Using images to extend smart object discovery in an Internet of Things scenario

Romulo Silva
Virtual University Institute
Federal University of Ceara
60451-970, Fortaleza, Ceará, Brazil
romulosn@alu.ufc.br

Windson Viana
Group of Computer Networks,
Software Engineering and Systems
Federal University of Ceara
60451-970, Fortaleza, Ceará, Brazil
windson@great.ufc.br

Paulo Filipe Dantas
Group of Computer Networks,
Software Engineering and Systems
Federal University of Ceara
60451-970, Fortaleza, Ceará, Brazil
paulodantas@great.ufc.br

ABSTRACT

The Internet of Things (IoT) represents a new paradigm in the Internet history and in the way people interact with everyday objects. Academics and several industry segments have been working over the past years to make this vision possible. It is estimated that there will be dozens or hundreds of devices simultaneously connected to the user's network (e.g., in their home) in the next years, which can make the discovery and interaction with smart object more complicated to IoT users increasingly. This work proposes a smart objects discovery approach using image recognition, which aims to make this task quicker and more selective from an user perspective. An initial assessment has shown that the proposed mechanism can reduce the discovery time in a scenario with several devices, and additionally ensure a good level of user satisfaction.

KEYWORDS

IoT, Smart object discovery, Pointing, image recognition

1 INTRODUCTION

The Internet emerged in the 60s with the primary goal of sharing information between large computers. Subsequently, the Internet connected not only machines but also people. Several technological solutions have been developed to simplify people's lives and computerize everyday tasks. Nowadays, the Internet is undergoing another paradigm shift, in which not only machines and human beings are connected, but everyday objects as well. [7] This new scenario is known as Internet of Things (IoT). Atzori et al [1] define IoT as the constant and invisible presence of a huge variety of computing devices in people's daily lives, which are capable of interacting with each other in order to cooperatively perform certain tasks. Those computing devices, which may vary terms of functionality and resources, are commonly called smart objects (SO). As examples of SOs we can mention smart TVs, Wi-Fi lamps, and connected thermostats.

Due to their growing popularity and easy market access, smartphones have become main way by which IoT users discover, connect to, and interact with SO's [5]. Smartphones can be seen as universal controls or environment browsers that work actively on the discovery of devices of interest to the user [4].

In: XV Workshop de Trabalhos de Iniciação Científica (WTIC 2018), Salvador, Brasil. Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pôsteres. Porto Alegre: Sociedade Brasileira de Computação, 2018.

ISBN 978-85-7669-435-9.

The most common smart object discovery approaches aim to discover and make all SO's within a network available to users. However, those approaches can become problematic when there are several SO's within reach. The Cisco Internet Business Solutions Group (Cisco IBSG) estimates that there will be 50 billion smart objects connected by 2020. The number of objects connected to the network can reach a mark of 6.58 per person. [3].

Under those circumstances, a more selective and appropriate smart object discovery approach is necessary to reduce the interest options of SO's in a network. The use of digital cameras – tools already available today on smartphones – can filter the options in the process of discovering and interacting with SO's. This work, therefore, aims to develop a smart object discovery approach using image recognition techniques. In order to achieve the main goal, the following specific goals were established: (i) analyze methods of interaction with smart objects currently used by mobile devices; (ii) identify smart object discovering challenges related to each method; (iii) evaluate a smart object discovery solution using smartphone cameras combined with a network-based discovery technique. Thus, this research is centered on the following question: how to cause the smart object discovery to be more selective and appropriate to the context of an IoT user?

2 INTERNET OF THINGS

The Internet of Things represents a vision in which the Internet extends to the real world and reaches everyday objects. These objects, now called smart objects, are capable of interacting and communicating with other SO's, users, and the environment in which they live. Thus, SO' are able to provide services and applications, autonomously and cooperatively [6]. IoT makes computing truly ubiquitous [10].

An important part of IoT systems is the discovery and configuration of SOs available to the users. In fact, SO discovery is a similar problem to the resource or the discovery challenge.

Service discovery is a known term from the distributed systems field. It is defined as the process of finding suitable services to perform a specific task [2]. These services may be presented in the form of available software abstractions – such as an e-mail server – or in the form of computing resources or devices (e.g., printers on a network). Dynamic Host Configuration Protocol (DHCP) and Simple Service Discovery Protocol (SSDP), Universal Plug and Play (UPnP), are examples of protocols that have service discovery mechanisms. Each service discovery protocol has characteristics that make it more suitable to a given domain.

