Interoperability Testing Guide for the Internet of Things

Karina da Silva Castelo Branco
Group of Computer Networks,
Software Engineering and Systems
(GREat)
Federal University of Ceará (UFC)
Fortaleza, Ceará, Brazil
karinascb@alu.ufc.br

Valéria Lelli Leitão Dantas
Computer Science Department
Group of Computer Networks,
Software Engineering and Systems
(GREat)
Federal University of Ceará (UFC)
Fortaleza, Ceará, Brazil
valerialelli@ufc.br

Liana Mara Carvalho
Group of Computer Networks,
Software Engineering and Systems
(GREat)
Federal University of Ceará (UFC)
Fortaleza, Ceará, Brazil
lianacdemenezes@gmail.com

ABSTRACT

The Internet of Things (IoT) has expanded the Internet by integrating smart objects, which when interconnected, can collect and share information to provide services. However, the intense data traffic and the diversity of interaction methods of smart objects, which vary based on the protocols and standards, bring several challenges for IoY Interoperability Testing. Such testing evaluates the capability of systems and devices to cooperate effectively. Regarding the challenges in IoT interoperability testing, we highlight the complexity of IoT architecture, the devices heterogeneity, and the guarantee of effective connectivity among the smart objects. In this context, this paper presents a interoperability testing guide for IoT. The guide was developed based on a literature review using systematic mapping and an analysis of real IoT environments. The guide's evaluation consisted of two steps: (1) a structural assessment using the Technology Acceptance Model (TAM), and (2) a controlled experiment applying the guide to test a real IoT application.

KEYWORDS

Interoperability, Internet of Things, Interoperability Testing

1 INTRODUCTION

Technology has significantly transformed human interactions with everyday objects, expanding their communication. Internet access has also evolved, becoming more accessible and faster [32], positively contributing to these objects' connectivity. This broad connectivity has driven the "Internet of Things (IoT)" to expand the Internet through the integration of smart objects. When interconnected, these objects have the capability to collect and share information, enabling them to provide several services.

The interconnection facilitated by IoT has outlined new perspectives regarding the smart objects that interact to automate various daily tasks. For example, refrigerators, air conditioners, smart locks in the context of a smart home, or even autonomous vehicles operating independently, guided by a variety of smart sensors [13, 27].

Nonetheless, the IoT scenario bring challenges for testing the quality characteristics of IoT applications [5, 6, 8, 12, 24]. Security, Interoperability, and Performance characteristics are identified as the most relevant, receiving considerable testing efforts [6].

Interoperability in IoT is a crucial facet to be addressed in the development of IoT systems [30]. It refers to the capability of two or more systems to communicate effectively while ensuring data integrity [38]. Therefore, IoT interoperability testing verifies the ability of systems to interact consistently and cohesively. Such testing also involves evaluating their efficiency in communicating and sharing information, ensuring resources can be accessed and used appropriately accross different systems and organizations [23].

In this context, the intense data traffic and the diverse interaction methods of smart objects, which vary depending on protocols (MQTT, HTTP, CoAP, Bluetooth, and Zigbee), pose considerable challenges in interoperability testing. It is worth mentioning that while Bluetooth and Zigbee are sometimes referred to as standards, they, like MQTT, HTTP, and CoAP, are also protocols that operate at different tier of the communication stack [15].

For example, considering an IoT smart home scenario with diverse devices (voice assistants, security cameras, thermostats, smart bulbs, and locks), promoting their communication and integration across different technologies for flawless operation poses challenges regarding the architectural complexity, the device heterogeneity, the effective connectivity, and the management of bandwidth and device resource limitations for real-time data processing.

Therefore, the goal of this paper is to present the *Interoperability Testing Guide for IoT applications.* The following research questions (RQ) are investigated in this work:

RQ1. How to evaluate the interoperability characteristic in IoT applications?

RQ2. What approaches are used to evaluate interoperability in IoT applications?

RQ3. What are the main challenges related to *Interoperability* testing in IoT applications?

The development of the guide follows the methodology proposed by [6], which suggests a general structure based on 11 key topics. Initially, we performed a literature review to develop the guide' content'according that topics. This review also aimed to investigate the interoperability testing in different application domains. Next, we focused on identifying interoperability subcharacteristics, standards and approaches used in IoT interoperability testing. The final version of the guide is structured in 12 topics, including an additional topic called "Challenges of Interoperability Testing".

To evaluate the guide, two evaluations were conducted: (i) guide evaluation using the Technology Acceptance Model [9] (TAM); and (ii) a controlled experiment [44] to assess the use of the guide for testing the interoperability of an IoT application.

In: Proceedings of the Brazilian Symposium on Multimedia and the Web (WebMedia'2024). Juiz de Fora, Brazil. Porto Alegre: Brazilian Computer Society, 2024.
© 2024 SBC – Brazilian Computing Society.
ISSN 2966-2753

The remaining structure of this paper is organized as follows: Section 2 presents a theoretical basis necessary to understand the research. Section 3 describes the methodology. Section 4 presents the structure and the content of the guide. Section 5 presents the guide evaluation and Section 6 addresses the research questions. Section 8 discusses related work, and finally, Section 9 presents conclusions and future work.

2 INTEROPERABILITY TESTING IN IOT

The Internet of Things (IoT) has transformed connectivity by linking devices globally, turning physical objects into smart, interconnected entities with advanced functions [28, 33, 34]. It facilitates universal interaction by keeping people and smart objects connected through various networks [39].

IoT can be defined from three perspectives: devices, which are the sensing elements; the Internet, which serves as the network framework; and semantics, which involves communication protocols [41]. It includes essential components such as data-collecting sensors, connectivity technologies like Wi-Fi and 5G, and systems for data processing and storage in the cloud, all working together to ensure efficient and secure information flow [2].

IoT encompasses various quality characteristics, with at least 27 identified, including interoperability, security, performance, availability, and maintainability [6]. Interoperability allows devices from different manufacturers to communicate effectively, though it can sometimes negatively impact security, especially concerning data encryption.

To evaluate these characteristics, they are often divided into subcharacteristics. For example, interoperability is broken down into four subcharacteristics: Communication Protocol, System Integration, Data Semantics, and Network Protocol, as suggested by [6, 23]. This division helps in assessing different aspects that affect interoperability.

