, conforme o exemplo citado.

Com a diversidade de dispositivos existentes em IoT é necessário compreender o fluxo dos dados dessas aplicações. O cenário IoT no qual este trabalho se baseia, (PATEL *et al.*, 2016), é formado por um fluxo que se inicia nos sensores no qual ocorre o sensoriamento do ambiente. Os dados do sensoriamento são enviados para aplicação onde são conduzidas as regras

e tomadas as decisões, com as decisões tomadas é acionada uma atuação no ambiente. O fluxo percorre três âmbitos: Sensores e Atuadores, Rede e Aplicação.

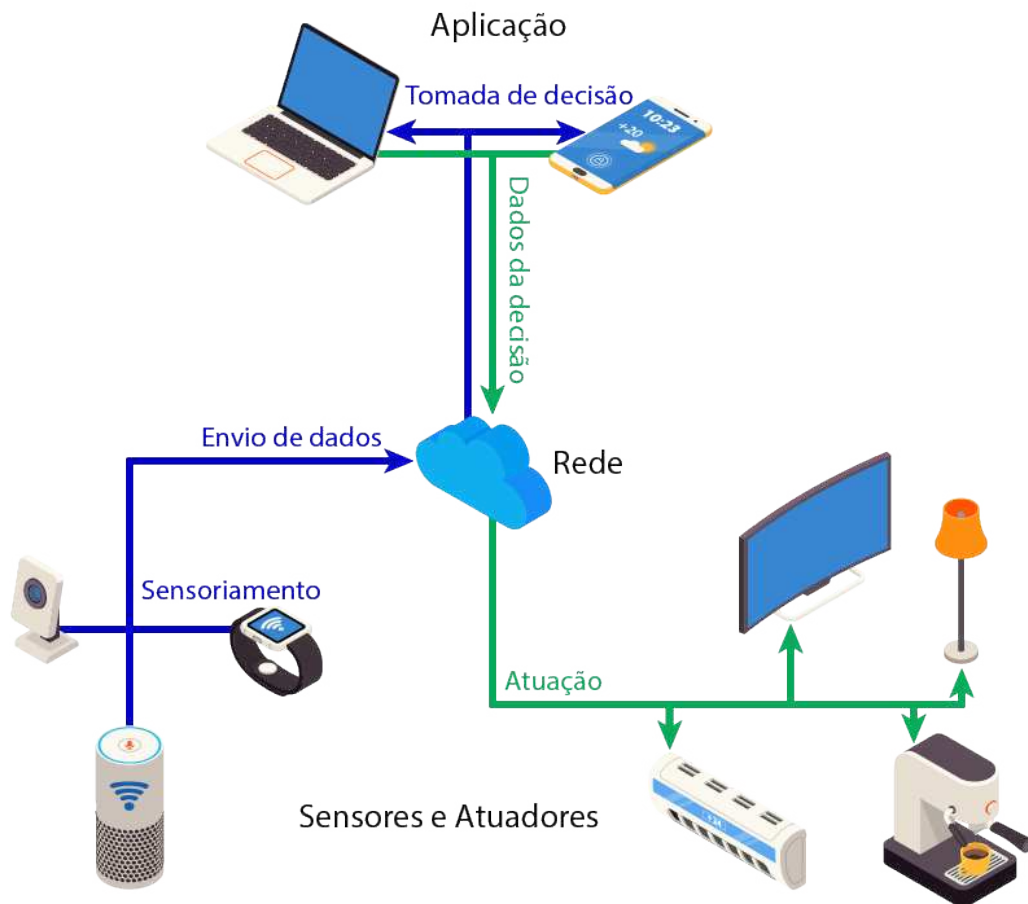


Figura 1 – Fluxo de dados das aplicações IoT. Fonte: O autor. Baseado em (PATEL *et al.*, 2016)

Os **Sensores e Atuadores** estão presentes nos “dispositivos inteligentes”, que são objetos conectados (do inglês *smart objects*) à rede com intuito de fornecer um serviço. Os objetos inteligentes realizam o sensoramento, coletando dados, tais como: temperatura, luminosidade, som, dentre outros, e realizam a atuação no ambiente. Esses objetos em sua grande maioria precisam de uma conexão com um *gateway* para criar uma ponte de comunicação, podendo ser feita de maneira local, rede LAN, Ethernet, WiFi, Bluetooth ou ZigBee.

Na **Rede**, os diversos dados gerados pelos sensores e aplicação são transportados. No transporte é necessário um bom desempenho, para que o sistema esteja pronto a rápidas mudanças de contexto, nesse aspecto começa a ficar mais claro o conceito presente em IoT, *thing-thing*, por meio da interação do envio de mensagem dos sensores e da aplicação através da rede.

Na **Aplicação**, os dados são processados, analisados e tomadas as decisões, gerando

as atuações necessárias.

Em uma aplicação prática de IoT, uma pesquisa abordou (GARRIDO-HIDALGO *et al.*, 2019) o problema de varejista no gerenciamento de cadeia de produção reversa, no qual os pedidos partem do cliente por razões diversas, por exemplo, produto incorreto, produto danificado, entre outros. Isso gerou uma problemática de gerenciamento dessa logística, visto que essa forma de produção pede atenção a novas variáveis, como: quantidade de produtos estornados, condições do produto devolvido, além de ter de acompanhar mais de perto a fabricação do produto. Como forma de solução foi proposto uma cooperação entre diversos protocolos que compõem a IoT, somados ao monitoramento através da nuvem, a fim de produzir um gerenciamento constante dos produtos e assim obter respostas quase em tempo real, evitando maiores perdas.

Observando contextos como esse, pode-se perceber, por exemplo, que a interconectividade em conjunto com a heterogeneidade evidenciadas na IoT, gera, um desafio no âmbito de Verificação & Validação: definir um conjunto de testes comuns para a validação dessas aplicações (AHMAD *et al.*, 2016). Surgindo assim, questionamentos sobre os testes desses sistemas, por exemplo:

- Que tipo de tecnologia pode ser usada para unificar as comunicações de dispositivos IoT?
- Que tipo de informações trafega entre os dispositivos IoT?
- Quais tipos de ferramentas utilizar a fim de obter melhores e mais rápidos resultados?
- Quais partes do software são necessárias testar?

2.2 Teste de software em IoT

Mesmo o teste sendo uma etapa fundamental na construção de um sistema (MYERS *et al.*, 2011), para aplicações IoT ainda há diversos obstáculos a serem explorados. Ao observar a etapa de planejamento dos testes, podemos perceber suas peculiaridades, como por exemplo, a definição do ambiente de teste para aplicações IoT e a geração de casos de testes, que por sua vez podem ser abstratos a fim de adaptar-se a diferentes contextos. Os casos de testes abstratos não possuem valores concretos, como dados de entradas ou resultados esperados, somente a partir da análise de cada contexto específico é que esses casos são preenchidos e concretizados (BOARD, 2018).

Além disso, para realização desses testes, por vezes, são necessárias validações em ambiente real, isso pode requerer testes que envolvam ambientes com pessoas, como vias de transporte, por exemplo. Por isso antes mesmo das aplicações passarem por esses testes em

ambientes reais, são necessários os testes controlados a fim de garantir que não haja acidentes, todo esse processo aumenta a complexidade e custo dos testes. Por exemplo, a troca de dados entre carros deve ser de alta velocidade, contudo garantir esse desempenho é um grande desafio. Por vezes, são necessários testes com grandes volumes de dados que devem ser processados em tempo real e esse processamento depende de hardware e do tipo de rede, logo manter o desempenho em um valor esperado não é simples.

Em um cenário IoT, apresentado na Figura 2, é trago esses conflitos existentes, que por vezes geram os desafios. Na Figura há vários dispositivos se comunicando (*e.g.*, bicicleta, luzes, veículos), mas com propriedades distintas, como, taxas de latências e grau de confiabilidade, porém, estas são propriedades com valores ideais que precisam funcionar juntas de maneira satisfatórias, gerando assim conflitos.

Deste modo é possível observar que as exigências para funcionamento de um sistema IoT são altas, bem como realizar a conexão desses objetos, sendo assim, as exigências de cada aplicação passam a influenciar em todas as aplicações, tornando um desafio suprir essas necessidades, a fim de fornecer um funcionamento ideal. Nesse aspecto residem os desafios de teste, pois validar essas aplicações com diferentes propriedades e manter o equilíbrio entre os valores se torna uma tarefa complexa. Mesmo com as dificuldades, segundo (MARINISSEN *et al.*, 2016), a IoT é um bem passível completamente de uso comercial, devido a sua fragmentação, podendo ser construída por meio de um único sensor, ou até milhares.

As abordagens já existentes da área de Verificação¹ estão sendo utilizadas como meio de avaliar o domínio IoT. Isso ocorre no trabalho de (ALMEIDA *et al.*, 2020), em que um *checklist*, técnica de inspeção guiada por listas de verificação, é proposto como ferramenta de apoio para avaliar como as características IoT afetam o comportamento do usuário. O *checklist* para avaliar a experiência do usuário é estruturado em duas categorias: *Human-Thing Interaction* e *Thing-Thing Interaction*, essas por sua vez se dividem em três subcategorias *General aspects of UX* que se relaciona a satisfação do usuário; *Context awareness* que se relaciona ao sistema fazer uso do contexto para fornecer informação relevante e/ou serviços; e *Programmability* que se relaciona a capacidade do sistema de assumir diversos comportamentos.

Mesmo com a IoT se expandindo, ao analisar o domínio, a partir da perspectiva dos testes, ainda é possível vislumbrar as dificuldades na realização dessa etapa. No Capítulo 3 de Desafios são abordadas as dificuldades presentes na área de testes IoT.

¹ A Verificação envolve a análise de um sistema para certificar se este atende aos requisitos funcionais e não funcionais. (SOMMERVILLE, 2011)

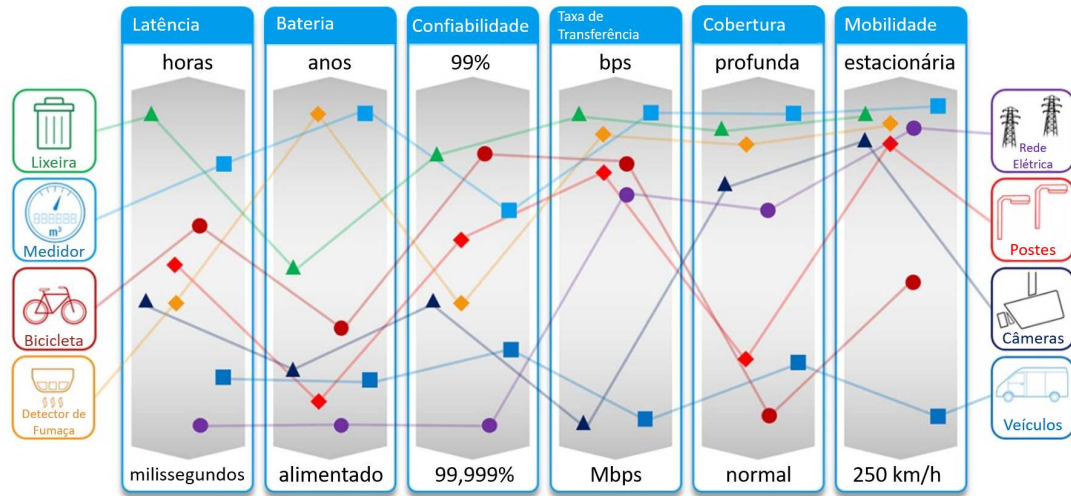


Figura 2 – Exigências de dispositivos para funcionamento ideal. Adaptada (INTERNATIONAL, 2014).

2.3 Teste de Desempenho de Aplicações IoT

A IoT lida com diversos dados que são visualizados, analisados, validados para que a partir disso possam ser feitas inferências. Resultados ágeis irão depender de diversos fatores, como, a velocidade de processamento desses dados, nesse aspecto o teste de desempenho de aplicações IoT é essencial.

De acordo com a (ISO/IEC 25010, 2011), Desempenho é a característica que representa a quantidade de recursos utilizados sob condições estabelecidas e ela se divide em três subcaracterísticas, conforme apresentado na Figura 3. A subcaracterística de “Comportamento Temporal” avalia se a taxa e tempo de resposta e processamento de um produto estão de acordo com os requisitos; a “Utilização de Recursos” verifica se a quantidade e os tipos de recursos utilizados estão de acordo com os requisitos; e a “Capacidade” valida o máximo em que os parâmetros e recursos do sistema estão de acordo com os requisitos.

Conforme apresentado na Seção 2.1, as características IoT possuem um papel relevante nas aplicações. Deste modo é essencial que elas sejam consideradas na condução dos testes. Para que tais características sejam consideradas é necessário haver um mapeamento das mesmas no que tange a aplicação a ser validada e sabendo que a depender da aplicação diferentes características podem apresentar maior ou menor impacto (PATEL *et al.*, 2016). Entretanto, avaliar baseado em características trazem desafios, por exemplo, para a característica de Intero-

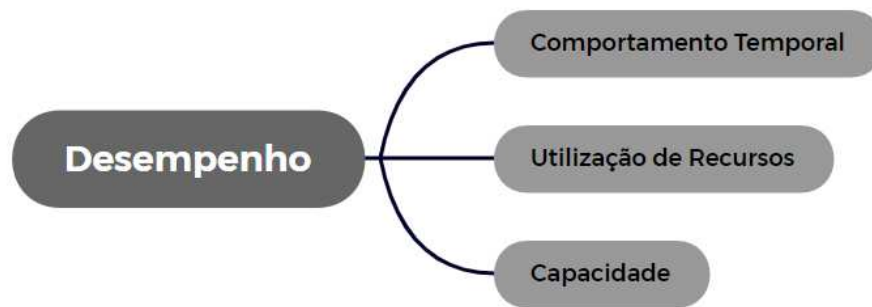


Figura 3 – Desempenho e suas subcaracterísticas. Fonte: O autor.

perabilidade é necessário haver uma abordagem que considere a comunicação das aplicações em diferentes contextos (KIM; ZIEGLER, 2017), assim como para Confiabilidade, no que diz respeito a garantir a integridade dos dados independente do caminho percorrido (AHMAD *et al.*, 2016). Portanto, é evidente que as características devem ser consideradas no âmbito de testes, porém é necessário construir soluções que abranjam os desafios presentes.

De acordo com o (GURIJALA, 2018), para que características IoT sejam avaliadas também é necessário compreender o valor de negócio do produto e realizar uma validação no ambiente real em que os dados transitam. Dessa forma, validar o desempenho em um ambiente IoT não funciona da mesma forma que em um ambiente tradicional, pois em IoT é necessário conhecer a infraestrutura de cada aplicação que fará parte da rede por meio de suas conexões e funções que exercem. Por exemplo, em um ambiente de casa inteligente, onde um assistente virtual se comunica com a TV, é necessário verificar que tipo de sensores há nos dispositivos, bem como sua forma de comunicação e ainda em que momento esses dispositivos deverão atuar, para que haja sincronização e a ação seja executada sem atraso. Nesse cenário, avaliar o desempenho requer um planejamento de Verificação & Validação elaborado, por exemplo, a definição de indicadores e medidas que avaliem as aplicações no contexto ponta a ponta e o desenvolvimento de ferramentas que possibilitem fazer essas medições (KIM; ZIEGLER, 2017), (BRADY *et al.*, 2017), (DAHMANN *et al.*, 2010), (WESSEL; MEYER, 2010).

Na Tabela 2 são sumarizadas as distinções entre aplicações tradicionais e aplicações IoT, por exemplo, em grande parte das simulações para testes tradicionais, os usuários finais são, comumente, responsáveis por validar as aplicações, no contexto de IoT, os dispositivos também terão de passar por simulações como meio de validação - interações *thing-thing*. Para avaliar as aplicações IoT é necessário, por vezes, efetuar testes nos ambientes reais (GURIJALA, 2018), custando mais que de aplicações tradicionais. A respeito dos protocolos ainda não

há um consenso sobre quais utilizar, existem protocolos mais comumente usados que tentam facilitar essa troca de dados, como no caso do MQTT e COAP que estão entre os protocolos IoT mais difundidos (HASSAN *et al.*, 2017), porém, não são os únicos protocolos usados e muitos adaptam protocolos para sua aplicação, gerando uma não-padronização, afetando ainda mais validação desses sistemas. Além disso, o uso de protocolos distintos e sem padronização dificulta a leitura das requisições e respostas recebidas, visto que o processo de troca de mensagem de cada protocolo é diferente. Por razões como estas surgiram inúmeros desafios nos testes de aplicações IoT que serão apresentados no Capítulo 3.

Tabela 1 – Diferença entre testes de desempenho tradicionais e testes de desempenho para IoT (GURIJALA, 2018)

| Crítérios | Teste de Desempenho Tradicionais | Teste de desempenho da IoT |
|-----------------------|---|--|
| Simulação | Simulação de usuários | Simulação de dispositivos/sensores |
| Custos dos testes | Custo padrão para testes de desempenho | Custos mais elevados, devido à configuração do ambiente de teste e a execução de simulações |
| Conjunto de Dados | Envia e recebe grande quantidade de dados por requisição | Envia e recebe o mínimo de dados por requisição, mas os dados são compartilhados continuamente |
| Protocolos | Protocolos padrão de comunicação | Protocolos não padronizados ou protocolos novos |
| Requisições/Respostas | Na maioria dos casos, os usuários criam requisições e recebem respostas | Geralmente dispositivos IoT criam requisições e recebem respostas, assim como recebem requisições e fornecem respostas |

2.4 Considerações finais

Neste capítulo foram apresentados os termos que contribuem para a compreensão deste trabalho. Foi definido o conceito de IoT e o fluxo de dados IoT que se baseia esta dissertação. Além disso, foi apresentado a importância das características IoT na condução dos testes. No capítulo também foi explicado sobre os testes em aplicações IoT, expondo as peculiaridades e especificidade para esses sistemas. Por fim, foi apresentada a característica de desempenho e suas subcaracterísticas, visto que a instanciación deste trabalho foi para a característica de desempenho, também foram elencadas as diferenças entre os testes tradicionais de desempenho e os testes em aplicações IoT. No próximo capítulo são apresentados os desafios para os testes dessas aplicações.

3 DESAFIOS PARA OS TESTES DE APLICAÇÕES IOT

Neste capítulo são apresentados os desafios de testes de aplicações IoT identificados através de um mapeamento sistemático (CARVALHO, 2018). Na Seção 3.1 são elencados os desafios no âmbito de testes IoT, na Subseção 3.1.1 são apresentados os desafios nos testes de desempenho de aplicações IoT, este trabalho propõem um guia de teste e a instanciação do mesmo ocorre para a característica de desempenho. Na Seção 3.2 são elencadas as lacunas, que se preenchidas, auxiliam na resolução dos desafios.

3.1 Categorização dos desafios de testes IoT

Vários estudos na literatura propõem estratégias para lidar com os problemas inerentes ao teste de aplicações IoT (BURES *et al.*, 2019), (SAND, 2015), (MARINISSEN *et al.*, 2016), há também estudos que abordam as relevâncias das correlações das características IoT (CARVALHO *et al.*, 2020), outros trabalhos propuseram um modelo de avaliação para sistemas ubíquos (SANTOS *et al.*, 2013), área que antecedeu a IoT (ANDRADE *et al.*, 2017). Como resultados preliminares da proposta deste trabalho, foram identificados os principais desafios e as características mais críticas¹ no teste de aplicações IoT através de um mapeamento sistemático (CARVALHO, 2018). Como resultado deste estudo, as características de “Segurança”, “Interoperabilidade” e “Desempenho”, nessa ordem, foram as que obtiveram mais desafios relacionados. Muitos dos desafios enfrentados estão relacionados diretamente a uma ou mais particularidades de IoT, dentre eles, pode-se citar:

- **Grande quantidade de dispositivos.** Uma aplicação sob teste (AUT) em IoT varia de um único dispositivo a infraestruturas e plataformas altamente dinâmicas. Nesse caso, as técnicas de teste precisam ser capazes de lidar com muitos dispositivos com interfaces abertas. Esses dispositivos são sensores, atuadores, microcontroladores ou *gateways*. Às vezes, eles são instalados em ambientes não confiáveis, além de trabalharem, por exemplo, no caso de sensores e microcontroladores sob limitação de recursos, como fornecimento de energia ou disponibilidade de rede. Dentre os desafios relacionados, têm-se: “*como orquestrar a quantidade de dispositivos IoT e as limitações existentes?*” (FARAHMANDPOUR *et al.*, 2017), (KRISTOFFERSEN *et al.*, 2017), (SCHIEFERDECKER *et al.*, 2017).
- **Diversidade de dispositivos distribuídos.** A configuração de dispositivos varia de acordo

¹ Entende-se por características críticas, as quais há um maior investimento de esforços em atividades de V&V, em termos de número de estudos publicados.

com o número de dispositivos que colaboram na rede, isso requer suporte automatizado para testes de interoperabilidade de vários dispositivos distribuídos, diferente do modelo tradicional que requer esforço humano para coordenar e configurar os testes para vários dispositivos. Dentre os desafios relacionados, têm-se: “*como gerenciar os testes, sabendo da distribuição de diferentes dispositivos na rede?*” (REETZ *et al.*, 2013), (KIM *et al.*, 2018), (GARCÍA *et al.*, 2018), (FARAHMANDPOUR *et al.*, 2017), (AMALFITANO *et al.*, 2017), (SCHIEFERDECKER *et al.*, 2017), (ROSENKRANZ *et al.*, 2015), (LIU *et al.*, 2013), (KOUICHE, 2012), (BERA; PASALA, 2012).

- **Falta de padronização das plataformas para software e hardware.** A maioria das ferramentas de testes tem dificuldades em interagir com os aplicativos IoT devido à heterogeneidade tanto de hardware quanto de software e, portanto, acarretando problemas para automatizar alguns tipos de testes. Dentre os desafios relacionados, têm-se: “*como promover, durante os testes, a interação entre as aplicações IoT, considerando a heterogeneidade de software e hardware?*” (KANSTRÉN *et al.*, 2018), (PONTES *et al.*, 2018a), (BURES *et al.*, 2019).
- **Detecção de faltas.** É uma tarefa difícil e complexa porque existem muitos fatores que influenciam os dados e podem causar falhas, além disso, elas dependem do tipo de aplicação e do sensor. Dentre os desafios relacionados, têm-se: “*como manipular os diferentes fatores que rodeiam a Internet das Coisas?*” (ALIPOUR, 2017), (LIU *et al.*, 2013).
- **Gerenciamento dos envolvidos.** O teste de dispositivos IoT envolve grupos de partes interessadas maiores e mais heterogêneos do que em outros domínios de software típicos. Nesse caso, estabelecer os critérios de validação é bem mais difícil. Dentre os desafios relacionados, têm-se: “*como promover a interação entre os envolvidos no domínio?*” (KIM *et al.*, 2018).
- **Manutenibilidade da suíte de testes.** Devido ao número crescente, de tamanho e complexidade das aplicações IoT, é imprescindível que os testes de sistemas aumentem os casos de testes funcionais, afetando diretamente na capacidade de manutenção da suíte de testes. Dentre os desafios relacionados, têm-se: “*como gerenciar o volume de suítes de testes para IoT?*” (HAMAD, 2018), (ALI *et al.*, 2012).
- **Integração de Testes.** As diferentes aplicações exigem diferentes tipos de testes. Quando essas aplicações se comunicam são necessários testes que validem essa integração. Dentre

os desafios relacionados, têm-se: “*como elaborar testes que integrem diferentes aplicações IoT?*” (HAMAD, 2018), (NEVES *et al.*, 2018), (BURES *et al.*, 2019).

- **Necessidade de conectividade intensa.** Os testes de aplicações IoT exigem comunicação e conectividade intensivas, por exemplo, na nuvem e em soluções críticas, como no caso de serviços médicos que exigem a garantia da confiabilidade do serviço. Dentre os desafios relacionados, têm-se: “*como garantir as conexões necessárias?*” (HAMAD, 2018), (BURES *et al.*, 2019).
- **Contextos heterogêneos.** A grande dificuldade é validar de forma homogênea, vários dispositivos com diferentes padrões de hardware e software. Dentre os desafios relacionados, têm-se: “*como projetar testes que possam ser usados para diferentes contextos e que envolva diferentes arquiteturas?*” (KIM; ZIEGLER, 2017), (ROSENKRANZ *et al.*, 2015), (CHEN *et al.*, 2013), (SAND, 2015), (AMALFITANO *et al.*, 2017), (GARCÍA *et al.*, 2018).
- **Manutenção da integridade dos dados.** É uma das características de IoT que mais apresenta desafios relacionados (CARVALHO, 2018). Testar a segurança em aplicações IoT deve abordar diversos aspectos comuns de segurança, tais como transição dos dados e criptografia com o intuito de diminuir a vulnerabilidade do sistema antes de ser executado. Dentre os desafios relacionados, têm-se: “*como manter a privacidade dos dados dos usuários (e.g., informações pessoais) e garantir que esses dados permanecerão inalterados e anônimos?*” (HAMAD, 2018), (GARCÍA *et al.*, 2018), (SACHIDANANDA *et al.*, 2017), (ALIPOUR, 2017), (VISOOTTIVISETH *et al.*, 2017), (NAIR *et al.*, 2017), (YAN *et al.*, 2012), (BURES *et al.*, 2019).
- **Limitação de recursos.** Os dispositivos IoT possuem recursos limitados, normalmente eles são alimentados por baterias e frequentemente adormecem para reduzir o seu consumo. Em IoT, as limitações de energia também podem ocasionar restrições de disponibilidade na rede e, conseqüentemente, perdas de pacotes podem ocorrer durante a transferência de dados. Dentre os desafios relacionados, têm-se: “*como orquestrar os recursos necessários e suas limitações?*” (ROSENKRANZ *et al.*, 2015), (KANSTRÉN *et al.*, 2018), (CHEN; VIHO, 2012), (GIMÉNEZ *et al.*, 2013), (CHEN *et al.*, 2013), (BURES *et al.*, 2019).
- **Grande esforço e alto custo.** O problema nos testes de aplicações IoT pode ocorrer se o número de interações entre os dispositivos for significativamente grande e, conseqüentemente, um número muito grande de casos de teste deve ser gerado. Nesse cenário, a

criação de um ambiente de teste real é difícil e custa caro devido à grande quantidade de esforço humano e investimento em hardware. Dentre os desafios relacionados, têm-se: *“Quando as atividades de Verificação & Validação devem ser realizadas? Elas deveriam ser realizadas cada vez que uma aplicação sofre alguma alteração? Isso custaria caro, porém, por outro lado, as mudanças podem incorrer em riscos para os diferentes participantes”* (KANSTRÉN *et al.*, 2018), (KIM *et al.*, 2018), (BRADY *et al.*, 2017), (NEVES *et al.*, 2018).

Dentre os desafios citados, também existem os relacionados à característica de Desempenho, que serão abordados na próxima subseção.

3.1.1 Desafios do testes de desempenho e IoT

Avaliar o desempenho de sistemas IoT envolve validar o comportamento individual das aplicações em um contexto ponta a ponta (*end-to-end*). A criação de um ambiente de teste com parâmetros ou propriedades que se adequem às circunstâncias específicas dos dispositivos quando estes interagem, por exemplo, em situações nas quais os dispositivos deveriam ter um gasto mínimo de energia e, com base nessa informação, os valores de recursos de hardware (*e.g.*, memória) deveriam ser otimizados para aumentar o desempenho no ambiente de execução. Por razões, como essas, surgiram diversos desafios nos testes de desempenho. Os principais desafios enfrentados na área de Desempenho de acordo com os autores (GURIJALA, 2018), (MARINISSEN *et al.*, 2016), (SAND, 2015), (BRADY *et al.*, 2017), (LIU *et al.*, 2013) são:

- **Ações em tempo real**

Com o surgimento da IoT houve a adição de funções em diversos produtos que estão inseridos em nosso dia a dia, com isso as respostas em tempo real se tornaram um requisito relevante nesses produtos. O monitoramento rápido surge como uma necessidade para agilizar as tomadas de decisões e gerar resultados em “tempo real”. Um dos desafios enfrentados é: *“como gerar resultados em tempo real, dado limitações como a rede e o grande volume de dados?”*.

- **Grande volume de dados**

Transmitir grandes quantidades de dados torna consideravelmente mais complexa a detecção de comportamentos anômalos dentro de uma aplicação IoT. Dentre os desafios relacionados, têm-se: *“como baratear o custo de testes em ambientes reais?”*, e *“como*

garantir uma previsibilidade em testes em ambiente reais?"

- **Protocolos**

A IoT trata a comunicação de diversos dispositivos, sendo um desafio unificar um protocolo comum para comunicação. Além disso, os protocolos existentes sofrem diversas alterações por serem novos e precisarem se adaptar aos novos contextos de aplicações, isso dificulta uma padronização e criação de ferramentas que auxiliem os testes dessas aplicações. Dentre os desafios enfrentados, pode-se citar: *“como padronizar os protocolos IoT?”*, e *“como garantir a funcionalidade desses protocolos?”*.

- **Simulação de ambientes e alto custo**

Para que a validação de uma aplicação IoT seja garantida, incluindo o desempenho, é necessário a simulação. Simular ambientes reais se torna custoso, não apenas por demandar um alto investimento financeiro, mas por ser difícil garantir a consistência dos dados nesses ambientes, tornando uma tarefa desafiadora. Dentre os desafios relacionados, têm-se: *“como baratear o custo de testes em ambientes reais?”*, e *“como garantir uma previsibilidade em testes em ambiente reais?”*

- **Distribuição geográfica**

Os dispositivos IoT estão espalhados pelo mundo todo, trocando informações. Esses dados precisam ser criptografados, enviados, decodificados, processados, de tal maneira a garantir respostas rápidas. Sendo necessário testar essas conexões a fim de verificar a velocidade das redes usadas. Nessa direção, *“como garantir respostas rápidas, dada a distância geográfica entre os dispositivos?”*

- **Fatores externos**

No contexto IoT há diversas aplicações que atuam no meio externo. No processo de verificação dessas aplicações deve haver uma maior atenção aos fatores externos que podem afetar o desempenho de suas aplicações, tais como ventos fortes, chuvas, altas temperaturas, que para determinado contexto podem ser fatores importantes. Um dos desafios se refere à *“como garantir total atenção a fatores ambientais, a fim de que os mesmos não afetem o desempenho da aplicação?”*

3.2 Lacunas identificadas nos testes IoT

A partir dos desafios elencados na Seção 3.1 foram identificadas lacunas que se tratadas poderiam vir a auxiliar na solução desses desafios.

- Ausência de soluções **estruturadas** para o teste de aplicações IoT (CARVALHO, 2018). A partir do desafio de “*Grande esforço e alto custo*” é possível identificar que o fornecimento de uma solução estruturada facilitaria a padronização no processo de teste dessas aplicações.
- Carência de abordagens **flexíveis** que possam validar diferentes características de IoT (SAND, 2015). Por a IoT ser composta de um ambiente heterogêneo, uma abordagem capaz de se adaptar a múltiplas características para o processo de validação auxiliaria, por exemplo, no desafio de “*Contextos heterogêneos*” ao propor uma solução que se adapta de maneira eficiente.
- Ausência de catalogação de **ferramentas e medidas** para o teste de aplicações IoT (PONTES *et al.*, 2018a). O fornecimento de ferramentas e medidas que auxiliam nos testes garantem uma padronização e automatização do processo, auxiliando na solução do desafio de “*Falta de padronização para plataformas de software e hardware*”.
- Falta de abordagens de teste que considerem o impacto da **correlação**² das características (CARVALHO *et al.*, 2020). Propor soluções que abrangem as correlações das características auxiliaria na detecção dos impactos que cada característica tem uma nas outras, contribuindo para a solução dos desafio de “*Gerenciamento dos envolvidos*”, além de auxiliar na condução da “*Integração dos testes*” de maneira mais eficaz, atingindo assim a outro desafio.

3.3 Considerações finais

Neste capítulo apresentamos os desafios gerais que permeiam as aplicações IoT, esses desafios foram categorizados de acordo com as particularidades de IoT. Dentre esses desafios se destaca o desempenho dessas aplicações. Os desafios de desempenho abrangem diversos aspectos desde a questão geográfica dos dispositivos até a diversidade de protocolos. Por fim apresentamos lacunas que se resolvidas auxiliam na solução de alguns dos desafios de IoT.

² As correlações representam as interações presentes entre as características de qualidade (CARVALHO *et al.*, 2020).

4 TRABALHOS RELACIONADOS

Neste Capítulo serão apresentados os trabalhos relacionados identificados na literatura, esses trabalhos estão divididos em Soluções gerais apresentados na Subseção 4.1.1 e em Guias para testes na Subseção 4.1.2, por fim é feita uma discussão sobre as análises dos estudos.

4.1 Trabalhos identificados

Mediante os desafios apresentados, buscou-se na literatura trabalhos que cobrissem as lacunas discutidas na Seção 3.2. Durante as revisões feitas na literatura não foram encontrados trabalhos que propunham, nesse mesmo contexto, guias. Entretanto, buscamos elaborar uma tabela comparativa entre os estudos selecionados de acordo com as lacunas apresentadas.

Na Tabela 2 são apresentados os trabalhos relacionados de acordo com as lacunas elencadas: se o trabalho propõe uma estrutura ou padronização da área; se a abordagem proposta é flexível e adaptável a diversas características IoT; se o estudo sugere métricas e ferramentas que auxiliem no processo de teste dessas aplicações; se é apresentada alguma correlação das características IoT; e a característica IoT abordada no trabalho.

Na Tabela 2, além dos trabalhos serem apresentados baseados nas lacunas, é feita a separação em dois grupos: trabalhos que apresentam soluções gerais na área de testes IoT e trabalhos que fornecem soluções por meio de guias ou padrões na área de testes IoT, ou demais áreas. Na última linha da tabela é apresentada a proposta deste trabalho como comparativo com as demais soluções.

Tabela 2 – Trabalhos realizados como soluções para teste de aplicações IoT

| ID | Trabalho | Estrutura | Flexibilidade | Ferramentas e Medidas | Correlação | Característica |
|--|-------------------------------------|------------|---------------|-----------------------|------------|-------------------|
| Trabalhos que apresentam soluções gerais | | | | | | |
| 1 | (RAMPRASAD <i>et al.</i> , 2018) | Não | Sim | Sim | Não | Heterogeneidade |
| 2 | (JEANNOTTE; TEKEOGLU, 2019) | Não | Sim | Sim | Não | Segurança |
| 3 | (LANZA <i>et al.</i> , 2018) | Sim | Não | Não | Não | - |
| 4 | (CARVALHO <i>et al.</i> , 2018) | Sim | Sim | Não | Não | Desempenho |
| Trabalhos que apresentam soluções por meio de padrões e guias | | | | | | |
| 5 | (PONTES <i>et al.</i> , 2018b) | Sim | Não | Não | Não | - |
| 6 | (DOD DIRECTOR <i>et al.</i> , 2008) | Sim | Não | Não | Não | - |
| 7 | (MEIER <i>et al.</i> , 2007) | Sim | Sim | Não | Não | Desempenho |
| 8 | (ROSA <i>et al.</i> , 2018) | Sim | Não | Sim | Não | Segurança |
| 9 | Proposta desta dissertação | Sim | Sim | Sim | Sim | Desempenho |

4.1.1 Soluções gerais

O Trabalho 1 (RAMPRASAD *et al.*, 2018) apresenta a produção de um *framework smart* para testes, onde através da execução de experimentos repetíveis é aprendido o comportamento do sistema a fim de melhor atuar nas mudanças de contexto. O trabalho apresenta teste de desempenho automatizado, com foco na característica de heterogeneidade, porém a solução não menciona a correlação das características IoT e não é fornecido um procedimento estruturado que auxilie a execução do processo.