2.1 UPnP

In our research we adopt UPnP as a starting point. UPnP[1] is a set of protocols that provide data description, easy configuration, automatic discovery of services and control networked devices. The UPnP architecture includes a combination of protocols such as HTTP (Hypertext Transfer Protocol), SOAP (Simplex Access Object), XML (Extensible Markup Language) and SSDP (Simple Service Discovery Protocol), which support integrated and seamlessly device discovery, connection, control and data exchange.

UPnP has two general classifications for devices: controlled devices and control points. A controlled device acts as a server, offering services and responding to control points requests. A control point, in the other hand, is a device that uses UPnP to consume services and control devices.

2.2 Physical Mobile Interactions

Other important aspect of our research is the user interaction with SOs. We follow some insights proposed by Enrico Rukzio et al in [9]. In that work, the authors proposed, evaluated and compared four types of physical mobile interactions using mobile devices: touch, pointing, scanning and user-mediated.

Touching is the interaction that involves an action where the user's device is put in physical contact with a device that it wants to interact with. In this type of interaction, the user needs to know the location of the object and to be aware of the possibility of performing this sort of interaction with the device. Examples of this type of interaction can be achieved using technologies such as Near Field Communication (NFC).

Pointing allows the user to interact with object by pointing to it with a mobile device. Pointing has been widely used in augmented reality applications through the mobile device camera. Applications that use pointing as a form of interaction perform visual recognition of bar codes, QR code, reference markers or patterns.

Scanning provides information from a geographical location or by connecting to a network, such as a list of available SO's in a given perimeter. One of the advantages of this approach is the possibility of interacting with SO's with no prior knowledge of their availability [9].

Finally, the user-mediated interaction is characterized by the user's action of entering some information previously provided by the SO in order to establish a connection with it, such as URLs printed in SO's that provide access to the object information and functions.

3 PROPOSAL

The main purpose of this work is to integrate and combine the pointing interaction with network based scanning mechanisms. As a result of doing so, we expect to improve smart object discovery services (e.g. UPnP) by using image recognition techniques.

3.1 Proof of Concept

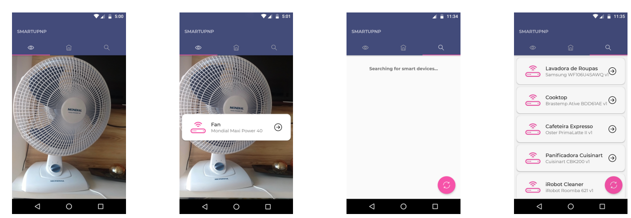
In order to illustrate the smart object discovery approach proposed by this work, a Proof of Concept (PoC) called SmartUPnP (Figure 1) was developed. This PoC consists of an Android[2] application, written in Java[3], which acts as a UPnP control point. It is capable of discovering devices within a network by using scanning and pointing physical interactions. TensorFlow[4] was used to add image classification features to SmartUPnP. TensorFlow uses neural convolutional networks to classify images, group them by similarity and identify objects within a given scene.

In order to be able to classify objects, 4809 public domain images from 29 different objects were collected. Those images were then used as inputs to create a data model with Tensorflow's MobileNets[5] architecture.

To simulate a real usage scenario of the PoC, a second Java application was also developed to simulates several smart objects that would be commonly found in smart homes, such as a refrigerator, a television and a vacuum cleaner. This simulating application also used UPnP to create those 17 devices, publish their services and receive commands sent by control points.

To implement the UPnP architecture on SmartUPnP, it was used Cling[6], which is an open source library – written in Java – with the goal of providing a programming interface compatible with the UPnP architecture layers.

When identifying an object through the camera, the SmartUPnP app searches for devices within reach that may correspond to the identified object name.



(a) Identifying an object (b) Object discovered by image recognition (c) Searching the network for objects (d) Objects found in the network

Figure 1: SmartUPnP interface

4 EVALUATION

In the presented context, the smart object discovery approaches using pointing and scanning techniques were evaluated by users. The experiment occurred in a residential setting, simulating a real use case.

4.1 Methodology

The experiment was carried out with a total of 9 participants, among university students and technology professionals. Although 9 people is a small number, researchers such as Nielsen [8] demonstrate that it is possible to find the majority of problems on a software product with only 5 users.