IoT interoperability testing checks if devices and systems from various manufacturers can work together seamlessly, exchanging data correctly and following specified behaviors. This testing is challenging due to the diverse range of devices, manufacturers, and protocols, as well as the dynamic environments in which IoT applications operate [10, 29, 35, 42].

3 METHODOLOGY

To develop the *Interoperability* Testing Guide, we followed the methodology proposed by Carvalho et al. [6]. The authors recommend a structure organized into 11 topics.

Once we defined the structure, we develop the guide’s content based on the instantiation methodology for an IoT characteristic [6], in our case, Interoperability. This methodology consist of six activities as shown in Figure 1.

The first activity of **“Literature Review”** was conducted following the guidelines of systematic mapping [25]. Thus, a search *string* focusing on the Interoperability characteristic was defined, as presented in Table 1. This string was formulated by combining keywords such as: “Internet of Things”, “interoperability testing”, “method”, “approach”, “challenge”, “tool”, among others.

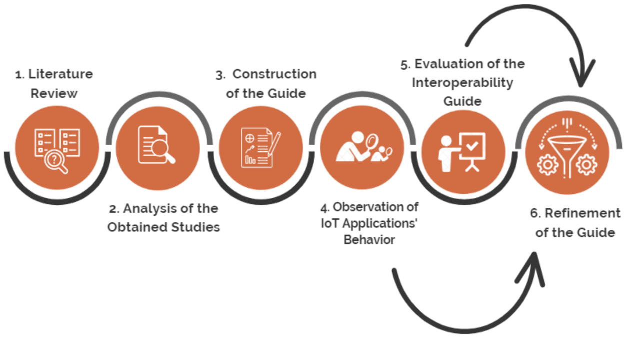
1. Literature Review
2. Analysis of the Obtained Studies
3. Construction of the Guide
4. Observation of IoT Applications' Behavior
5. Evaluation of the Interoperability Guide
6. Refinement of the Guide

Figure 1: Methodology for instantiation. Adapted from [6].

Table 1: String de busca

Search string
(*“internet of things*”) AND (*“interoperability test” or (“interoperability testing”) AND (method OR approach OR challenge OR framework OR tool OR architecture OR framework)*

Table 2: Inclusion and Exclusion Criteria

ID	Description
IC1	Studies related to interoperability testing in IoT apps
IC2	Studies that present IoT testing guide or similar artifacts
EC1	Studies that do not address interoperability testing in IoT
EC2	Studies for which the full text is not accessible
EC3	Studies that are shortened versions of others

In the second activity, **“Analysis of the obtained studies”**, we used the online tool Parsifal[1] to analyze the data and organize the protocol elements such as research questions, search terms and selection criteria. The analysis of the studies occurred in two phases. In the first round, the titles and abstracts of the 681 identified studies were read, resulting in the selection of 102 preliminary studies. Of these, 50 were from ACM, 24 from Scopus and 28 from IEEE.

The data extraction focused on the following aspects: definitions of interoperability; correlations of interoperability with other characteristics; challenges related to IoT interoperability; configuration requirements for IoT test environments; subdivision of interoperability into subcharacteristics; reported metrics and their calculation methods for evaluating interoperability; properties used to assess these subcharacteristics; base test cases; cost-benefit analyses; and tools employed in the studies to assess interoperability.

The third activity, **“Construction of the guide”**, consisted of developing the content of the guide regarding the 12 topics. The content was provided using the data obtained in the previous activity. For example, the topic named “Challenges of IoT Interoperability Testing” presents several challenges identified in the literature review related to IoT architecture complex, Communication standards, Device heterogeneity and communication. Additionally, other topics were enriched with examples of test cases, explanations of metric formulas, and suggested tools for automating the measurement.

[1] https://parsif.al/login/

In the fourth activity, **"Observation of IoT applications' behavior"**, the use of the guide was analyzed in an IoT application. This app aimed to provide information to public transport users (bus schedules and route details) through a mobile device that collects information using GPS. Using the guide to test this application allowed us to identify challenges in interoperability testing as reported in the guide, as well as create new test cases.

In the fifth activity, **"Evaluation of the Interoperability Guide"**, the guide was evaluated in two steps to ensure its quality. Initially, an expert who had developed a similar guide for another IoT characteristic assessed our guide, leading to several improvements. We also conducted two additional evaluations with undergraduate and graduate students into a V&V course at a university: one using the TAM model and another through a controlled experiment testing a real IoT application.

In the last activity, **"Refinement of the guide"**, we improved the guide based on the analysis and the evaluations conducted in activities 4 and 5 to enhance its utility and effectiveness.

The materials used for the conception and evaluation of the guide are available in the repository of this study[2].

4 INTEROPERABILITY TESTING GUIDE

The guide proposed in this paper was developed to cover a wide variety of testing scenarios related to *Interoperability* in IoT. The structure of the guide is based on the strucuture of 11 topics proposed by [6]. Thus, our guide addresses the following topics: "Characteristic definition", "Correlation of Characteristcs", "Challenges of IoT Testing Interoperability", "Test Environment configuration", "Impact of Subcharacteristics", "Cost-Benefit", "Tool Suggestions" and "Example of Guide Use". The guide also includes the recommended topics such as "Introduction", "Instructions for Using the Guide" and "References".

As recommended by the authors, a characteristic should be divided into one or more susubcharacteristics. In the case of Interoperability, we divided it into four subcharacteristics according to ISO 30141:2018 [23]: "Data Semantics", "Communication Protocol", "System Integration", "Network Protocol". For each subcharacteristic, the guide includes topics such as "Definition", "Contextualization", "Abstract Test Cases", and "Measurements".

Furthermore, we provided an extra topic named "Challenges of IoT Interoperability Testing". Therefore, the guide is structured into 12 distinct topics organized in sections. The full version of the guide is available in the repository of this study[3].

The following subsections introduce the Interoperability Testing Guide, with each one corresponding to a section of the guide.