O Trabalho 2 (JEANNOTTE; TEKEOGLU, 2019) apresenta uma solução para detectar com maior facilidade as vulnerabilidades comuns de dispositivos IoT. A solução realiza a junção de aplicações (Nikto, OWASP-Zap, Nmap e Hydra) que escaneiam a vulnerabilidade de sistemas, baseado no modelo top 10 da OWASP¹. O modelo apresenta boas práticas para educar desenvolvedores, projetistas, arquitetos, gestores e organizações sobre as consequências das mais importantes vulnerabilidades de segurança de aplicações web, a solução propõe cobrir as categorias 1, 2, 3, 4 e 9 do modelo. Semelhante ao trabalho anterior, a proposta de solução não apresenta passos estruturados e não correlaciona as demais características IoT com a de Segurança abordada no estudo.

O Trabalho 3 (LANZA *et al.*, 2018) abordou a problemática de escassez de infraestrutura de grande escala que possibilitem a avaliação de soluções desenvolvidas em âmbito mundial, essas infraestruturas quando disponíveis, não são acessíveis aqueles dispostos a testar suas inovações. A solução é uma plataforma implementada para permitir a *Everything as a Service*² (EaaS) em várias *testbeds* de IoT simples, a fim de facilitar o processamento, executando-o na nuvem. A solução exige a implementação das semânticas e padrões próprios da solução, não sendo uma abordagem flexível, além de não tratar de ferramentas e métricas para o auxílio nos testes e não mencionar as correlações de características IoT.

O Trabalho 4 (CARVALHO *et al.*, 2018) apresenta o LoRaWAN³, uma tecnologia de referência, por ser de natureza aberta e possuir uma variedade de módulos disponíveis, no entanto, não há muitas discussões relacionadas à infraestrutura de atrasos. Neste contexto, esse trabalho propõe a elaboração de uma metodologia de medição para avaliar experimentalmente o desempenho, relacionado ao tempo de execução de toda uma infraestrutura real do LoRaWAN. A metodologia não sugeriu melhoria no desempenho, quanto ao crescimento exponencial do

¹ <<https://owasp.org/www-project-top-ten/>>

² <<https://tinyurl.com/3y7udsfj>>

³ <<https://loro-alliance.org/about-lorawan>>

LoRaWAN em meio ao grande volume de dados. No trabalho os autores não sugerem ferramentas para automatizar o processo e a característica de desempenho na qual o trabalho aborda não foi correlacionada com as demais características.

4.1.2 Guias para testes

Outros trabalhos identificados na literatura, propuseram padrões e guias na área de teste IoT e em demais áreas, estes estudos são apresentados a seguir.

O Trabalho 5 (PONTES *et al.*, 2018b) apresenta uma solução para a falta de padronização nos testes IoT. Neste trabalho um conjunto de estratégias de teste é associado a um padrão de IoT e com isso cinco diferentes padrões de teste foram identificados: *Test Periodic Reading*; *Test Triggered Reading*; *Test Alerts*; *Test Actions*; e *Test Actuators*. Esse trabalho trouxe semelhança ao guia proposto, pois apresenta uma estrutura para a condução dos teste por meio de padrões, mas o contexto em questão são os testes funcionais da aplicação, não se aplicando a características específicas de IoT, por exemplo, o padrão de *Test Actuators* é utilizado para verificar se o atuador está executando a sua função de acordo com o esperado. Os autores mencionam, como trabalho futuro, propor alguma solução para testes não-funcionais, mas até o presente momento essa solução não foi proposta. O trabalho também não menciona ferramentas e métricas que auxiliem no processo de teste IoT.

O Trabalho 6 (DOD DIRECTOR *et al.*, 2008) apresenta uma estrutura inicial de um guia desenvolvido pelo Departamento de Defesa dos Estados Unidos. No guia é apresentada uma visão geral sobre “Sistemas de Sistemas” (SoS) e as abordagens utilizadas, comumente, para a construção desses sistemas. O guia traz uma compreensão sobre a área de SoS e como esta se relaciona com as normas do departamento de Defesa do país, porém a versão disponível do Guia não abrange aspectos práticos, são apenas propostas definições e dada uma contextualização da área. A respeito das lacunas o trabalho cobre apenas o fornecimento da estrutura, as demais lacunas não são cobertas.

O Trabalho 7 (MEIER *et al.*, 2007) propõe um guia de desempenho para aplicações Web. No guia são apresentadas questões como: as principais atividades do teste de desempenho, como garantir o desempenho em aplicações Web, entre outras. Por mais que esse guia tenha semelhança ao proposto neste trabalho, ele não aborda IoT. Os campos fornecidos no guia são gerais, como planejamento dos testes, execução dos testes e análises. No trabalho não são elencadas ferramentas ou métricas e não é mencionado as correlações de características com

desempenho. O objetivo principal do trabalho é definir um modelo de teste de desempenho para aplicações Web, diferindo assim, em foco e abrangência ao objetivo deste trabalho.

Trabalho 8 (ROSA *et al.*, 2018) apresenta heurísticas para avaliação da Segurança em IoT. Nesse trabalho é definido formalmente o termo segurança e são apresentadas 11 propriedades. A partir das heurísticas é possível selecionar e priorizar avaliações a serem realizadas na segurança de aplicações IoT. Os autores não apresentam uma estrutura clara para condução da avaliação, mas ao ler o trabalho é possível identificar um procedimento metodológico, em que primeiro são apresentadas as definições de segurança; em seguida as suas 11 propriedades; logo depois a heurística para a seleção dos aspectos que serão avaliados em segurança; e as métricas que poderão ser executadas durante a avaliação. A segurança na qual se baseia o trabalho é fundamentada na definição da ISO 27001, porém de acordo com os autores a abordagem pode ser adaptada ao contexto de segurança IoT de cada aplicação. Apesar dos autores defenderem a proposta de um guia adaptável, no trabalho não é fornecida nenhuma instrução que facilite esse processo, além de não ser apresentada a correlação da característica de segurança com as demais de IoT.

4.2 Discussão

Após a análise dos trabalhos relacionados, foi possível identificar que no contexto da área de testes em aplicações IoT, não há trabalhos que satisfaçam por completo as lacunas apresentadas (ver Tabela 2). No entanto, a importância do tratamento destas lacunas já foram evidenciadas, conforme apresentado na Seção 3.2. Portanto, trabalhos que atuem na busca de soluções para essa área são relevantes e oportunos. Deste modo, conforme apresentado na última linha da Tabela 2, este trabalho se propõe a satisfazer todas as lacunas apresentadas, por meio de uma abordagem de guias estruturados, sendo essa abordagem instanciada para a característica IoT de Desempenho. A metodologia para condução da solução proposta será apresentada no capítulo seguinte.

4.3 Considerações finais

Esse capítulo tratou dos trabalhos relacionados obtidos a partir de uma revisão da literatura, em que foram analisados 8 trabalhos que abordavam as lacunas apresentadas na seção 3.2. Entretanto, não encontramos nenhum estudo que se referia de fato a um guia de teste

IoT, além do que os estudos encontrados não satisfaziam todas as lacunas. Porém, o estudo aqui proposto busca cobrir todas as lacunas observadas. No próximo Capítulo será apresentada a abordagem utilizada.

5 ABORDAGEM DE TESTE BASEADA EM GUIAS ESTRUTURADOS

Conforme mencionado na introdução desta dissertação, este trabalho propõe a elaboração de uma abordagem de teste IoT que é apresentada em modelo de um guia. Entende-se como guia uma sequência de tópicos que permite facilitar o entendimento e a execução de atividades para o âmbito ao qual é designado. Além disso, a abordagem de teste proposta nesta dissertação é para a avaliação de características IoT, deste modo, o modelo de guia proposto, ou seja, os tópicos sugeridos, tem como foco as características IoT, sendo assim capaz de ser aplicado a qualquer características IoT. Será tratado no Capítulo 6 sobre como aplicar o guia à qualquer característica IoT, a esse processo é dado o nome de "instanciação".

Para a elaboração da abordagem foram seguidos dois passos macros: (i) construir os tópicos para compor o guia e (ii) estruturar a sequência em que esses tópicos são apresentados no guia. Esse primeiro passo foi nomeado de "construção", como resultado desse passo macro são obtidos os tópicos que irão compor o guia, e o segundo passo foi nomeado de "estruturação", como resultado desse passo macro é gerado a sequência estrutural em que são organizados os tópicos obtidos, ou seja, é gerada a estrutura do guia.

No atual Capítulo são apresentadas a construção e estruturação do guia de teste. Deste modo, na Seção 5.1 é apresentada a metodologia para a construção e estruturação do guia de teste e na Seção 5.2 a própria estrutura do guia de teste.

5.1 Construção e estruturação do guia

O guia foi elaborado a partir de 5 atividades conforme apresentado na Figura 4, essas atividades consistiam em:

1. realizar uma revisão da literatura utilizando uma *string* de busca;
2. analisar dois repositórios de *bugs* de aplicações IoT;
3. extrair dados obtidos na revisão e de outras fontes;
4. construir a estruturação inicial do guia;
5. refinar a estrutura do guia.

Para a construção do guia foram inicialmente conduzidas duas atividades em paralelo: “Realizar uma revisão da literatura”¹ e “Analisar repositórios”. Na primeira atividade foi elaborado o protocolo de revisão, baseado no método proposto por (KITCHENHAM *et al.*,

¹ Revisão para construção do guia: <<https://tinyurl.com/22kzabxk>>

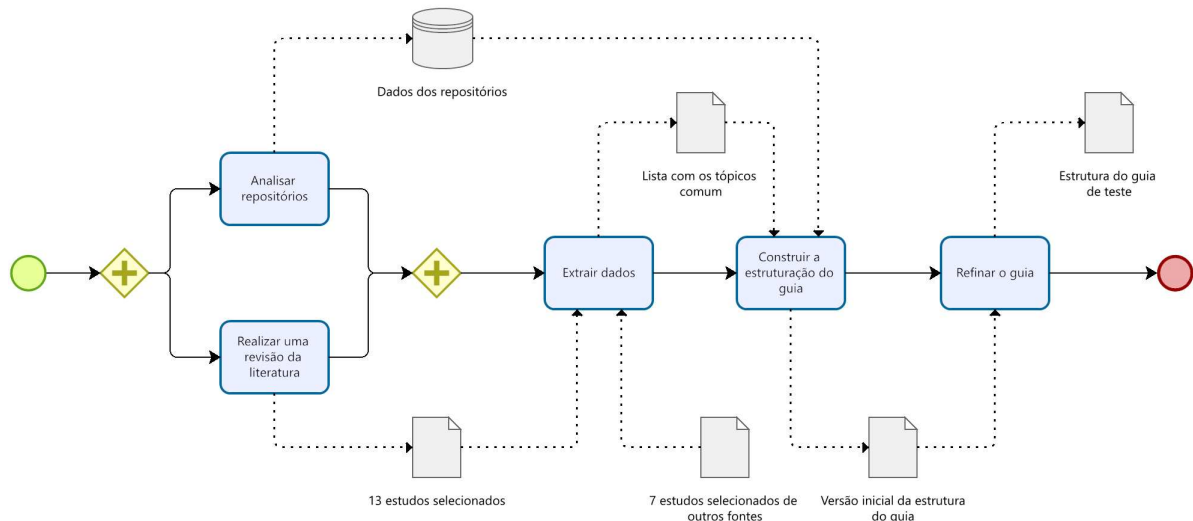


Figura 4 – Procedimento metodológico para a construção e estruturação do guia. Fonte: O autor.

2009) (ver Apêndice A). O protocolo apresenta as questões de pesquisa e a *string* de busca para identificar estudos que pudessem auxiliar na construção da estrutura do guia, tais como: catálogos, *checklists* e diretrizes. A *string* é apresentada na Tabela 3 e foi executada na base de dados *Scopus*, por se tratar de uma base indexadora onde acopla outras demais bases. Como resultado, foram obtidos 707 estudos e, destes, selecionados 13 (LEAL *et al.*, 2014; BURATTI *et al.*, 2015; KIM *et al.*, 2016; KOVACS *et al.*, 2016; FUCHS *et al.*, 2016; BYAMUKAMA *et al.*, 2017; GEIGER; HEADRICK, 2017; WU; CHIU, 2017; FEHLMANN; KRANICH, 2017; COSTANTINO *et al.*, 2018; LOWENSTEIN; SLATER, 2018; FAHMI; ABDUROHMAN, 2017; PONTES *et al.*, 2018a). Esses estudos são listados em “Estudos Selecionados na Revisão” no Apêndice B. Além desses estudos, outros 7 estudos relevantes relacionados com IoT (MAIA *et al.*, 2016; AHMAD *et al.*, 2016; MATHEU-GARCÍA *et al.*, 2019; CARVALHO *et al.*, 2020; GUIDE... , 2016; SANTOS *et al.*, 2013; ISO/IEC, 2004) foram utilizados para compreender as características de IoT e as suas correlações. A lista desses estudos está em “Estudos Selecionados de outras fontes” no Apêndice B. Na segunda atividade, “Analisar Repositórios”, foram selecionados, a partir do *GitHub*, dois repositórios de *bugs* IoT². O principal objetivo foi compreender as falhas mais comuns que ocorrem durante os testes das aplicações IoT para melhor fundamentar a estrutura do guia proposto.

Para “Extrair dados” obtidos na revisão e de outras fontes, foi utilizado um Formulário de extração de dados: com objetivo de extrair os dados e auxiliar na aceleração desse processo (ver Apêndice C). Na atividade de “Construir a estrutura do guia”, foram identificados, por meio dos 20 estudos (ver estudos no Apêndice B), tópicos que eram relevantes para a validação de

² Exemplo do repositório selecionado: <<https://github.com/taosdata/TDEngine/issues>>

Tabela 3 – Construção da string de busca

| Termos | Termos Relacionados |
|---|---|
| Internet of Things | - |
| Test | Testing |
| Guideline | Standard, Catalog, Checklist, Guide, Guidance |
| String Final | |
| ("internet of things") AND (test OR testing) AND (guideline OR standard OR catalog OR checklist OR guide OR guidance) | |

aplicações IoT, por exemplo:

- Casos de teste;
- Ferramentas;
- Identificação das correlações entre as características;
- Métricas; dentre outros.

5.1.1 Identificação dos tópicos a partir da literatura

Os tópicos estruturam o guia e servem de base para facilitar os testes das aplicações da IoT e são relevantes para fornecer uma melhor compreensão do âmbito dos testes. Na Tabela 4 é apresentado os aspectos que derivaram cada tópico a partir dos estudos da literatura. Inicialmente, na Tabela são apresentados os ids de cada item, foram elencados 7 aspectos durante a leitura dos estudos. Na segunda coluna são listados os aspectos relevantes para os testes. Na terceira coluna é sintetizado o contexto do qual derivou os tópicos. E na quarta coluna são elencados os nomes iniciais dados aos tópicos. Um exemplo apresentado na Tabela é o aspecto de Ferramentas do qual derivou o tópico, que inicialmente recebeu o nome de Ferramentas, apoiado nos trabalhos citados (COSTANTINO *et al.*, 2018), (BURATTI *et al.*, 2015), (GEIGER; HEADRICK, 2017), visto que pode auxiliar no processo de teste IoT.

Resumindo, a proposta inicial de cada tópico. O primeiro de “Correlação de Características” deveria ser abordado as relações presentes entre a característica alvo e as demais; seguindo do “Custo-benefício” em que, por meio de uma fórmula proposta, buscou-se definir um custo-benefício para a realização dos testes. No terceiro tópico “Casos de testes abstratos e medições”, dado a relevância dos dois aspectos observados na literatura, ficou evidente a importância de tê-los na abordagem, definidos por meios da busca de métricas na literatura e da construção de cenários de testes passíveis de adaptação. No quarto tópico “Ferramentas” o objetivo proposto é de recolher ferramentas que auxiliassem na aceleração e automação de parte do processo de testes. O quinto tópico “Impacto das subcaracterísticas” faria um papel

Tabela 4 – Construção inicial dos tópicos da estrutura do guia

| ID | Aspectos Relevantes | Contexto | Nomes iniciais do Tópicos |
|----|-------------------------------|--|-------------------------------------|
| 1 | Correlação | As correlações trazem impactos sobre as características. (MAIA <i>et al.</i> , 2016),(CARVALHO <i>et al.</i> , 2020) | Correlação das Características |
| 2 | Custo benefício | Análise dos custos para execução de testes em aplicações IoT são essenciais.(WU; CHIU, 2017),(COSTANTINO <i>et al.</i> , 2018) | Custo-benefício |
| 3 | Casos de testes e Métricas | Construção e elaboração de cenários de testes para fazer as validações em conjunto com a coleta de métricas, a fim de facilitar a identificação de falhas nos sistemas. (KOVACS <i>et al.</i> , 2016),(PONTES <i>et al.</i> , 2018a),(LOWENSTEIN; SLATER, 2018),(KIM <i>et al.</i> , 2016) | Casos de Teste Abstratos e Medições |
| 4 | Ferramentas | Ferramentas auxiliam na automatização de processos custosos de testes em sistemas IoT. (COSTANTINO <i>et al.</i> , 2018),(BURATTI <i>et al.</i> , 2015),(GEIGER; HEADRICK, 2017) | Ferramentas |
| 5 | Conflitos de interesse | A mesma aplicação pode buscar diferentes interesses que conflitam entre si. (LEAL <i>et al.</i> , 2014),(KIM <i>et al.</i> , 2016) | Impacto das Subcaracterísticas |
| 6 | Ambiente necessário | Definição clara dos recursos necessários para a execução da validação.(GEIGER; HEADRICK, 2017),(BURATTI <i>et al.</i> , 2015),(LEAL <i>et al.</i> , 2014) | Ambiente de Teste |
| 7 | Divisão em subcaracterísticas | A análise do sistema feita de maneira particionada pode garantir de forma mais segura a cobertura total. (FUCHS <i>et al.</i> , 2016),(FEHLMANN; KRANICH, 2017) | Subcaracterísticas |

semelhante ao da correlação, porém internamente, o usuário do guia elencaria as relações entre as subcaracterísticas. O sexto tópico “Ambiente de teste” foi criado com intuito de expressar quais recursos deveriam ser configurados para execução dos testes. O sétimo tópico de “Subcaracterísticas” teve como foco gerar separações dentro do guia para cada uma das subcaracterísticas da característica alvo.

Esses tópicos compuseram a estrutura inicial do guia. Essa versão foi analisada por um especialista em Teste de software e Computação móvel, o qual foram feitas sugestões, como, a separação dos tópicos Casos de Teste Abstratos e Medições que, de início, haviam sido organizados juntos. Foi sugerido também a criação de tópicos de definições, como foi o caso de Definição da Característica e Definição da Subcaracterística, um exemplo de uso do guia, além de mudanças nos nomes de alguns tópicos e melhorias internas. A partir das sugestões foi

conduzida a última atividade de “Refinar o guia” em que essas mudanças foram feitas e assim foi gerada a estrutura final do guia.

5.2 Estrutura do Guia

A estrutura final do guia é composta por 11 tópicos que assumem papéis de seções e/ou subseções a depender da característica selecionada, os quais são:

1. Definição da Característica
2. Correlação de Características
3. Configuração do Ambiente
4. Definição da Subcaracterística
5. Contextualização
6. Casos de Teste Abstratos
7. Medições
8. Impacto das Subcaracterísticas
9. Custo-benefício
10. Sugestões de Ferramentas
11. Exemplo de Uso do Guia

No primeiro tópico “*Definição da Característica*” devem ser elencadas as definições da característica alvo, que será testada pelo guia. Em seguida é apresentada a “*Correlação das Características*” em que são definidas as correlações entre a característica alvo e as demais características IoT. Na “*Configuração do Ambiente*” deve ser descrito o ambiente IoT necessário para testar a característica alvo, tais como os tipos de dispositivos e o software ou hardware externo necessários.

Na “*Definição da Subcaracterística*” devem ser apresentadas definições das subcaracterísticas, quando existentes. A “*Contextualização*” deve apresentar as propriedades que caracterizam a característica alvo. No “*Casos de Teste Abstratos*” devem ser fornecidos passos genéricos para orientar os testes da característica alvo. A “*Medições*” devem ser elencadas métricas que podem ser aplicadas a característica alvo.

O “*Impacto das Subcaracterísticas*” deve apresentar as relações entre as subcaracterísticas existentes, baseado nas propriedades da Contextualização.

No “*Custo-benefício*” é fornecido um cálculo para a relação de custo-benefício da realização dos testes. Esse cálculo é proposto baseado nas correlações presentes tópico de

“*Correlações das características*”, nos casos de testes e nas métricas selecionadas. A fórmula para calcular o custo-benefício é definida por:

$$IC = \frac{ORC}{RC}$$

ORC é o número de características correlacionadas priorizadas na aplicação e RC é o número total de características correlacionadas com a característica alvo. Quanto ao impacto é definido pelos valores de tempo e custo para a condução dos testes e das métricas. Então esses dois valores, impacto e esforço, podem ser avaliados seguindo um modelo de priorização proposto que os categoriza em: Grupo I - alto esforço e baixo impacto. Alto custo e baixo benefício = baixa prioridade; Grupo II: baixo esforço e baixo impacto. Baixo custo e baixo benefício = prioridade média; Grupo III: alto esforço e alto impacto. Custo alto, mas alto benefício = alta prioridade; e Grupo IV: baixo esforço e alto impacto. Baixo custo e alto benefício = muito alta prioridade. No guia fornecemos todos os passos para calcular o custo-benefício e esse cálculo também é explanado na Subseção 6.3.6.

No tópico de “*Sugestões de Ferramentas*” devem ser feitas sugestões de ferramentas elencadas da literatura que podem ser utilizadas no processo de teste a fim de automatizá-lo. Por fim, é apresentado o “*Exemplo de Uso do Guia*” em que deve conter um exemplo de utilização do guia no contexto da característica alvo.

A estrutura do guia é proposta para características IoT com ou sem subcaracterísticas. Portanto, para se adequar a esses diferentes contextos, a organização dos tópicos do guia pode ser feita de duas maneiras: (organização 1) quando a característica a ser testada pode ser subdividida em subcaracterísticas; (organização 2) quando a característica a ser testada não é dividida em subcaracterísticas. Na primeira organização (1), cada subcaracterística se torna uma seção nova no guia e os tópicos de “Definições da subcaracterística”, “Contextualização”, “Casos de Teste Abstratos” e “Medições” são apresentados na forma de subseção para cada subcaracterística da característica alvo. Na organização 2, os tópicos ocupam exclusivamente papéis de seções no guia e não há nenhum tópico que foi proposto para subcaracterísticas, como é o caso do tópico “Definição da subcaracterística” (presente na organização 1).

Na Figura 5 é possível observar as organizações estruturais do guia com os tópicos seguindo os papéis de seção e subseção. Ela apresenta as duas organizações possíveis que os tópicos podem assumir, para a composição do guia: *Organização 1* - quando há, por exemplo, 2 subcaracterísticas, a estrutura terá 11 seções nas quais duas delas são as próprias subcaracterísti-

cas, conforme apresentado à esquerda na figura, e para cada subcaracterística haverá 4 subseções que contemplam, respectivamente, os tópicos Definição da Subcaracterística, Contextualização, Casos de Teste Abstratos e Medições; *Organização 2* - quando não há subcaracterística, a estrutura será composta pelas mesmas seções da Organização 1, exceto os tópicos Impacto das Subcaracterísticas e Definição da Subcaracterística, além disso, os tópicos Contextualização, Casos de Teste Abstratos e Medições passam a ser seções relacionados agora a característica diretamente e não mais a subcaracterística. Para mais de duas subcaracterísticas segue-se o modelo de Organização 1 e adiciona-se 4 subseções - definidas pelos tópicos: Definição da Subcaracterística, Contextualização, Casos de Teste Abstratos e Medições - para cada subcaracterística da característica alvo.

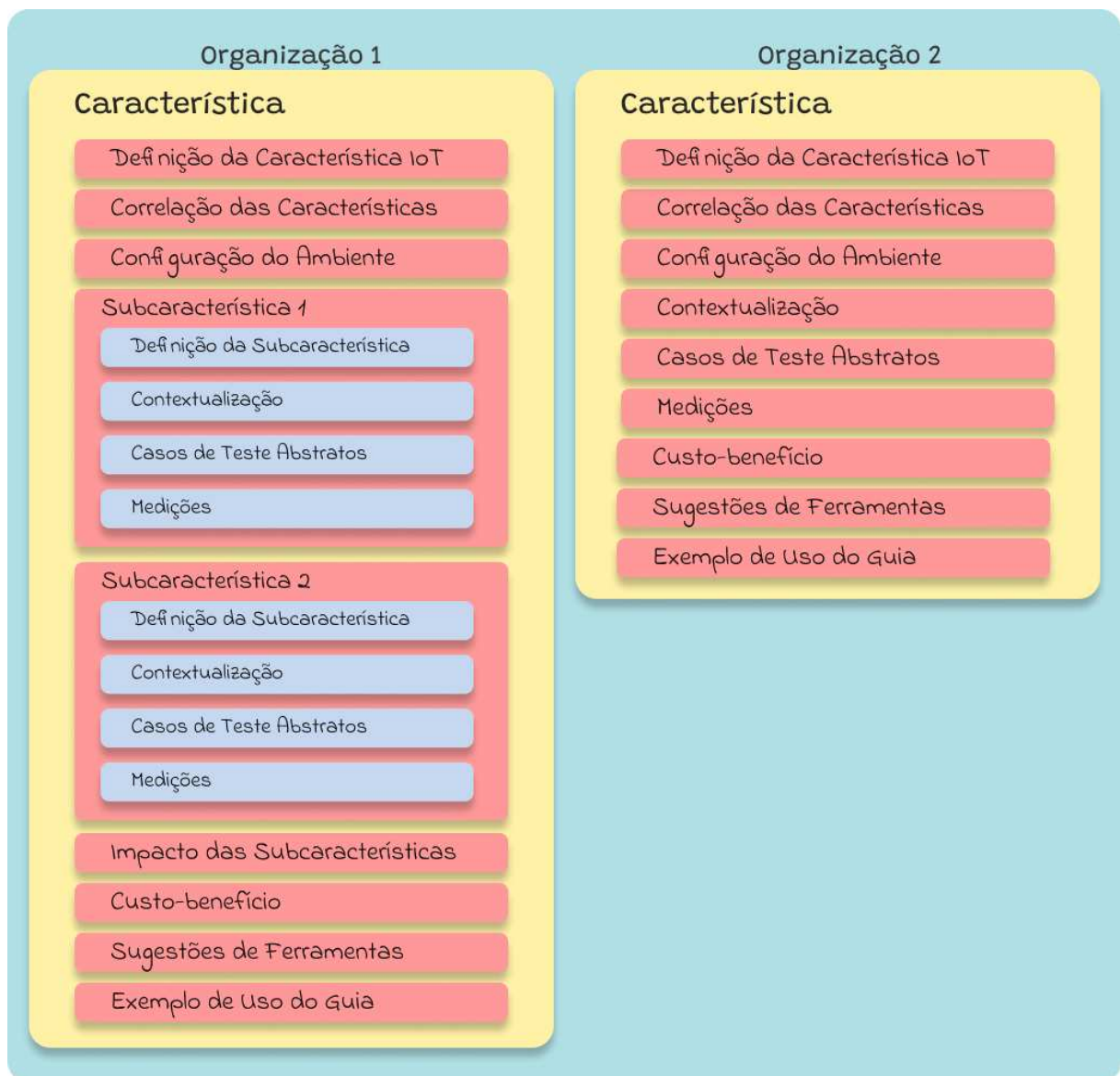


Figura 5 – Possíveis organizações estruturais do guia. Fonte: O autor.

5.3 Considerações finais

Neste capítulo foi apresentado o procedimento metodológico executado para a construção e estruturação do guia.

Para construir o guia foi conduzida uma revisão da literatura em conjunto com os estudos vindos de outras fontes, além de uma análise de repositórios de bugs. A partir desses resultados foram extraídos dados imprescindíveis na execução de testes IoT, como, estrutura de casos de testes, padrões de testes, falhas comuns e boas práticas. Esses dados passaram por um refinamento com um especialista para então ser finalizada a concepção da estrutura do guia. O guia elaborado possui 11 tópicos baseados nos estudos e no refinamento feito. Esses tópicos abrangem métricas, casos de testes, configuração de ambiente, correlações, entre outros. No próximo capítulo é apresentado o processo metodológico para a instanciação do guia, bem como, um exemplo dessa instanciação para a característica de desempenho.

6 INSTANCIACÃO DO GUIA DE TESTE

A partir da estrutura do guia, apresentada no Capítulo 5, é possível utilizar a mesma para avaliar diferentes características IoT, esse processo é chamado de instanciação. Este capítulo apresenta na Seção 6.1 o procedimento metodológico para a instanciação do guia. Na Seção 6.2 é apresentado um exemplo de instanciação do guia para a característica de desempenho, e o detalhamento do próprio guia de desempenho na Seção 6.3 e ao final, na Seção 6.4, é apresentada uma ferramenta de apoio criada para o uso do guia de desempenho.

6.1 Procedimento metodológico de instanciação do guia

A metodologia definida para instanciar o guia de teste para avaliar uma característica IoT consiste em seis atividades, conforme apresentado na Figura 6.

A primeira atividade da metodologia é “Realizar a revisão da literatura sobre a característica IoT”. Nessa atividade deve ser feita uma revisão da literatura utilizando uma *string* de busca definida de acordo com a área de foco, no caso IoT, adicionando os termos que dizem respeito à característica selecionada. A *string* deve ser executada em uma base de interesse, também de acordo com a característica a ser avaliada. A fim de buscar estudos relevantes para o contexto a ser investigado pode-se pedir indicações de especialistas da área para que estudos imprescindíveis sejam adicionados para a fase extração. Na seleção devem ser aceitos artigos que tragam abordagens de teste relacionadas à característica em questão, bem como métricas, ferramentas e desafios. Deve ser dada preferência a estudos completos e no idioma inglês.

A atividade seguinte é a de “Extrair dados” que irão povoar as seções e subseções do guia. Na extração pode ser utilizado um formulário de extração de dados para conduzir essa

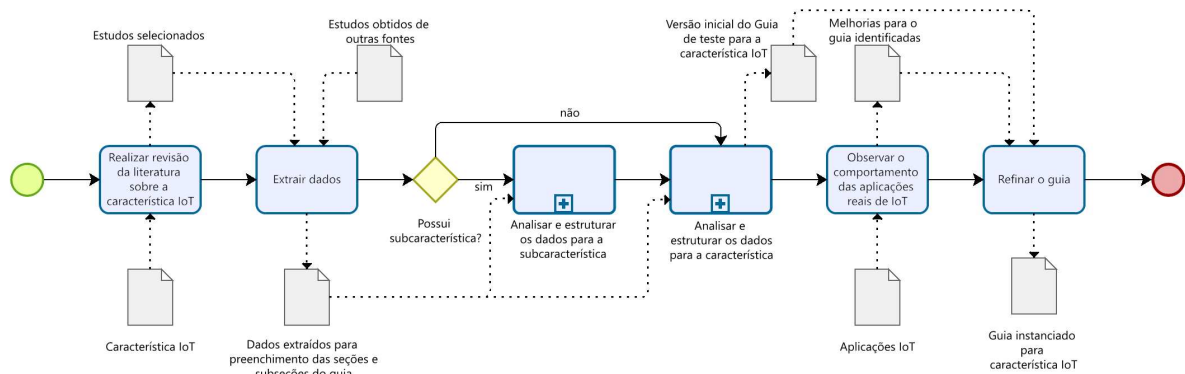


Figura 6 – Procedimento metodológico para instanciação do guia para uma característica. Fonte: o autor.

etapa.

A próxima atividade dependerá se a característica selecionada possui ou não subcaracterísticas, em caso afirmativo, a próxima atividade é “Analisar e estruturar os dados para a subcaracterística”, caso contrário deve ser “Analisar e estruturar os dados para a característica”.

Em caso da característica possuir subcaracterísticas, o objetivo primeiramente deve ser analisar e organizar os dados para povoar as seções e subseções do guia específicas para as subcaracterísticas. Devem ser observadas as seguintes informações:

- Definições de subcaracterísticas;
- Propriedades que compõem as subcaracterísticas;
- Como são feitos os testes;
- As métricas utilizadas; e
- Como as subcaracterísticas se correlacionam.

A estruturação desses dados relacionados às subcaracterísticas é conduzida baseada em 5 artefatos:

1. A lista das definições das subcaracterísticas;
2. A lista das propriedades que caracterizam a subcaracterística. Cada subcaracterística está relacionada com um conjunto de propriedades que a definem, por exemplo, o *Comportamento Temporal* é definido como o tempo de resposta da aplicação e reconexão.
3. A especificação de casos de teste abstratos para cobrir as subcaracterísticas e as suas propriedades;
4. A lista de métricas para apoiar a validação da característica. Cada métrica é analisada considerando os casos de teste abstratos, as propriedades e a definição da subcaracterística; e
5. A relação do impacto das subcaracterísticas da característica selecionada é construída a partir da identificação nas relações existentes entre as propriedades, casos de testes abstratos e métricas.

Após ser conduzida a atividade de “Analisar e estruturar os dados para a subcaracterística”, atividade mandatória para características que possuem subcaracterísticas, é iniciada a atividade de “Analisar e estruturar os dados para a característica”. Com todas as informações sobre as subcaracterísticas preenchidas, os esforços devem ser direcionados para as seções relativas à característica alvo: Definição da Característica, Correlação das Características, Configuração de Ambiente, Custo-benefício, Sugestões de Ferramentas e Exemplo de Uso do Guia. Quando

uma característica não tem subcaracterística, todos os tópicos do guia são preenchidos, exceto *Definição de subcaracterística e Impacto das subcaracterísticas*, que são específicas para as subcaracterísticas.

O preenchimento das seções do guia - Definição da Característica, Correlação das Características, Configuração de Ambiente, Custo-benefício, Sugestões de Ferramentas e Exemplo de Uso do Guia, referentes à característica propriamente dita é baseado nos 6 seguintes artefatos:

1. Lista das definições da *Característica*;
2. Lista das características correlacionadas;
3. Catalogação de ferramentas que são utilizadas para realizar os testes;
4. Adaptação do cálculo para custo-benefício;
5. Descrição do ambiente de IoT necessário para testar a característica; e
6. Um exemplo de utilização do *Guia de Teste para a característica IoT*.