[1] UPnP: https://openconnectivity.org/developer/specifications/upnp-resources
[2] Android: https://developer.android.com/about/
[3] Java: https://www.java.com/about/
[4] TensorFlow: https://www.tensorflow.org/
[5] MobileNets: https://ai.googleblog.com/2017/06/mobilenets-open-source-models-for.html
[6] Cling: https://github.com/4thline/cling/

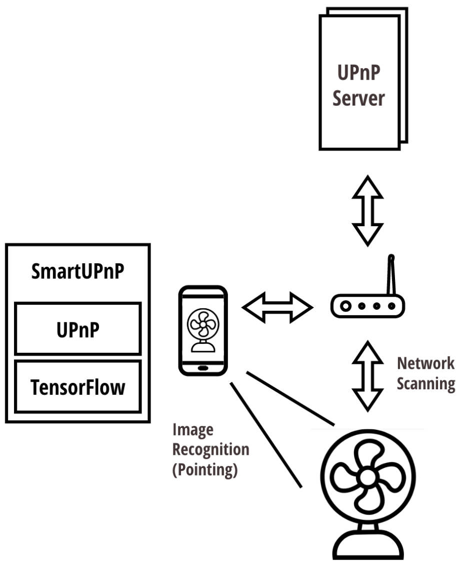


Figure 2: SmartUPnP Architecture

Four images of possible intelligent objects were fixed in a room: a table lamp, a fan, a refrigerator and a vacuum cleaner. All users were familiar with the use of smartphones. Firstly, a pre-test questionnaire was applied in order to identify the participants' prior perception of how interaction between smartphones and real-world objects could be performed.

Next, the experiment scenario was presented to each participant: a smart home with several devices connected to the Internet. Each participant was then instructed to find each of the four objects using the network search method and then the image recognition approach. The entire experiment was timed in order to collect the time taken by each user to discover each object using each of the two types of discovery approach.

After the test, another questionnaire was applied in order to evaluate the participants' experience with each of the discovery methods.

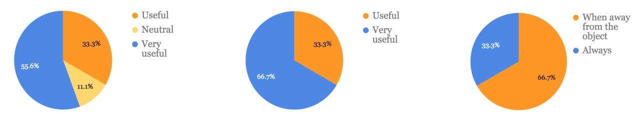
4.2 Results

4.2.1 User profile

All participants in the experiment were smartphone users and already had some knowledge about the term "smart object". 6 of them reported that they currently use or have used a SO before. All participants related SO's to devices that are able connect to the Internet or perform tasks autonomously.

When asked about their opinion on the use of smartphones to interact with SO's, 8 participants considered it very useful or useful to use a smartphone to access information about the state of an object, such as the temperature of the oven or if the room lights are on (Figure 3). When asked how useful they considered using a smartphone to control an SO, such as adjusting the brightness of an environment or locking the doors of a car, all participants considered it very useful or useful.

Participants' views on when they would use their smartphones were also requested. 6 participants answered that they would use their smartphones only when they were away from the object, while only 3 said they would always use them.



(a) Use a smartphone to obtain information of an object

(b) Use a smartphone to control an object

(c) When they would use a smartphone

Figure 3: Use of smartphones to interact with objects

Before beginning the experiment, participants were asked to indicate their preferences as to the type of interaction they would find most interesting in an IoT context (Figure 4). Only 6 claimed to be interested in using the camera to point to an object, while 8 would choose a network search approach.

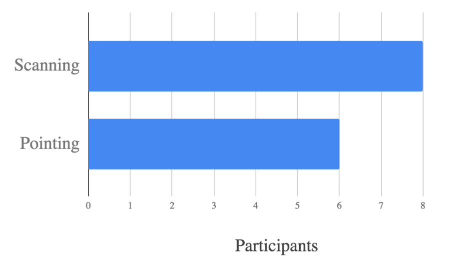


Figure 4: Participants' preferences as to the type of interaction

4.2.2 Experiments

During the experiment, all participants had some difficulties with both discovery approaches. However, after the test, they all said they were generally satisfied with the way the objects were discovered.

Using the network scanning approach, users took considerable time to locate at least one of the objects. Search times were higher usually for the first objects of the experiment, reaching 31.13 seconds. This was probably due to the fact that users were still familiarizing themselves with the application interface and the list of available connected objects. The discovery mean time using the network scanning approach was 9.51 seconds [7.53-11.49s], meanwhile the use of image recognition had an average time of 6.53 seconds [5, 22-7.84s]. That shows evidence that the image recognition discovery method cab be faster than the network scanning method in a scenario with several connected devices.

On a scale of 1 to 5, where 1 represents not satisfied and 5 represents very satisfied, the network scanning method obtained an average of 3.6 points (Figure 6). Participants argued that they found this type of discovery more suitable to use in a context with few connected objects. They would also like that the application interface provided some options for filtering the available objects. This filtering could be done by name or device type, or even by usage context, such as the user location and time of day.

On a scale of 1 to 5, where 1 is not satisfied and 5 is very satisfied, the image recognition discovery approach has averaged 4.3 points (Figure 6). Despite the long response time in some cases, and the fact that the application would sometimes mistakenly show more than one object, the participants stated that they found this approach to be faster, simpler and more natural. However, they also commented that they would like the application to be more precise in terms of identification.