4.1 Interoperability Definition

This section of the guide defines the characteristic of "Interoperability". The goal is to standardize the knowledge about what will be tested and facilitate understanding for software engineers and professionals from various fields. The section includes six definitions sourced from the results of the literature review and ISO/IEC 30141:2018 [23]. Below, we presented two of these definitions:

(1) Interoperability is *"the ability of a system to exchange data and information with other systems without loss or corruption of information."* [20]
(2) Interoperability is *"the ability of different systems and organizations to work together (exchange of information and actions) effectively and efficiently."*[23]

4.2 Correlation of Characteristics

The systematic mapping presented by Carvalho et al. [6] identified a set of 27 quality characteristics related to IoT. From this set, 14 characteristics were selected based on their correlation with Interoperability, as illustrated in Figure 2. The Correlation of Characteristics section is essential, as it clarifies how Interoperability is influenced by other IoT characteristics, guiding definitions, optimizations, and improvements in testing strategies.

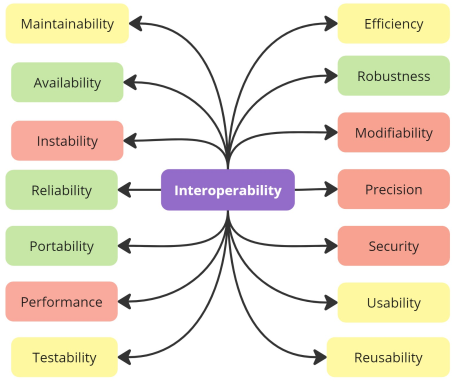

Figure 2: Characteristics correlations with interoperability

The correlations of IoT characteristics are organized into three groups: (i) positive (green rectangles), which indicate a favorable influence on interoperability; (ii) negative (red rectangles), which indicate the opposite; and (iii) variable (yellow rectangles), which depend on the context of a applications under test. As illustrated in Figure 2, we identified four positive correlations, five negative correlations, and five variable correlations regarding "Interoperability". "Availability", for instance, may enhance 'Interoperability' with sufficient servers but can have a negative effect otherwise. Similarly, "Instability" can have a negative impact on Interoperability by causing frequent communication failures, hindering the exchange of information between IoT devices. Another example is 'High 'Performance", which in certain situations can ensure smoother and more efficient communication between devices, but may also result in system overload and reduced performance.

The guide provides all definitions of the IoT characteristics correlated with "Interoperability", which were extracted from ISO standards [21, 22]. Below, we present the definitions of the four characteristics with positive correlations:

(1) **Availability:** refers to the system's ability to be operational and accessible when needed, minimizing interruptions or failures.
(2) **Instability:** relates to the system's propensity for failures or unexpected crashes, resulting in inconsistent operation.

[2] https://drive.google.com/drive/folders/1y4wVQTvfIxoO0t0NG9tAF0_lwa5mwGFk?usp=drive_link
[3] https://drive.google.com/drive/folders/1DozFXdxNVxTbI5fs3pI81T0uQ2JEclJU?usp=sharing

(3) **Performance:** concerns the system’s ability to effectively respond to requests and operate within established limits, ensuring acceptable response times.

(4) **Portability:** refers to how easily a system can be transferred or adapted to different environments or platforms without significant loss of functionality.

4.3 Challenges of IoT Interoperability Testing

The challenges regarding the “Interoperability Testing” in IoT are illustrated in Figure 3. In our research, we have identified 20 challenges, which are categorized into three groups: (i) *challenges most critical mentioned by literature* (red rows); (ii) *challenges most cited* in the literature (green rows), and (iii) *challenges observed in pratical* IoT applications (yellow rows).

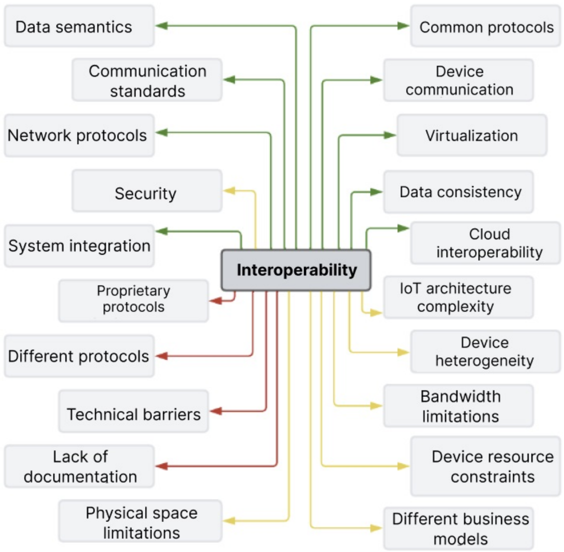

Figure 3: Main challenges in interoperability testing

Below, we present an example of a challenge per group:

(1) **Communication Standards**. Common standards facilitate integration and communication between different devices and systems, promoting interoperability. In IoT, heterogeneous devices operate using several protocols (MQTT, HTTP, and CoAP) and standards (Bluetooth and Zigbee). This diversity impacts the complexity of testing activity, for instance, most testing tools cannot interact properly with IoT applications, leading to challenges in test automation. One relevant issue posed by diverse communication standards is “How to ensure that IoT systems work correctly across all platforms and technologies?” [3, 4].

(2) **Security**. Interoperability must ensure that communication between devices is secure and reliable, adhering to security standards to protect the information exchanged. The challenge posed by security in IoT Devices is: “How to guarantee security in communication between IoT devices?” [36].

(3) **Proprietary Protocols**. These types of protocols pose challenges to interoperability with devices from other manufacturers, creating technical barriers. In this context, testing activities must address limited compatibility with standard protocols; restricted technical information; and higher costs and complexity associated with customizing testing procedures. The main challenge is: “How to overcome the technical barriers imposed by proprietary protocols?” [14, 15].

4.4 Test Environment Configuration

The environment setup for IoT interoperability testing encompasses diverse devices, protocols, manufacturers, and network conditions. This section outlines the essential elements:

IoT Devices: sensors and control devices compatible with various communication protocols.

Network Infrastructure: configuration of wireless networks (Wi-Fi, Bluetooth), switches, routers, and firewalls to ensure secure and reliable communication.

Actuators: devices that perform actions based on sensor data or external commands.

Decision and Command Application: platform that coordinates devices from different manufacturers and protocols to ensure system interoperability.