Após a estruturação dos dados da subcaracterística e/ou da característica é fornecida uma primeira versão do guia instanciado. A próxima atividade é “Observar o comportamento de aplicações reais de IoT”. O intuito dessa atividade é identificar aspectos, que não foram atingidos por meio da literatura, que também possam auxiliar no preenchimento dos tópicos do guia. Para a condução dessa atividade devem ser selecionadas aplicações IoT, se possível de contextos diferentes, a fim utilizá-las para observar cenários de testes, métricas, ambiente de teste e outros aspectos que não foram completamente extraídos na literatura.

A última atividade consiste em “Refinar o Guia” com base nas observações das aplicações reais de IoT. Essa atividade é feita a partir da adição das melhorias identificadas na atividade anterior. O artefato de saída final é o *Guia de Teste instanciado para a característica IoT*.

6.2 Exemplo de instanciamento do Guia - característica de Desempenho

Dada a relevância da característica de *Desempenho* no domínio IoT e pelos desafios apresentados no Capítulo 3, a mesma foi selecionada para a instanciamento da primeira versão do guia.

A metodologia para a instanciamento do guia foi apresentada no início deste Capítulo. Na Figura 7 é apresentada a metodologia instanciamento concretizada para a característica de *Desempenho*.

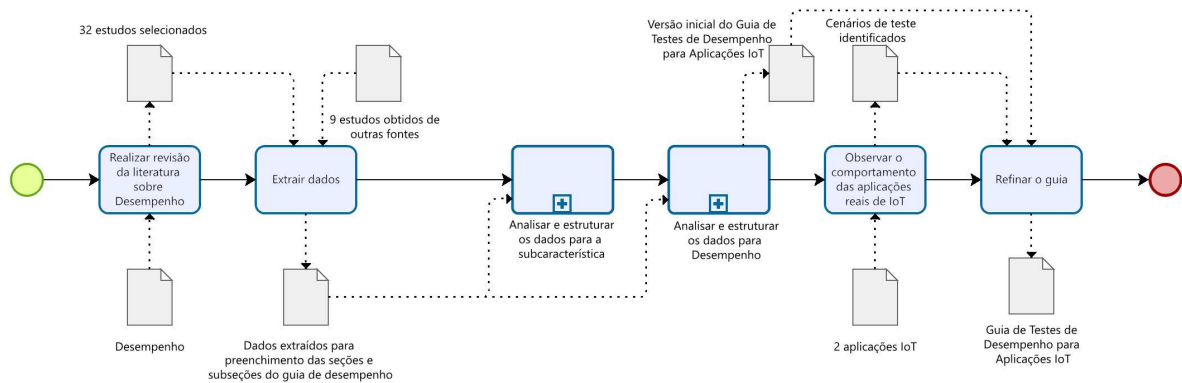


Figura 7 – Procedimento metodológico concretizado para Desempenho, seguindo o modelo da Figura 6. Fonte: O autor.

A primeira atividade conduzida para a instanciação foi a de “Realizar revisão da literatura sobre Desempenho¹”. Essa revisão concentrou-se nos testes de desempenho para IoT. Para a busca foi definida uma *string*, baseada nas questões pesquisas da revisão, elaboradas a partir da estratégia PICO.

O **PICO** para essa revisão teve como: **P**opulação a Internet das Coisas, pois se trata do contexto do qual essa pesquisa está inserida; a **I**ntervenção da pesquisa foram os testes de desempenho, sabendo que a instanciação do guia é para testes IoT no âmbito da característica de *Desempenho*; **C**omparação, para essa revisão não se aplicou a utilização de comparação; **O**utcome, como resultado foi esperado a obtenção de soluções e artefatos que eram utilizadas nos testes de desempenho IoT.

Todo o protocolo da revisão está disponível no Apêndice D.

A partir da elaboração do PICO, foram derivados os sinônimos e então definidas as questões de pesquisa, que foram:

- Quais sistematizações são utilizadas para o teste de desempenho em aplicações IoT?
- Quais propriedades e características rodeiam o âmbito de desempenho em aplicações IoT?
- Quais métricas e ferramentas são utilizadas para avaliar o desempenho de aplicações IoT?

Com as questões de pesquisa elaboradas, foi definida a *string* de busca, que está ilustrada na Tabela 5, onde contém os termos e os termos relacionados. A *string* foi executada nas fontes de pesquisa *Scopus*, *IEEE*, *ACM* e *Science Direct*.

A partir dessa execução, foram obtidos como retorno 592 estudos. Esses estudos foram analisados e selecionados a partir dos critérios de seleção. Os critérios se dividiram entre: critérios de inclusão e critérios de exclusão. Para o primeiro foram definidos critérios, como, “os

¹ Revisão para a instanciação do guia: <<https://tinyurl.com/yt62tuh4>>

Tabela 5 – String de busca - Desempenho

| Termos | Termos Relacionados |
|---|--|
| Internet of Things | - |
| Performance Test | Performance Testing, Load Testing, Load Test, Stress Testing, Stress Test, Workload Testing, Workload Test |
| Standard | Model, Catalog, Metric, Measure, Tool, Test Case, Property, Correlation |
| String Final | |
| ("internet of things") AND ("performance test"OR "performance testing"OR "load testing"OR "load test"OR "stress testing"OR "stress test"OR "workload testing"OR "workload test") AND (standard OR model OR catalog OR metric OR measure OR tool OR property OR "test case"OR correlation) | |

Tabela 6 – Estudos levantados na revisão de desempenho

| Base | Total de estudos extraídos | Total de estudos selecionados |
|----------------|-----------------------------------|--------------------------------------|
| ACM | 157 | 6 |
| IEEE | 53 | 5 |
| Science Direct | 134 | 1 |
| Scopus | 248 | 20 |
| Total | 592 | 32 |

estudos deveriam ser escritos em inglês”, “apresentarem soluções para o teste de desempenho IoT” e “apresentarem métricas ferramentas, propriedades ou características relacionadas ao teste de desempenho em aplicações IoT”. Os critérios de exclusão foram, por exemplo, “estudos que não eram de Computação ou áreas relacionadas”, “estudos com versão já publicada, era selecionada a mais completa” e “o estudo é indexado, tutorial ou *keynote*” (mais detalhes, ver Apêndice D). Conforme apresentado na Tabela 6, após a aplicação dos critérios de seleção, restaram 32 estudos. Além dos 32 estudos selecionados, foram utilizados 9 estudos, (NORBERTO *et al.*, 2019; LANGSARI *et al.*, 2018; SOARES *et al.*, 2014; LU *et al.*, 2000; ZHANG *et al.*, 2018; THIMMANNAGARI, 2004; MEIER *et al.*, 2007; ZHANG; CHAKRABARTY, 2004; JAIN, 1991) de outras fontes, uma vez que discutem o desempenho em outros domínios relacionados com a IoT, tais como o desempenho em aplicações móveis.

A atividade seguinte foi a de “Extrair dados”. Nessa fase foram extraídos os dados relativos a *Desempenho* dos 41 estudos selecionados (32 selecionados das bases + 9 de outras fontes, ver Apêndice E). Para essa etapa foi utilizado um Formulário de extração de dados para Desempenho: com objetivo de extrair os dados e como benefício promover a aceleração desse processo (ver Apêndice F).

No caso da característica de *Desempenho*, esta pode ser dividida em três subcaracterísticas conforme apresentado na Seção 2.3, portanto, a próxima atividade conduzida foi a de “Analisar e estruturar os dados para a subcaracterística”.

Os artefatos 1, 2, 3 e 4, citados na Seção 6.1 (respectivamente, a lista das definições das subcaracterísticas, a lista das propriedades, a especificação de casos de teste abstratos e a lista de métricas) foram obtidos a partir dos 41 estudos selecionados.

Os formulários de extração preenchidos auxiliaram nessa etapa de forma que caso fosse identificada um estudo que possuía uma métrica, por exemplo, relacionada com uma subcaracterística (*e.g.*, Comportamento Temporal), esta era extraída para a seção no guia de *Medições*. Do mesmo modo, essa extração sucedeu para cenários de testes identificados e as propriedades, apresentados, respectivamente, nas seções “Casos de Teste Abstratos” e “Contextualização”. Com base nesses 4 artefatos e nos outros dados obtidos na extração, o artefato 5 (a relação do impacto das subcaracterísticas) foi proposto. A partir desses artefatos, as seções do guia, “Definição da Subcaracterística”, “Contextualização”, “Casos de Teste Abstratos”, “Medições” e “Impacto das Subcaracterísticas” foram completamente preenchidas.

A quarta atividade foi “Analisar e estruturar os dados para a característica” de Desempenho. Uma vez fornecidas todas as informações sobre as subcaracterísticas, o foco se destina ao preenchimento das seções do guia relativas à característica propriamente dita, as quais foram: “Definição da Característica”, “Correlação das Características”, “Configuração do Ambiente”, “Custo-benefício”, “Sugestões de Ferramentas” e “Exemplo de Uso do Guia”. A sequência de atividades é semelhante as seguidas nas subcaracterísticas. Sendo elas, lista de definições da característica, lista de características correlacionadas, catalogação das ferramentas, adaptação da fórmula de custo-benefício, descrição do ambiente e a descrição de um exemplo de uso do guia.

Baseado nos estudos da literatura foi construída a lista das definições de *Desempenho* e catalogadas as ferramentas que são utilizadas para conduzir os testes. A partir da revisão e extração de dados foram gerados resultados como, por exemplo, Diagramas de termos relacionados as características² e uma Planilha de relação das características (Apêndice G) que colaboraram para o desenvolvimento da lista das características correlacionadas com *Desempenho*, conseqüentemente, a adaptação da fórmula proposta para Custo-benefício. As seções do guia de “Configuração do Ambiente”, necessário para testar a característica de Desempenho, e o “Exemplo de Uso do Guia”, foram esboçados, entretanto, só foram completamente consolidados, após a realização da atividade de “Observar o comportamento de aplicações IoT reais”.

Ao final da atividade de “Analisar e estruturar os dados para a característica”, a

² Termos relacionados as características: <<https://tinyurl.com/5296rjrm>>

primeira versão do *Guia de Teste de Desempenho para Aplicações IoT* foi gerada.

A próxima atividade foi “Observar o comportamento de aplicações IoT reais”. Para essa atividade foram selecionadas duas aplicações IoT: uma *Horta Inteligente*³ e um *Relógio Inteligente*⁴. A primeira aplicação tem como função principal regar uma horta, para isso sempre que o sensor de umidade detecta o valor limite definido pelo usuário, o atuador executa a sua tarefa de regar. Na segunda aplicação, *Relógio Inteligente*, foi observada a funcionalidade de “exercício”. Nessa função o relógio envia uma notificação a um usuário que está inerte, para que o mesmo realize um exercício de alongamento; caso o usuário faça o exercício, o relógio capta o tempo do movimento e armazena. No final do dia, a aplicação fornece o tempo total que o usuário se exercitou.

A partir dessa atividade de observação das aplicações, foram identificados cenários de teste que ainda não estavam especificados nos casos de teste abstratos. Por exemplo, durante a análise dos 41 estudos, foi identificado que deveria haver um caso de teste para validar as informações que são enviadas a um atuador. Contudo, quando foram observadas as duas aplicações de IoT, foi identificado também outros cenários, tais como a necessidade de um caso de teste para validar o comportamento do atuador quando a aplicação envia muitas requisições em um período de pico. Dessa forma, com base nessas observações, foi conduzida a atividade “Refinar o guia”, por exemplo, com a adição e estruturação no guia dos cenários de testes identificados na atividade anterior.

Como artefato de saída desta atividade foi gerado o *Guia de Teste de Desempenho para Aplicações IoT*, este guia já passou por duas avaliações que serão apresentadas no Capítulo 7. A versão completa do guia com as melhorias incorporadas após as avaliações está disponível no repositório deste trabalho⁵. Nas próximas subseções o *Guia de Teste de Desempenho para Aplicações IoT* é detalhado.

6.3 Visão geral do Guia Desempenho

O *Guia de Teste de Desempenho para Aplicações IoT* consiste em uma introdução e 10 seções. Na introdução há uma breve descrição em que é apresentado o fluxo de dados no qual se aplica o guia (conforme ilustrado na Figura 3). Em seguida são definidas instruções de como usar o guia, onde é explicada de forma genérica como cada seção e subseção deve ser aplicada e

³ Horta Inteligente - Disponível em: <<https://tinyurl.com/s2uxm6jj>>

⁴ Relógio Inteligente- Disponível em: <<https://tinyurl.com/3znncxz4>>

⁵ *Guia de Teste de Desempenho para Aplicações IoT* : <<https://tinyurl.com/ym4zzrnt>>

como executar e selecionar os testes e métricas. As demais seções são referentes aos tópicos definidos para o guia. Por *Desempenho* ser uma característica que contém 3 subcaracterísticas, a estrutura do guia seguiu um modelo semelhante ao de Organização 1, conforme apresentado na Figura 5 do Capítulo 5, em que todos os tópicos estão presentes no guia e para cada subcaracterística são fornecidos os tópicos de “Definição da Subcaracterística”, “Contextualização”, “Casos de Teste Abstratos” e “Medições”.

A seguir cada um das seções e subseções são detalhadas e exemplificadas conforme apresentado no guia.

6.3.1 Definição de Desempenho

Na primeira seção do Guia são fornecidas cinco “Definições de Desempenho” extraídas da literatura, elas são listadas de forma sequencial e com seus respectivos autores, por exemplo:

“Desempenho é o tempo gasto para executar um serviço, a taxa em que o serviço é executado, e os recursos consumidos durante a execução do serviço.” (JAIN, 1991).

“O desempenho está preocupado com a qualidade da resposta do software quando um evento ocorre. Para avaliar se um sistema tem bom desempenho, o tempo entre o evento e a resposta pode ser medido primeiro e depois comparado com uma restrição de tempo determinada anteriormente.” (LANGSARI et al., 2018).

Estas definições ajudam os engenheiros de software, quando utilizam o guia, a compreender em que definições de desempenho o guia se baseia.

6.3.2 Correlação das Características

Ao analisar o mapeamento sistemático de (CARVALHO, 2018) é possível identificar ao menos 27 características IoT. A fim de verificar as relações de desempenho com as demais características IoT, na seção de “Correlação das Características” são apresentadas 18 características relacionadas com o desempenho, conforme a Figura 8.

Essas relações se classificam em relações positivas, cuja característica influencia o desempenho de uma forma positiva (retângulo verde); relações negativas (retângulo vermelho) que são o oposto; e relações que podem ser positivas ou negativas dependendo do contexto em

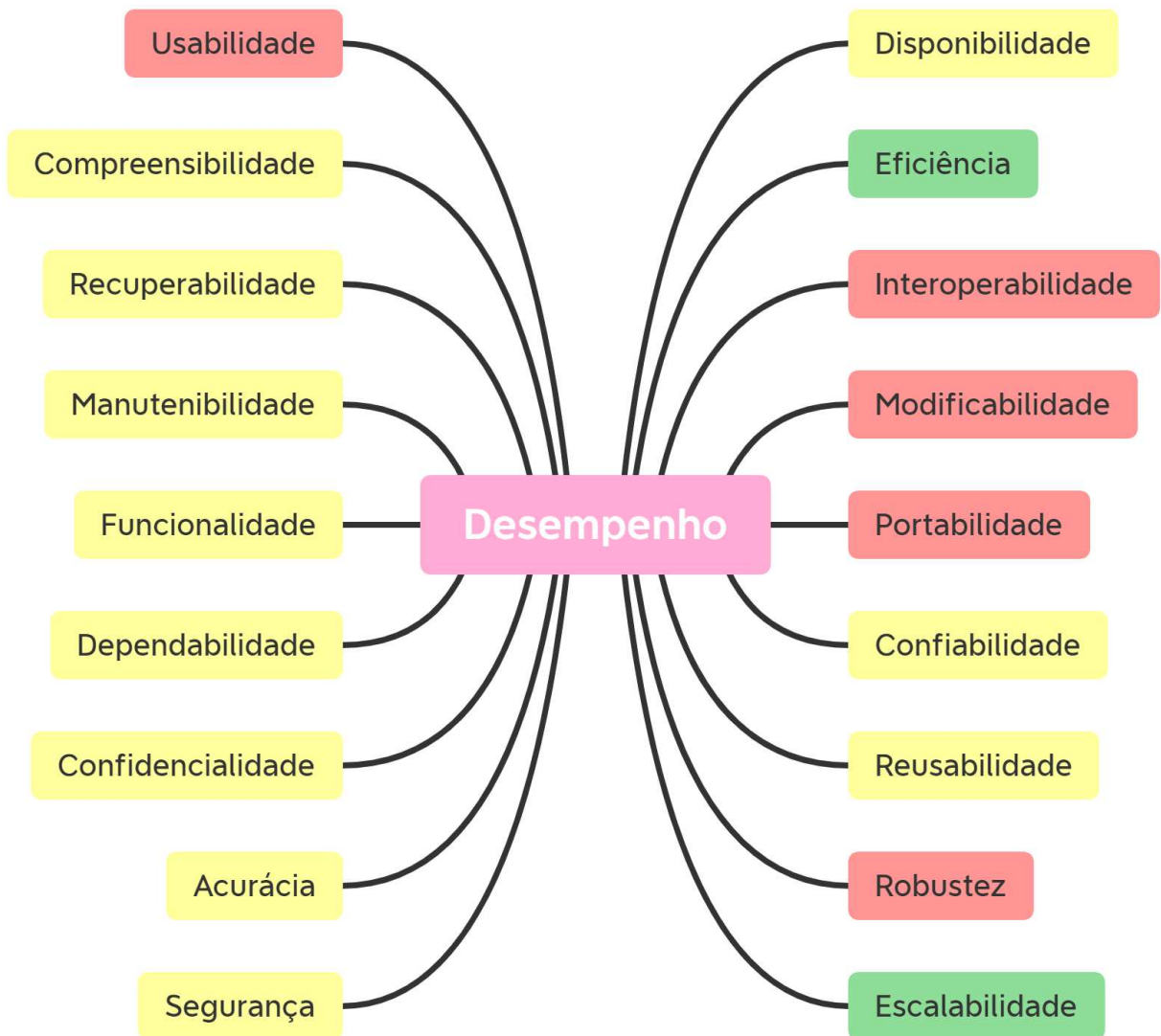


Figura 8 – Relações das demais características com desempenho. Fonte: O autor.

que ocorrem (retângulo amarelo). Por exemplo, a característica *Disponibilidade* pode ter uma influência positiva no *Desempenho* - se houver servidores suficientes para satisfazer uma procura e, assim, acelerar o processo de troca de mensagens; caso contrário, pode ter uma influência negativa. Essas correlações ajudam a compreender possíveis conflitos de requisitos que possam existir em aplicações IoT (CARVALHO *et al.*, 2020).

6.3.3 Configuração do Ambiente

Na Seção do guia de “Configuração do Ambiente” são detalhadas as configurações necessárias para utilizar o guia, ou seja, o usuário do guia deve organizar a execução dos testes, garantindo que o sistema a ser testado se adequa ao ambiente necessário. Como exemplo de configuração do ambiente para o guia de desempenho temos que o ambiente pode ser composto por:

- um ou mais sensores;
- um mais atuadores;
- uma aplicação;
- um medidor de energia para habilitar alguns cenários de teste, por exemplo, os casos de teste necessários para testar a subcaracterística “Utilização de Recursos”.

6.3.4 Subcaracterísticas

As seções 4, 5 e 6 do guia dizem respeito às subcaracterísticas de “Comportamento Temporal”, “Utilização de Recursos” e “Capacidade”. A fim de orientar os testes da característica de desempenho nessas seções estão presentes os tópicos de: “Definição da Subcaracterística”, “Contextualização”, “Casos de Teste Abstratos” e “Medições”, relacionados a cada subcaracterística. As subseções 4(a), 5(a) e 6(a) do guia fornecem as definições para cada subcaracterística, todas as definições são extraídas da (ISO/IEC 25010, 2011), por exemplo:

“Comportamento Temporal é o nível em que a resposta e o tempo de processamento e as taxas de transferência de um produto ou sistema, ao executar suas funções, atendem aos requisitos”;

“Utilização de Recursos é o grau em que as quantidades e tipos de recursos usados por um produto ou sistema, ao desempenhar suas funções, cumprem as exigências”;

“Capacidade é o grau em que os limites máximos de um produto ou parâmetro do sistema atende aos requisitos”.

De forma similar, a “Contextualização” das subcaracterísticas supracitadas é apresentada nas subseções 4(b), 5(b) e 6(b) do guia. Essas subseções apresentam as propriedades obtidas na revisão da literatura que representam cada subcaracterística. A Tabela 7 fornece uma visão geral das subcaracterísticas e suas propriedades. O guia descreve 8 propriedades para o “Comportamento Temporal”, 8 para “Utilização de Recursos” e 6 propriedades para a “Capacidade”. Ao utilizar o guia um profissional e/ou usuário mais experiente pode selecionar as propriedades que deseja avaliar, a opção de escolha também se aplica aos casos de testes, métricas e até mesmo as subcaracterísticas, entretanto isso pode impactar diretamente na avaliação da aplicação no que se refere à característica de Desempenho.

As subseções 4(c), 5(c) e 6(c) do guia fornecem os “Casos de teste abstratos” para

Tabela 7 – Subcaracterísticas e suas propriedades

| Subcaracterística | Propriedade |
|--------------------------|--|
| Comportamento Temporal | Tempo de despacho; Tempo de Execução; Tempo de Transmissão de mensagem; Tempo mínimo de espera; Tempo de reconexão; Tempo de resposta; Tempo de carregamento; e Tempo de adaptação |
| Utilização de Recursos | Disponibilidade da CPU; Consumo da CPU; Consumo de energia; Eficiência energética; Consumo de memória; Economia de energia; Tempo de uso; e Consumo de dados |
| Capacidade | Taxa de Download/Upload; Throughput; Tamanho da mensagem; Uso de rede; e Largura de banda. |

cada subcaracterística. Ao total foram especificados 23 *Casos de Teste Abstratos* para as subcaracterísticas de *Desempenho*, dos quais 11 casos de testes abstratos foram especificados para a subcaracterística de “Comportamento Temporal”, 8 para “Utilização de Recursos” e 4 para “Capacidade. Os casos de teste seguem a estrutura ilustrada na Tabela 8. O título refere-se ao que é tratado no caso de teste. O ambiente do teste indica o tipo de configuração necessária dos dispositivos para a execução. As precondições informam o estado necessário para a execução de um teste. O passo a passo fornece as instruções para a execução de um teste e as pós-condições indicam o estado alcançado após a execução de um teste. O caso de teste da Tabela 8 avalia o envio de um comando ao atuador. No cenário deve haver ao menos um sensor, um atuador e uma aplicação em comunicação, a fim de que o sensor envie um comando ao atuador e o esperado é que o mesmo execute uma ação condizente ao comando enviado, para que assim, se obtenha sucesso no caso de teste.

Tabela 8 – Exemplo de um caso de teste de “Comportamento temporal”

| Caso de Teste 01 - CTO1 | |
|--------------------------------|--|
| Título: | Enviar comando ao atuador |
| Ambiente de Teste: | N atuadores e 1 aplicação |
| Precondição: | O atuador deve estar em funcionamento |
| Passo a passo: | 1 - Através da aplicação enviar o comando desejado 2 - Verificar comportamento do atuador |
| Pós-condição: | O atuador executou o comando enviado |

As subseções 4(d), 5(d) e 6(d) do guia fornecem as métricas para cada subcaracterística. Além dos casos de teste, foram elencadas 22 métricas da literatura para ajudar na avaliação do *Desempenho*, 8 métricas da subcaracterística de “Comportamento Temporal”, 12 de “Utilização de Recursos” e 2 de “Capacidade”. Essas métricas são estruturadas conforme

apresentado na Tabela 9. A estrutura contempla o propósito da métrica; o método utilizado para aplicar a métrica; a medida para realizar a medição; e a referência bibliográfica do qual foi extraída a métrica. Na métrica da Tabela 9 é proposto medir o tempo que se leva para executar uma tarefa, para isso, é necessário marcar o tempo em que se inicia a tarefa e subtrair do tempo em que a tarefa é concluída, obtendo assim o tempo de execução, de acordo com a definição de (NAIR, 2015).

Tabela 9 – Exemplo de uma métrica para o comportamento temporal

| Tempo de Execução - M02 | |
|--------------------------------|---|
| Propósito: | Avaliar o tempo de execução de uma tarefa. |
| Método: | Acompanhar o início da execução da tarefa e compará-la com o momento em que foi concluída. |
| | $X = t2 - t1$ |
| Medição: | X = tempo de execução t1 = tempo do início da execução t2 = tempo em que a tarefa foi concluída |
| Referência: | (NAIR, 2015) |

6.3.5 Impacto das Subcaracterísticas

Os impactos das subcaracterísticas estabelecem conexões entres as propriedades de cada uma delas, a fim de auxiliar na detecção de fricções que possam vir a existir na aplicação. A Seção 7 do guia apresenta os “Impactos entre as subcaracterísticas” conforme as suas propriedades.

Nas Figuras 9, 10 e 11 são apresentados o impacto de cada uma das subcaracterísticas. Na cor verde estão as propriedades relacionadas a “Comportamento Temporal”, na cor amarela as propriedades relacionadas a “Utilização de Recursos” e na cor azul as propriedades relacionadas a “Capacidade”. Cada subcaracterística é apresentada ao centro e seguida por cada uma de suas propriedades. Nas ramificações das propriedades, são, então, apresentadas as demais propriedades relacionadas.

Na Figura 9 é apresentada a relação da subcaracterística de “Comportamento Temporal” com as demais. Por exemplo, a propriedade “Tempo de resposta” da subcaracterística de *Comportamento Temporal* é afetada pela propriedade “Consumo de CPU” da subcaracterística de *Utilização de Recursos*, porque se o consumo de CPU for demasiado elevado, o processamento da informação para enviar a resposta será elevado.

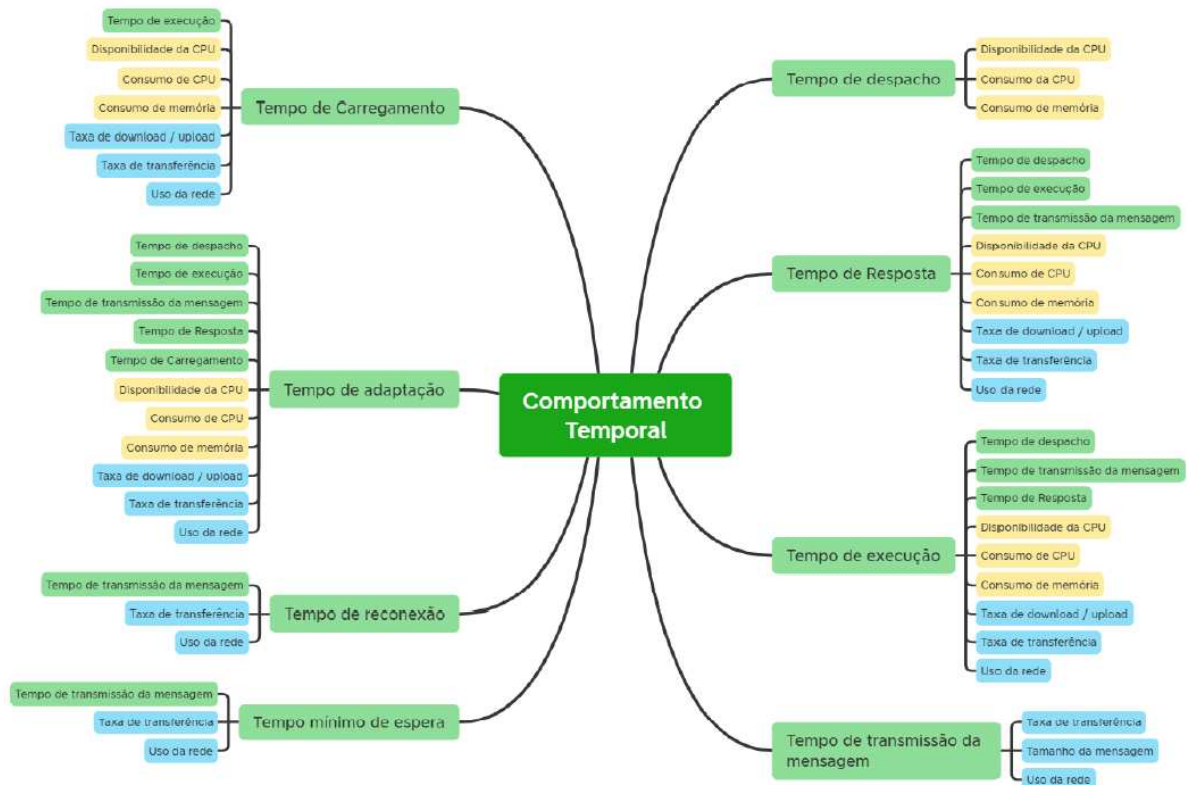


Figura 9 – Relação das propriedades de Comportamento Temporal com as demais propriedades.
Fonte: O autor.

Na Figura 10 é apresentada a relação da subcaracterística de “Utilização de Recursos” com as demais. Por exemplo, a propriedade “Consumo de energia” da subcaracterística de *Utilização de Recursos* se relaciona com “Uso da rede” da subcaracterística de *Capacidade*, visto que, quão maior o uso da rede, maior será o consumo de energia, assim como, quão menor o uso de rede, menor será o consumo de energia.

Na Figura 11 é apresentada a relação da subcaracterística de “Capacidade” com as demais. Por exemplo, a propriedade “Uso da rede”, da subcaracterística de *Capacidade*, se relaciona com “Consumo de dados”, pertencente a *Utilização de Recursos*, apresentando uma relação de proporcionalidade, onde, quanto maior o tempo de uso da aplicação, maior será o consumo de dados.

Exemplos como esses, supracitados, auxiliam a verificar e evitar possíveis problemas e conflitos que possam ocorrer na aplicação IoT.

6.3.6 Custo-benefício

Para definir o impacto e esforço gasto para execução dos testes de desempenho, foram propostas duas fórmulas principais. Conforme apresentadas abaixo.

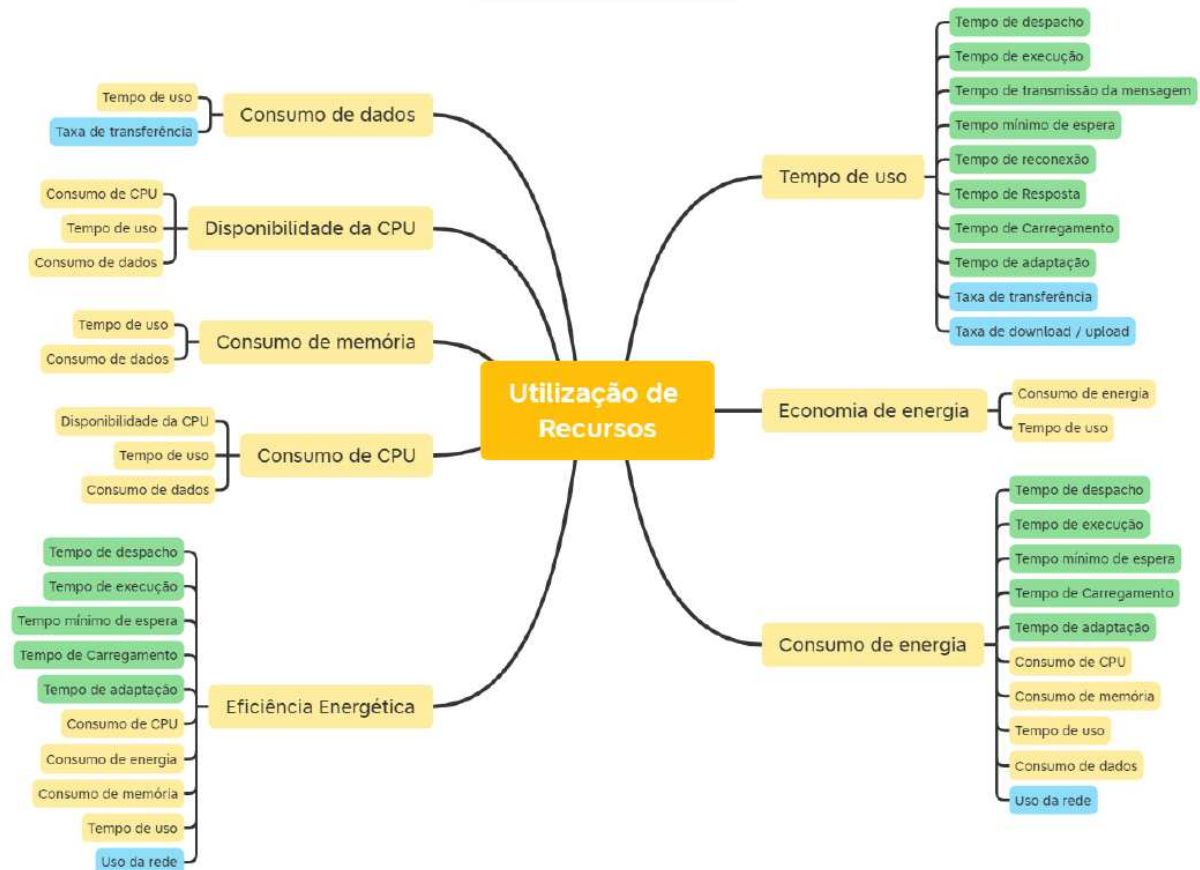


Figura 10 – Relação das propriedades de Utilização de Recursos com as demais propriedades.
Fonte: O autor.

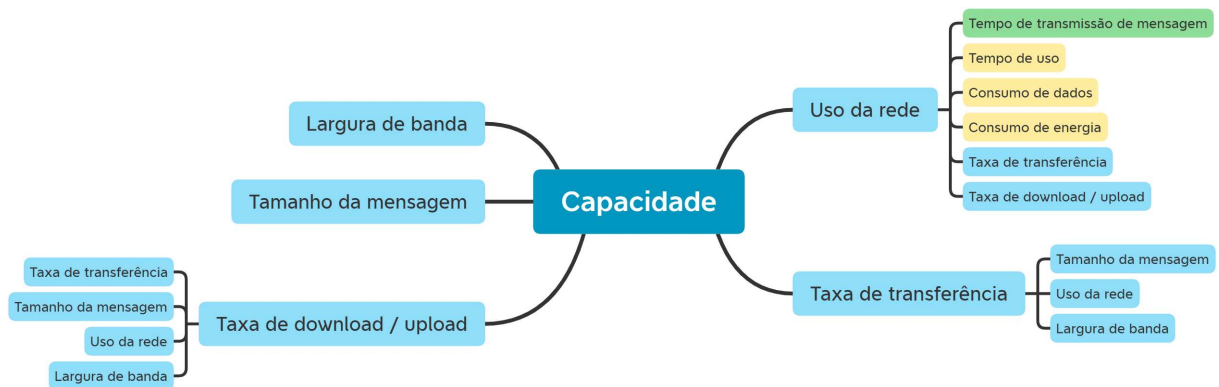


Figura 11 – Relação das propriedades de Capacidade com as demais propriedades. Fonte: O autor.