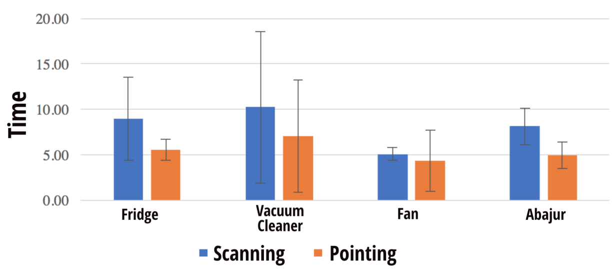


Figure 5: Smart objects discovery times

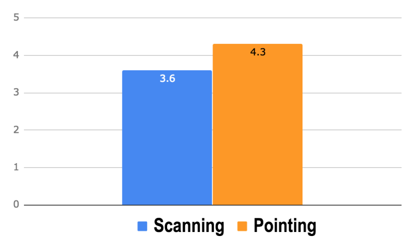


Figure 6: User satisfaction score

Participants were generally satisfied with this work proposal. From the positive comments and points of satisfaction obtained, it is possible to validate the smart objects discovery by image recognition in an IoT context. The image recognition method is as satisfactory as the network scanning method according to the users and the time performance test. The difficulties and issues found during the experiment can be addressed by improving the usability of the graphical user interface and improving the image recognition precision.

5 FINAL CONSIDERATIONS

With the rise of the Internet of Things, it is expected that there will be dozens of devices connected to the same network. In this context, several aspects related to the discovery of smart objects are still to be addressed. This work presents an smart object discovery approach that uses image recognition in order to make the discovery processes faster and more selective from the point of view of an IoT user.

A proof of concept has been developed using technologies already known and used by academia and industry (i.e TensorFlow and UPnP). This solution was based on two types of physical mobile interaction proposed by [9], scanning and pointing.

An experiment evaluated the users' perception and the effectiveness of a smart object discovery approach using image recognition. The results showed that the proposal is feasible and can be used in a real context. It was also possible to conclude that users were satisfied with the way objects were discovered using a smartphone camera. In most cases tested, in a scenario with many connected devices, users were able to discover a connected object using the camera of their smartphones more quickly when compared to a network scanning approach.

During the research, some future works were identified that could improve the proposal of this work. The limited integration between the TensorFlow image recognition and the UPnP discovery mechanism showed the need to create a software component that would more specifically support smart object discovery by image recognition.

It would also be interesting to conduct new experiments with different machine learning algorithms to find out if other approaches or set of approaches could minimize time and error while detecting objects in a scene. Image recognition algorithms are often computationally expensive and, consequently, consume mobile devices energy resources quickly. Therefore, it would also be relevant to investigate the energy efficiency of the this work and what other approaches could offer a better cost-benefit in terms of energy.

REFERENCES

[1] Luigi Atzori, Antonio Iera, and Giacomo Morabito. 2010. The internet of things: A survey. *Computer networks* 54, 15 (2010), 2787–2805.
[2] Marco Crasso, Alejandro Zunino, and Marcelo Campo. 2008. Easy web service discovery: A query-by-example approach. *Science of Computer Programming* 71, 2 (2008), 144–164.
[3] Evans Dave. 2011. The internet of things: how the next evolution of the internet is changing everything. *CISCO white paper* (2011).
[4] Nayeem Islam and Roy Want. 2014. Smartphones: Past, present, and future. *IEEE Pervasive Computing* 13, 4 (2014), 89–92.
[5] Feida Lin and Weiguo Ye. 2009. Operating system battle in the ecosystem of smartphone industry. In *Information Engineering and Electronic Commerce, 2009. IEEC'09. International Symposium on.* IEEE, 617–621.
[6] 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.
[7] NIC.br. 2014. A Internet das Coisas, explicada pelo NIC.br. (July 2014). https://www.youtube.com/watch?v=jlkvzcG1UMk
[8] Jakob Nielsen and Thomas K. Landauer. 1993. A Mathematical Model of the Finding of Usability Problems. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems (CHI '93).* ACM, New York, NY, USA, 206–213. https://doi.org/10.1145/169059.169166
[9] Enrico Rukzio, Gregor Broll, Karin Leichtenstern, and Albrecht Schmidt. 2007. Mobile interaction with the real world: An evaluation and comparison of physical mobile interaction techniques. In *European Conference on Ambient Intelligence.* Springer, 1–18.
[10] Mark Weiser. 1999. The computer for the 21st century. *Mobile Computing and Communications Review* 3, 3 (1999), 3–11.