4.5 Interoperability Subcharacteristics

According to ISO/IEC 30141:2018[23], we divided Interoperability into four characteristics: (1) “Data Semantics”, which refers to the ability to interpret data, enabling systems and devices to share and use data efficiently; (2) ““Communication Protocols”, which concerns how devices must communicate with each other, ensuring the efficient exchange of information and interoperability; (3) “System Integration”, is the process of standardizing the way different systems connect and communicate with each other; and (4) “Network Protocol”, which allows communication and coordination between devices, services and applications on a network, defining rules for data exchange and synchronization.

Sections 5 to 8 of the guide address these subcharacteristics with the following topics:

- *Definition* presents the explanation of each subcharacteristic.
- *Contextualization* describes the properties to evaluate each subcharacteristic, extracted from the literature.
- *Abstract Test Cases* provides structured and implementation-independent steps to test a subcharacteristic. We define 25 test cases covering four subcharacteristics: six for “Data Semantics”, seven for “Communication Protocols”, five for “System Integration”, and seven for “Network Protocols”. An example of a test case is illustrated in Table 3.
- *Measurements* describes the methods and metrics to quantify specific aspects of the system under evaluation. In addition, 19 metrics were provided to assist the evaluation of *Interoperability*. Table 4 shows an example of a metric.

4.6 Impact of Subcharacteristics

The four subcharacteristics of Interoperability may impact each other. In the scope of validation, it is crucial to evaluate the impact among interoperability subcharacteristics to prevent incompatible systems, , those that cannot communicate with each other.

Table 3: Example of an abstract test case

Test Case 01 - TC01	
Title	Data Reading
Test Environment	A network of heterogeneous IoT devices
Precondition	Devices connected to the same Wi-Fi
Step-by-Step	1 - Start the mobile application 2 - Select reading of data 3 - Verify the data displayed
Postcondition	The data displayed on the mobile device screen should correspond to the same data requested by the actuator

Table 4: Example of a metric

Device capacity - M01	
Purpose	Evaluate the ability of different IoT devices to interact with each other effectively.
Method	Perform interoperability tests with different IoT device firmwares.
Measure	**Success Rate** = (Number of Successful Interactions / Total Number of Interactions) x 100
Explanation	Calculates the success rate as percentage (%), where the number of successful interactions is divided by the total number of interactions and multiplied by 100 to obtain the % representation.
Reference	[23][2]

The influence of an subcharacteristic is contextual and depends on the application under test. Our guide provides an overview of the correlations among the four subcharacteristics of Interoperability.

Figure 4 shows the correlations between “Data Semantics” and the other three subcharacteristics. The colors represent the properties of each characteristic: (i) yellow for “Data Semantics“; (ii) green for System Integration”; (iii) blue for Communication Protocol”; and (iv) red for Network Protocol”. We identified 25 properties to these characteristics: eight for “Data Semantics”, seven for “Communication Protocol”, five for “System Integration”, and five for “Network Protocol”. The figure illustrates that when we evaluate a specific property of “Data Semantics”, we must consider related properties of the other subcharacteristics. For example, assessing “Common Interpretation” property involves considering “Protocol Compatibility”, “Data “Compatibility”, “Consistency” and “Compatibility”. Similarly, evaluating “Semantic Compatibility” implies considering “Reliability” and “Scalability”.

4.7 Cost-Benefit

The cost-benefit (CB) calculation is based on the formula proposed by [6]. According the authors, the “cost-benefit” evaluates the testing effort based on correlations of IoT characteristics. The CB calculation considers the impact of an IoT characteristic, associated the tests’ cost using specific parameters. Thus, the CB formula can be applied to evaluate the Interoperability characteristic, in which we identified 14 correlated characteristics (see subsection 4.2). The calculation of the CB formula is described below.

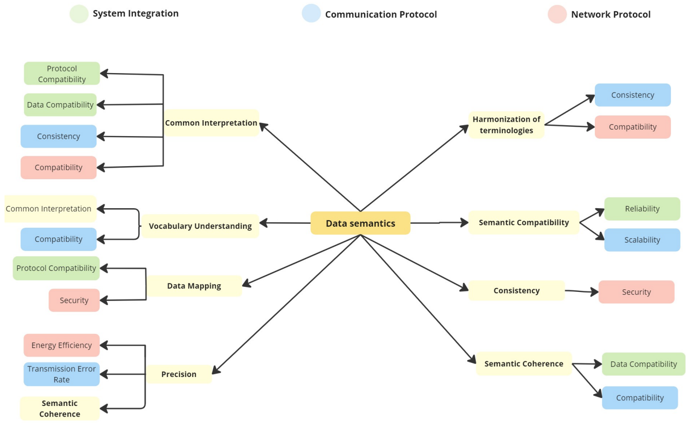

Figure 4: Impact of ’Data Semantics’ on subcharacteristics

$$CI = ORC/RC \tag{1}$$

- ORC: number of characteristics correlated to interoperability prioritized in the application
- RC: total number of characteristics related to interoperability

Impact (CI) and effort metrics are used to calculate the cost-benefit ratio, aiding in test prioritization. The calculation consists of the estimated cost of each test case based on the average execution time and the tester’s hourly rate, using the following formula:

$$CTi = TCi * VHCi \tag{2}$$

- CTi: Estimated cost to execute the test case
- TCi: Average time of a tester to execute a test case
- VHCi: Value of the time of a tester who will execute the test case

When all CTs are completed, the maximum value found is obtained as follows:

$$MCT = max(CT) \tag{3}$$

- MCT: highest cost for performing the test case
- CT: all cost estimates

Normalizes the average costs of the test cases as follows:

$$CCT = (\sum_{i=1}^{n} CT_i/MCT)/n \tag{4}$$

- CCT: average cost of standardized test cases
- CTi/MCT: estimated value of the cost of test case i normalized to the highest cost
- n: number of test cases

Repeat the process for the measurements to obtain the CMD:

$$CMD = averagecostofstandardizedmeasurements. \tag{5}$$

Thus, the Effort (ESF) is defined as follows:

$$ESF = (CCT + CMD)/2 \tag{6}$$

The result is analyzed from a Cartesian plane that varies from 0 to 1 and depends on the quadrant. In the CB of interoperability, the x-axis represents Impact (CI) and the y-axis represents Effort (ESF). Tests are prioritized as follows: Group I (High Effort, Low Impact) has low priority; Group II (Low Effort, Low Impact) has medium priority; Group III (High Effort, High Impact) has high priority; and Group IV (Low Effort, High Impact) has very high priority.