Cálculo do impacto (benefício):

$$IC = \frac{ORC}{RC}$$

Cálculo do esforço (custo):

$$ESF = \frac{CCT + CMD}{2}$$

A variável IC equivale à presença das características correlacionadas a desempenho na aplicação, entendendo, que, quão maior esse número, mais da característica de desempenho está presente na aplicação, gerando maior impacto. O IC é calculado a partir de RC - número total de características correlacionadas com o desempenho, esse valor é fixo e igual à 18, conforme apresentado na Figura 8. A segunda variável é ORC - número das características correlacionadas ao desempenho, imprescindíveis⁶ na aplicação. Sabendo do valor de RC, o valor ORC é definido de modo empírico pelo usuário do guia, baseado nos requisitos da aplicação que será testada. Com as duas variáveis obtidas, basta dividir ORC por RC e assim gerar o impacto - IC.

A outra fórmula proposta trata do esforço - ESF, gasto para a execução dos testes e métricas do guia. Para o cálculo é necessário o CCT - “custo médio de casos de teste normalizados” e o CMD - “custo médio de métricas normalizadas”. O CCT é obtido a partir da soma e normalização do “custo médio para a realização de cada caso de teste” - CT, para isso deve ser multiplicado “o tempo médio do profissional para realizar um caso de teste” - TC e VHC - “valor do tempo do profissional que realizará o caso teste”, conforme apresentado a seguir.

Cálculo para o custo médio para a realização de cada caso de teste:

$$CT_i = TC_i * VHC_i$$

Essa fórmula é aplicada para todos os casos de testes selecionados e depois os resultados são somados e normalizados, obtendo assim o CCT.

O CMD é gerado seguindo o mesmo processo do CCT, aplicando apenas a métricas ao invés de casos de testes. Ao obter o CCT e CMD é aplicado a fórmula do esforço e gerado o resultado.

O resultado gerado é analisado a partir de um plano cartesiano que varia de 0 à 1, a depender do quadrante em que o resultado se encontra é definida uma maior ou menor prioridade na execução dos testes, baseado no custo-benefício. Na Figura 12 em que no eixo x (horizontal) é avaliado o impacto e no eixo y (vertical) o esforço. Cada grupo dos quadrantes gera uma interpretação.

• Grupo I: Alto esforço e baixo impacto. Alto custo e baixo benefício = baixa prioridade.

O esforço para execução dos testes é muito alto e há poucas características correlacionadas a desempenho que são de fato imprescindíveis para a aplicação. Com isso, o indicado

⁶ Entende-se por característica imprescindível aquela cuja importância é evidenciada nos requisitos da aplicação.

é ser avaliado a importância da característica de Desempenho para a aplicação, caso a característica seja fundamental, mesmo com o alto esforço, devem ser priorizados a condução dos testes, fazendo uso das ferramentas e das tabelas de relações para diminuir o esforço durante a execução, e até mesmo rever os casos de testes e métricas selecionadas, a fim de ao menos, atingir o Grupo II de prioridade. Caso a característica de Desempenho não seja essencial ao sistema, não há necessidade de uma execução imediata dos testes.

- **Grupo II: Baixo esforço e baixo impacto. Baixo custo e baixo benefício = média prioridade.** O esforço para execução dos testes é baixo e há poucas características correlacionadas a desempenho que são de fato imprescindíveis para a aplicação. O indicado é ser avaliado a importância da característica de Desempenho para a aplicação, caso a característica seja fundamental devem ser priorizados a condução dos testes, caso contrário não há necessidade de uma execução imediata dos testes, porém visto que para a execução dos testes é exigido um baixo esforço, em um momento oportuno esses testes podem ser realizados facilmente. Caso a característica de Desempenho não seja essencial ao sistema, não há necessidade de uma execução imediata dos testes.
- **Grupo III: Alto esforço e alto impacto. Alto custo, porém, grande benefício = alta prioridade.** O esforço para execução dos testes é alto e o impacto da característica de Desempenho, também, é alto, os testes devem ser priorizados e deve-se utilizar as estratégias do guia, como ferramentas e as tabelas de relações para diminuir o esforço na execução dos testes.
- **Grupo IV: Baixo esforço e alto impacto. Baixo custo e grande benefício = alta prioridade.** O esforço para execução dos testes é baixo e o impacto da característica de Desempenho é alto na aplicação, os testes devem ser conduzidos e por envolver um baixo esforço pode-se, inclusive, adicionar mais propriedades a serem avaliadas por meio da seção Impacto das subcaracterísticas e até mesmo rever os testes e métricas selecionados a fim de adicionar outros caso necessário.

6.3.7 Sugestões de Ferramentas

Outra seção do guia é a de “Sugestões de Ferramentas”. Nela são apresentadas todas as ferramentas obtidas da literatura que podem auxiliar no teste de desempenho para IoT. Para essa seção foram catalogadas 6 ferramentas, em que 2 são de código aberto. Na Tabela 10 são elencadas as ferramentas de acordo com seus nomes, a descrição fornecendo detalhes, o

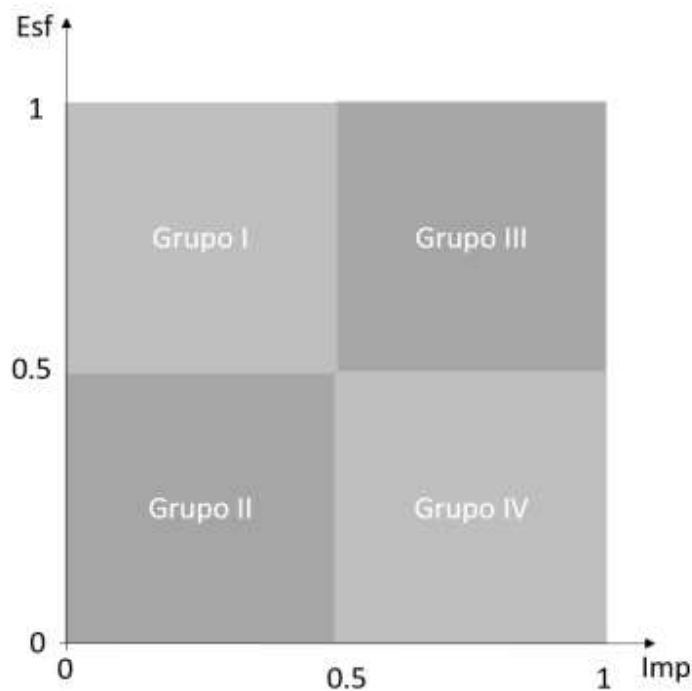


Figura 12 – Interpretação dos resultados do custo-benefício. Fonte: O autor.

ambiente de teste da ferramenta, a licença de uso (se é código aberto ou não) e a referência. Uma das ferramentas é um software instalável chamado *Wireshark*, sistema que faz a análise do tráfego da rede, podendo coletar métricas como, Tempo de execução, Tempo de transmissão da mensagem, Tempo mínimo de espera, entre outros.

6.3.8 Exemplo de Uso do Guia

A última seção do guia fornece um passo a passo baseado em um exemplo de um cenário de IoT. O exemplo trata-se de uma microempresa de produtos orgânicos, gerenciada por João, um engenheiro experiente. A empresa monitora, de forma simples, a umidade do solo em que são cultivados seus produtos. Na Figura 13 é descrito, de forma resumida, o passo a passo executado pelo João para a utilização do guia. A sequência está enumerada e a cada passo é apresentada uma descrição de como sucedeu cada etapa, como, por exemplo, no passo 3 é apresentado o ambiente de teste que João possui em seu sistema, 3 sensores de umidade, 4 aspersores para atuação, uma aplicação para o gerenciamento e João adicionou um medidor de energia para utilizar nos casos de testes que serão executados.

Além das seções e subseções apresentadas, o guia contém todas as referências

Tabela 10 – Ferramentas para teste de desempenho de IoT

| Ferramenta | Descrição | Ambiente de Teste | Licença | Referência |
|-------------------|---|--------------------------|----------------|-------------------|
| Neotys | Plataforma de teste para computação em nuvem. Executa análises, testes de carga. Permite que você avalie o tempo de resposta, fornecendo dados de desempenho. | Plataforma | Fechado | (NEOTYS, 2022) |
| Wireshark | Um sistema que analisa o tráfego na rede, podendo monitorar a entrada e saída de dados em diferentes protocolos. | Software Instalável | Código aberto | (WIRESHARK, 2016) |
| LoadUI Pro | Plataforma de teste para computação em nuvem. Executa testes de carga, testes funcionais, entre outros. Fornece cenários de teste em que o usuário pode fazer uso de diretamente ou adaptá-lo a sua aplicação ou até mesmo criar seus próprios casos de teste. Além disso, a ferramenta realiza monitoramento do tráfego na rede. | Plataforma | Fechado | (LOADUI, 2022) |
| IoTIFY | Oferece um laboratório virtual, com dispositivos virtuais para o teste de desempenho, de segurança e outros. | Simulador | Fechado | (IOTIFY, 2022) |
| Tcpdump | Semelhante ao Wireshark, monitora o tráfego de pacotes na rede. É possível identificar o tráfego em portas específicas, tráfego de um remetente específico, e destinatários, e outras funções. | Software Instalável | Código aberto | (TCPDUMP, 2010) |
| SOASTA CloudTest | Ferramenta usada no teste de desempenho e teste de carga em Sistemas WEB, permite simular um ambiente com vários dispositivos, com análise em tempo real. | Software Instalável | Fechado | (CLOUDTEST, 2022) |

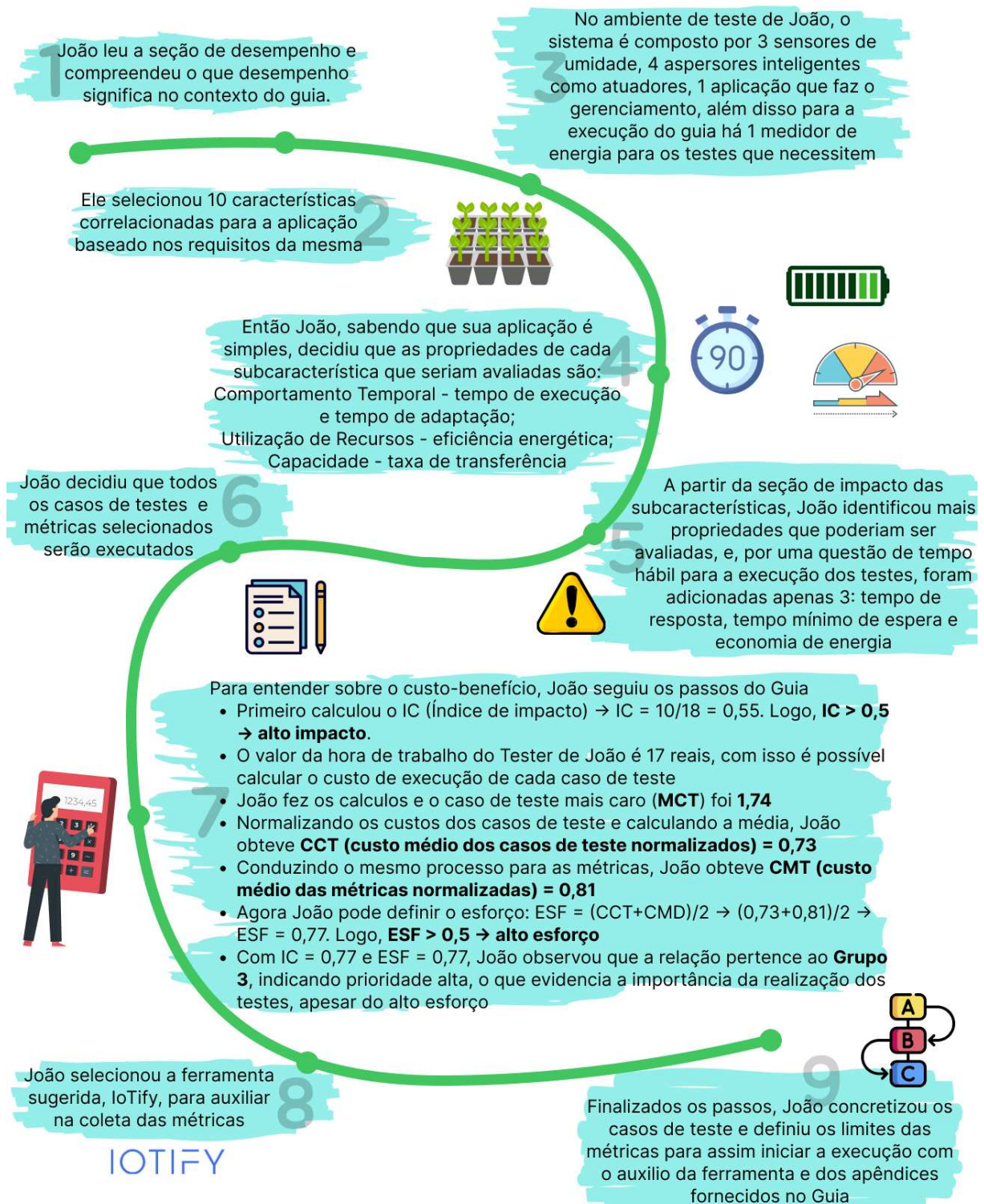


Figura 13 – Resumo da aplicabilidade do guia na empresa do João. Fonte: O autor.

que foram utilizadas para a sua elaboração e contém dois apêndices. No primeiro apêndice é apresentada uma tabela com a relação entre propriedades, casos de testes abstratos e métricas (ver Apêndice H). Essa tabela busca auxiliar no processo de execução dos testes ao selecionar as propriedades. Por exemplo, por meio da tabela é possível identificar que no momento da

execução do caso de teste CT12 - “Verificar quanto o sistema consome de energia ao atuar no ambiente”, é factível a realização da coleta da métrica M17 - “Avaliar o consumo médio de energia do sistema”, ajudando, assim, a otimizar o processo de teste.

O segundo apêndice traz a tabela de relação entre métricas, casos de testes abstratos e ferramentas (ver Apêndice I), busca auxiliar na automação de algumas medições e testes, por meio das ferramentas. Por exemplo, através da tabela é possível identificar que as métricas M03 - “Tempo de transmissão da mensagem” e M06 - “Tempo de Resposta”, podem ter suas coletas automatizadas por meio da ferramenta *IoTify*.

Ambos os apêndices foram adicionados ao guia de desempenho a partir de sugestões de especialistas obtidas através das avaliações que serão apresentadas no Capítulo 7.

6.4 Ferramenta de apoio ao uso do Guia de Desempenho

A estrutura do guia é dividida em seções e subseções para facilitar o uso, entretanto, foi desenvolvida uma ferramenta de apoio para o uso do guia de desempenho, a fim de agilizar ainda mais o processo de teste. Dessa forma, foi implementada uma Wiki⁷ que se concentra na característica *Desempenho* e cobre todas as seções e subseções do guia apresentadas anteriormente e ao final é gerado um plano de teste que pode ser baixado pelo usuário. Na Figura 14 é apresentada a tela inicial ao entrar na Wiki.

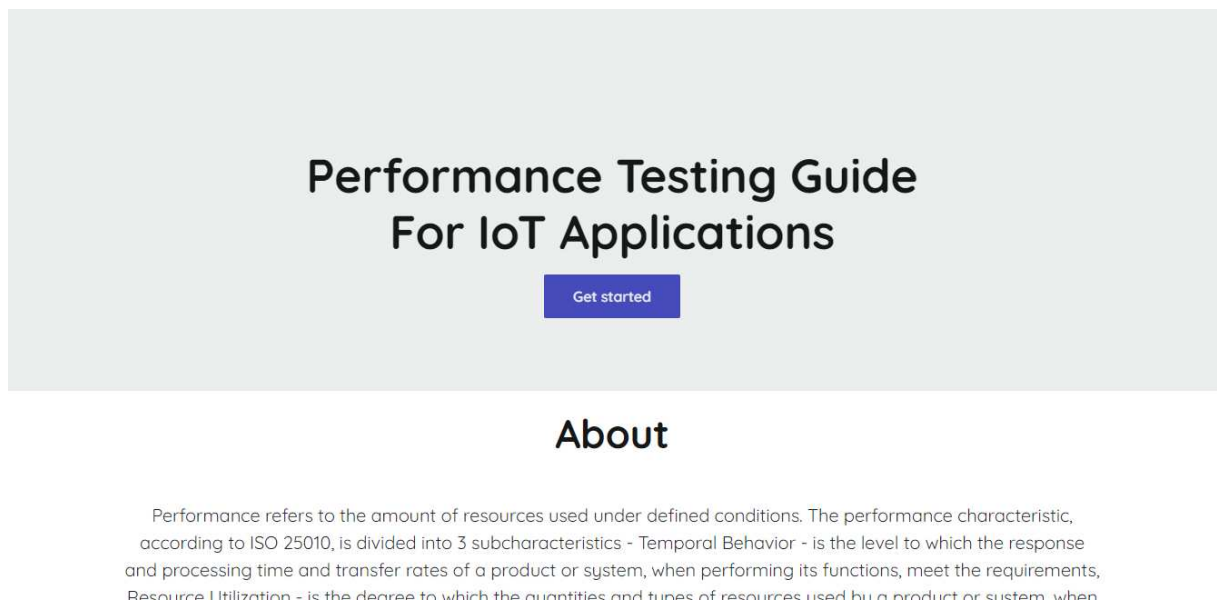


Figura 14 – Tela inicial da Wiki. Fonte: O autor.

⁷ Wiki : <<https://tinyurl.com/4pa8m52a>>

Na Wiki, o usuário do guia:

1. Seleciona a característica de Desempenho, podendo escolher definições dentre as fornecidas para personalizar seu plano de teste, conforme apresentado na Figura 15.
2. Seleciona quais as características, relacionadas a desempenho, são imprescindíveis⁶ na avaliação da aplicação. Para isso são apresentadas as definições de cada uma das 18 características relacionadas a *Desempenho* (ver Figura 16);

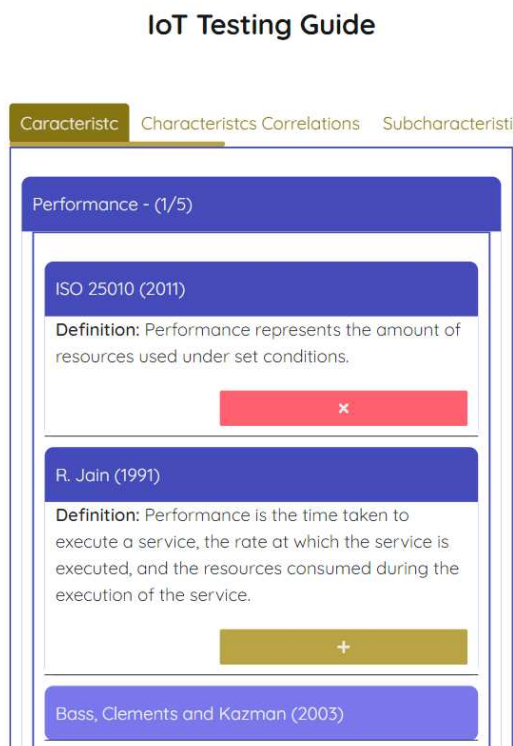


Figura 15 – Seleção das definições

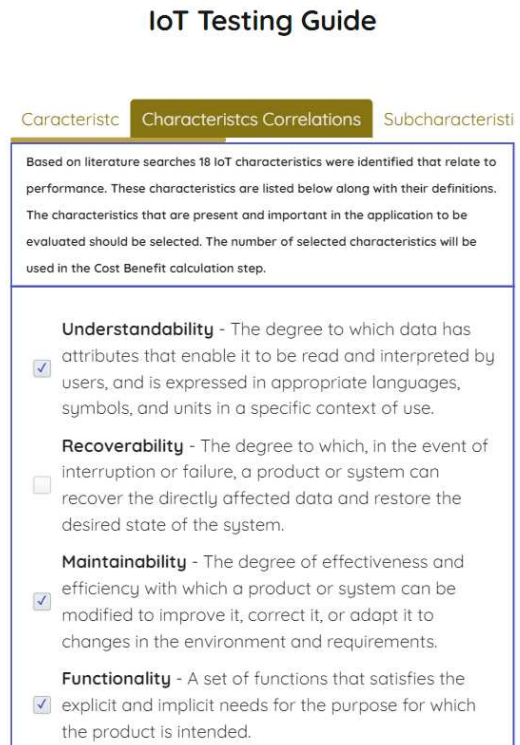


Figura 16 – Seleção das características correlacionadas

3. Seleciona as subcaracterísticas de desempenho das quais quer validar, conforme apresentado na Figura 17. Essa opção de seleção é útil ao se tratar de um engenheiro de software mais experiente, pois o mesmo pode fazer a escolha de testar um aspecto específico na sua aplicação ao invés de tudo relacionado a desempenho;
4. Seleciona as propriedades que pretende avaliar. À medida que são selecionadas as propriedades, caso seja selecionada uma propriedade que impacte outras, a Wiki sinaliza, sendo assim, útil para que seja evidenciada a importância de outras propriedades também sejam validadas. Na Figura 18, no canto superior direito, é apresentado o alerta no momento da seleção das propriedades;
5. Após a seleção das propriedades a própria Wiki recomenda quais casos de testes deveriam

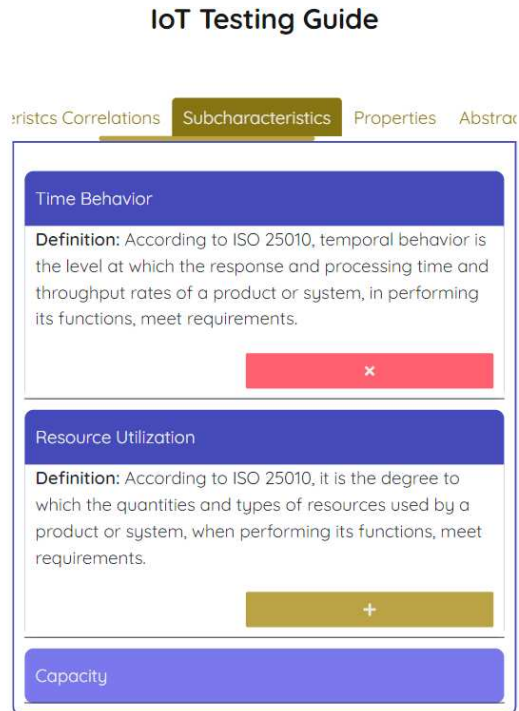


Figura 17 – Seleção das subcaracterísticas

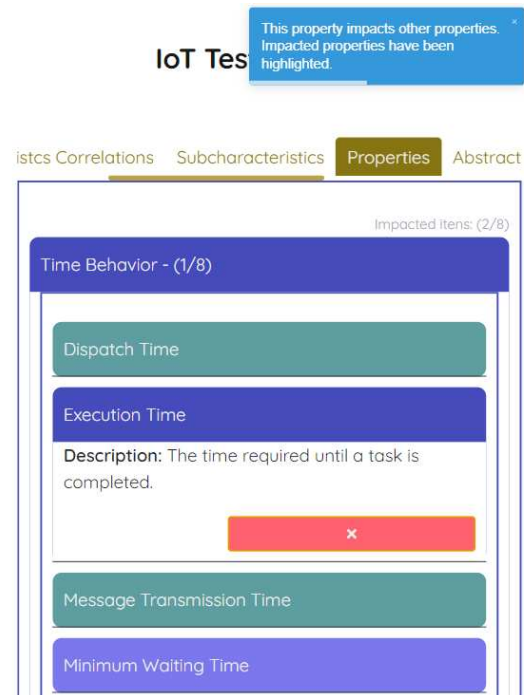


Figura 18 – Seleção das propriedades

ser selecionados, na Figura 19 essas recomendações são apresentadas nos casos de testes destacados no tom amarelo na imagem. Essas sugestões podem ser utilizadas, bem como pode-se editar, adicionando ou retirando casos de testes;

6. Do mesmo modo como os casos de testes, as métricas também são sugeridas e pode-se editar, adicionando ou retirando métricas, conforme apresentado na Figura 20. Essa flexibilidade para métricas e casos de teste pode ser utilizada pelos profissionais mais experientes, que buscam analisar aspectos específicos na aplicação e já está convicto de quais métricas e casos de testes lhe serão úteis.
7. Com os casos de testes e métricas, a Wiki calcula o custo-benefício a partir de valores médios de esforços, apresentados na Figura 21, esses valores também podem ser editados. Após a definição dos valores pode ser gerado o resultado e dado em que quadrante se encaixa o resultado (ver Figura 22);
8. As ferramentas também já são sugeridas na Wiki com base nas seleções de propriedades, métricas e casos de testes feitos, conforme apresentado na Figura 23;
9. Na Wiki é possível a qualquer momento ver um *preview* do plano de teste que está sendo criado e baixá-lo no formato PDF (ver Figura 24).

IoT Testing Guide

Properties Abstract Test Cases Metrics Cost Benefit

It is recommended that test cases that have been preselected by properties remain selected. Otherwise, test completeness is not guaranteed.

Time Behavior - (10/11)

Test Case 01 - Send command to actuator

Test Case 02 - Send command to actuator via external network

Test Case 03 - Send commands simultaneously

Test Case 04 - Send several of the same commands

Test Case 05 - Send command to the actuator at peak time

Test Case 06 - Receive sensor reading

Figura 19 – Seleção dos casos de teste

IoT Testing Guide

Properties Abstract Test Cases Metrics Cost Benefit

It is recommended that metrics that have been preselected by properties remain selected. Otherwise, measurement completeness is not guaranteed.

Time Behavior - (1/8)

M01 - Dispatch Time

M02 - Execution Time

M03 - Message Transmission Time

Purpose: Evaluate the transfer time of a message

Method: Consider the time when the message was sent and compare it to the time when the message was received

Measure:

$$X = t2 - t1$$

Y = message transmission time

Figura 20 – Seleção das métricas

IoT Testing Guide

Properties Abstract Test Cases Metrics Cost Benefit Suggested

The cost benefit provides insight into how high a priority it is to run the selected test cases and metrics. The calculation is based on the hourly value of the tester, the quantities of selected features in correlations, and the time to execute each test and metric.

* Average hourly wage of a North American "Tester" in the year 2021.

**The suggested time value of metrics and test cases is based on an execution performed by an experienced Tester on an application that has 2 sensors and 1 actuator and did not use any tools in the process.

☒ Use suggested hourly wage* (\$18/hour)

Professional hour value

18

☒ Use suggested times for test cases and metrics**

Test Cases

| TC01 | TC02 | TC03 | TC04 | TC05 | TC06 | TC07 |
|------------|------------|------------|------------|------------|------------|------------|
| 2 | 2 | 3 | 3 | 2 | 1 | 1 |
| in minutes | in minutes | in minutes | in minutes | in minutes | in minutes | in minutes |

Figura 21 – Cálculo de custo-benefício

| TC08 | TC09 | TC10 | TC11 |
|------------|------------|------------|------------|
| 2 | 2 | 1 | 2 |
| in minutes | in minutes | in minutes | in minutes |

Metrics

| M01 | M02 | M03 |
|------------|------------|------------|
| 3 | 5 | 2 |
| in minutes | in minutes | in minutes |

Calculate

Impact: 0.65
Effort: 0.82

High effort and high impact, that is, high cost but high benefit = high priority. The effort to execute the tests is very high and not executing impacts many correlated characteristics, the tests must be conducted and must use the strategies in the guide, such as tools and the relationship tables to decrease the effort in executing the tests.

Figura 22 – Resultado de custo-benefício

6.5 Considerações finais

Neste capítulo foi apresentado o procedimento metodológico para a instanciação do guia, bem como, a instanciação para a característica de *Desempenho* e uma ferramenta de apoio criada para auxiliar no uso do guia.

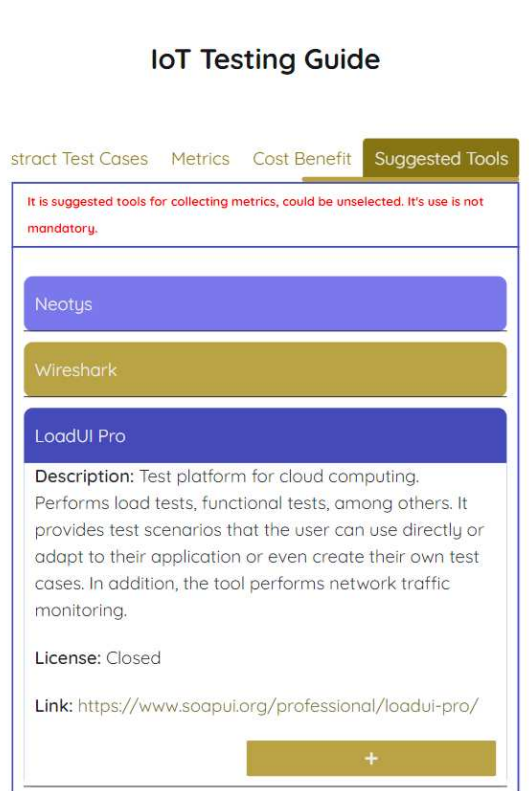


Figura 23 – Seleção de ferramentas

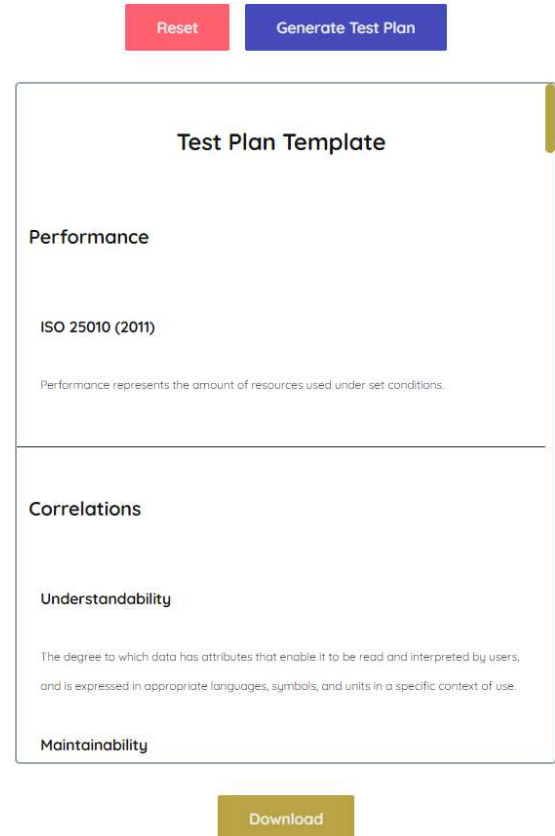


Figura 24 – Plano de teste

O procedimento metodológico para a instanciação do guia consiste na realização de revisão da literatura, análise e estruturação dos dados relacionados a característica e subcaracterísticas, observação de aplicações reais e refinamento. Durante a instanciação são elaborados artefatos como: lista de definições da característica, catalogação de ferramentas utilizadas no teste da característica, descrição ambiente necessário para os testes, dentre outros. Esses artefatos compõem a instanciação do guia, ao final do processo o guia instanciado deve passar por um refinamento e assim ser consolidado.

Para a instanciação para desempenho foram seguidos os 6 passos, listados. Ao final, todos os 11 tópicos estão presentes, as seções relacionadas às subcaracterísticas no guia são replicadas para as 3 subcaracterísticas existentes em *Desempenho*. Cada uma das subcaracterísticas possuem as subseções de Definição da subcaracterística, Contextualização, Casos de teste abstratos e Medições. No guia de desempenho são fornecidos 23 casos de testes, 22 métricas, obtidas a partir das 21 propriedades, também apresentadas no guia e são sugeridas 6 ferramentas para avaliar o *Desempenho*.

Para a utilização do guia de desempenho de forma prática foi implementada uma ferramenta de apoio, a Wiki, em que contempla todas as seções do guia. Nela é possível

selecionar subcaracterísticas e propriedades a serem avaliadas, a partir dessas seleções, são feitas sugestões de casos de testes, métricas e ferramentas, além disso, o custo-benefício é calculado automaticamente e ao final é possível gerar um plano de teste que pode ser baixado e utilizado pelo usuário.

No próximo capítulo serão apresentadas as avaliações aplicadas ao guia e os seus resultados.

7 AVALIAÇÃO

Neste capítulo são apresentadas as avaliações conduzidas para avaliar a instância do guia gerado pela abordagem proposta. Para a avaliação foi utilizado o *Guia de Teste de Desempenho para Aplicações IoT*, apresentado no Capítulo 6. A avaliação do guia foi conduzida em duas etapas: (i) do ponto de vista de especialista com o intuito de avaliar a estrutura geral do guia e o seu conteúdo; e através (ii) de um experimento controlado utilizando o guia para avaliar uma aplicação IoT. No capítulo também são apresentados os resultados e análises, bem como, as respostas às questões de pesquisa e ameaças à validade do estudo.

7.1 Avaliação com especialistas

A avaliação com os especialistas seguiu a metodologia descrita abaixo:

- **Objetivo.** O objetivo foi avaliar a estrutura e o conteúdo do *Guia de Teste de Desempenho para Aplicações IoT* do ponto de vista de especialistas.
- **Contexto.** A avaliação foi conduzida em dois dias com sete especialistas. Foi fornecida uma lista de verificação (*i.e checklist*) com perguntas para auxiliar na avaliação do guia.
- **Instrumentação.** O instrumento utilizado para o estudo é um *checklist* (ver Apêndice J) composto por 72 perguntas divididas em três partes:
 - (i) 6 perguntas sobre o perfil do especialista;
 - (ii) 5 perguntas relacionadas com a estrutura do guia; e
 - (iii) 61 perguntas dizem respeito ao conteúdo do guia.

Cada pergunta tem as seguintes opções: “Sim”, “Não”, “Não se aplica (N/A)” e “Observação”. Os especialistas podem utilizar essa última opção para explicar as razões da sua resposta, neste caso, irá ajudar a compreender, por exemplo, quando eles estão parcialmente de acordo em uma pergunta. Para avaliar o nível de concordância entre os especialistas, foi aplicado o método Kappa (FK) (FLEISS; COHEN, 1973) Esse método define um coeficiente de concordância que varia de 0 a 1. Se o coeficiente estiver mais próximo de 1, este valor indica que tem uma concordância maior entre os especialistas, e mais próximo de 0 indica que a concordância é aleatória.

- **Participantes.** Os participantes do estudo foram especialistas de áreas acadêmicas e/ou profissionais da Computação com experiência em Teste de software e/ou IoT. A Tabela 11 apresenta uma visão geral do perfil dos especialistas. Quanto aos seus conhecimentos,

quatro deles (1, 3, 4 e 6) têm experiência em Teste de Software, um especialista (2) com doutorado tem experiência na área de IoT e dois especialistas (5 e 7) possuem doutorado e experiência em ambas as áreas.

Tabela 11 – Perfil dos especialistas

| Especialista | Área | Ocupação | Título | Conhecimento | Experiência |
|---------------------|--------------|----------------------------------|------------------------|---------------------|--------------------|
| 1 | Profissional | Analista de Teste | Bacharel | Teste | 5 anos |
| 2 | Academia | Professor / Pesquisador | Doutor | IoT | 5 anos |
| 3 | Ambos | Teste Analista | Estudante de mestrado | Teste | 3 anos |
| 4 | Ambos | Gerente de uma fábrica de testes | Doutor | Teste | 13 anos |
| 5 | Academia | Analista de sistemas | Estudante de doutorado | Ambos | 5 anos |
| 6 | Profissional | Gerente de projetos | Mestre | Teste | 4 anos |
| 7 | Ambos | Analista de teste | Estudante de doutorado | Ambos | 9 anos |

7.1.1 Resultados e Discussão

A análise e resultados da avaliação dos especialistas são apresentados a seguir em duas partes: Avaliação da Estrutura do Guia e Avaliação do Conteúdo do Guia.

7.1.1.1 Avaliação da Estrutura do Guia

Para a avaliação foram elaboradas cinco perguntas para compreender se o guia tem uma estrutura padronizada e as suas seções seguem uma ordem lógica. As perguntas foram:

- EQ1. A estrutura proposta para o guia auxilia no seu uso?
- EQ2. As seções propostas são suficientes para propiciar a avaliação de uma característica IoT?
- EQ3. A ordem das seções segue uma sequência lógica?
- EQ4. Todas as seções possuem uma explicação do que se trata?
- EQ5. As informações apresentadas nas seções do guia, quando aplicável, possuem um estrutura padronizada (e.g., casos de teste, métricas)?

A Figura 25 mostra a concordância entre os especialistas. Cada linha representa um especialista e cada espaço horizontal delimitado representa possíveis respostas: “Sim”, “Não”,

“Parcialmente”, obtidas a partir das informações extraídas do campo de “Observações”, e “Não se aplica”.

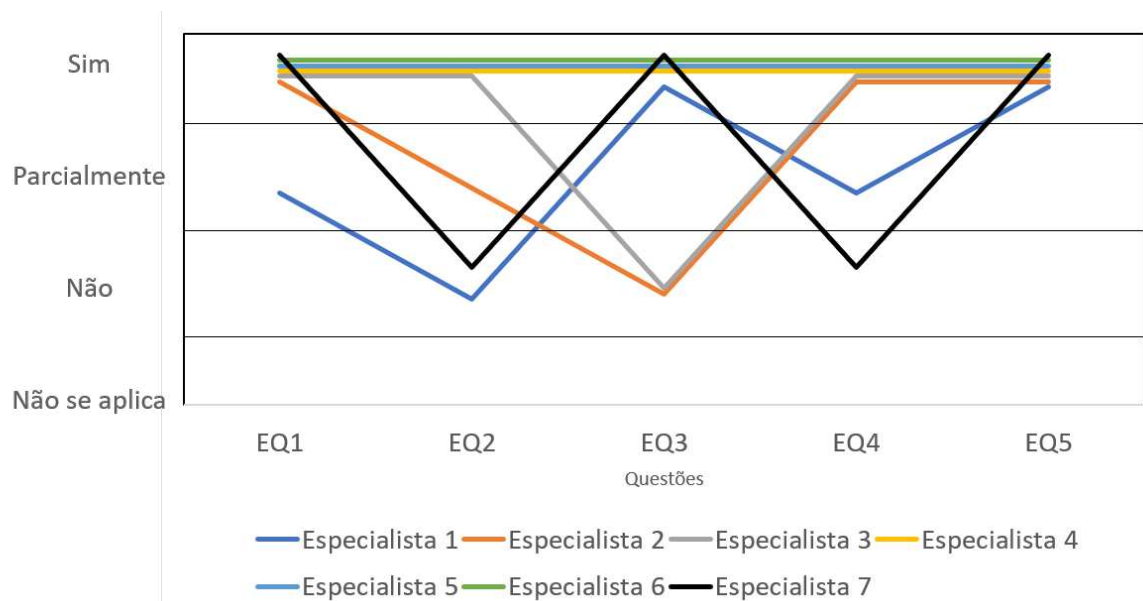


Figura 25 – Concordância entre especialistas

No eixo x são apresentadas as perguntas, EQ1, EQ2, EQ3, EQ4 e EQ5. Com base nas respostas foi observado que os Especialistas 4, 5 e 6 responderam a todas as perguntas da mesma forma, mostrando concordância entre as suas respostas. Dois deles (Especialistas 4 e 6) têm experiência na área de Teste de Software e o outro (Especialista 5) tem experiência em ambas as áreas: Testes de Software e IoT. Por outro lado, o Especialista 3 apenas discordou desses especialistas na pergunta EQ3, que se trata sobre o guia seguir uma sequência lógica. O Especialista 3 disse que a seção referente ao tópico de “Configuração de Ambiente” não têm uma ligação lógica com os tópicos “Definições da Características”, “Correlações da Característica” e “Definições das Subcaracterísticas”, que estão mais relacionadas com definições.

Para avaliar a concordância entre os especialistas sobre a estrutura do guia, o coeficiente FK foi calculado sobre as respostas dos especialistas nas cinco perguntas supracitadas. O valor do coeficiente FK obtido foi de 0,610, este valor apresenta uma *Concordância Substancial* entre os especialistas sobre a estrutura do guia, o que mostra uma boa indicação de concordância entre os especialistas.

Para compreender se a concordância entre os especialistas é positiva ou negativa, avaliou-se as respostas individuais sobre cada pergunta. Observou-se que a pergunta EQ2 teve uma maior taxa de discordância entre os especialistas. Por exemplo, embora o Especialista 4 tenha respondido “Sim” para EQ2, o mesmo disse que teria de utilizar o guia para afirmar que

as seções do guia são suficientes. A colocação do especialista deve-se ao fato de que alguns aspectos são mais difíceis de serem analisados sem a utilização do guia, como por exemplo, a “Configuração do Ambiente” ou o “Custo-benefício”.

As perguntas EQ1, EQ4 e EQ5, relacionadas com a facilidade de utilização do guia, apresentaram poucas divergências. Para a EQ3, o Especialista 3 disse que a seção “Configuração do Ambiente” deveria estar disposta em outra parte do guia, uma vez que as primeiras seções são mais gerais e, do ponto de vista desse Especialista, essa seção é mais específica. Entretanto, a seção de “Configuração do Ambiente” é definida e apresentada no guia como uma seção geral, uma vez que nela é apresentado o ambiente necessário (*e.g.*, sensores e atuadores) para a condução dos testes em todo o guia.

7.1.1.2 Avaliação do Conteúdo do Guia

O conteúdo do guia foi avaliado pelos especialistas com base em 61 perguntas disponíveis no *checklist* apresentado no Apêndice J.

Estas perguntas se subdividem, as primeiras são 7 perguntas gerais, por exemplo:

- CQ2 - “*A ordem das seções segue uma sequência lógica de apresentação?*”
- CQ4 - “*O título do guia é claro, conciso e sem ambiguidade?*”

As demais perguntas são específicas sobre cada um dos tópicos do guia, por exemplo para o tópico de “Definição da Característica” foi elaborada a seguinte pergunta: CQ8 - “*As definições de Desempenho apresentadas no guia estão claras?*”, relacionado ao tópico de “Configuração do Ambiente” há três perguntas, entre elas a CQ14 - “*Toda configuração de ambiente necessária para execução dos testes do guia estão listados?*”. Com relação as subcaracterísticas de *Desempenho*, foram conduzidas as mesmas perguntas para “Comportamento Temporal”, “Utilização de Recursos” e “Capacidade”, respectivamente, como por exemplo:

- CQ17/CQ29/CQ41 - “*Existem propriedades redundantes para essa subcaracterística?*”
- CQ23/CQ35/CQ47 - “*Cada métrica está descrita com clareza, concisão e sem ambiguidade?*”
- CQ18/CQ30/CQ42 - “*As definições apresentadas para cada propriedade estão claras?*”.

Além dessas perguntas foram feitas outras relacionadas aos demais tópicos como: CQ59 - “*Esta seção auxilia na identificação de ferramentas que possam ser úteis durante a validação com o guia?*”, referente ao tópico de “Sugestões de Ferramentas” e CQ60 - “*O texto apresentado na seção é suficiente para compreender como deve ser o uso do guia?*”, referente à

“Exemplo de Uso do Guia”.

A Figura 26 apresenta a visão geral das respostas dos 7 especialistas para cada pergunta. Observou-se que os especialistas concordaram em 19 perguntas (CQ4, CQ12, CQ15, CQ18, CQ20, CQ27, CQ28, CQ29, CQ35, CQ36, CQ38, CQ39, CQ42, CQ44, CQ46, CQ48, CQ50, CQ54 e CQ59). Essas perguntas se relacionam com as definições, descrições e título de cada seção do guia como, por exemplo, as perguntas: CQ04 - *“O título do guia é claro, conciso e sem ambiguidade?”*, CQ15 - *“A definição de comportamento temporal está clara?”* e CQ18 - *“As definições apresentadas para cada propriedade estão claras?”*, para todas essas perguntas as respostas dos especialistas se mostraram positivas, concordando entre si. Houve perguntas também relacionadas a redundância e clareza das seções como, por exemplo, CQ20 - *“Existem casos de testes redundantes? (e.g., , casos de testes com o mesmo objetivo)”*, CQ35 - *“Cada métrica está descrita com clareza, concisão e sem ambiguidade?”* e CQ48 - *“Existem métricas redundantes? (e.g., casos de testes com o mesmo objetivo)”*, todos os especialistas, novamente, responderam de forma positiva a essas perguntas, fornecendo resultados favoráveis sobre as seções de Medições e Casos de Teste Abstratos. Além dessas perguntas, houve outras relacionadas aos objetivos das seções, por exemplo, CQ54 - *“Esta seção (Impacto das Subcaracterísticas) auxilia na identificação de possíveis casos de teste ou métricas de subcaracterísticas distintas que possam ser feitas de maneira conjunta?”* e CQ59 - *“Esta seção (Sugestões de Ferramentas) auxilia na identificação de ferramentas que possam ser úteis durante a validação com o guia?”*, como resposta a essas perguntas todos os especialistas concordaram entre si, apresentando feedbacks positivos reconhecendo o auxílio prático fornecido por essas seções.

Os especialistas 2 e 4 tiveram dúvidas sobre algumas propriedades das subcaracterísticas de “Comportamento Temporal” e “Utilização de Recursos” nas perguntas, CQ18 e CQ30, respectivamente. Por exemplo, o Especialista 2 falou em relação às propriedades de “Comportamento Temporal” que *“Apesar de fornecer uma resposta positiva, gostaria de entender melhor a diferença entre o tempo de execução e o tempo de resposta”* e o Especialista 4 sugeriu a adição da propriedade “Consumo de dados” na subcaracterística “Utilização de Recursos”, estando inclusa na versão final do *Guia de Teste de Desempenho para Aplicações IoT*.

Seis especialistas concordaram que o tópico de correlação ajuda a identificar os conflitos que podem existir entre *Desempenho* e as outras características da IoT, de acordo com as respostas à pergunta CQ11 (*“Esta seção auxilia na identificação de possíveis conflitos entre*

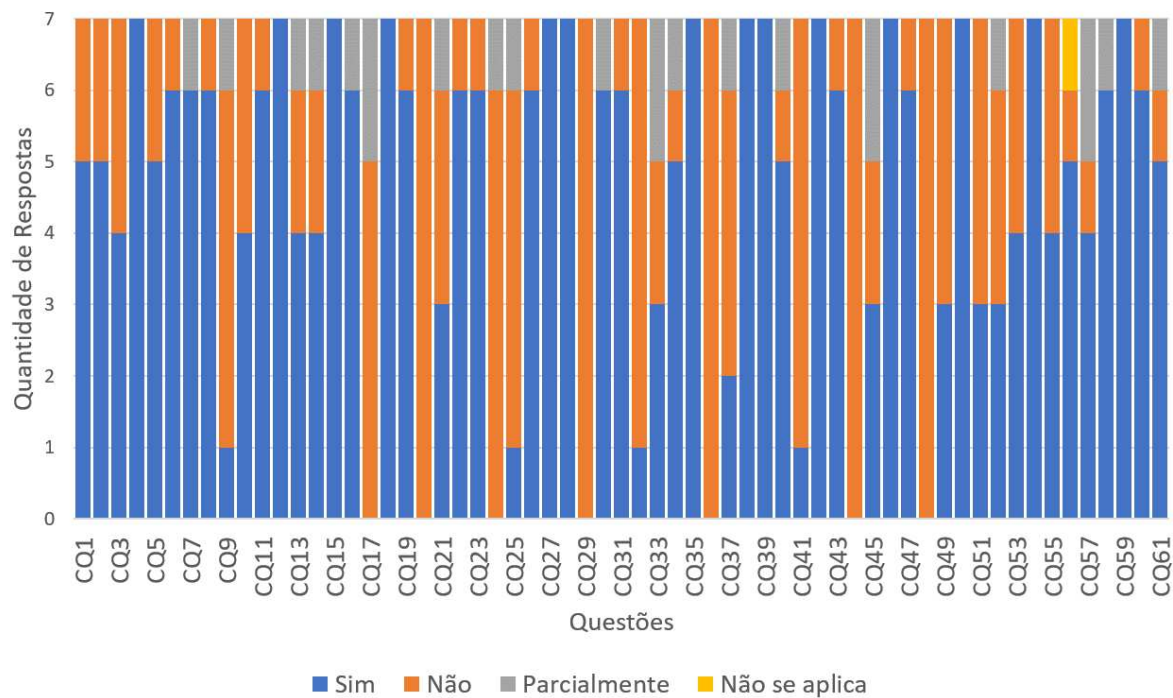


Figura 26 – Respostas dos especialistas para as perguntas sobre a avaliação do conteúdo do guia

características de IoT e Desempenho?”). O Especialista 7, que discordou desses especialistas, mencionou que deveria haver informação sobre que tipos de métricas poderiam ser automatizadas.

Sobre “Configuração do Ambiente”, os especialistas 2 e 3 não compreenderam as informações adicionais relacionadas com o contexto da seção. Por exemplo, o Especialista 3 salientou que - “a informação adicional parece ser genérica e não está relacionada com a Configuração do Ambiente”.

Na avaliação do “Custo-benéfico”, o Especialista 2 sugeriu mais detalhes sobre os custos e a explicação sobre a forma de analisar os resultados obtidos. Seis especialistas concordaram que o conteúdo apresentado na seção “Exemplo de Uso do Guia” é suficiente para compreender a utilização do guia. Por outro lado, o Especialista 1 disse que o passo a passo fornecido nessa seção não era suficiente para ajudar à execução do guia.

O valor do coeficiente FK calculado para as respostas obtidas referente à avaliação do conteúdo do guia foi de 0,306, este valor indica uma *Concordância Justa* entre os especialistas. Este valor foi inferior em comparação com a análise da estrutura do guia, contudo, acredita-se que o maior número de perguntas (61 contra 5) pode gerar uma maior probabilidade nas discordâncias entre os especialistas. Além disso, a avaliação do conteúdo é mais complexa sem a utilização do guia numa aplicação real. Assim, observou-se que alguns especialistas não conseguiram compreender tão claramente o conteúdo de alguns tópicos.

As principais discordâncias entre os especialistas foram nas seções de “Correlação das Características” e “Impacto das Subcaracterísticas”. Por exemplo, apenas o especialista 5 afirmou ser possível estabelecer as correlações através da seção “Correlação das Características”, por outro lado, os outros especialistas não acharam as informações suficientes para ajudar os usuários do guia a definir as correlações entre as características. É possível que o mal-entendido tenha sido causado por essa seção não haver um passo a passo detalhando como outra característica poderia ser relacionada com *Desempenho*. Em relação à seção “Impacto das subcaracterísticas”, três especialistas (2, 4 e 7) mencionaram que a sua explicação era confusa. Essa reação foi causada pela falta de clareza na forma como os impactos apresentados podem ser utilizados de forma prática. Ambas as seções, “Correlação das Características” e “Impacto das subcaracterísticas” foram reescritas e melhoradas as figuras.

Na avaliação das seções de “Contextualização”, “Casos de Testes Abstratos” e “Medições”, relacionados a cada subcaracterística, houveram poucas discordâncias entre os especialistas. Por exemplo, nas perguntas CQ25, CQ37, CQ49 - “*Cada métrica fornece detalhamento suficiente para ser entendida e executada?*” - relacionadas a cada subcaracterística, os especialistas apresentaram dificuldades em fornecer uma resposta, afirmaram que a resposta seria mais assertiva, caso fizessem o uso das métricas, na prática. As perguntas CQ22 e CQ34 - “*Os campos dos casos de teste são preenchidos corretamente? (e.g., cada caso de teste reflete o seu título?)*”, relacionadas as subcaracterísticas de Comportamento Temporal e Utilização de Recursos, respectivamente, os especialistas sugeriram para os campos relacionados aos casos de testes, “ambiente de teste” e “precondições”, que fossem fornecidos maiores detalhes, como melhorias conduzidas a partir desse *feedback*, foram adicionados como recursos necessários, para alguns casos de testes, medidores de energia, além disso, as precondições foram mais detalhadas, adicionando definições aos termos utilizados, tais como, horário de pico e rede local e externa.

Todos os artefatos dessa avaliação estão disponíveis no repositório de avaliação com especialistas deste trabalho¹.

7.2 Experimento controlado

A segunda avaliação, foi conduzida com foco na obtenção de resultados no âmbito mais prático sobre o *Guia de Teste de Desempenho para Aplicações IoT*.

O método de avaliação seguiu o modelo de um Experimento Controlado baseado

¹ Repositório de avaliação dos especialistas: <<https://tinyurl.com/yxx8k8rt>>

no processo definido por (WOHLIN *et al.*, 2012) (ver design no Apêndice K). O experimento controlado é uma metodologia de pesquisa, na qual os sujeitos experimentais são definidos de forma aleatória e seu objetivo principal é manter variáveis fixas e manipular outras, para então os resultados da manipulação serem analisados. Para a condução do experimento foram seguidas 9 atividades, conforme apresentadas a seguir.

1. Definição da ideia do experimento;
2. Definição do objetivo;
3. Definição das questões de pesquisa;
4. Definição das hipóteses;
5. Definição da aplicação;
6. Seleção dos participantes;
7. Definição da instrumentação;
8. Execução do experimento;
9. Análise e resultados.

A “Definição da ideia do experimento” parte da necessidade de avaliação do *Guia de Teste de Desempenho para Aplicações IoT* de forma prática, criando assim, grupos de controle (participantes utilizando o guia e participantes sem utilizar o guia), a fim de que esses participantes realizem a avaliação de desempenho de uma aplicação IoT, previamente definida, desde a etapa de planejamento dos testes, à execução e geração de falhas.

No que diz respeito a “Definição do objetivo” para esse experimento, o objetivo principal foi verificar se o *Guia de Teste de Desempenho para Aplicações IoT* oferece os meios necessários para detecção de falhas de desempenho existentes em aplicações IoT e se o mesmo, reduz o esforço na realização dos testes. Para isso, como condução da segunda atividade, foram definidas as seguintes questões de pesquisa:

- **Q1** - O uso do guia reduz o esforço destinado à condução dos testes de desempenho?
- **Q2** - Há uma maior eficácia nos casos de testes gerados por meio do uso do guia?
- **Q3** - O uso do guia amplia a detecção de falhas IoT?

A partir das questões de pesquisas foi conduzida a “Definição das hipóteses”. As hipóteses foram divididas em nulas, que afirmam não haver uma relação sobre o que está sendo tratado, e as alternativas, afirma uma relação da qual é desejada que seja verdadeira. As hipóteses representam as questões de pesquisa, por exemplo, as hipóteses $H_{0,0}$ e $H_{1,1}$ se relacionam com a Q1, as hipóteses $H_{0,1}$ e $H_{1,2}$ se relacionam com a Q2 e as hipóteses $H_{0,3}$ e $H_{1,3}$ se relacionam

com a Q3.

- **Hipóteses Nulas:**

H_{0,0} - A abordagem baseada em guias estruturados para conduzir atividades de teste de desempenho requer o mesmo esforço de teste que os testes de desempenho tradicionais.

H_{0,1}: A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho produz casos de teste com eficácia igual aos testes de desempenho tradicionais.

H_{0,2} - A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho detecta o mesmo número de falhas IoT que os testes de desempenho tradicionais.

- **Hipóteses Alternativas:**

H_{1,1}: A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho reduz o esforço de teste se comparado com os testes de desempenho tradicionais. $H_{1,1}: \text{Esforço (Com o guia)} < \text{Esforço (Sem o guia)}$.

H_{1,2}: A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho produz casos de teste mais eficazes do que os testes de desempenho tradicionais. $H_{1,2}: \text{Eficácia dos casos de teste (Com o guia)} > \text{Eficácia dos casos de teste (Sem o guia)}$.

H_{1,3}: A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho detecta mais falhas na IoT do que os testes de desempenho tradicionais. $H_{1,3}: \text{Número de falhas da IoT (Com o guia)} > \text{Número de falhas da IoT (Sem o guia)}$.

A avaliação de cada hipótese se dará da seguinte forma:

Para verificação das hipóteses nulas será aplicado o teste estatístico *T-Student's* (FIENBERG; LAZAR, 2001). No *T-Student's* a decisão de rejeição ou não das hipóteses nulas é baseada no *p-valor* - valor da probabilidade, e no nível de significância - valor base utilizado para comparação, caso *p-valor* for menor que o nível de significância a hipótese nula é rejeitada. Normalmente, nos testes estatísticos, usa-se um nível de significância igual a 0,05.

Sabendo que, para este trabalho, o esforço de teste é definido pelo tempo gasto no planejamento dos testes, para a hipótese $H_{1,1}$ será contabilizado o tempo de planejamento no que se refere ao tempo gasto para planejar cenários de teste, para selecionar as métricas e para configuração do ambiente e feito a média desse tempo, para cada grupo. Esse tempo será recolhido de cada grupo e comparado os resultados, a fim de verificar se há diferença no esforço da condução dos testes utilizando o guia e sem utilizar o guia.

Sabendo que, para este trabalho, a eficácia de um caso de teste é definida pela sua capacidade de detecção de falhas. Com isso, para a hipótese $H_{1,2}$ serão contabilizados os casos de testes criados por cada grupo e feito a média para cada grupo (i) e também serão contabilizados, do total, os casos de testes que detectaram falhas e feito a média para cada grupo (ii), será dividido segundo valor pelo primeiro. O valor obtido define o quão eficaz são os casos de testes de cada grupo. Os resultados dos grupos serão comparados entre si, a fim de verificar se há diferença na eficácia dos testes.

Para a hipótese $H_{1,3}$ será contabilizado a quantidade de falhas detectadas relacionadas à IoT de cada grupo e feito uma média. Esse valor será comparado, a fim de verificar se há diferença na quantidade de falhas IoT detectadas ao utilizar o guia.

Na atividade de “Definição da aplicação” foi selecionado o Automa GREat (ANDRADE *et al.*, 2017). O Automa GREat é um aplicativo IoT que possui como objetivo gerenciar o controle de ar condicionados e lâmpadas. Essa aplicação foi desenvolvida pelo laboratório GREat. No Automa GREat os usuários podem ligar e desligar as luzes e modificar a cor e intensidade. Na Figura 27 é apresentada a tela para o gerenciamento das luzes na aplicação. Para o ar condicionado o funcionamento é semelhante, é possível ligar e desligar e aumentar ou reduzir a temperatura. Na aplicação o usuário pode interagir de dois modos, o manual e o automático. No modo manual, o usuário controla todos os objetos, e no modo automático, o aplicativo liga e desliga os objetos por meio do sensor de movimento.



Figura 27 – Tela de gerenciamento das luzes no Automa GREat. Fonte: O autor

Devido à pandemia da COVID-19² o experimento foi conduzido de forma remota, para isso foi necessário realizar ajustes na aplicação que viabilizassem o uso. Portanto, foi utilizado apenas a função do gerenciamento das luzes, no que diz respeito a ligar e desligar, a fim de facilitar para os usuários. Como forma de simular a atuação, do acender das lâmpadas, visto que os usuários não possuíam luzes inteligentes, foi desenvolvido um alerta flutuante para ser exibido na tela da aplicação, conforme apresentado na parte inferior da Figura 27, exibindo o texto “Luz 3 está desligada”.

Na atividade de “Seleção dos participantes”, foram definidos 12 participantes diferentes organizados em dois grupos: G1 - 6 participantes utilizando o guia; e G2 - 6 participantes sem utilizar o guia. Na Tabela 12 é apresentado o perfil de cada participante, de acordo com o grupo. Nos grupos há 1 doutor, 2 mestres, 6 bacharéis e 3 estudantes da área de computação. Os participantes se dividem entre a área acadêmica e profissional, e entre eles há atuantes em ambas, como o caso dos participantes 1 e 5 do grupo G1. Na tabela também são apresentadas as funções exercidas por cada participante e as áreas de conhecimento, podendo ser IoT, Teste ou ambas, por exemplo, os participantes 1, 3 e 4 do grupo G2 possuem experiência em ambas as áreas. É importante salientar que todos os participantes, possuem alguma experiência na avaliação de requisitos não-funcionais.

Para a atividade de “Definição da instrumentação” foram elaborados vídeos, planilhas e documentos para dar suporte a execução do experimento. Como primeiro item da instrumentação foi fornecido um vídeo tratando sobre teste de desempenho e IoT, explicando conceitos básicos, a fim de garantir um conhecimento alinhado entre todos. Para que os participantes compreendessem sobre a aplicação, também foi elaborado um vídeo, detalhando seu funcionamento e auxiliando a como fazer a configuração da aplicação para a execução do experimento, todos os participantes tiveram acesso a esse vídeo. Exclusivo para os participantes do G1, foi disponibilizado um vídeo explicativo sobre a Wiki e fornecido o *Guia de Teste de Desempenho para Aplicações IoT*.

Com a intenção de recolher as falhas na aplicação detectadas pelos participantes, foi fornecido, para todos os participantes, um conjunto de documentos: (i) um *template* para relatório de falhas, (ii) definição das características e (iii) definição de níveis de gravidade das falhas. O *template* para relatório de falhas apresenta os seguintes campos:

- **Id** - identificação fornecida para falha detectada;

² Pandemia causada pelo vírus SARS-CoV-2 altamente contagioso, iniciou-se em 31 de dezembro de 2019.

Tabela 12 – Perfil dos participantes do experimento

| Grupo 1 - G1 (com o Guia) | | | | |
|----------------------------------|------------------------|--------------|-----------------------------------|---------------------|
| Participante | Título | Área | Ocupação | Conhecimento |
| 1 | Doutor | Ambas | Professor | Ambas |
| 2 | Estudante de graduação | Acadêmica | Aluno de Ciência da Computação | Teste |
| 3 | Estudante de mestrado | Ambas | Mestrando em Computação | Ambas |
| 4 | Estudante de graduação | Acadêmica | Aluno de Engenharia da Computação | Teste |
| 5 | Bacharel | Ambas | Mestrando em Computação | Teste |
| 6 | Analista de sistemas | Ambas | Analista | Ambas |
| Grupo 2 - G2 (sem o Guia) | | | | |
| 1 | Bacharel em Computação | Profissional | Gerente de tecnologia | Ambas |
| 2 | Bacharel | Acadêmica | Mestrando em Computação | Teste |
| 3 | Analista de sistemas | Profissional | Analista | Ambas |
| 4 | Mestre | Ambas | Doutorando em Computação | Ambas |
| 5 | Mestre | Profissional | Líder de TI | Teste |
| 6 | Estudante de graduação | Ambas | Programador | Teste |

- **Descrição** - descrição da falha que foi detectada;
- **Caso de teste/Métrica** - caso de teste ou métrica do qual o participante detectou a falha;
- **Características** - características relacionadas a falha detectada;
- **Gravidade** - nível de gravidade da falha detectada.

Como suporte para o preenchimento do relatório de falhas são fornecidos a definição das características (Apêndice L) e a definição de níveis de gravidade das falhas. No primeiro documento são elencadas as 18 definições das características correlacionadas com desempenho. No segundo documento são apresentados os níveis de gravidade e suas definições, seguindo o modelo GUT (CARDOSO *et al.*, 2021). Este modelo permite classificar as falhas em três categorias de severidade: pouco grave, grave e muito grave. Em que, Pouco Grave acontece quando há uma interrupção no fluxo de operação, mas ainda é possível concluir a ação por meio de tentativas ou *work around*; Grave acontece quando há uma interrupção no fluxo de operação, mas ainda é possível contornar ou remediar a situação, imediatamente, com a intervenção de terceiros; e Muito Grave acontece quando há uma interrupção no fluxo de operação que causa indisponibilidade de dados/operações e só é possível remediar a situação com mudanças

no sistema realizadas pela equipe técnica. A partir dessas definições cada participantes pode preencher o campo “Gravidade” do relatório de falhas.

Para obter os dados da execução do experimento, como, duração, documentos utilizados e *feedbacks*, foi dado aos participantes um *checklist* (ver Apêndice M) para ser preenchido ao final do experimento. No *checklist* do grupo G1 havia mais perguntas, essas relacionadas a WiKi (ver Apêndice N). Foram 38 perguntas, iguais, para ambos os grupos, em que 7 foram sobre o perfil dos participantes, que geraram a Tabela 12, e 31 sobre o experimento em si. Para os participantes do grupo G1 foram 22 perguntas a mais, sobre o uso da WiKi.

As 31 perguntas, feitas para ambos os grupos, foram divididas em 6 categorias: 6 perguntas da categoria Geral - que diz respeito à preparação para execução dos testes, como, por exemplo: “Foi necessário assistir mais de uma vez algum vídeo? Se sim, quais?”; a segunda categoria contém 3 perguntas relacionadas aos artefatos gerados, por exemplo, “Você utilizou algum *template* para a geração dos artefatos, se sim, qual?”; a terceira categoria relacionada à Execução contém uma pergunta, “Você utilizou alguma ferramenta para automatizar a execução de alguma etapa do teste de desempenho? Se sim em qual a etapa e quais ferramentas usou e, porque optou por essas?”; a quarta categoria é a de Esforço, que busca obter os valores dos tempos gastos em cada etapa, ela contém 18 perguntas, um exemplo de pergunta dessa categoria é, “Quanto tempo foi gasto na etapa de planejamento de testes?”; a quinta categoria é focada nas Dificuldades durante o experimento, contém 4 perguntas, por exemplo, “Você teve dificuldades na execução dos testes? Se sim, quais?”; por último, a sexta categoria são as Considerações finais que contém duas perguntas, como, “Há alguma sugestão ou mudança que você indicaria na organização experimento?”.

As 22 perguntas feitas, exclusivamente, para o grupo que utilizou o guia - G1, seguiu as mesmas categorias das perguntas supracitadas. Na categoria de Geral foi definida uma pergunta, “Você seguiu todo o processo sugerido no guia para a condução dos testes?”; na categoria Artefato, também foi definida uma pergunta “Você utilizou algum *template* além do fornecido pela wiki para a geração dos artefatos, se sim, qual e por quê?”; na categoria Execução foram elaboradas 8 perguntas, entre elas, “Você utilizou todas as etapas da WiKi na construção do plano de teste, se não quais não utilizou e por quê?”; na categoria Esforço foram definidas 5 perguntas, como, por exemplo, “Quanto tempo foi gasto na geração do plano de teste na WiKi?”; na categoria Dificuldades foram geradas 3 perguntas, por exemplo, “Você teve alguma dificuldade com a utilização da Wiki se sim quais foram as dificuldades?”; e na categoria de

Considerações finais, foram definidas 3 perguntas, entre elas, “Todos os campos que você usou da WiKi para geração do plano de teste foram realmente úteis e funcionais? Caso não, explicar o porquê.”.

Além de fornecidos os artefatos supracitados, foi dado aos participantes o link para download da aplicação e solicitado que cada participante gerasse, ao menos, um plano de teste, casos de testes e métricas (que podiam está presentes do plano de teste), além dos demais documentos já citados (o checklist e o relatório de falhas). Além disso, foi disponibilizado a cada participante um e-mail e telefone para contato, em caso de dúvidas. Toda a instrumentação utilizada para o experimento está disponível na nota de rodapé do repositório de instrumentação³.

Para a “Execução do experimento” cada participante teve em média 2 semanas. Os participantes que utilizaram o guia seguiram 8 passos e os demais seguiram 6 passos. Cada passo é detalhado abaixo.

1. Assistir o vídeo sobre teste de desempenho e IoT para um maior esclarecimento;
2. Para os participantes que utilizaram o guia. Assistir o vídeo sobre a WiKi e o guia para facilitar a utilização;
3. Baixar e instalar a aplicação;
4. Assistir o vídeo sobre a aplicação para um melhor entendimento e fazer a configuração seguindo o vídeo;
5. Para os participantes que utilizaram o guia. Acessar o link da WiKi para iniciar a avaliação;
6. Conduzir o planejamento, construção e execução dos testes, incluindo o preenchimento do relatório de falhas;
7. Ao final, responder checklist do experimento;
8. Enviar todos os artefatos via e-mail.

A última atividade da condução do experimento, “Análise e resultados”, é apresentada na subseção a seguir. Nela são detalhados os dados obtidos e cálculos estatísticos feitos, bem como o resultado das hipóteses e respostas as questões de pesquisa elencadas para o experimento.

7.2.1 Análise e resultados

Na Tabela 13 é apresentada a síntese dos dados utilizados para responder às questões de pesquisas e avaliar as hipóteses. A tabela está dividida segundo os grupos G1 e G2, em que para cada participante é apresentado: o ID, que equivale à coluna “Participante” da Tabela 12,

³ Experimento controlado: <<https://tinyurl.com/2s35dujt>>

o tempo gasto no planejamento em minutos; a quantidade de casos de teste especificados, a quantidade de casos de testes que detectaram falhas, quantidade falhas detectadas e a quantidade de falhas IoT detectadas, ao final de cada grupo é apresentado o total das informações. Na Tabela pode-se observar que, o participante 1 do grupo G1, doutor, com conhecimento em Testes de software e IoT, levou 50 minutos no planejamento, especificou 6 casos de testes, destes 5 detectaram falhas, totalizando 3 falhas, em que as 3 foram falhas IoT. Foi considerado falha IoT, toda falha que possa vir a indicar problemas nas variáveis que permeiam a IoT (*e.g.*, sensores, atuadores, a rede).

Tabela 13 – Visão geral dos dados do experimento por grupo

| Grupo 1 - G1 (com o Guia) | | | | | |
|----------------------------------|--|---------------------------|------------------------------------|------------------|----------------------|
| ID | Tempo de Planejamento (minutos) | Casos de Testes(#) | Casos de Testes - Falhas(#) | Falhas(#) | Falhas IoT(#) |
| 1 | 50 | 6 | 5 | 3 | 3 |
| 2 | 70 | 11 | 9 | 10 | 7 |
| 3 | 90 | 6 | 3 | 3 | 3 |
| 4 | 80 | 14 | 11 | 10 | 6 |
| 5 | 90 | 6 | 2 | 2 | 2 |
| 6 | 50 | 6 | 4 | 3 | 2 |
| Total | 430 | 49 | 34 | 31 | 23 |
| Grupo 2 - G2 (sem o Guia) | | | | | |
| 1 | 90 | 7 | 3 | 3 | 1 |
| 2 | 189 | 5 | 2 | 4 | 2 |
| 3 | 80 | 4 | 2 | 3 | 1 |
| 4 | 120 | 6 | 2 | 3 | 2 |
| 5 | 90 | 6 | 4 | 5 | 2 |
| 6 | 100 | 2 | 1 | 2 | 2 |
| Total | 660 | 30 | 14 | 20 | 10 |

Ao analisar a Tabela 13 é possível observar que todos os valores totais do G1 são superiores ao de G2, exceto pelo Tempo de planejamento. Um dos fatores que pode vir a ser a causa disso é o fato de que durante o planejamento, os participantes do grupo G2, entraram em contato a fim de confirmar se de fato não haveria nenhum *template* de suporte para o planejamento dos testes. Foi explicado que todo o suporte já havia sido fornecido por meio dos vídeos e planilhas, (os participantes do G1 não entraram em contato). Os participantes do G2 compreenderam, mas para eles o processo de testes desempenho, feito do zero, refletia em algo custoso, talvez por essa razão o tempo de planejamento desses participantes tenham sido maiores. Consequentemente os participantes do grupo G2, por o processo ser custoso, acabariam por gerar um número menor de casos de testes e esse fator impactar nos demais dados. Aspectos

como esse, evidenciam a necessidade e a relevância em haver uma abordagem estruturada para os testes das características IoT.