4.8 Tool Suggestions

In the guide, we catalog a list of eight tools to test interoperability in IoT applications. These tools include Eclipse IoT [17], OpenIoT[18], Wireshark [43], IoTIFY [19], CoAPthon [7], FreeRTOS [11], Tasmota [40], and Home Assistant [1]. Each tool is detailed with aspects such as description, testing methodology, testing environment, test execution, type of license, and availability. Notably, IoTIFY is the only proprietary tool, while the others are open source.

4.9 Example of Guide Use

An example use case for the Interoperability Testing Guide in an IoT application is provided. The use of the guide aims to improve student travel planning and reduce waiting times at bus stops on a university campus. To ensure effective interoperability, several test scenarios address different aspects of this application. The steps to conduct interoperability testing on this application include:

(1) **Definitions of Interoperability:** Understand the fundamental definitions of interoperability outlined in the guide.
(2) **Characteristic Selection:** Identify the key characteristics relevant to the IoT application under test.
(3) **Environment Setup:** Prepare the test environment according to specified requirements.
(4) **Subcharacteristics and Properties:** Explore relevant subcharacteristics and select their properties for evaluation.
(5) **Impact of Subcharacteristics:** Assess how the chosen subcharacteristics might be affected by various decisions.
(6) **Metric Selection:** Choose appropriate metrics for evaluating the selected subcharacteristics.
(7) **Cost-Benefit Calculation:** Perform a cost-benefit analysis to justify the necessary investments in testing.
(8) **Tool Utilization:** Consider the recommended tools for effective metrics collection.
(9) **Test Execution:** Finalize and execute the abstract test cases with the provided data as planned.

The key interoperability characteristics prioritized for that IoT application under test (AUT) include availability, performance, security, portability, and systems integration. They are crucial for effective operation in transport scheduling and real-time updates.

The test environment configuration meet AUT requirements, involving smart devices, actuators, and an external application for real-time location commands.

The selected metrics evaluate properties such as response time, adaptability to different transport systems, security in data exchange, ease of integration with third-party systems, and platform portability.

The Cost-Benefit (CB) calculation uses hypothetical values to justify investment in interoperability tests.

The above example illustrates the application of the Interoperability guide in specific scenarios, emphasizing key characteristics, metrics, and cost-benefit analysis.

5 GUIDE EVALUATION

To evaluate the proposed guide, we conducted two evaluations:

(1) A analysis of the guide's structure using the Technology Acceptance Model (TAM); and
(2) A controlled experiment with 18 students, 16 undergraduate and 2 graduate, in a Software Verification and Validation (V&V) course at a university.

Regarding experience levels, 10 students have worked in both industry and academia, while 8 students are dedicated exclusively to academia. In terms of interoperability testing knowledge, four students had prior experience, whereas 14 students had experience in testing non-functional requirements.

The evaluation was conducted in the last two face-to-face classes of the V&V course, following the completion of the Validation module (unit, functional, and non-functional testing). To standardize the students' knowledge of interoperability testing, the first class covered theoretical concepts and a practical application using real IoT devices like Amazon Alexa. In the second class, both evaluations (1 and 2) were conducted in a real IoT application designed to assist students plan their trips and reduce their waiting times at bus stops on a university campus. The app features include route and schedule visualization; real-time tracking; arrival estimation; and stop location.

Next, we presented the two evaluations.

5.1 Evaluation using the TAM model

The TAM model was used to evaluate the structure, acceptance, and adaptation of users to the Interoperability Testing Guide. First, we presented the IoT application under test and the Guide in PDF format. Six of the 18 students in the second class evaluation did not attend the first leveling class and were invited only for the TAM evaluation. They were organized into pairs to evaluate the guide using the IoT application. After completing the evaluation, the students filled out the TAM questionnaire, consisting of 14 Likert Scale questions, covering 5 categories: *Perceived Usefulness* (PU), which measures users' perceptions of the guide's utility for effective testing; *Perceived Ease of Use* (PEOU), which assesses the guide's ease of understanding and learning; *Intention to Use in the Future* (IU), which checks users' intentions to future adoption and recommendations of the guide; *Impact on Test Efficiency* (ITE), which examines the guide's contributions to test efficiency; and *Overall Satisfaction* (OS).

5.2 Results of the TAM model

The TAM model results are detailed in Figure 5[4]. Using a scale where "strongly disagree" is 1, "'neutral" is 3 and "strongly agree" is 5, the global mode was 5. Thus, the "Strongly Agree" response was most frequent for most questions, indicating high acceptance and satisfaction with the Interoperability Testing Guide, demonstrating its effectiveness and utility.

In the *Perceived Usefulness*, all students fully agreed that the guide is useful for conducting tests (Q1) and effective in planning and specifying tests (Q2) in IoT. However, six students fully or partially agreed on the guide's effectiveness during test execution

[4] Chart generated with Likertplot tool . Available on: https://www.likertplot.com/