A seguir são apresentados os resultados do experimento controlado, baseado nas questões de pesquisa.

7.2.1.1 Q1 - O uso do guia reduz o esforço destinado à condução dos testes de desempenho?

Para responder a essa questão de pesquisa foram avaliadas a hipótese nula $H_{0,0}$ - A abordagem baseada em guias estruturados para conduzir atividades de teste de desempenho requer o mesmo esforço de teste que os testes de desempenho tradicionais; e a hipótese alternativa $H_{1,1}$ - A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho reduz o esforço de teste se comparado com os testes de desempenho tradicionais.

Para a avaliação da hipótese nula foi executado o teste *T-Student's* (FIENBERG; LAZAR, 2001) com nível de equivalência igual a 0,05. Para o cálculo foram considerados o tempo gasto (em minutos) no planejamento dos testes, desde o planejamento dos cenários, definição de métricas até a configuração do ambiente. Esses dados foram coletados a partir das respostas fornecidas pelos participantes no *checklist* do experimento. Os valores de tempo utilizados estão discriminados na coluna de “Tempo de Planejamento (minutos)” da Tabela 13. A partir desses valores foi calculado a média de tempo gasto por cada grupo. Na Tabela 14 são detalhados os resultados, são apresentados os ids das hipóteses relacionadas aos dados, a comparação utilizada para avaliar a hipótese alternativa, nesse caso, é baseado no esforço, a média de tempo gasto por cada grupo, o valor da probabilidade e o resultado estatístico, que pode ser “Diferença estatística significativa” ou “Não houve diferença estatística significativa”, correspondendo ao quão o resultado difere da hipótese nula, a ponto de que a mesma possa ser rejeitada.

Tabela 14 – Resultados da hipótese $H_{1,1}$

| $H_{x,y}$ | Comparação | Média G1 | Média G2 | Probabilidade | Resultado |
|-----------------------------|---|----------|----------|-----------------|-------------------------------------|
| $H_{0,0}$ e $H_{1,1}$ | Esforço: Tempo gasto (em minutos) no planejamento | 71,1 | 110 | p-value = 0,030 | Diferença estatística significativa |

Na Tabela 14 o valor do *p-value* obtido foi de 0,03, representando um valor menor comparado com o nível de equivalência igual a 0,05, gerando um resultado de “Diferença estatística significativa”, logo a hipótese nula $H_{0,0}$ - A abordagem baseada em guias estruturados

para conduzir atividades de teste de desempenho requer o mesmo esforço de teste que os testes de desempenho tradicionais, deve ser rejeitada.

A partir da rejeição da hipótese nula $H_{0,0}$ é possível avaliar a hipótese alternativa $H_{1,1}$ - A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho reduz mais o esforço de teste do que os testes de desempenho tradicionais. Conforme apresentado na Seção 7.2 ao detalhar as hipóteses, para a hipótese alternativa $H_{1,1}$ é comparado se o *Esforço (Com o guia)* < *Esforço (Sem o guia)*, dado que o esforço é medido pelo tempo gasto (em minutos) no planejamento. Seguindo o que é apresentado na Tabela 14, temos que o tempo médio de esforço gasto pelo grupo G1 para o planejamento foi de 71,1 minutos, enquanto para o grupo G2 foi de 110 minutos, resultando na seguinte sentença:

$$\text{Esforço (Com o guia)} < \text{Esforço (Sem o guia)} \equiv 71,1 \text{ minutos} < 110 \text{ minutos?}$$

A sentença apresentada é verdadeira, o esforço médio gasto para as atividades de testes de desempenho é reduzida com o uso do guia, conclui-se que a hipótese $H_{1,1}$ deve ser aceita.

7.2.1.2 Q2 - Há uma maior eficácia nos casos de testes gerados por meio do uso do guia?

Para responder a essa questão de pesquisa foram avaliadas a hipótese nula $H_{0,1}$ - A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho produz casos de teste com eficácia igual aos testes de desempenho tradicionais e a hipótese alternativa $H_{1,2}$ - A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho produz casos de teste mais eficazes do que os testes de desempenho tradicionais.

Para a avaliação da hipótese nula foi executado o teste *T-Student's* (FIENBERG; LAZAR, 2001) com nível de equivalência igual a 0,05. Para o cálculo foram considerados a quantidade de casos de testes utilizados por cada participantes e quantidade de casos de testes que detectaram falhas, esses dados equivalem à “Casos de Testes” e “Casos de Testes - falhas” da Tabela 13, respectivamente. Esses dados foram coletados a partir do relatório de falhas fornecidas pelos participantes. A partir desses valores foi calculado a média de casos de testes utilizados e a média de casos de testes que detectaram falhas, de cada grupo. Na Tabela 15 são detalhados os resultados, são apresentados os ids das hipóteses relacionadas aos dados, a comparação utilizada para avaliar a hipótese alternativa, nesse caso, é baseado na eficácia, a média de tempo gasto por cada grupo, o valor da probabilidade e o resultado estatístico, que pode ser “Diferença estatística

significativa” ou “Não houve diferença estatística significativa”, diz respeito a quão diferente o resultado difere da hipótese nula, podendo gerar assim, a rejeição da mesma.

Tabela 15 – Resultados da hipótese $H_{1,2}$

| $H_{x,y}$ | Comparação | Média G1 | Média G2 | Probabilidade | Resultado |
|-----------------------------|--|----------|----------|-----------------|-------------------------------------|
| $H_{0,1}$ e $H_{1,2}$ | <i>Eficácia</i> : Casos de testes que detectaram falhas / Total de casos de testes | 0,65 | 0,47 | p-value = 0,043 | Diferença estatística significativa |

Na Tabela 15 o valor do *p-value* obtido foi de 0,0018, representando um valor menor comparado com o nível de equivalência igual a 0,05, gerando um resultado de “Diferença estatística significativa”, logo a hipótese nula $H_{0,1}$ - A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho produz casos de teste com eficácia igual aos testes de desempenho tradicionais, deve ser rejeitada.

A partir da rejeição da hipótese nula $H_{0,1}$ é possível avaliar a hipótese alternativa $H_{1,2}$ - A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho produz casos de teste mais eficazes do que os testes de desempenho tradicionais. Conforme apresentado na Seção 7.2 ao detalhar as hipóteses, para a hipótese alternativa $H_{1,2}$ é comparado se o *Eficácia dos casos de teste (Com o guia)* > *Eficácia dos casos de teste (Sem o guia)*, dado que a eficácia é medida pela divisão dos casos de testes que detectaram falhas pelo total de casos de testes. Seguindo o que é apresentado na Tabela 15, temos que a eficácia média do grupo G1 foi de 0,65, ou seja, a cada 10 casos, em média, 6 detectaram falhas, enquanto para o grupo G2 a média foi de 0,47, ou seja, a cada 10 casos, em média, 4 detectaram falhas, resultando na seguinte sentença:

***Eficácia dos casos de teste (Com o guia)* > *Eficácia dos casos de teste (Sem o guia)* \equiv 0,65
casos de testes que detectam falhas > 0,47 casos de testes que detectam falhas?**

A sentença apresentada é verdadeira, a eficácia nos casos de testes gerados por meio do guia é superior aos demais casos de testes gerados de forma tradicional, conclui-se que a hipótese $H_{1,2}$ deve ser aceita.

7.2.1.3 Q3 - O uso do guia amplia a detecção de falhas IoT?

Para responder a essa questão de pesquisa foram avaliadas a hipótese nula $H_{0,3}$ - A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho

detecta o mesmo número de falhas IoT que os testes de desempenho tradicionais e a hipótese alternativa $H_{1,3}$ - A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho encontra mais falhas na IoT do que os testes de desempenho tradicionais.

Para a avaliação da hipótese nula foi executado o teste *T-Student's* (FIENBERG; LAZAR, 2001) com nível de equivalência igual a 0,05. Para o cálculo foram consideradas os números de falhas IoT detectadas. Esses dados foram coletados a partir do relatório de falhas fornecidas pelos participantes, realizando a análise de todas as falhas foram contabilizadas aquelas consideradas falhas IoT, como falhas IoT, considera-se toda falha que possa vir a indicar problemas nas variáveis que permeiam a IoT (*e.g.*, sensores, atuadores, a rede). Os valores da quantidade total das falhas e das falhas IoT, estão discriminados na Tabela 13, nas colunas Falhas e Falhas IoT, respectivamente. A partir dos valores das falhas IoT foi calculado a média de falhas IoT detectadas por cada grupo. Na Tabela 16 são detalhados os resultados, são apresentados os IDs das hipóteses relacionadas aos dados, a comparação utilizada para avaliar a hipótese, nesse caso, é baseado nas falhas IoT, a média de falhas IoT detectadas por cada grupo, o valor da probabilidade e o resultado estatístico, que pode ser “Diferença estatística significativa” ou “Não houve diferença estatística significativa”, correspondendo ao quão o resultado difere da hipótese nula, a ponto de que a mesma possa ser rejeitada.

Tabela 16 – Resultados da hipótese $H_{1,3}$

| $H_{x,y}$ | Comparação | Média G1 | Média G2 | Probabilidade | Resultado |
|-----------------------------|--|----------|----------|-----------------|-------------------------------------|
| $H_{0,2}$ e $H_{1,3}$ | Falhas IoT: Números de falhas detectadas | 3,8 | 1,6 | p-value = 0,027 | Diferença estatística significativa |

Na Tabela 16 o valor do *p-value* obtido foi de 0,0030, representando um valor menor comparado com o nível de equivalência igual a 0,05, gerando um resultado de “Diferença estatística significativa”, logo a hipótese nula $H_{0,3}$ - A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho detecta o mesmo número de falhas IoT que os testes de desempenho tradicionais, é rejeitada.

A partir da rejeição da hipótese nula $H_{0,3}$ é possível avaliar a hipótese alternativa $H_{1,3}$ - A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho encontra mais falhas na IoT do que os testes de desempenho tradicionais. Conforme apresentado na Seção 7.2 ao detalhar as hipóteses, para a hipótese alternativa $H_{1,3}$ é comparado se o *Número de falhas da IoT (Com o guia) > Número de falhas da IoT (Sem o guia)*. Seguindo o que é

apresentado na Tabela 16, temos que a quantidade média de falhas IoT detectadas pelo grupo G1 foi de 3,8 falhas IoT, enquanto para o grupo G2 foi de 1,6 falhas IoT, resultando na seguinte sentença:

Número de falhas da IoT (Com o guia) > Número de falhas da IoT (Sem o guia) \equiv 3,8 falhas IoT > 1,6 falhas IoT?

A sentença apresentada é verdadeira, a quantidade média de falhas IoT detectadas é superior com o uso do guia, conclui-se que a hipótese $H_{1,3}$ é aceita.

Ainda em relação à questão de pesquisa Q3, além de haver sido coletado as falhas, os participantes classificaram a gravidade das mesmas seguindo a matriz GUT (CARDOSO *et al.*, 2021). A classificação da gravidade nos testes estatísticos foi apenas como um fator comparativo entre os resultados, uma vez que através da gravidade é possível identificar erros que geram problemas críticos no sistema. Este modelo permite classificar as falhas em três categorias de severidade: pouco grave, grave e muito grave. A distribuição da gravidade de 23 falhas IoT detectadas por G1 são: 5 foram muito graves; 6 foram graves; e 12 foram pouco graves; enquanto em G2, das 10 falhas IoT, 2 foram muito graves; 3 foram graves; e 5 pouco graves.

Na contabilização houveram falhas iguais que foram identificadas por: (i) participantes que pertenciam ao mesmo grupo; (ii) participantes que perteciam a grupos distintos.

No primeiro caso essas falhas foram tidas como duplicadas, sendo assim, contabilizadas apenas uma única vez. No segundo caso de falhas que foram identificadas por participantes de grupos distintos, foi contabilizado a mesma falha para cada grupo. Por exemplo, o G1 indentificou 5 falhas muito graves, e o G2 indentificou 2 falhas muito graves, totalizando 7 falhas muito graves identificadas pelos grupos. Dessas 7 falhas, duas delas foram identificadas por ambos os grupos, deste modo, no sistema como um todo foram identificadas 5 falhas IoT muito graves, entretanto, contabilizamos para os cálculos do experimento 7 falhas muito graves dado que avaliamos cada grupo individualmente.

Na Tabela 17 é possível observar todas as falhas IoT que os participantes classificaram como muito graves no sistema. Na primeira coluna estão os ids, na segunda coluna as falhas identificadas e na terceira coluna por quais grupos foi identificada cada falha. Todos as falhas IoT, sejam muito graves, graves ou pouco graves, detectadas por G2, também foram detectadas por G1, o contrário não é verdade.

Tabela 17 – Falhas IoT categorizadas como muito graves

| Id | Falha | Grupo |
|----|---|---------|
| 1 | Perda de conexão constante, mesmo com dispositivos estando próximos | G1 e G2 |
| 2 | Quando o usuário liga todas as lâmpadas ocorre um timeout entre as lâmpadas | G1 |
| 3 | Após ligar no mínimo 5x a mesma lâmpada o sistema não reconhece que a lâmpada está ligada | G1 |
| 4 | Ao ter mais de uma sala conectada o sistema não reconhece mais as mudanças do ambiente | G1 |
| 5 | Toast de atuação envia mensagem duplicadas mesmo ambiente não tendo mudanças | G1 e G2 |

7.3 Ameaças à validade

As ameaças à validade estão divididas em interna e externa (WOHLIN *et al.*, 2012). A primeira diz respeito ao quão confiável e livre de viés é a avaliação, a segunda corresponde à capacidade da replicabilidade da avaliação. A seguir são elencadas as ameaças.

Validade interna. De forma geral, uma ameaça à validade, enfrentada, foi a de que a abordagem de guias estruturados, propriamente dita, não passou por uma avaliação sistemática. Em ambas as avaliações a abordagem estava instanciada para desempenho. Entretanto, os *drafts* iniciais da abordagem de guias estruturados são apresentados a especialistas desde a etapa de qualificação deste trabalho, quando foi apresentado a primeira vez os tópicos da estrutura do guia. A estrutura do guia, também, foi analisada por um especialista em Teste de software e Computação móvel, conforme apresentado ao final da metodologia de construção do guia, e esse especialista sugeriu melhorias que foram aplicadas. Além disso, a estrutura do guia instanciado pela abordagem proposta foi avaliada por especialistas conforme apresentado na Seção 7.1.

Na primeira avaliação, em que foi validado a estrutura e o conteúdo, a principal ameaça diz respeito aos especialistas. Alguns deles não possuíam conhecimentos em ambas as áreas: Testes de Software e IoT. Para minimizar esta ameaça, foi fornecido um *checklist* com 72 perguntas para apoiar a avaliação.

Na segunda avaliação, o experimento controlado, não houve a condução de um piloto, entretanto o *design* do experimento vem sendo apresentado desde a proposta deste trabalho, foram recebidos *feedbacks* e conduzidas constantes melhorias. Outra ameaça, é o fato que de alguns participantes do experimento não possuem conhecimento em ambas as áreas: Testes de Software e IoT, entretanto a aplicação fornecida para a avaliação era simples e foi dado todo o suporte, tendo à disposição comunicação constante para esclarecer qualquer dúvida durante

a execução. Relacionado ao experimento ser remoto não afetou os resultados, uma vez que foi fornecido um roteiro para execução e dado vídeos para explicar os conceitos envolvidos no experimento. Outra ameaça diz respeito à utilização de apenas uma aplicação no experimento. Para minimizar esta ameaça, foi selecionada uma aplicação que cobre as funcionalidades básicas da IoT e nos permite testar a característica *Desempenho*, bem como o guia. A última ameaça interna trata-se do *checklist* ter sido respondido ao final do experimento e talvez os participantes não tenham se atentado na coleta de tempo de forma precisa, entretanto, todos os participantes tinham alguma experiência em avaliações e a diferença de tempo obtido entre os grupos foi um valor expressivo.

Validade externa. Para a condução do experimento controlado de forma remota, foi necessário realizar um ajuste na aplicação criando um alerta flutuante para simular a atuação, esse ajuste pode dificultar a replicação da avaliação. Entretanto, a mudança realizada foi demasiadamente simples e a mesma, não altera a sequência das atividades de condução do experimento.

7.4 Considerações Finais

Este capítulo apresentou as avaliações conduzidas para validar a instância do guia gerado pela abordagem proposta e as ameaças à validade. Foram conduzidas duas avaliações. A primeira foi um estudo empírico com sete especialistas para avaliar a estrutura e o conteúdo do guia. Para essa avaliação foi aplicado o método Fleiss' Kappa para analisar a concordância entre os especialistas. As respostas da maioria dos especialistas sobre a estrutura do guia foram positivas, como, por exemplo, 5 especialistas concordaram que a estrutura proposta pelo guia auxilia no uso do mesmo. As maiores discordâncias dos especialistas foi sobre o conteúdo do guia, uma vez que alguns aspectos são difíceis de compreender sem a utilização do guia, como as métricas. Deste modo, foi executado um experimento controlado com 12 participantes, para usarem o guia na prática e os resultados mostraram os benefícios do guia no que diz respeito à redução do esforço nos testes, o aumento da eficácia dos casos de teste, e a detecção de falhas IoT.

8 DISCUSSÃO

Neste capítulo são discutidas as questões de pesquisa investigadas neste trabalho:

- **QP1.** Como o processo de teste para aplicações IoT deve ser organizado?
- **QP2.** Como avaliar as características das aplicações IoT?
- **QP3.** Como facilitar a detecção de falhas IoT nas aplicações?

O capítulo se divide em três seções principais, na Seção 8.1 é discutida a questão QP1, na Seção 8.2 é discutida a questão QP2 e na Seção 8.3 é discutida a questão QP3.

8.1 QP1. Como deve ser organizado o processo de teste da IoT?

Sabendo que avaliar aplicações IoT envolve diversos desafios (CARVALHO, 2018), o objetivo dessa questão de pesquisa foi compreender como melhor pode ser estruturado a execução os testes dessas aplicações, a fim de propor uma solução para colaborar com uma organização mais sistemática e padronizada para os testes IoT.

Ao realizar a análise da literatura foram identificadas as 4 lacunas apresentadas na Seção 3.2, são elas:

- Ausência de **soluções estruturadas**;
- Carência de **abordagens flexíveis**;
- Ausência de **catalogação de ferramentas e medidas**, e;
- Carência de abordagens que considerem o **impacto das correlações das características**.

Com essas lacunas, é sugestivo que a organização do processo de teste IoT deva ser baseado em uma “**solução estruturada e flexível** que apresente **ferramentas e métricas** para ser utilizadas no processo e que considere o **impacto das correlações das características IoT**”.

Portanto, a abordagem de teste baseado em guias estruturados para avaliar características IoT, apresentada neste trabalho, visa fornecer uma solução que colabore para organização do processo de teste IoT, fundamentado nas lacunas elencadas.

A abordagem elaborada seguiu um modelo de tópicos, provendo assim uma solução estruturada para os testes. Entre os tópicos está Sugestões de ferramentas, Medições e Correlação das características. A estrutura do guia é adaptável para a qualquer característica IoT, considerando, características que possuem ou não subcaracterísticas. Para a adaptação basta ser seguida a metodologia de instanciação do guia, apresentada na Figura 6.

A partir da estrutura flexível provida pela abordagem, foi gerada a instância do

guia para a característica *Desempenho*. O *Guia de Teste de Desempenho para Aplicações IoT* recebeu feedbacks positivos. A organização estrutural foi mencionada, pelos especialistas, como padronizada e as sequências dos tópicos foram vistas como lógicas.

Como análise final, o guia contempla todas as lacunas citadas no início da seção e a sua estrutura se mostrou favorável para a organização do processo de teste IoT.

8.2 QP2. Como avaliar as características de aplicações de IoT?

Visto que muitos dos desafios de teste são relacionados às características de IoT, esta questão de pesquisa buscou investigar os métodos utilizados para avaliá-las e, assim, propor uma solução que se baseia nessas características para testar aplicações IoT.

Ao analisar a literatura foram identificados estudos que tratavam das características IoT em diversos contextos, conforme apresentado na Subseção 5.1.1, três desses contextos são:

- Correlação;
- Conflitos de interesse, e;
- Divisão em subcaracterísticas.

Em correlações é evidenciado os impactos que as características IoT podem ter uma sobre as outras. Em conflitos de interesses são discutidos os conflitos que possam vir a existir entre as subcaracterísticas da característica em questão, e na divisão em subcaracterísticas é defendido a quebra de características em subcaracterísticas para garantir uma maior cobertura da aplicação. Tendo em vista esses contextos, a solução proposta neste trabalho, buscou contemplá-los.

O guia pode ser dividido de acordo com as subcaracterísticas, que por sua vez, cada subcaracterística se divide em propriedades que gerarão as métricas. Esse detalhamento auxilia na captura de aspectos mais específicos a serem avaliados. No guia há o tópico de Correlações da característica, em que é possível estabelecer conexão com o Custo-benefício e definir prioridades para a execução dos testes. Além disso, as correlações podem ser positivas e negativas, saber quais são as características que impactam positiva ou negativamente a característica a ser avaliada, auxilia na antecipação e mitigações para possíveis problemas que venham a existir. No guia é definido um tópico de Impacto das subcaracterísticas, em que as propriedades, são relacionadas a partir do impacto existente entre elas. Os participantes do experimento do grupo G1 fizeram uso da Wiki e utilizaram esse impacto, em que a própria Wiki sugere de forma automática outras propriedades a serem avaliadas (Figura 18), seguindo essas sugestões os participantes do G1,

obtiveram resultados positivos, inclusive gerando casos de testes mais eficazes (ver Subsubseção 7.2.1.2).

O *Guia de Teste de Desempenho para Aplicações IoT* em que estão presentes a Correlações da característica, Contextualização e Impacto das subcaracterísticas, obteve excelentes resultados ao avaliar o *Desempenho*. O participante 1 do grupo G1 do experimento, afirmou, que o guia é suficiente para a condução dos testes de desempenho de aplicações IoT. Com os resultados provenientes das avaliações, é possível notar que um guia centrado nas características IoT pode ser mais eficaz para avaliar os aspectos intrínsecos da área. As métricas e casos de testes presentes no guia também são totalmente responsáveis pela avaliação das características, tendo em vista que são desses tópicos que provém as falhas e os resultados das medições. Na próxima seção esses dois tópicos serão fundamentais para a discussão da questão de pesquisa.

8.3 QP3. Como facilitar a detecção de falhas IoT nas aplicações?

Dado os desafios nos testes de aplicações IoT, com essa questão de pesquisa buscou investigar meios que viabilizassem e facilitassem a detecção de falhas IoT, a fim de incorporá-los na solução proposta.

Baseado-se na Subseção 5.1.1, foi atestado na literatura a relevância dos casos de testes e métricas para a detecção de falhas. A abordagem fornece uma metodologia que colabora para construção de casos de testes focados na característica alvo e nos aspectos que rodeiam a IoT - sensores e atuadores. Na própria instanciamento do guia para uma característica que possui subcaracterísticas, como no caso de *Desempenho*, os casos de testes abstratos são divididos entre cada uma das subcaracterísticas, a fim de facilitar a detecção das falhas.

As métricas, são provenientes das propriedades e semelhantes aos casos de testes, cada subcaracterística possuem as suas propriedades. As métricas são do contexto IoT e metrifica aspectos da característica alvo, por essa razão os resultados obtidos têm maior probabilidade de capturar falhas.

No experimento controlado utilizando o *Guia de Teste de Desempenho para Aplicações IoT*, os participantes do grupo G1, que utilizaram o guia, detectaram mais falhas IoT, do que o grupo que não utilizou o guia (23 contra 10). Todos os participantes do grupo G1 detectaram a falhas, utilizando, unicamente, os casos de testes e métricas provenientes do guia.

8.4 Considerações Finais

Nesse capítulo foram feitas discussões acerca das questões de pesquisa deste trabalho.

Para a Q1 - Como deve ser organizado o processo de teste da IoT? - foi identificado que lacunas provenientes dos desafios sugeriam uma possível organização para o processo de teste. As lacunas serviram como base para estruturação da abordagem dessa pesquisa que apresentou resultados positivos no processo de testes de aplicações IoT.

Para a Q2 - Como avaliar as características de aplicações de IoT? - a partir dos estudos utilizados para construção do guia foi evidenciada aspectos fundamentais na avaliação das características IoT, esses são representados pelos tópicos de Correlação das características, Contextualização (propriedades) e Impacto da subcaracterísticas, esses tópicos são úteis para avaliar as características IoT, tendo em vista que são elencadas as correlações que interagem de forma positiva ou negativa com característica. A característica, quando possível, se divide em subcaracterísticas e são listados os impactos existentes entre as subcaracterísticas e isso pode vir a garantir um menor índice de problemas futuros. Os tópicos de métricas e casos de testes também contribuem no processo de avaliação de uma característica IoT.

Para a Q3 - Como facilitar a detecção de falhas IoT nas aplicações? - a partir dos estudos utilizados para construção do guia foi evidenciada aspectos fundamentais na avaliação das características IoT, esses são representados pelos tópicos de Medições e Casos de teste abstratos. As métricas e casos de testes provenientes do guia são facilitadores na detecção de falhas, por haverem sido definidos baseados em estudos no contexto de testes IoT e em aplicações IoT, portanto, provendo cenários e medições que contemplem o contexto.

9 CONCLUSÕES E TRABALHOS FUTUROS

Nesta dissertação foi apresentada uma abordagem de teste para aplicações IoT baseado em características. Ao analisar a literatura foi identificados desafios nos testes, por exemplo, uma falta de padronização nos testes de aplicações IoT para essas características (CORTÉS *et al.*, 2019),(BURES *et al.*, 2019). A partir dos desafios foram detectadas lacunas (soluções estruturadas, abordagens flexíveis, catalogação de ferramentas e medidas, correlação das características), as quais fundamentou uma revisão da literatura a fim de buscar estudos que as preenchessem. Na revisão da literatura não identificamos nenhum estudo que preenchessem todas as lacunas, então propomos uma solução baseada em guias estruturadas para testes de aplicações IoT que cobre todas as lacunas identificadas.

Primeiramente foi conduzida uma revisão na literatura, com base nos estudos selecionados, foi proposto um guia. Como primeiro passo foi criada uma estrutura baseada em tópicos e depois foi selecionada uma característica crítica para o teste de IoT de acordo com o mapeamento conduzido, a característica escolhida foi Desempenho. Após esse passo foi realizada uma outra revisão da literatura com foco na característica de Desempenho, a partir dessa revisão o guia foi instanciado para essa característica. Além disso, foi elaborado uma Wiki para automação na utilização do guia de desempenho. Para validar o guia, foram realizadas duas avaliações. A primeira avaliação foi conduzida com 7 especialistas em teste de software, IoT ou ambas as áreas, para avaliar a estrutura e o conteúdo do guia. A segunda avaliação foi um experimento controlado com 12 participantes, dividido em dois grupo, um que utilizou o guia e outro não. Os resultados de ambos as avaliações se mostraram positivas, evidenciando o auxílio do guia na execução dos testes de aplicações IoT e a sua estrutura mostrou facilitar o processo de teste por meio da sistematização fornecida pela abordagem proposta. Além disso, nos resultados do experimento foram observadas diminuição no esforço de execução e aumento na eficiência dos testes, bem como, maior detecção de falhas IoT.

9.1 Resultados Alcançados

São seis os principais resultados desta dissertação. O primeiro trata-se do *conjunto de desafios de teste relacionados com as características de qualidade relevantes para IoT* (i), em que no Capítulo 3 são listados e categorizados todos os desafios identificados na literatura durante a revisão. Foram esses desafios que geraram as lacunas apresentadas, dando origem a proposta

deste trabalho. Outro importante resultado desta dissertação refere-se ao próprio *modelo de um guia estruturado em tópicos, relevantes para o contexto de aplicações IoT* (ii), na Seção 5.1 é apresentada toda abordagem utilizada para elaboração do guia em que se baseia em 5 atividades desde a revisão da literatura até o refinamento da estrutura. Um terceiro resultado deste trabalho é a *a metodologia de instaciação do guia para uma característica IoT* (iii), descrita na Seção 6.1, a metodologia consiste em seis atividades que são baseadas na carcaterística selecionada, é feita a revisão da literatura, a extração e estruturação do dos dados, a observação de aplicações reais e o refinamento do guia. O *Guia de Teste de Desempenho para Aplicações IoT* (iv), é outro importante resultado deste trabalho, por meio da metodologia de instanciação fornecida o guia foi construído para a característica de desempenho, tendo em vista sua relevância. O guia possui todos os 11 tópicos preenchidos e contempla todas as subcaracterísticas de Desempenho - Comportamento Temporal, Utilização de Recursos e Capacidade. Este trabalho trouxe também como resultado a *WiKi para automação do guia* (v), um meio para fazer o uso do *Guia de Teste de Desempenho para Aplicações IoT*, de forma mais simples e rápida. Na Wiki há todos os tópicos do guia para manusear facilmente. Por fim como resultado, a pesquisa para a concepção do guia de desempenho, no que diz respeito a sua instanciação e o próprio guia em si, foi publicado na *International Conference on Enterprise Information Systems - ICEIS 2022* (vi), sendo inclusive indicado a *best paper* (CARVALHO *et al.*, 2022).

9.2 Trabalhos Futuros

Como trabalho futuros será feita a instanciação da abordagem para a característica de Interoperabilidade, tendo visto seus desafios em sistemas IoT. Também está previsto compilar essa pesquisa de uma maneira detalhada a fim realizar uma publicação em um *journal*. Com o avanço da pesquisa e novas instanciações sendo feitas é vislumbrado a possibilidade de expansão da WiKi para abranger as demais características. E por fim, outras pesquisas com o uso do guia de desempenho, também, estão sendo desenvolvidas na área de aprendizagem de máquina.

REFERÊNCIAS

- AHMAD, A.; BOUQUET, F.; FOURNERET, E.; GALL, F. L.; LEGEARD, B. Model-based testing as a service for iot platforms. In: SPRINGER. **Leveraging Applications of Formal Methods, Verification and Validation: Discussion, Dissemination, Applications: 7th International Symposium, ISoLA 2016, Imperial, Corfu, Greece, October 10-14, 2016, Proceedings, Part II** 7. Corfu, 2016. p. 727–742.
- ALI, N. B.; PETERSEN, K.; MÄNTYLÄ, M. Testing highly complex system of systems: an industrial case study. In: **Proceedings of the ACM-IEEE international symposium on Empirical software engineering and measurement**. Lund Sweden: ACM, 2012. p. 211–220. ISBN 9781450310567.
- ALIPOUR, M. A. Fault injection in the internet of things applications. In: **Proceedings of the 1st ACM SIGSOFT International Workshop on Testing Embedded and Cyber-Physical Systems**. Santa Barbara CA USA: ACM, 2017. p. 9–11. ISBN 9781450351126.
- ALMEIDA, R. L. A.; ANDRADE, R. M. C.; DARIN, T. G. R.; PAIVA, J. O. V. Chase: checklist to assess user experience in iot environments. In: **Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering: New Ideas and Emerging Results**. New York, NY, USA: Association for Computing Machinery, 2020. (ICSE-NIER '20), p. 41–44. ISBN 9781450371261.
- AMALFITANO, D.; AMATUCCI, N.; SIMONE, V. D.; RICCIO, V.; RITA, F. A. Towards a thing-in-the-loop approach for the verification and validation of iot systems. In: **Proceedings of the 1st ACM Workshop on the Internet of Safe Things**. New York, NY, USA: Association for Computing Machinery, 2017. (SafeThings'17), p. 57–63. ISBN 9781450355452.
- ANDRADE, R. M. C.; CARVALHO, R. M.; ARAÚJO, I. L. de; OLIVEIRA, K. M.; MAIA, M. E. F. What changes from ubiquitous computing to internet of things in interaction evaluation? In: **Distributed, Ambient and Pervasive Interactions: 5th International Conference, DAPI 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9–14, 2017, Proceedings**. Berlin, Heidelberg: Springer-Verlag, 2017. p. 3–21. ISBN 978-3-319-58696-0.
- BERA, P.; PASALA, A. A framework for optimizing effort in testing of system of systems. In: IEEE. **2012 Third International Conference on Services in Emerging Markets**. Massachusetts Ave., NW Washington, DC United States, 2012. p. 136–141.
- BOARD, I. S. T. Q. Standard glossary of terms used in software testing, . version 3.2, 2018. **ISTQB**, 01 2018.
- BRADY, S.; HAVA, A.; PERRY, P.; MURPHY, J.; MAGONI, D.; PORTILLO-DOMINGUEZ, A. O. Towards an emulated iot test environment for anomaly detection using nemu. **2017 Global Internet of Things Summit (GIoTTS)**, p. 1–6, 2017.
- BURATTI, C.; STAJKIC, A.; GARDASEVIC, G.; MILARDO, S.; ABRIGNANI, M. D.; MIJOVIC, S.; MORABITO, G.; VERDONE, R. Testing protocols for the internet of things on the euwin platform. **IEEE Internet of Things Journal**, IEEE, v. 3, n. 1, p. 124–133, 2015.
- BURES, M.; CERNY, T.; AHMED, B. S. Internet of things: Current challenges in the quality assurance and testing methods. In: KIM, K. J.; BAEK, N. (Ed.). **Information Science and Applications 2018**. Singapore: Springer Singapore, 2019. p. 625–634. ISBN 978-981-13-1056-0.

BYAMUKAMA, M.; NANNONO, J. N.; RUHINDA, K.; PEHRSON, B.; NSABAGWA, M.; AKOL, R.; OLSSON, R.; BAKKABULINDI, G.; KONDELA, E. Design guidelines for ultra-low power gateways in environment monitoring wireless sensor networks. In: **2017 IEEE AFRICON**. Cape Town: IEEE, 2017. p. 1472–1478. ISBN 9781538627754.

CARDOSO, G.; FILHO, F.; BARBOSA, R.; KOWAL, J. Business process management (bpm) and gut matrix in support to the information technology. **SSRN Electronic Journal**, 01 2021.

CARVALHO, D. F.; FERRARI, P.; FLAMMINI, A.; SISINNI, E. A test bench for evaluating communication delays in lorawan applications. In: IEEE. **2018 Workshop on Metrology for Industry 4.0 and IoT**. Brescia, Italy, 2018. p. 248–253.

CARVALHO, L. **Um Mapeamento Sistemático de Testes para Aplicações IoT**. 2018. Monografia (Bacharel em Engenharia de Software), UFC (Universidade Federal do Ceará), Russas, Brasil.

CARVALHO, L.; LELLI, V.; ANDRADE, R. Performance testing guide for iot applications:. In: **Proceedings of the 24th International Conference on Enterprise Information Systems**. Online Streaming: SCITEPRESS - Science and Technology Publications, 2022. p. 667–678. ISBN 9789897585692.

CARVALHO, R. M.; ANDRADE, R. M. d. C.; LELLI, V.; SILVA, E. G.; OLIVEIRA, K. Marçal de. What About Catalogs of Non-Functional Requirements? In: **REFSQ 2020. Joint Proceedings of REFSQ-2020 Workshops, Doctoral Symposium, Live Studies Track, and Poster Track co-located with the 26th International Conference on Requirements Engineering: Foundation for Software Quality (REFSQ 2020), CEUR Workshop Proceedings 2584**. Pisa, Italy, 2020.

CHEN, N.; VIHO, C. A passive interoperability testing approach applied to the constrained application protocol (coap). In: ICTSS 2012. **CFIP/NOTERE**. Berlin, Heidelberg, 2012.

CHEN, N.; VIHO, P. C.; BAIRE, A.; HUANG, X.; ZHA, J. Ensuring interoperability for the internet of things: Experience with coap protocol testing. **Automatika**, Taylor Francis, v. 54, n. 4, p. 448–458, 2013.

CLOUDTEST. **CloudTest: ferramenta de teste de carga em tempo real | Akamai**. 2022. Acesso em: 22 mar. 2022. Disponível em: <<https://www.akamai.com/pt/products/cloudtest>>.

CORTÉS, M.; SARAIVA, R.; SOUZA, M.; MELLO, P.; SOARES, P. Adoption of software testing in internet of things: A systematic literature mapping. In: **Proceedings of the IV Brazilian Symposium on Systematic and Automated Software Testing**. New York, NY, USA: Association for Computing Machinery, 2019. (SAST '19), p. 3–11. ISBN 9781450376488.

COSTANTINO, D.; MALAGNINI, G.; CARRERA, F.; RIZZARDI, A.; BOCCADORO, P.; SICARI, S.; GRIECO, L. A. Solving interoperability within the smart building: A real test-bed. In: IEEE. **2018 IEEE International Conference on Communications Workshops (ICC Workshops)**. Kansas City, MO, USA, 2018. p. 1–6.

DAHMAN, J.; LANE, J. A.; REBOVICH, G.; LOWRY, R. Systems of systems test and evaluation challenges. In: IEEE. **2010 5th International Conference on System of Systems Engineering**. Loughborough, UK, 2010. p. 1–6.

DOD DIRECTOR, S.; (ACQUISITION, S. E. D. U. S. of D.; (ACQUISITION, T. Technology); Office of the Under Secretary of D.; LOGISTICS). Systems engineering guide for systems of systems. **Department of Defense**, v. 1, n. 1, p. 148, 2008.

FAHMI, F. Z.; ABDUROHMAN, M. Performance testing of m2m middleware platform. In: IEEE. **2017 3rd International Conference on Science in Information Technology (ICSITech)**. Bandung, Indonesia, 2017. p. 378–382.

FARAHMANDPOUR, Z.; VERSTEEG, S.; HAN, J.; KAMESWARAN, A. Service virtualisation of internet-of-things devices: techniques and challenges. In: IEEE. **2017 IEEE/ACM 3rd International Workshop on Rapid Continuous Software Engineering (RCoSE)**. Buenos Aires, Argentina, 2017. p. 32–35.

FEHLMANN, T.; KRANICH, E. Autonomous real-time software systems testing. In: **Association for Computing Machinery**. New York, NY, USA: [s.n.], 2017. (IWSM Mensura '17), p. 54–63. ISBN 9781450348539.

FIENBERG, S. E.; LAZAR, N. William sealy gosset. In: _____. **Statisticians of the Centuries**. New York, NY: Springer New York, 2001. p. 312–317. ISBN 978-1-4613-0179-0.

FLEISS, J. L.; COHEN, J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. **Educational and Psychological Measurement**, v. 33, n. 3, p. 613–619, 1973.

FUCHS, S.; SCHMIDT, H.-P.; WITTE, S. Test and on-line monitoring of real-time ethernet with mixed physical layer for industry 4.0. In: IEEE. **2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)**. Berlin, Germany, 2016. p. 1–4.

GARCÍA, S. N. M.; HERNANDEZ-RAMOS, J. L.; SKARMETA, A. F. Test-based risk assessment and security certification proposal for the internet of things. In: IEEE. **2018 IEEE 4th World Forum on Internet of Things (WF-IoT)**. Singapore, 2018. p. 641–646.

GARRIDO-HIDALGO, C.; OLIVARES, T.; RAMIREZ, F. J.; RODA-SANCHEZ, L. An end-to-end internet of things solution for reverse supply chain management in industry 4.0. **Computers in Industry**, v. 112, p. 103127, 2019. ISSN 0166-3615.

GEIGER, C.; HEADRICK, W. J. The thingosity of automated test systems in the internet of things. In: IEEE. **2017 IEEE AUTOTESTCON**. Schaumburg, IL, USA, 2017. p. 1–4.

GIMÉNEZ, P.; MOLÍNA, B.; PALAU, C. E.; ESTEVE, M. SWE simulation and testing for the iot. In: IEEE. **2013 IEEE International Conference on Systems, Man, and Cybernetics**. Manchester, UK, 2013. p. 356–361.

GIUSTO, D.; IERA, A.; MORABITO, G.; ATZORI, L. **The Internet of Things**. New York: Springer-Verlag, 2010.

GUIDE for writing standards taking into account the needs of micro, small and medium-sized enterprises. Geneva, CH, 2016. v. 1.

GURIJALA, Y. R. **Performance Testing Internet of Things**. 2018. Accessed: 02 jan. 2020. Disponível em: <<https://www.infosys.com/de/documents/performance-testing-iot.pdf>>.

HAMAD, R. M. H. Automation testing and monitoring lab on the cloud for iot smart fleet system (atml sfs). In: **Proceedings of the Fourth International Conference on Engineering MIS 2018**. New York, NY, USA: Association for Computing Machinery, 2018. (ICEMIS '18). ISBN 9781450363921.

HASSAN, Q.; KHAN, A.; MADANI, S. **Internet of Things: Challenges, Advances, and Applications**. New York: CRC Press, 2017. (Chapman & Hall/CRC Computer and Information Science Series). ISBN 9781351651059.

INTERNATIONAL, R. . S. **Testes de dispositivos de Internet das Coisas (IoT)**. 2014. Accessed: 22 mar. 2021. Disponível em: <https://www.rohde-schwarz.com/br/solucoes/test-and-measurement/wireless-communication/iot-m2m/iot-m2m-device-testing/tema-testes-de-dispositivos-m2m-iot_233892.html>.

IOTIFY. **IoTIFY – Cloud native IoT Simulator – IoT Functional Performance testing platform**. 2022. Accessed: 17 oct. 2022. Disponível em: <<https://iotify.io/>>.

ISO/IEC. **Standardization and related activities - General vocabulary**. 8. ed. Geneva, 2004. Provides general terms and definitions concerning standardization and related activities.

ISO/IEC 25010. **ISO/IEC 25010:2011, Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models**. 2011.

JAIN, R. **The art of computer systems performance analysis - techniques for experimental design, measurement, simulation, and modeling**. New York: Wiley, 1991. I-XXVII, 1-685 p. (Wiley professional computing). ISBN 978-0-471-50336-1.

JEANNOTTE, B.; TEKEOGLU, A. Artorias: Iot security testing framework. In: IEEE. **2019 26th International Conference on Telecommunications (ICT)**. Hanoi, Vietnam, 2019. p. 233–237.

KANSTRÉN, T.; MÄKELÄ, J.; KARHULA, P. Architectures and experiences in testing iot communications. In: IEEE. **2018 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)**. Västerås, Sweden, 2018. p. 98–103.

KIM, E. E.; ZIEGLER, S. Towards an open framework of online interoperability and performance tests for the internet of things. In: IEEE. **2017 Global Internet of Things Summit (GloTS)**. Geneva, Switzerland, 2017. p. 1–6.

KIM, H.; AHMAD, A.; HWANG, J.; BAQA, H.; GALL, F. L.; ORTEGA, M. A. R.; SONG, J. Iot-taas: Towards a prospective iot testing framework. **IEEE Access**, v. 6, p. 15480–15493, 2018.

KIM, J.; YUN, J.; CHOI, S.-C.; SEED, D. N.; LU, G.; BAUER, M.; AL-HEZMI, A.; CAMPOWSKY, K.; SONG, J. Standard-based iot platforms interworking: implementation, experiences, and lessons learned. **IEEE Communications Magazine**, IEEE, v. 54, n. 7, p. 48–54, 2016.

KITCHENHAM, B.; BRERETON, O. P.; BUDGEN, D.; TURNER, M.; BAILEY, J.; LINKMAN, S. Systematic literature reviews in software engineering—a systematic literature review. **Information and software technology**, Elsevier, v. 51, n. 1, p. 7–15, 2009.

KOUCHE, A. E. Towards a wireless sensor network platform for the internet of things: Sprouts wsn platform. In: IEEE. **2012 IEEE International Conference on Communications (ICC)**. Ottawa, ON, Canada, 2012. p. 632–636.

KOVACS, E.; BAUER, M.; KIM, J.; YUN, J.; GALL, F. L.; ZHAO, M. Standards-based worldwide semantic interoperability for iot. **IEEE Communications Magazine**, IEEE, v. 54, n. 12, p. 40–46, 2016.

KRISTOFFERSEN, M.; LÖTBÄCK, C. P.; SKOUSEN, D. Reverberation chambers for flexible over-the-air testing of internet-of-things devices. In: IEEE. **2017 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting**. San Diego, CA, USA, 2017. p. 1937–1938.

LANGSARI, K.; ROCHIMAH, S.; AKBAR, R. J. Measuring performance efficiency of application applying design patterns and refactoring method. **IPTEK Journal of Proceedings Series**, n. 1, p. 149–155, 2018.

LANZA, J.; SANCHEZ, L.; SANTANA, J. R.; AGARWAL, R.; KEFALAKIS, N.; GRACE, P.; ELSALEH, T.; ZHAO, M.; TRAGOS, E.; NGUYEN, H. *et al.* Experimentation as a service over semantically interoperable internet of things testbeds. **IEEE Access**, IEEE, v. 6, p. 51607–51625, 2018.

LEAL, A. G.; SANTIAGO, A.; MIYAKE, M. Y.; NODA, M. K.; PEREIRA, M. J.; AVANÇO, L. Integrated environment for testing iot and rfid technologies applied on intelligent transportation system in brazilian scenarios. In: IEEE. **2014 IEEE Brasil RFID**. Sao Paulo, Brazil, 2014. p. 22–24.

LIU, Y.; YANG, Y.; LV, X.; WANG, L. A self-learning sensor fault detection framework for industry monitoring iot. **Mathematical Problems in Engineering**, v. 2013, 10 2013.

LOADUI. **Ensure High Performance APIs in Less Time | SoapUI**. 2022. Accessed: 17 oct 2022. Disponível em: <<https://www.soapui.org/professional/loadui-pro/>>.

LOWENSTEIN, D.; SLATER, C. Management of test utilization, optimization, and health through real-time data. In: IEEE. **2018 IEEE AUTOTESTCON**. MD, USA, 2018. p. 1–6.

LU, C.; STANKOVIC, J. A.; ABDELZAHER, T. F.; TAO, G.; SON, S. H.; MARLEY, M. Performance specifications and metrics for adaptive real-time systems. In: IEEE. **Proceedings 21st IEEE Real-Time Systems Symposium**. Orlando, FL, USA, 2000. p. 13–23.

MAIA, P.; CAVALCANTE, E.; GOMES, P.; BATISTA, T.; DELICATO, F. C.; PIRES, P. F. On the development of systems-of-systems based on the internet of things: A systematic mapping. In: ECSAW '14. **Proceedings of the 2014 European Conference on Software Architecture Workshops**. Viena, Áustria, 2014. p. 1–8.

MAIA, R.; ANDRADE, R.; OLIVEIRA, K.; SANTOS, I. D. S.; BEZERRA, C. I. M. Quality characteristics and measures for human–computer interaction evaluation in ubiquitous systems. **Software Quality Journal**, 07 2016.

MARINISSEN, E. J.; ZORIAN, Y.; KONIJNENBURG, M.; HUANG, C.-T.; HSIEH, P.-H.; COCKBURN, P.; DELVAUX, J.; ROŽIĆ, V.; YANG, B.; SINGELÉE, D. *et al.* Iot: Source of test challenges. In: IEEE. **2016 21th IEEE European test symposium (ETS)**. Amsterdam, Netherlands, 2016. p. 1–10.

- MATEEN, Q.; SIRSHAR, M. *et al.* Software quality assurance in internet of things. **Int. J. Comput. Appl.**, v. 109, n. 9, p. 16–24, 2015.
- MATHEU-GARCÍA, S. N.; HERNÁNDEZ-RAMOS, J. L.; SKARMETA, A. F.; BALDINI, G. Risk-based automated assessment and testing for the cybersecurity certification and labelling of iot devices. **Computer Standards Interfaces**, v. 62, p. 64 – 83, 2019. ISSN 0920-5489.
- MEIER, J.; FARRE, C.; BANSODE, P.; BARBER, S.; REA, D. **Performance Testing Guidance for Web Applications: Patterns Practices**. USA: Microsoft Press, 2007. ISBN 9780735625709.
- MURAD, G.; BADARNEH, A.; QUSEF, A.; ALMASALHA, F. Software testing techniques in iot. In: IEEE. **2018 8th International conference on computer science and information technology (CSIT)**. Amman, Jordan, 2018. p. 17–21.
- MYERS, G. J.; SANDLER, C.; BADGETT, T. **The art of software testing**. United States: John Wiley & Sons, 2011.
- NAIR, K. K.; DUBE, E.; LEFOPHANE, S. Modelling an iot testbed in context with the security vulnerabilities of south africa. In: IEEE. **2017 3rd IEEE International Conference on Computer and Communications (ICCC)**. Chengdu, China, 2017. p. 244–248.
- NAIR, T. G. **Advanced Real Time Systems**. USA: IGI Global, 2015. ISBN 10.4018/978-1-4666-5888-2.ch689.
- NEOTYS. **The Load Testing Platform Accelerating DevOps | NeoLoad by Neotys**. 2022. Accessed: 17 oct. 2022. Disponível em: <<https://www.neotys.com/>>.
- NEVES, V. D. O.; BERTOLINO, A.; ANGELIS, G. D.; GARCÉS, L. Do we need new strategies for testing systems-of-systems? In: **Proceedings of the 6th International Workshop on Software Engineering for Systems-of-Systems**. New York, NY, USA: Association for Computing Machinery, 2018. (SESoS '18), p. 29–32. ISBN 9781450357470.
- NORBERTO, M.; GAEDICKE, L.; BERNARDINO, M.; LEGRAMANTE, G.; BASSO, F. P.; RODRIGUES, E. M. Performance testing in mobile application: A systematic literature map. In: **Proceedings of the XVIII Brazilian Symposium on Software Quality**. New York, NY, USA: Association for Computing Machinery, 2019. (SBQS'19), p. 99–108. ISBN 9781450372824.
- PATEL, K.; PATEL, S.; SCHOLAR, P.; SALAZAR, C. **Internet of Things-IOT: Definition, Characteristics, Architecture, Enabling Technologies, Application Future Challenges**. 2016.
- PONTES, P. M.; LIMA, B.; FARIA, J. a. P. Izinto: A pattern-based iot testing framework. In: **Companion Proceedings for the ISSSTA/ECOOP 2018 Workshops**. New York, NY, USA: Association for Computing Machinery, 2018. (ISSSTA '18), p. 125–131. ISBN 9781450359399.
- PONTES, P. M.; LIMA, B.; FARIA, J. a. P. Test patterns for iot. In: **Proceedings of the 9th ACM SIGSOFT International Workshop on Automating TEST Case Design, Selection, and Evaluation**. New York, NY, USA: Association for Computing Machinery, 2018. (A-TEST 2018), p. 63–66. ISBN 9781450360531.
- RAMPRASAD, B.; MUKHERJEE, J.; LITOIU, M. A smart testing framework for iot applications. In: IEEE. **2018 IEEE/ACM International Conference on Utility and Cloud Computing Companion (UCC Companion)**. Zurich, Switzerland, 2018. p. 252–257.

RAZZAQUE, M. A.; MILOJEVIC-JEVRIĆ, M.; PALADE, A.; CLARKE, S. Middleware for internet of things: a survey. **IEEE Internet of things journal**, IEEE, v. 3, n. 1, p. 70–95, 2015.

REETZ, E. S.; KUEMPER, D.; MOESSNER, K.; TÖNJES, R. How to test iot-based services before deploying them into real world. In: VDE. **European Wireless 2013; 19th European Wireless Conference**. Guildford, UK, 2013. p. 1–6.

ROSA, F. D. F.; JINO, M.; BUENO, P. M. S.; BONACIN, R. Coverage-based heuristics for selecting assessment items from security standards: a core set proposal. In: IEEE. **2018 Workshop on Metrology for Industry 4.0 and IoT**. Brescia, Italy, 2018. p. 192–197.

ROSENKRANZ, P.; WÄHLISCH, M.; BACCELLI, E.; ORTMANN, L. A distributed test system architecture for open-source iot software. In: **IoT-Sys '15: Proceedings of the 2015 Workshop on IoT challenges in Mobile and Industrial Systems**. New York, NY, USA: [s.n.], 2015. p. 43–48.

SACHIDANANDA, V.; SIBONI, S.; SHABTAI, A.; TOH, J.; BHAIKAV, S.; ELOVICI, Y. Let the cat out of the bag: A holistic approach towards security analysis of the internet of things. In: **Proceedings of the 3rd ACM International Workshop on IoT Privacy, Trust, and Security**. New York, NY, USA: Association for Computing Machinery, 2017. (IoTPTS '17), p. 3–10. ISBN 9781450349697.

SAND, B. Iot testing-the big challenge why, what and how. In: SPRINGER. **International Internet of Things Summit**. Rome, Italy, 2015. p. 70–76.

SANTOS, R. M.; OLIVEIRA, K. M. de; ANDRADE, R. M.; SANTOS, I. S.; LIMA, E. R. A quality model for human-computer interaction evaluation in ubiquitous systems. In: SPRINGER. **Latin American conference on human computer interaction**. Carrillo, Costa Rica, 2013. p. 63–70.

SCHIEFERDECKER, I.; KRETZSCHMANN, S.; RENNOCH, A.; WAGNER, M. Iot-testware - an eclipse project. In: IEEE. **2017 IEEE International Conference on Software Quality, Reliability and Security (QRS)**. Prague, Czech Republic, 2017. p. 1–8.

SOARES, L.; POTENA, P.; MACHADO, I.; CRNKOVIC, I.; ALMEIDA, E. Analysis of non-functional properties in software product lines: A systematic review. In: IEEE. **Proceedings - 40th Euromicro Conference Series on Software Engineering and Advanced Applications, SEAA 2014**. Verona, Italy, 2014.

SOMMERVILLE, I. Engenharia de software-8ª edição (2007). **Ed Person Education**, 2011.

TCPDUMP. **Home | TCPDUMP LIBPCAP**. 2010. Accessed: 17 oct. 2022. Disponível em: <<http://www.tcpdump.org/>>.

THIMMANNAGARI, C. **CPU design: answers to frequently asked questions**. New York: Springer Science & Business Media, 2004.

UCHÔA, A. G.; LIMA, L. P.; BEZERRA, C. I. M.; MONTEIRO, J. M.; ANDRADE, R. M. C. Dymmer-nfp: Modeling non-functional properties and multiple context adaptation scenarios in software product lines. In: BOTTERWECK, G.; WERNER, C. (Ed.). **Mastering Scale and Complexity in Software Reuse**. Cham: Springer International Publishing, 2017. p. 175–183.

VISOOTTIVISETH, V.; AKARASIRIWONG, P.; CHAIYASART, S.; CHOTIVATUNYU, S. Pentos: Penetration testing tool for internet of thing devices. In: IEEE. **TENCON 2017-2017 IEEE Region 10 Conference**. Penang, Malaysia, 2017. p. 2279–2284.

WESSEL, J. T.; MEYER, B. L. Assessing system software performance in complex system of systems environments. In: IEEE. **2010-MILCOM 2010 MILITARY COMMUNICATIONS CONFERENCE**. San Jose, CA, USA, 2010. p. 2310–2315.

WIRESHARK. **Wireshark**. 2016. Accessed: 22 jan. 2021. Disponível em: <<https://www.wireshark.org/>>.

WOHLIN, C.; RUNESON, P.; HÖST, M.; OHLSSON, M. C.; REGNELL, B.; WESSLÉN, A. **Experimentation in software engineering**. Berlin, Heidelberg: Springer Science & Business Media, 2012.

WU, J.-J.; CHIU, M.-C. An innovative interface design and customized usability testing method: Case study of internet of things integration platform interface. In: IOS PRESS. **Transdisciplinary Engineering: A Paradigm Shift: Proceedings of the 24th ISPE Inc. International Conference on Transdisciplinary Engineering, July 10-14, 2017**. [S.l.], 2017. v. 5, p. 365.

YAN, C.; FENG, K.; ZHANG, Z. Impact of internet of things on security grade evaluation. In: IEEE. **2012 International Conference on Computer Science and Electronics Engineering**. Hangzhou, China, 2012.

ZHANG, K.; LENG, S.; HE, Y.; MAHARJAN, S.; ZHANG, Y. Mobile edge computing and networking for green and low-latency internet of things. **IEEE Communications Magazine**, v. 56, p. 39–45, 2018.

ZHANG, Y.; CHAKRABARTY, K. Dynamic adaptation for fault tolerance and power management in embedded real-time systems. **ACM Transactions on Embedded Computing Systems (TECS)**, ACM New York, NY, USA, v. 3, n. 2, p. 336–360, 2004.

APÊNDICE A – PROTOCOLO PARA REVISÃO DA CONSTRUÇÃO DA ESTRUTURA DO GUIA

Objetivo: Elencar catálogos, checklists, guidelines que tenham relações com teste de IoT

| PICO | |
|--------------------|---|
| População | <i>Internet of Things</i> |
| Intervenção | <i>Test</i> |
| Comparação | Não aplicável |
| Resultado | <i>Guideline</i> |
| Contexto | Dada a dificuldade dos testes em IoT, é viável a criação de um guia que dê suporte a essa atividade |

| Termos relacionados | |
|---------------------------|---|
| <i>Internet of Things</i> | - |
| <i>Test</i> | <i>Testing</i> |
| <i>Guideline</i> | <i>Standard, Catalog, Checklist, Guide and Guidance</i> |

| Questões de Pesquisa |
|---|
| Existem estratégias documentadas para o teste de aplicações IoT? |
| Como são documentadas as estratégias de teste de aplicações IoT? |
| Quais são os principais elementos estruturais dessas estratégias? |
| Quais são as características de IoT que são contempladas por essas estratégias? |

| String de Busca |
|--|
| <i>("internet of things") AND (test OR testing) AND (guideline OR standard OR catalog OR checklist OR guide OR guidance)</i> |

| String de Busca |
|-----------------|
| <i>Scopus</i> |

| Critério de Seleção | |
|-----------------------|---|
| Critérios de Inclusão | Estudos escritos em inglês |
| | Estudos que apresentam guias, checklists ou catálogos para testes IoT |
| Critérios de Exclusão | Não é possível acessar o estudo |
| | O estudo é indexado, tutorial ou <i>keynote</i> |
| | O estudo não possui guias, checklists ou catálogos para testes IoT |
| | O estudo não é de Computação ou áreas relacionadas |
| | Para estudos com versão já publicada, é selecionada a mais completa |

**APÊNDICE B – RESULTADO DA REVISÃO PARA CONSTRUÇÃO DA
ESTRUTURA DO GUIA**

| Id | Título | Autores | Veículo | Ano |
|---|--|--|--|------------|
| Estudos selecionados a partir da revisão | | | | |
| 1 | <i>Integrated environment for testing lot and RFID technologies applied on intelligent transportation system in Brazilian scenarios</i> | <i>Leal, A.G. and Santiago, A. and Miyake, M.Y. and Noda, M.K. and Pereira, M.J. and Avanyo, L.</i> | <i>2014 IEEE Brazil RFID</i> | 2014 |
| 2 | <i>Testing protocols for the internet of things on the EuWin platform</i> | <i>Buratti, C. and Stajkic, A. and Gardasevic, G. and Milardo, S. and Abrignani, M.D. and Mijovic, S. and Morabito, G. and Verdone, R.</i> | <i>IEEE Internet of Things Journal</i> | 2016 |
| 3 | <i>Standard-based IoT platforms interworking: Implementation, experiences, and lessons learned</i> | <i>Kim, J. and Yun, J. and Choi, S.-C. and Seed, D.N. and Lu, G. and Bauer, M. and Al-Hezmi, A. and Campowsky, K. and Song, J.</i> | <i>IEEE Communications Magazine</i> | 2016 |
| 4 | <i>Standards-Based Worldwide Semantic Interoperability for IoT</i> | <i>Kovacs, E. and Bauer, M. and Kim, J. and Yun, J. and Le Gall, F. and Zhao, M.</i> | <i>IEEE Communications Magazine</i> | 2016 |

Continua na próxima página

Continuação da tabela

| Id | Título | Autores | Veículo | Ano |
|-----------|---|--|--|------------|
| 5 | <i>Test and on-line monitoring of real-time Ethernet with mixed physical layer for Industry 4.0</i> | <i>Fuchs, S. and Schmidt, H.-P. and Witte, S.</i> | <i>IEEE International Conference on Emerging Technologies and Factory Automation, ETFA</i> | 2016 |
| 6 | <i>Design guidelines for ultra-low power gateways in environment monitoring wireless sensor networks</i> | <i>Byamukama, M. and Nannono, J.N. and Ruhinda K. and Pehrson, B. and Nsabagwa, M. and Akol, R. and Olsson, R. and Bakkabulindi, G and Kondela, E.</i> | <i>2017 IEEE AFRICON: Science, Technology and Innovation for Africa, AFRICON 2017</i> | 2017 |
| 7 | <i>The thingosity of automated test systems in the Internet of Things</i> | <i>Geiger, C. and Headrick, W.J.</i> | <i>AUTOTESTCON (Proceedings)</i> | 2017 |
| 8 | <i>An innovative interface design and customized usability testing method: Case study of internet of things integration platform interface</i> | <i>Wu, J.-J. and Chiu, M.-C.</i> | <i>ISPE Inc. International Conference on Transdisciplinary Engineering</i> | 2017 |

Continua na próxima página

Continuação da tabela

| Id | Título | Autores | Veículo | Ano |
|-----------|---|--|---|------------|
| 9 | <i>Performance testing of M2M middleware platform</i> | <i>Fahmi, F.Z. and Abdurohman, M.</i> | <i>Proceeding - 2017 3rd International Conference on Science in Information Technology: Theory and Application of IT for Education, Industry and Society in Big Data Era, ICSITech 2017</i> | 2017 |
| 10 | <i>Autonomous real-time software & systems testing</i> | <i>Fehlmann, T. and Kranich, E.</i> | <i>ACM International Conference Proceeding Series</i> | 2017 |
| 11 | <i>Solving interoperability within the smart building: A real test-bed</i> | <i>Costantino, D. and Malagnini, G. and Carrera, F. and Rizzardi, A. and Boccadoro, P. and Sicari, S. and Grieco, L.A.</i> | <i>2018 IEEE International Conference on Communications Workshops, ICC Workshops 2018 - Proceedings</i> | 2018 |
| 12 | <i>Management of Test Utilization, Optimization, and Health through Real-Time Data</i> | <i>Lowenstein, D. and Slater, C.</i> | <i>AUTOTESTCON (Proceedings)</i> | 2018 |

Continua na próxima página

Continuação da tabela

| Id | Título | Autores | Veículo | Ano |
|--|---|--|--|------------|
| 13 | <i>Izinto: A pattern-based IoT testing framework</i> | <i>Pontes, P.M. and Lima, B. and Faria, J.P.</i> | <i>Companion Proceedings for the ISSTA/ ECOOP 2018 Workshops</i> | 2018 |
| Estudos obtidos a partir de outras fontes | | | | |
| 14 | <i>Guide 2:2004 Standardization and related activities — General vocabulary</i> | <i>International Organization for Standardization</i> | - | 2004 |
| 15 | <i>A Quality Model for Human-Computer Interaction Evaluation in Ubiquitous Systems.</i> | <i>Rainara M. Santos, Káthia M. Oliveira, Rossana M. Andrade, Ismayle S. Santos, and Edmilson R. Lima.</i> | <i>Proceedings of the 6th Latin American Conference on Human Computer Interaction - Volume 8278 (CLIHC 2013)</i> | 2013 |
| 16 | <i>Guide 17:2016 Guide for writing standards talking into account the needs of micro, small and medium-sized enterprises</i> | <i>International Organization for Standardization</i> | - | 2016 |

Continua na próxima página

Continuação da tabela

| Id | Título | Autores | Veículo | Ano |
|-----------|--|---|---|------------|
| 17 | <i>Model-Based Testing as a Service for IoT Platforms</i> | <i>Ahmad, Abbas, Bouquet, Fabrice, Fournieret, Elizabeta, Le Gall, Franck, Legeard, Bruno, Margaria, Tiziana, Steffen, Bernhard</i> | <i>Leveraging Applications of Formal Methods, Verification and Validation: Discussion, Dissemination , Applications - Springer International Publishing</i> | 2016 |
| 18 | <i>Quality characteristics and measures for human-computer interaction evaluation in ubiquitous systems</i> | <i>Carvalho, R.M., de Castro Andrade, R.M., de Oliveira, K.M. et al.</i> | <i>Software Quality Journal 25.3 (2017): 743-795.</i> | 2017 |
| 19 | <i>Risk-based automated assessment and testing for the cybersecurity certification and labelling of IoT devices</i> | <i>Sara N. Matheu- García, José L. Hernández-Ramos, Antonio F. Skarmeta, Gianmarco Baldini</i> | <i>Computer Standards & Interfaces 62 (2019): 64-83</i> | 2019 |
| 20 | <i>"What About Catalogs of Non-Functional Requirements?"</i> | <i>Carvalho, Rainara. M, Andrade, Rossana, Lelli, Valéria, Silva, Erika, and Oliveira, Káthia.</i> | <i>Proceedings of REFSQ-2020 Workshops, Doctoral Symposium, Live Studies Track, and Poster Track</i> | 2020 |

Fim da tabela

APÊNDICE C – FORMULÁRIO DE EXTRAÇÃO DE DADOS

| Item | Description | Note |
|---|-------------|------|
| Year | | |
| Author | | |
| Title | | |
| Source (ACM, Compendex, Scopus, IEEE, ...) | | |
| Journal/Conference name | | |
| Characteristics? | | |
| Guidelines? Guides? | | |
| Standarts? Process? | | |
| Catalogs? | | |
| Lessons learned? | | |
| Others? | | |
| Study Quality | | |
| Are the aim(s) of study clearly stated?(Yes/No) | | |
| Is there any discussion about the results of the study?(Yes/No/Partially) | | |

APÊNDICE D – PROTOCOLO PARA REVISÃO DE DESEMPENHO

Objetivo: Elencar as soluções aplicadas nos testes IoT no âmbito de desempenho

| PICO | |
|--------------------|--|
| População | <i>Internet of Things</i> |
| Intervenção | <i>Performance Test</i> |
| Comparação | Não aplicável |
| Resultado | <i>Standard</i> |
| Contexto | Testes de Desempenho para aplicações IoT |

| Termos relacionados | |
|---------------------------|--|
| <i>Internet of Things</i> | - |
| <i>Performance Test</i> | <i>Performance Testing, Load Testing, Load Test, Stress Testing, Stress Test, Workload Testing and Workload Test</i> |
| <i>Standard</i> | <i>Model, Catalog, Metric, Measure, Tool, Property, Test Case and Correlation</i> |

| Questões de Pesquisa |
|--|
| Quais sistematizações são utilizadas para o teste de desempenho em aplicações IoT? |
| Quais propriedades e características rodeiam o âmbito de desempenho em aplicações IoT? |
| Quais métricas e ferramentas são utilizadas para avaliar o desempenho de aplicações IoT? |

| String de Busca |
|---|
| <i>("internet of things") AND ("performance test"OR "performance testing"OR "load testing"OR "load test"OR "stress testing"OR "stress test"OR "workload testing"OR "workload test") AND (standard OR model OR catalog OR metric OR measure OR tool OR property OR "test case" OR correlation)</i> |

| Base de Dados |
|---|
| <i>Scopus, IEEE, ACM e Science Direct</i> |

| Critérios de Seleção | |
|-----------------------|---|
| Critérios de Inclusão | Estudos escritos em inglês |
| | Estudos que apresentem soluções para o teste de desempenho em aplicações IoT |
| | Estudos que apresentem métricas ou ferramentas ou propriedades ou características relacionadas ao teste de desempenho em aplicações IoT |
| Critérios de Exclusão | Não é possível acessar o estudo |
| | O estudo é indexado, tutorial ou <i>keynote</i> |
| | Estudo não apresenta uma solução para o teste de desempenho em aplicações IoT |
| | O estudo não é de Computação ou áreas relacionadas |
| | Para estudos com versão já publicada, é selecionada a mais completa |

APÊNDICE E – RESULTADOS DA REVISÃO DE DESEMPENHO

| Id | Título | Autores | Veículo | Ano |
|---|--|--|---|------------|
| Estudos selecionados a partir da revisão | | | | |
| 1 | <i>Testing of Sensor Observation Services: A Performance Evaluation</i> | <i>Poorazizi, Mohammad Ebrahim and Liang, Steve H. L. and Hunter, Andrew J. S.</i> | <i>SWE '12: Proceedings of the First ACM SIGSPATIAL Workshop on Sensor Web Enablement</i> | 2012 |
| 2 | <i>A Passive Testing Approach for Protocols in Internet of Things</i> | <i>X. Che and S. Maag</i> | <i>IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber Physical and Social Computing</i> | 2013 |
| 3 | <i>Conception of ID Layer Performance at the Network Level for Internet of Things</i> | <i>Kovacs, E. and Bauer, M. and Kim, J. and Yun, J. and Le Gall, F. and Zhao, M.</i> | <i>Personal Ubiquitous Comput.</i> | 2016 |