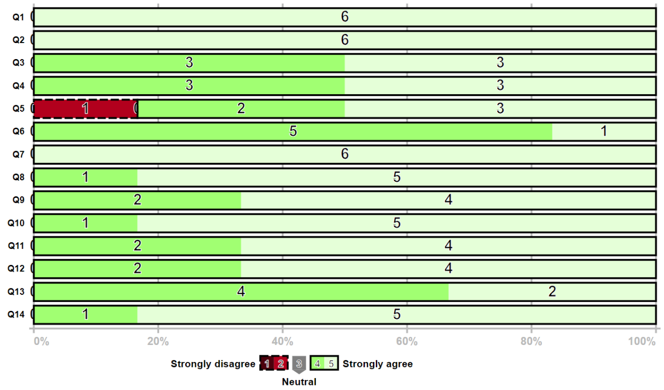

Figure 5: Results per question of TAM model

(Q3). Regarding *Perceived Ease of Use* of the guide, six students fully or partially agreed that the structure and instructions are easy to understand (Q4). Three students fully and two partially agreed that the learning curve was smooth (Q5), while one student partially disagreed. Four students fully and two partially agreed that the organization of the topics and their sequence were clear (Q6). All six students agreed that guide facilitates the testing planning and specification (Q7), whereas five students fully and one partially agreed that the guide facilitates the testing exectuion in IoT (Q8). In the *Intention to Use in the Future (IU)*, four students expressed their intention to use the guide in future IoT test projects (Q9) and five would recommend it to their colleagues (Q10). In the terms of *Impact on Test Efficiency*, four students fully and two partially agreed that the use of the guide contributed to the overall effectiveness of interoperatibitly tests in IoT (Q11). Regarding *Overall Satisfaction*, four students fully and two partially agreed that the guide's instructions and approach to test planning were clear and understandable (Q12), while for tests execution, two fully students and four partially agreed. Additionally, five students fully and one partially agreed on their satisfactions with using the Interoperability Test Guide (Q14).

5.3 Controlled Experiment

We conducted the experiment with 12 students who attended the first theoretical-practical class. They received materials, including a manual, presentation, videos, experiment design, failure report template, testing plan examples, and an explanation of the IoT app under test. In the second class, students were organized into pairs and divided into two groups: G1 (using the guide) and G2 (without the guide). They individually answered a pre-test questionnaire to assess their understanding of interoperability testing. The hypotheses of the experiment are as follows:

- **Null Hypotheses. $\mathbf{H_{0,0}}$** - The structured guide-based approach to conducting interoperability testing activities requires the same effort as traditional interoperability testing. $\mathbf{H_{0,1}}$ - The structured guide-based approach to conducting interoperability testing activities detects the same number of IoT failures as traditional interoperability testing.
- **Alternative Hypotheses. $\mathbf{H_{1,1}}$** - The structured guide-based approach to conducting interoperability testing activities reduces testing effort more than traditional interoperability testing. $H_{1,1}$: *Effort with the guide < Effort without the guide.* $\mathbf{H_{1,2}}$ - The structured guide-based approach to conducting interoperability testing activities produces more effective test cases than traditional interoperability testing. $H_{1,2}$: *Effectiveness of test cases with the guide > Effectiveness of test cases without the guide.* $\mathbf{H_{1,3}}$ - The structured guide-based approach to conducting interoperability testing activities finds more IoT failures than traditional interoperability testing. $H_{1,3}$: *Number of IoT failures with the guide > Number of IoT failures without the guide.*
- **Dependent Variables:** Test cases
- **Independent Variables:** Specific failures for IoT interoperability, effort in planning and executing tests.

5.4 Results of the Controlled Experiment

Table 5 gives an overview of the experiment results regarding the groups that used the guide (CG) and those that did not (SG). The figure shows the ID, planning time, number of test cases and reported IoT failures for each group. Planning time refers to the effort spent on setting up the test environment, devising test scenarios, choosing metrics, and defining the test plan scope.

Table 5: Experiment's results per group

ID	Time (min)	Test Cases (#)	IoT Failures (#)
		Group 1 - CG	
CG1	50	8	3
CG2	45	10	2
CG3	40	6	0
		Group 2 - SG	
SG1	90	4	0
SG2	60	3	0
SG3	50	6	2

Based on the analysis of the experiment data, the hypotheses were evaluated using the Student's T-test [44]. The objective of the hypothesis analysis is to verify if there is a significant difference (p-value < 0.05) in the effort to plan the tests, the effectiveness of test cases, and the number of IoT failures between the participants who used the guide and those who did not.

Regarding **hypothesis $\mathbf{H_{1,1}}$**, planning time was collected to compare the efforts between the two groups (GC and SG). The comparison showed a statistically significant difference in planning effort (p-value = 0.035), leading to the rejection of the null hypothesis $H_{0,1}$. Thus, the alternative hypothesis $H_{1,1}$ is accepted, indicating that the guide-based approach significantly reduces the testing effort. For **hypothesis $\mathbf{H_{1,2}}$**, the effectiveness of the test cases generated by each group was analyzed. The analysis revealed a statistically significant difference (p-value = 0.023) between GC and SG, leading to the rejection of the null hypothesis $H_{0,2}$. The alternative hypothesis $H_{1,2}$ is accepted, indicating that the guide-based approach generates more effective test cases. Regarding **hypothesis**

$\mathbf{H_{1,3}}$, which addresses the identification of IoT faults, the number of faults reported by each group was analyzed. The comparison between CG and SG revealed statistically significant differences (p-value = 0.03), leading to the rejection of the null hypothesis $H_{0,3}$. Thus, the alternative hypothesis $H_{1,3}$ is accepted, indicating that the guide-based approach detects more IoT faults.

Based on the results of the statistical analysis, the null hypotheses $H_{0,0}$, $H_{0,1}$, $H_{0,2}$ were rejected in favor of the alternative hypotheses $H_{1,1}$, $H_{1,2}$, $H_{1,3}$ respectively. This result indicated that the structured guide-based approach is more efficient in terms of effort, test case effectiveness, and fault detection in IoT compared to traditional interoperability testing.

6 DISCUSSION

In this section, the research questions are discussed.

RQ1. How to evaluate the *Interoperability* characteristic in IoT applications?

To evaluate *Interoperability* in IoT applications is crucial to carefully plan the tests. Thus, the test plan should provide the characaracteristics correlated to IoT, their impact in Interoperability, abstract test cases, and properties and metrics specific for IoT. Moreover, the plan should guide the proper configuration of the test environment to replicate real-world conditions. By covering these aspects accurately, it is possible to conduct an effective evaluation of interoperability in IoT applications.

RQ2. What are the testing approaches used to evaluate *Interoperability* in IoT applications? Approaches to evaluate interoperability in IoT applications were found, including a checklist model proposed by [5], framework-based evaluation as per [26], and an automated test generation framework by [31]. However, none of these approaches are specifically tailored for IoT interoperability according to ISO [23] standards. This gap motivated this work to focus on interoperability in specific IoT contexts, guided by a dedicated framework developed for this purpose.

RQ3. What are the main challenges related to testing *Interoperability* in IoT applications?

In our literature review, we identified 20 challenges of interoperability testing in IoT. Based on these findings, we included a new section titled "Challenges of Interoperability Testing" in our guide to explain the types of challenges found in interoperability testing. This section aims to assist users in identifying potential issues and plan alternative solutions. Additionally, this section strengthens the guide by becoming more tailored and comprehensive, specifically addressing the specific challenges present in IoT environments.

7 THREATS TO VALIDITY

In our research, we identified threats to validity related to the creation and evaluation of the guide, such as potential issues with its generalization, participants' varying levels of prior knowledge, and the limited number of participants (six for TAM and 12 in the controlled experiment). To mitigate these threats, we developed the guide based on a comprehensive literature review using systematic mapping guidelines. The two evaluations were conducted after students completed a V&V course. We employed two evaluation methods: TAM model focused on students who did not attend the first leveling class on Interoperability testing, and a controlled experiment involving diverse student profiles (undergraduate and graduates students, and professionals) organized in two groups.

8 RELATED WORK

Given the challenges outlined in Section 4, a literature review was conducted to explore related studies addressing gaps in IoT interoperability testing.

Zaid et al. [45] present a methodology based on contextual signatures for testing IoT interoperability. This approach focuses on protocol layers and interoperability properties through event correlation and signature verification. While their study includes environment configuration and test execution, it primarily emphasizes test execution for interoperability. In contrast, our guide offers comprehensive steps for test planning, specification, and execution.

Other studies, such as those by Caldas [5] and Silva et al. [37], focus on checklists for evaluating IoT interoperability. Caldas proposes a checklist for smart home devices, identifying interoperability levels and common devices. Silva et al. introduce ScenarIoT, a checklist for evaluating interactions in various environments, covering IoT components, requirements, and device interactions. Our guide differs by providing a detailed framework of 12 topics, including abstract test cases, properties, and metrics, and addresses correlations between interoperability and other IoT characteristics.

Gunathilaka et al. [16] propose a smart grid testing system for evaluating the interoperability of security solutions in IoT. Their model focuses on message translation and communication at gateways but lacks structured steps and does not correlate interoperability with other IoT characteristics.

Carvalho et al. [6] present a structured approach for testing IoT characteristics, specifically for performance. We adapted this guide-based approach to develop our guide for testing interoperability, including an additional topic on the challenges faced in this area.

9 CONCLUSION AND FUTURE WORK

In this article, we present the IoT interoperability testing guide. This guide was developed based on literature reviews and ISO/IEC 30141:2018 [23]. In our research, we conducted two literature reviews: the first aimed to broadly understand and identify the challenges in IoT interoperability testing, while the second focused on developing the proposed guide.

The guide was created following the methodology proposed by [6], which recommends a structure based on 11 topics. Our guide covers 12 topics, including one specifically for "IoT Interoperability Testing Challenges". These topics are organized into sections that define Interoperability and address its four sub-characteristics: *Data Semantics*; *Communication Protocols*; *System Integration*; and *Network Protocols.* For each feature, the guide provides related abstract test cases, property measurements.

We evaluated the guide through two evaluations (TAM and experiment) conducted during a VV course. The results showed the usefulness of the guide in assisting users with interoperability testing, identifying IoT failures in this context.

As future work, we plan to conduct evaluations with industry experts to refine the practical use of the guide. We also intend to develop a wiki to facilitate the use of the guide by automatically providing a test plan.

REFERENCES

[1] Home Assistant. 2024. Awaken your home. https://www.home-assistant.io/

[2] Luigi Atzori, Antonio Iera, and Giacomo Morabito. 2010. The internet of things: A survey. *Computer networks* 54, 15 (2010), 2787–2805.

[3] Luigi Atzori, Antonio Iera, and Giacomo Morabito. 2010. The internet of things: A survey. *Computer networks* 54, 15 (2010), 2787–2805.

[4] Miroslav Bures, Bestoun S. Ahmed, Vaclav Rechtberger, Matej Klima, Michal Trnka, Miroslav Jaros, Xavier Bellekens, Dani Almog, and Pavel Herout. 2021. PatrIoT: IoT Automated Interoperability and Integration Testing Framework. In *2021 14th IEEE Conference on Software Testing, Verification and Validation (ICST)*. 454–459. https://doi.org/10.1109/ICST49551.2021.00059

[5] Eduardo Alves Lima Caldas. 2023. Checklist para avaliação da interoperabilidade em dispositivos iot com foco em casas inteligentes. (2023).

[6] Liana M Carvalho, Valéria Lelli, and Rossana MC Andrade. 2022. Performance Testing Guide for IoT Applications.. In *ICEIS (1)*. 667–678.

[7] CoAPthon. 2024. https://github.com/Tanganelli/CoAPthon

[8] Mariela Cortés, Raphael Saraiva, Marcia Souza, Patricia Mello, and Pamella Soares. 2019. 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*. 3–11.

[9] Fred D. Davis. 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly* 13, 3 (1989), 319–340.

[10] Alexandra Desmoulin and César Viho. 2009. Formalizing interoperability for test case generation purpose. *International journal on software tools for technology transfer* 11, 3 (2009), 261–267.

[11] freeRTOS. 2024. Simplifying Authenticated Cloud Connectivity for Any Device. https://www.freertos.org/index.html

[12] Sara Nieves Matheu García, José Luis Hernández-Ramos, and Antonio F. Skarmeta. 2018. Test-based risk assessment and security certification proposal for the Internet of Things. *2018 IEEE 4th World Forum on Internet of Things (WF-IoT)* (2018), 641–646.

[13] Daniel Giusto, Antonio Iera, Giacomo Morabito, and Luigi Atzori. 2010. *The internet of things: 20th Tyrrhenian workshop on digital communications*. Springer Science & Business Media.

[14] Jayavardhana Gubbi, Rajkumar Buyya, Slaven Marusic, and Marimuthu Palaniswami. 2013. Internet of Things (IoT): A vision, architectural elements, and future directions. *Future generation computer systems* 29, 7 (2013), 1645–1660.

[15] Jon Atle Gulla, Stein L Tomassen, and Darijus Strasunskas. 2006. Semantic Interoperability in the Norwegian Petroleum Industry.. In *ISTA*. 81–93.