Continua na próxima página

Continuação da tabela

| Id | Título | Autores | Veículo | Ano |
|-----------|---|---|---|------------|
| 4 | <i>D-P2P-Sim+: A novel distributed framework for P2P protocols performance testing</i> | <i>S. Sioutas and E. Sakkopoulos and A. Panaretos and D. Tsoumakos and P. Gerolymatos and G. Tzimas and Y. Manolopoulos</i> | <i>Journal of Systems and Software</i> | 2015 |
| 5 | <i>Performance management of IoT/M2M platforms</i> | <i>Kuo, H.C. and Lin, F.J.</i> | <i>2016 IEEE 6th International Conference on Communications and Electronics, IEEE ICCE 2016</i> | 2016 |
| 6 | <i>Measurements, performance and analysis of LoRa FABIAN, a real-world implementation of LPWAN</i> | <i>T. Petrić and M. Goessens and L. Nuaymi and L. Toutain and A. Pelov</i> | <i>2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio</i> | 2016 |
| 7 | <i>Performance Prediction of IoT Application: An Experimental Analysis</i> | <i>Duttagupta, Subhasri and Kumar, Mukund and Ranjan, Ritesh and Nambiar, Manoj</i> | <i>IoT'16: Proceedings of the 6th International Conference on the Internet of Things</i> | 2016 |

Continua na próxima página

Continuação da tabela

| Id | Título | Autores | Veículo | Ano |
|-----------|--|---|---|------------|
| 8 | <i>Good Vibes: The Impact of Haptic Patterns on Stress Levels</i> | <i>Kelling, Chelsea and Pitaro, Daniella and Rantala, Jussi</i> | <i>The 20th International Academic Mindtrek Conference</i> | 2016 |
| 9 | <i>Towards a Thing-In-the-Loop Approach for the Verification and Validation of IoT Systems</i> | <i>Amalfitano, Domenico and Amatucci, Nicola and De Simone, Vincenzo and Riccio, Vincenzo and Rita, Fasolino Anna</i> | <i>SafeThings'17: Proceedings of the 1st ACM Workshop on the Internet of Safe Things</i> | 2017 |
| 10 | <i>Performance evaluation of IoT mesh networking technology in ISM frequency band</i> | <i>Vondrous, O. and Kocur Z. and Hegr, T. and Slavicek, O.</i> | <i>Proceedings of the 2016 17th International Conference on Mechatronics Mechatronics, ME 2016</i> | 2017 |
| 11 | <i>F-Interop - Online platform of interoperability and performance tests for the Internet of Things</i> | <i>Ziegler, S. and Baron, L. and Vermeulen, B. and Fdida, S.</i> | <i>Building the Future Internet through FIRE: 2016 FIRE Book: a Research and Experimentation based Approach English</i> | 2017 |

Continua na próxima página

Continuação da tabela

| Id | Título | Autores | Veículo | Ano |
|-----------|---|--|---|------------|
| 12 | <i>Performance testing of M2M middleware platform</i> | <i>Fahmi, F.Z. and Abdurohman, M.</i> | <i>Proceeding - 2017 3rd International Conference on Science in Information Technology: Theory and Application of IT for Education, Industry and Society in Big Data Era, ICSITech 2017</i> | 2017 |
| 13 | <i>Towards an open framework of online interoperability and performance tests for the Internet of Things</i> | <i>Kim, E.E. and Ziegler, S.</i> | <i>GIoTS 2017 - Global Internet of Things Summit, Proceedings</i> | 2017 |
| 14 | <i>Theoretical analysis and performance testing of LoRa technology</i> | <i>Sun, Y. and Hu, J. and Liu, Y. and Tian, Z.</i> | <i>Proceedings - 2017 International Conference on Computer Technology, Electronics and Communication, ICCTEC 2017</i> | 2017 |

Continua na próxima página

Continuação da tabela

| Id | Título | Autores | Veículo | Ano |
|-----------|--|--|---|------------|
| 15 | <i>IoT Sensor Based Mobility Performance Test-Bed for Disaster Response Robots</i> | <i>Y. Kim and S. Jung and D. Gu and H. Kim and C.Song</i> | <i>2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)</i> | 2017 |
| 16 | <i>In-Test Adaptation of Workload in Enterprise Application Performance Testing</i> | <i>Kaczmariski, Maciej and Perry, Philip and Murphy John and Portillo-Dominguez, A. Omar</i> | <i>ICPE '17 Companion: Proceedings of the 8th ACM/SPEC on International Conference on Performance Engineering Companion</i> | 2017 |
| 17 | <i>Performance testing of an internet of things platform</i> | <i>Esquiagola, J. and Costa, L. and Calcina P. and Fedrecheski, G. and Zuffo, M.</i> | <i>IoTBDs 2017 - Proceedings of the 2nd International Conference on Internet of Things, Big Data and Security</i> | 2017 |

Continua na próxima página

Continuação da tabela

| Id | Título | Autores | Veículo | Ano |
|-----------|---|--|--|------------|
| 18 | <i>Performance Evaluation of IoT Network Management Platforms</i> | <i>De C. Silva, J. and Pereira, P.H.M. and De Souza, L.L. and Marins, C.N.M. and Marcondes, G.A.B. and Rodrigues, J.J.P.C.</i> | <i>2018 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2018</i> | 2018 |
| 19 | <i>Research On Performance Testing For Urban Scenario Based On Terminal Equipment Of LTE-V2X Vehicle Network</i> | <i>Y. Lv and Y. Wang and X. Liu and R. Xu and J. Fang and X. Peng</i> | <i>2018 14th IEEE International Conference on Signal Processing (ICSP)</i> | 2018 |
| 20 | <i>The platform prototype for testing heterogeneous networks in IoT environment</i> | <i>Shmatkov, V. and Spynu, S. and Pimenov, V.</i> | <i>CEUR Workshop Proceedings</i> | 2019 |
| 21 | <i>OPC UA versus ROS, DDS, and MQTT: Performance evaluation of industry 4.0 protocols</i> | <i>Profanter, S. and Tekat, A. and Dorofeev, K. and Rickert, M. and Knoll, A.</i> | <i>Proceedings of the IEEE International Conference on Industrial Technology</i> | 2019 |
| 22 | <i>OTA Measurement for IoT Wireless Device Performance Evaluation: Challenges and Solutions</i> | <i>Shen, P. and Qi, Y. and Yu, W. and Fan, J. and Li, F.</i> | <i>IEEE Internet of Things Journal</i> | 2019 |

Continua na próxima página

Continuação da tabela

| Id | Título | Autores | Veículo | Ano |
|-----------|--|--|--|------------|
| 23 | <i>Research on LoRa Communication Performance in Manhole Cover Monitoring</i> | <i>L. Xuan and Z. Hesheng and L. Lei</i> | <i>2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)</i> | 2019 |
| 24 | <i>Long-range body-to-body LoRa link at 868 MHz</i> | <i>Van Torre, P. and Ameloot, T. and Rogier, H.</i> | <i>13th European Conference on Antennas and Propagation, EuCAP 2019</i> | 2019 |
| 25 | <i>LoRa-based Measurement Station for Water Quality Monitoring: Case of Botanical Garden Pool</i> | <i>Ngom, B. and Diallo, M and Gueye, B. and Marilleau, N.</i> | <i>SAS 2019 - 2019 IEEE Sensors Applications Symposium, Conference Proceedings</i> | 2019 |
| 26 | <i>InterOpT: A new testing platform based on oneM2M standards for IoT systems</i> | <i>Demirel, S.T. and Demirel, M. and Dogru, I. and Das, R.</i> | <i>2019 International Symposium on Networks, Computers and Communications, ISNCC 2019</i> | 2019 |

Continua na próxima página

Continuação da tabela

| Id | Título | Autores | Veículo | Ano |
|-----------|---|--|---|------------|
| 27 | <i>Performance assessment of software defined networks management protocol in real environments</i> | <i>Da Costa, T.S. and Silva, J.D.C. and Rodrigues, J.J.P.C. and Rabelo, R.A.L. and Kumar, N. and Solic, P.</i> | <i>2019 4th International Conference on Smart and Sustainable Technologies, SpliTech 2019</i> | 2019 |
| 28 | <i>A model based test pattern generation and testing framework for IoT applications</i> | <i>Sathyavathy, V. and Shanmuga Priyaa, D.</i> | <i>International Journal of Innovative Technology and Exploring Engineering</i> | 2019 |
| 29 | <i>Performance escalation and optimization of overheads in the advanced underwater sensor networks with internet of things</i> | <i>Pradeep, S. and Sharma, Y.K.</i> | <i>International Journal of Innovative Technology and Exploring Engineering</i> | 2019 |
| 30 | <i>Adoption of software testing in internet of things: A systematic literature mapping</i> | <i>Cortés, M. and Saraiva, R. and Souza, M. and Mello, P. and Soares, P.</i> | <i>ACM International Conference Proceeding Series</i> | 2019 |
| 31 | <i>Performance evaluation of FIWARE: A cloud-based IoT platform for smart cities</i> | <i>Araujo, V. and Mitra, K. and Saguna, S. and Åhlund, C.</i> | <i>Journal of Parallel and Distributed Computing</i> | 2019 |

Continua na próxima página

Continuação da tabela

| Id | Título | Autores | Veículo | Ano |
|--|---|--|--|------------|
| 32 | <i>An end-to-end Internet of Things solution for Reverse Supply Chain Management in Industry 4.0</i> | <i>Garrido-Hidalgo, C. and Olivares, T. and Ramirez, F.J. and Roda-Sanchez, L.</i> | <i>Computers in Industry</i> | 2019 |
| Estudos obtidos a partir de outras fontes | | | | |
| 1 | <i>The art of computer systems performance analysis - techniques for experimental design, measurement, simulation, and modeling</i> | <i>Jain, Raj</i> | <i>Wiley professional computing</i> | 1991 |
| 2 | <i>Performance specifications and metrics for adaptive real-time systems</i> | <i>Lu, Chenyang and Stankovic, John A and Abdelzaher, Tarek F and Tao, Gang and Son, Sang Hyuk and Marley, Michael</i> | <i>Proceedings 21st IEEE Real-Time Systems Symposium</i> | 2000 |
| 3 | <i>Dynamic adaptation for fault tolerance and power management in embedded real-time systems</i> | <i>Zhang, Ying and Chakrabarty, Krishnendu</i> | <i>ACM Transactions on Embedded Computing Systems (TECS)</i> | 2004 |
| 4 | <i>CPU design: answers to frequently asked questions</i> | <i>Thimmannagari, Chandra</i> | <i>Springer Science & Business Media</i> | 2004 |
| 5 | <i>Performance Testing Guidance for Web Applications: Patterns & Practices</i> | <i>Meier, J. and Farre, Carlos and Bansode, Prashant and Barber, Scott and Rea, Dennis</i> | <i>Microsoft Press</i> | 2007 |

Continua na próxima página

Continuação da tabela

| Id | Título | Autores | Veículo | Ano |
|-----------|---|---|--|------------|
| 6 | <i>Analysis of Non-functional Properties in Software Product Lines: A Systematic Review</i> | <i>Soares, Larissa and Potena, Pasqualina and Machado, Ivan and Crnkovic, Ivica and Almeida, Eduardo</i> | <i>Proceedings - 40th Euromicro Conference Series on Software Engineering and Advanced Applications, SEAA 2014</i> | 2004 |
| 7 | <i>Measuring Performance Efficiency of Application applying Design Patterns and Refactoring Method</i> | <i>Langsari, Kholed and Rochimah, Siti and Akbar, Rizky</i> | <i>IPTEK Journal of Proceedings Series</i> | 2018 |
| 8 | <i>Mobile Edge Computing and Networking for Green and Low-Latency Internet of Things</i> | <i>Ke Zhang and Supeng Lengd Yejun He and Sabita Maharjan and Yan Zhang</i> | <i>IEEE Communications Magazine</i> | 2018 |
| 9 | <i>Performance Testing in Mobile Application: a Systematic Literature Map</i> | <i>Norberto, Marcus and Gaedicke, Lukas and Bernardino, Maicon and Legramante, Guilherme and Basso, Fabio Paulo and Rodrigues, Elder Macedo</i> | <i>SBQS'19: Proceedings of the XVIII Brazilian Symposium on Software Quality</i> | 2004 |

Fim da tabela

APÊNDICE F – FORMULÁRIO DE EXTRAÇÃO DE DADOS PARA DESEMPENHO

| Item | Description | Note |
|---|-------------|------|
| Year | | |
| Author | | |
| Title | | |
| Source (ACM, Compendex, Scopus, IEEE, ...) | | |
| Journal/Conference name | | |
| Characteristics? | | |
| Metrics? | | |
| Tests cases? Standart? | | |
| Tools? | | |
| Lessons learned? | | |
| Others? | | |
| Study Quality | | |
| Are the aim(s) of study clearly stated?(Yes/No) | | |
| Is there any discussion about the results of the study?(Yes/No/Partially) | | |

[illegible]

[illegible]

APÊNDICE J – CHECKLIST DA AVALIAÇÃO COM ESPECIALISTAS

Sobre o Perfil

| | Perfil do Especialista |
|-----|---|
| PE1 | Qual sua formação acadêmica? |
| PE2 | Sua ocupação atualmente é em áreas profissionais, acadêmicas ou ambas? |
| PE3 | Qual função exerce no meio que você atua? (e.g., aluno de mestrado, programador) |
| PE4 | Em qual área das seguintes você possui experiência: Teste, IoT, ambas? |
| PE5 | Quanto tempo de experiência possui na área? |
| PE6 | Qual sua experiência na avaliação de trabalhos da área de teste, IoT ou ambas (e.g., alta, avalio muitos trabalhos na área de teste)? |

Sobre a estrutura do Guia

| | Estrutura |
|-----|--|
| EQ1 | A estrutura proposta para o guia auxilia no seu uso? |
| EQ2 | As seções propostas são suficientes para propiciar a avaliação de uma característica IoT? |
| EQ3 | A ordem das seções segue uma sequência lógica? |
| EQ4 | Todas as seções possuem uma explicação do que se trata? |
| EQ5 | As informações apresentadas nas seções do guia, quando aplicável, possuem um estrutura padronizada (e.g., casos de teste, métricas)? |

Sobre o conteúdo do Guia

| | Geral |
|------|---|
| CQ1 | É claro a qual domínio o guia é direcionado? |
| CQ2 | A ordem das seções segue uma sequência lógica de apresentação? |
| CQ3 | As seções presentes no guia são suficientes para validar uma aplicação IoT? |
| CQ4 | O título do guia é claro, conciso e sem ambiguidade? |
| CQ5 | Cada título de seção e subseção está descrito com clareza, concisão e sem ambiguidade? |
| CQ6 | Cada seção e subseção está descrita com clareza, concisão e sem ambiguidade? |
| CQ7 | A introdução apresenta de maneira clara o propósito do guia? |
| | Definição da Característica |
| CQ8 | As definições de Desempenho apresentadas no guia estão claras? |
| | Correlação das Característica |
| CQ9 | O texto apresentado na seção é suficiente para a compreensão das correlações? |
| CQ10 | A figura representa de maneira clara as relações de Desempenho com as demais características? |
| CQ11 | Esta seção auxilia na identificação de possíveis conflitos entre características de IoT e Desempenho? |
| | Configuração do Ambiente |
| CQ12 | O título da seção está condizente com o conteúdo apresentado? |
| CQ13 | O texto apresentado na seção é suficiente para a compreensão do ambiente de teste? |
| CQ14 | Toda configuração de ambiente necessária para execução dos testes do guia estão listados? |
| | Subcaracterística - Comportamento Temporal |
| CQ15 | A definição de comportamento temporal está clara? |

| | |
|---|--|
| CQ16 | As propriedades apresentadas são suficientes para caracterizar comportamento temporal? |
| CQ17 | Existem propriedades redundantes? |
| CQ18 | As definições apresentadas para cada propriedade estão claras? |
| CQ19 | Cada caso de teste está descrito com clareza, concisão e sem ambiguidade? |
| CQ20 | Existem casos de testes redundantes? (e.g., casos de testes com o mesmo objetivo) |
| CQ21 | Cada caso de teste fornece detalhamento suficiente para ser entendido e executado? |
| CQ22 | Os campos dos casos de teste são preenchidos corretamente? (e.g, cada caso de teste reflete o seu título?) |
| CQ23 | Cada métrica está descrita com clareza, concisão e sem ambiguidade? |
| CQ24 | Existem métricas redundantes? (e.g., casos de testes com o mesmo objetivo) |
| CQ25 | Cada métrica fornece detalhamento suficiente para ser entendida e executada? |
| CQ26 | Os campos das métricas são preenchidos corretamente? (e.g, cada métrica reflete o seu título?) |
| Subcaracterística - Utilização de Recursos | |
| CQ27 | A definição de utilização de recursos está clara? |
| CQ28 | As propriedades apresentadas são suficientes para caracterizar utilização de recursos? |
| CQ29 | Existem propriedades redundantes? |
| CQ30 | As definições apresentadas para cada propriedade estão claras? |
| CQ31 | Cada caso de teste está descrito com clareza, concisão e sem ambiguidade? |
| CQ32 | Existem casos de testes redundantes? (e.g., casos de testes com o mesmo objetivo) |
| CQ33 | Cada caso de teste fornece detalhamento suficiente para ser entendido e executado? |
| CQ34 | Os campos dos casos de teste são preenchidos corretamente? (e.g, cada caso de teste reflete o seu título?) |
| CQ35 | Cada métrica está descrita com clareza, concisão e sem ambiguidade? |
| CQ36 | Existem métricas redundantes? (e.g., casos de testes com o mesmo objetivo) |
| CQ37 | Cada métrica fornece detalhamento suficiente para ser entendida e executada? |
| CQ38 | Os campos das métricas são preenchidos corretamente? (e.g, cada métrica reflete o seu título?) |
| Subcaracterística - Capacidade | |
| CQ39 | A definição de capacidade está clara? |
| CQ40 | As propriedades apresentadas são suficientes para caracterizar capacidade? |
| CQ41 | Existem propriedades redundantes? |
| CQ42 | As definições apresentadas para cada propriedade estão claras? |
| CQ43 | Cada caso de teste está descrito com clareza, concisão e sem ambiguidade? |
| CQ44 | Existem casos de testes redundantes? (e.g., casos de testes com o mesmo objetivo) |
| CQ45 | Cada caso de teste fornece detalhamento suficiente para ser entendido e executado? |
| CQ46 | Os campos dos casos de teste são preenchidos corretamente? (e.g, cada caso de teste reflete o seu título?) |
| CQ47 | Cada métrica está descrita com clareza, concisão e sem ambiguidade? |
| CQ48 | Existem métricas redundantes? (e.g., casos de testes com o mesmo objetivo) |
| CQ49 | Cada métrica fornece detalhamento suficiente para ser entendida e executada? |
| CQ50 | Os campos das métricas são preenchidos corretamente? (e.g, cada métrica reflete o seu título?) |

| | Impacto das Subcaracterísticas |
|------|--|
| CQ51 | O texto apresentado na seção é suficiente para a compreensão as relações? |
| CQ52 | As figuras representam de maneira clara as relações? |
| CQ53 | Esta seção auxilia na identificação de possíveis conflitos entre as subcaracterísticas de Desempenho? |
| CQ54 | Esta seção auxilia na identificação de possíveis casos de teste ou métricas de subcaracterísticas distintas que possam ser feitas de maneira conjunta? |
| | Custo Benefício |
| CQ55 | O texto apresentado na seção é suficiente para a relação de custo-benefício? |
| CQ56 | Esta seção auxilia de maneira clara no cálculo do custo-benefício da validação do desempenho de aplicações IoT? |
| CQ57 | A figura auxilia na compreensão da seção? |
| | Sugestões de Ferramentas |
| CQ58 | O texto apresentado na seção é suficiente para uma compreensão básica de cada ferramenta? |
| CQ59 | Esta seção auxilia na identificação de ferramentas que possam ser úteis durante a validação com o guia? |
| | Exemplo de Uso do Guia |
| CQ60 | O texto apresentado na seção é suficiente para compreender como deve ser o uso do guia? |
| CQ61 | Cada passo é descrito com clareza, concisão e sem ambiguidade? |

APÊNDICE K – DESIGN DO EXPERIMENTO

Ideia do experimento:

Parte da necessidade de avaliação do guia de desempenho de forma prática, separando em grupos de participantes que utilizando o guia e participantes sem utilizar o guia, a fim de que eles realizem a avaliação de desempenho de uma aplicação IoT, previamente definida, desde a etapa de planejamento dos testes, à execução e geração de falhas.

1. OBJETIVO

Objetivo global

Avaliar se o guia de teste de desempenho para IoT oferece os meios necessários para detecção de falhas de desempenho existentes em aplicações IoT e se o mesmo, reduz o esforço na realização dos testes.

Objetivo de estudo

Analisar o guia de teste de desempenho IoT

Com propósito de avaliar

Com respeito à detecção de falhas e esforço utilizado na condução dos testes

Do ponto de vista do responsável pela execução dos testes

No contexto de aplicações IoT

Questões

Q1 - O uso do guia reduz o esforço destinado à condução dos testes de desempenho?

Q2 - Há uma maior eficácia nos casos de testes gerados por meio do uso do guia?

Q3 - O uso do guia amplia a detecção de falhas IoT?

2. PLANEJAMENTO

Definição das Hipóteses

Hipóteses nulas:

(Avaliadas por meio de T-Student's)

H_{0,0}: A abordagem baseada em guias estruturados para conduzir atividades de teste de desempenho requer o mesmo esforço de teste que os testes de desempenho tradicionais.

H_{0,1}: A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho produz casos de teste com eficácia igual aos testes de desempenho

tradicionais.

H_{0,2}: A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho detecta o mesmo número de falhas IoT que os testes de desempenho tradicionais.

Hipóteses Alternativas:

H_{1,1}: A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho reduz mais o esforço de teste do que os testes de desempenho tradicionais.

Esforço (Com o guia) < Esforço (Sem o guia).

H_{1,2}: Abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho produz casos de teste mais eficazes do que os testes de desempenho tradicionais.

Eficácia dos casos de teste (Com o guia) > Eficácia dos casos de teste (Sem o guia).

H_{1,3}: A abordagem baseada em guias estruturados para conduzir atividades de testes de desempenho encontra mais falhas na IoT do que os testes de desempenho tradicionais.

Número de falhas da IoT (Com o guia) > Número de falhas da IoT (Sem o guia).

Variáveis

Variáveis dependentes:

- Casos de testes

Variáveis independentes:

- Falhas específicas para o desempenho IoT
- Esforço no planejamento e execução dos testes

Descrição da instrumentação

A instrumentação do experimento será dividida em itens utilizado pelos dois grupos e itens utilizado exclusivamente pelo grupo com o guia:

- Itens utilizados pelo dois grupos
 - Vídeos sobre testes de desempenho
 - Vídeos sobre o funcionamento da aplicação
 - *Template* para relatório de falhas
 - Definição das características
 - Definição de níveis e gravidade de falhas
 - Checklist do experimento
 - Link para download da aplicação
 - E-mail e telefone para contato
- Itens utilizados exclusivamente pelo grupo com o guia
 - Vídeo explicativo sobre a Wiki
 - Guia de Teste de Desempenho para Aplicações IoT
 - Perguntas adicionadas ao checklist do experimento sobre a Wiki

Aplicações IoT

Automa GREat: sistema IoT para controlar ar condicionados e lâmpadas

Seleção dos Participantes

Os participantes serão alunos de graduação e pós-graduação, doutores e profissionais da área de tecnologia.

3. EXECUÇÃO DO EXPERIMENTO

Os participantes que utilizaram o guia seguiram 8 passos e os demais seguiram 6 passos. Cada passo é detalhado abaixo:

1. Assistir o vídeo sobre teste de desempenho e IoT para um maior esclarecimento;
2. Para os participantes que utilizaram o guia. Assistir o vídeo sobre a WiKi e o guia para facilitar a utilização;
3. Baixar e instalar a aplicação;
4. Assistir o vídeo sobre a aplicação para um melhor entendimento e fazer a configuração seguindo o vídeo;
5. Para os participantes que utilizaram o guia. Acessar o link da WiKi para iniciar a avaliação;
6. Condução do planejamento, construção e execução dos testes, incluindo o preenchimento do relatório de falhas;
7. Ao final, responder checklist do experimento;
8. Enviar todos os artefatos via e-mail.

APÊNDICE L – DEFINIÇÃO DAS CARACTERÍSTICAS

| Característica | Definição |
|-----------------------|---|
| Compreensibilidade | grau em que os dados têm atributos que permitem sua leitura e interpretação pelos usuários, e são expressos em idiomas, símbolos e unidades apropriados em um contexto específico de uso. |
| Recuperabilidade | grau no qual, em caso de interrupção ou falha, um produto ou sistema pode recuperar os dados diretamente afetados e restabeleça o estado desejado do sistema. |
| Manutenibilidade | grau de eficácia e eficiência com que um produto ou sistema pode ser modificado para melhorá-lo, corrigi-lo ou adaptá-lo às mudanças no ambiente e nas exigências. |
| Funcionalidade | conjunto de funções que satisfaz as necessidades explícitas e implícitas para a finalidade a que se destina o produto. |
| Dependabilidade | qualidade do serviço fornecido por um dado sistema e a confiança depositada no serviço fornecido. |
| Confidencialidade | grau em que os dados têm atributos que garantem que sejam acessíveis e interpretáveis apenas por usuários autorizados em um contexto de uso específico. |
| Acurácia | grau em que os dados têm atributos que representam corretamente o verdadeiro valor do atributo pretendido de um conceito ou evento em um contexto de uso específico. |
| Usabilidade | grau em que um produto ou sistema pode ser usado por usuários específicos para atingir objetivos específicos com eficácia, eficiência e satisfação em um contexto de uso especificado. |
| Segurança | grau em que um produto ou sistema protege informações e dados para que pessoas ou outros produtos ou sistemas tenham o grau de acesso aos dados apropriado para seus tipos e níveis de autorização. |
| Disponibilidade | grau em que os dados têm atributos que permitem que sejam recuperados por usuários e / ou aplicativos autorizados em um contexto de uso específico. |
| Eficiência | grau em que os dados têm atributos que podem ser processados e fornecem os níveis esperados de desempenho usando as quantidades e tipos de recursos apropriados em um contexto de uso específico. |
| Interoperabilidade | grau em que dois ou mais sistemas, produtos ou componentes podem trocar informações e usar as informações que foram trocadas. |
| Modificabilidade | grau no qual um produto ou sistema pode ser modificado de forma eficaz e eficiente sem introduzir defeitos ou degradar a qualidade do produto existente. |
| Portabilidade | grau em que os dados possuem atributos que permitem que sejam instalados, substituídos ou movidos de um sistema para outro, preservando a qualidade existente em um contexto de uso específico. |
| Confiabilidade | grau em que um sistema, produto ou componente executa funções especificadas sob condições especificadas por um período especificado. |
| Reusabilidade | grau em que um recurso pode ser usado em mais de um sistema, ou na construção de outros recursos. |
| Robustez | habilidade de um sistema, produto ou componente funcionar mesmo em condições anormais. |
| Escalabilidade | facilidade com que uma aplicação ou componente pode ser modificado para expandir suas capacidades existentes, incluindo a capacidade de acomodar grandes volumes de dados. |

Definições obtidas em: (ISO/IEC 25010, 2011).

APÊNDICE M – CHECKLIST DO EXPERIMENTO

Perfil

| | Perfil |
|-----|---|
| PE1 | Qual sua formação acadêmica? |
| PE2 | Como você considera seu nível inglês: iniciante, intermediário, avançado? |
| PE3 | Sua ocupação atualmente é em áreas profissionais, acadêmicas ou ambas? |
| PE4 | Qual função exerce no meio que você atual? (e.g., aluno de mestrado, programador) |
| PE5 | Em qual área das seguintes você possui experiência: Teste, IoT, ambas? |
| PE6 | Quanto tempo de experiência possui na(s) área(s)? |
| PE7 | Você tem experiência em testar requisitos não-funcionais, se sim, quais requisitos não-funcionais você já avaliou? (e.g. Desempenho, Segurança, Usabilidade, etc) |

Perguntas

| ID | Perguntas |
|------------------|---|
| GERAL | |
| P1 | Você seguiu algum processo para a realização dos testes, se sim qual? Descreva o passo a passo deste o planejamento até a execução. Caso não tenha utilizado nenhum processo, escreva de maneira informal como você executou os testes. |
| P2 | Você realizou os testes em uma única vez ou vários dias? |
| P3 | Você assistiu todas as instruções fornecidas em vídeo? Caso não, quais você não assistiu e por que não? |
| P4 | Foi necessário assistir mais de uma vez algum vídeo? Se sim, quais? |
| P5 | Os vídeos fornecidos como insumos cumpriu seu objetivo de maneira clara e foi suficiente para compreender o assunto do qual ele abordava? Caso não, explicar o porquê. |
| P6 | Os vídeos e todos os documentos fornecidos para a condução do experimento cumpriram seu objetivo de maneira clara e foi suficiente para razão da qual ele foi criado? Caso não, explicar o porquê. |
| ARTEFATOS | |
| P7 | Quais artefatos de testes você gerou durante o experimento? (Ex plano de testes) |
| P8 | Você utilizou algum template para a geração dos artefatos, se sim, qual e por quê? |
| P9 | Caso tenha utilizado algum template para a geração dos artefatos de teste, você utilizou todas as seções deste template? Caso não, informe as seções que não usou e a razão. |
| P10 | Você conseguiu criar todos os artefatos que planejou no início do experimento? Caso não, informe quais você não criou e a razão pela qual você não conseguiu. |
| EXECUÇÃO | |

| | |
|-----------------------------|--|
| P11 | Você utilizou alguma ferramenta para automatizar a execução de alguma etapa do teste de desempenho? Se sim, em qual a etapa e quais ferramentas usou e por que optou por essas? |
| ESFORÇO | |
| P12 | A execução dos testes foi totalmente manual, automática ou ambas? |
| P13 | Quanto foi gasto no total do experimento considerando que temos 5 etapas que requerem esforços: etapa de treinamento (vídeos) e compreensão da aplicação, etapa de configuração - da aplicação e de ferramentas de testes, etapa de planejamento dos testes, etapa de especificação dos testes e etapa de execução dos testes. |
| P14 | Quanto tempo foi gasto na etapa de treinamento? |
| P15 | Quanto tempo foi gasto na etapa de configuração? |
| P16 | Quanto tempo foi gasto para na etapa de planejamento de testes? |
| P17 | Quanto tempo foi gasto na geração de cada artefato nas etapas de planejamento e/ou especificação e/ou execução? |
| P18 | Quantos casos de testes foram planejados? |
| P19 | Quanto tempo foi gasto para na etapa de especificação? |
| P20 | Todos os casos de testes que foram planejados, foram especificados? Caso não, informe a razão. |
| P21 | Quanto tempo foi gasto para na etapa de execução? |
| P22 | Todos os casos de teste que foram especificados, foram executados? Caso não, informe a razão. |
| P23 | Fez uso de métricas? Se sim, quantas foram coletadas? |
| P24 | Todas as métricas que foram planejados foram coletadas de fato? Caso não, informe a razão. |
| P25 | Caso tenha utilizado ferramentas, quais foram as vantagens e desvantagens no uso de cada delas? |
| DIFICULDADES | |
| P26 | Você teve dificuldades na execução dos testes? Se sim quais? |
| P27 | Caso tenha tido dúvidas na execução do experimento, todas foram sanadas até o fim do experimento? Caso não, quais das dificuldades não foram sanadas? |
| P28 | Você teve alguma dificuldade com a aplicação utilizada no experimento, se sim quais foram as dificuldades? |
| P29 | Caso tenha encontrado falhas nas aplicações você teve dificuldade em classificar e fornecer gravidade (severidade) a essas falhas, se sim quais foram as dificuldades? |
| CONSIDERAÇÕES FINAIS | |

| | |
|-----|--|
| P30 | Seguindo a escala de Likert de 1 a 5 sendo: 1) discordo totalmente, 2) discordo, 3) indiferente (ou neutro), 4) concordo e 5) concordo totalmente. Qual o seu nível de concordância com essa frase - "O nível de experiência nas áreas que envolvem o experimento - teste de desempenho e IoT - são essenciais para a execução do experimento."? |
| P31 | Há alguma sugestão ou mudança que você indicaria na organização experimento? |

APÊNDICE N – PERGUNTAS ADICIONADAS AO CHECKLIST DO GRUPO G1

| ID | Perguntas |
|---------------------|---|
| GERAL | |
| G1 | Você seguiu todo o processo sugerido no guia para a condução dos testes? |
| ARTEFATOS | |
| G2 | Você utilizou algum template além do fornecido pela wiki para a geração dos artefatos, se sim, qual e por quê? |
| EXECUÇÃO | |
| G3 | A geração do plano de teste na WiKi foi feita uma única vez? Caso não, por que foi necessário gerar mais que uma vez e quantas vezes foram gerados? |
| G4 | Você teve a necessidade de adicionar alguma etapa a mais no seu plano de teste além dos que são fornecidas pela Wiki, se sim quais? |
| G5 | Você utilizou todas as etapas da WiKi na construção do plano de teste, se não quais não utilizou e por quê? |
| G6 | Foi necessário acessar o guia em PDF? Se sim quantas vezes? E por quê? |
| G7 | Você utilizou alguma outra definição além das sugeridas na WiKi, se sim, quais e por quê? |
| G8 | Você utilizou alguma outra caso de teste além dos sugeridos na WiKi, se sim, quais e por quê? |
| G9 | Você utilizou alguma outra métrica além das sugerida na Wiki, se sim, quais e por quê? |
| G10 | Você utilizou alguma outra ferramenta além das sugeridas na WIKI, se sim, quais e por quê? |
| G11 | Você utilizou uma ou mais ferramentas para automatizar a execução de alguma etapa do teste de desempenho? Se sim, em qual a etapa e quais ferramentas usou e por que optou por essas? |
| ESFORÇO | |
| G12 | Quanto tempo foi gasto na compreensão da WiKi? |
| G13 | Quanto tempo foi gasto na geração do plano de teste na WiKi? |
| G14 | Todos os casos de testes que foram gerados no plano de teste da WiKi foram concretizados? Caso não, informe a razão. |
| G15 | Todos os casos de testes que foram gerados no plano de teste da WiKi foram executados? Caso não, informe a razão. |
| G16 | Todas as métricas que foram geradas no plano de teste da WiKi foram coletadas? Caso não, informe a razão. |
| DIFICULDADES | |
| G17 | Você teve alguma dificuldade com a utilização da Wiki se sim quais foram as dificuldades? |

| | |
|-----------------------------|---|
| G18 | Você teve alguma dificuldade com as ferramentas sugeridas no guia, se sim quais foram as dificuldades e em quais ferramentas? |
| G19 | Caso você tenha acessado o guia em PDF você teve alguma dificuldade com a utilização/compreensão, se sim, quais foram as dificuldades? |
| CONSIDERAÇÕES FINAIS | |
| G20 | Seguindo a escala de Likert de 1 a 5 sendo: 1) discordo totalmente, 2) discordo, 3) indiferente (ou neutro), 4) concordo e 5) concordo totalmente. Qual o seu nível de concordância com essa frase - "A WiKi é facilitadora para a execução do teste de desempenho IoT."? |
| G21 | Todos os campos que você usou da WiKi para geração do plano de teste foram realmente úteis e funcionais? Caso não, explicar o porquê. |
| G22 | Há alguma sugestão ou mudança que você indicaria, seja na organização experimento, na WiKi ou no Guia? |