[16] Prageeth Gunathilaka, Daisuke Mashima, and Binbin Chen. 2016. Softgrid: A software-based smart grid testbed for evaluating substation cybersecurity solutions. In *Proceedings of the 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy*. 113–124.

[17] Eclipse IoT. 2024. Open Source for IoT. Eclipse IoT technologies power the world's leading commercial IoT solutions. https://iot.eclipse.org/

[18] Open IoT. 2024. Open IoT Org. https://github.com/OpenIotOrg/openiot

[19] IoTIFY. 2024. IoTIFY Network Simulator. https://docs.iotify.io/

[20] ISO 15926. 2011. ISO 15926 - Industrial automation systems and integration - Integration of life-cycle data for process plants including oil and gas production facilities. Geneva: ISO. https://www.iso.org/standard/50694.html Accessed on 28 Mar. 2023.

[21] ISO 25012. 2008. ISO/IEC 25012:2008 - Software Engineering - Software Product Quality Requirements and Evaluation (SQuaRE) - Data Quality Model. International Organization for Standardization. https://www.iso.org/standard/35736.html

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

[23] ISO/IEC 30141. 2018. ISO/IEC 30141:2018 - Systems and software engineering – Content of systems and software life cycle process information products (Documentation). International Standard. https://www.iso.org/standard/65132.html

[24] Eunsook Eunah Kim and Sebastien Ziegler. 2017. Towards an open framework of online interoperability and performance tests for the Internet of Things. In *2017 Global Internet of Things Summit (GIoTS)*. 1–6. https://doi.org/10.1109/GIOTS.2017.8016248

[25] Barbara Kitchenham, O Pearl Brereton, David Budgen, Mark Turner, John Bailey, and Stephen Linkman. 2009. Systematic literature reviews in software engineering–a systematic literature review. *Information and software technology* 51, 1 (2009), 7–15.

[26] Maciej Kuzniar, Peter Peresini, Marco Canini, Daniele Venzano, and Dejan Kostic. 2012. A soft way for openflow switch interoperability testing. In *Proceedings of the 8th international conference on Emerging networking experiments and technologies*. 265–276.

[27] Friedemann Mattern and Christian Floerkemeier. 2010. From the Internet of Computers to the Internet of Things. In *From active data management to event-based systems and more*. Springer, 242–259.

[28] Javier Miranda, Niko Mäkitalo, Jose Garcia-Alonso, Javier Berrocal, Tommi Mikkonen, Carlos Canal, and Juan M Murillo. 2015. From the Internet of Things to the Internet of People. *IEEE Internet Computing* 19, 2 (2015), 40–47.

[29] Rebeca Campos Motta, Káthia Marçal De Oliveira, and Guilherme Horta Travassos. 2017. Rethinking interoperability in contemporary software systems. In *2017 IEEE/ACM Joint 5th International Workshop on Software Engineering for Systems-of-Systems and 11th Workshop on Distributed Software Development, Software Ecosystems and Systems-of-Systems (JSOS)*. IEEE, 9–15.

[30] Rebeca C Motta, Káthia M de Oliveira, and Guilherme H Travassos. 2019. A conceptual perspective on interoperability in context-aware software systems. *Information and Software Technology* 114 (2019), 231–257.

[31] Srikanth Mujjiga and Srihari Sukumaran. 2007. Modelling and test generation using SAL for interoperability testing in Consumer Electronics. In *Proceedings of the second workshop on Automated formal methods*. 32–40.

[32] Mohammad Abdur Razzaque, Marija Milojevic-Jevric, Andrei Palade, and Siobhán Clarke. 2015. Middleware for internet of things: a survey. *IEEE Internet of things journal* 3, 1 (2015), 70–95.

[33] S Revell. 2013. Internet of things (IoT) and machine to machine communications (M2M) challenges and opportunities. *Final Paper, London, UK Google Scholar* (2013).

[34] Bruno P Santos, Lucas A Silva, CSFS Celes, João B Borges, Bruna S Peres Neto, Marcos Augusto M Vieira, Luiz Filipe M Vieira, Olga N Goussevskaia, and Antonio Loureiro. 2016. Internet das coisas: da teoria à prática. *Minicursos SBRC-Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuıdos* 31 (2016), 16.

[35] Luis Fernando Sayão and Carlos Henrique Marcondes. 2008. O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. *Transinformação* 20 (2008), 133–148.

[36] Sabrina Sicari, Alessandra Rizzardi, Luigi Alfredo Grieco, and Alberto Coen-Porisini. 2015. Security, privacy and trust in Internet of Things: The road ahead. *Computer networks* 76 (2015), 146–164.

[37] Valéria Martins da Silva. 2019. ScenarIoT: support for scenario specification of internet of things-based software systems. (2019).

[38] Delfina de Sá Soares. 2010. Interoperabilidade entre sistemas de informação na Administração Pública. (2010).

[39] Harald Sundmaeker, Patrick Guillemin, Peter Friess, Sylvie Woelfflé, et al. 2010. Vision and challenges for realising the Internet of Things. *Cluster of European research projects on the internet of things, European Commision* 3, 3 (2010), 34–36.

[40] Tasmota. 2024. Open source firmware for ESP devices. https://tasmota.github.io/docs/

[41] Souvik Pal Valentina Emilia Balas. 2020. *Healthcare Paradigms in the Internet of Things Ecosystem*. Academic Press; 1st edition.

[42] Leila Cristina Weiss et al. 2019. Interoperabilidade semântica: uma análise sob a perspectiva da abordagem ontológica de Willard van Orman Quine. (2019).

[43] Wireshark. 2024. The world's most popular network protocol analyzer. https://www.wireshark.org/

[44] Claes Wohlin, Per Runeson, Martin Höst, Magnus C Ohlsson, Björn Regnell, and Anders Wesslén. 2012. *Experimentation in Software Engineering*. Springer Science & Business Media.

[45] Fatiha Zaidi, Emmanuel Bayse, and Ana Cavalli. 2009. Network protocol interoperability testing based on contextual signatures and passive testing. In *Proceedings of the 2009 ACM symposium on Applied Computing*. 2–7.