Computer science research for the Internet of Things

Miguel Pardal
Instituto Superior Técnico
Department of Information Systems and Computer Engineering
miguel.pardal@dei.ist.utl.pt

Abstract

The forthcoming widespread use of smart things, like RFID tags and sensors, along with omnipresent wireless networks, will create an Internet of Things (IoT), where most everyday objects will be interconnected and part of an universal-purpose system akin to the Internet.

This position paper looks at the IoT in a computer science research perspective. It identifies the main challenges to address and points to a set of possible research paths. It also provides some insight into what will change in the way information systems are designed and used in the IoT era.

1. Introduction

This position paper is about the Internet of Things (IoT) in a computer science research perspective. It identifies the main problems to address and references top academic and industry players in the field.

The computer science aspects of IoT research are related to information systems, ubiquitous computing and enterprise middleware, among other topics.

1.1. Scope

The scope of the research is restricted to Computer Science and Engineering, leaving out Electronics. In other words, we are looking mostly at the software research problems and not at the hardware.

Fortunately, we didn’t have to start from scratch when trying to address such a vast field. The starting point was a CERP’s report [18] that outlines major problems and application areas using a very thorough model and a research needs list, represented in figure 1.

The scope of research is focused on items:

- 2 - Distributed decision making;
- 4 - User acceptance and privacy
- 5 - Complex interaction modeling and new RFID based interaction models;
- 6 - Standardized data model for B2B exchange;
- 7 - Application integration;
- 13 - Access control and security policies for data exchange;
- 19 - User interaction

1.2. Paper outline

The paper is organized in sections.

The first section defines the motivation for new research, describing the most promising features of the new technologies.

The second section defines an information systems model, its basic assumptions and the changes brought about by new technologies.

The third section states the most important and specific problems of Internet of Things systems.

The fourth and final section draws the paper conclusions and highlights interesting research paths.
2. Motivation

“The whole world is made of change” - The words of the sixteenth century Portuguese poet Luís de Camões still hold true today, and the information technology industry is a prime example. Few have evolved so rapidly and so driven by innovation. The Internet, as a global-scale public network, allows for new ways to cooperate and create value in an open and dynamic environment [14].

Another example of innovation is the computing devices themselves. They continue to become smaller (and cheaper) in such a way that it’s now possible to have smart things i.e. very small computers with data processing, storage and communication capabilities that can react to external stimuli [1]. Two good examples of smart things are smartcards [29] and radio-frequency identification (RFID) tags [12].

The forthcoming widespread use of smart things, along with omnipresent wireless networks, will create an Internet of Things (IoT), where most everyday objects will be interconnected and part of an universal-purpose information system 1. Some applications of the IoT are still beyond our imagination, but one thing is certain: it will introduce changes to society in an unprecedented scale, as it will impact nearly everything we do as human beings.

In the following subsections, smart things technologies will be analyzed separating the aspects of automatic identification and automatic sensing. The purpose of this analysis is the assessment of its strengths and weaknesses i.e. capabilities and limitations.

2.1. Automatic identification

Automatic Identification (Auto-ID) is a set of technologies that enable the identification of objects without human intervention.

An Auto-ID system needs, at least, two distinct components: the ID and the Reader. The ID is a device that holds an identifier and is attached to the object being identified. The Reader is a device that can communicate with an ID and read the identifier value.

The identification data set is represented in Figure 2. The core data is the identifier. The ID can also include status data. The Reader-ID relationship adds location and time attributes 2. The identifier can identify a unique class of object (e.g. a bottle of soda of brand X) or it can identify a unique instance of a class of object (e.g. bottle of soda of brand X with serial number 220706).

The ID can be a simple barcode or a smart thing, like a RFID tag.

The barcode is a 30-year-old technology in widespread use. The ID is a printed label. The Reader is an infrared scanner that needs a line of sight in a short distance with label to read the printed barcode. The key features of barcodes are simplicity and low-cost. According to GS1 estimates [10], an average of 6 RFID is considered the state-of-the-art technology for Auto-ID. The ID is an electronic tag. The Reader sends a radio-frequency request signal to the tag and receives a response signal containing data. The key features of RFID are instance identification and longer reading distances and no line-of-sight requirement between the Reader and the ID. RFID is very appealing to applications like supply chain management [3]. However it still has to achieve widespread adoption due to costs and some technical obstacles.

Barcodes and RFID are further explained and compared by Hunt et al [12].

2.1.1. RFID trade-offs RFID is not a single technology, but a suite of technologies. The choice is determined by the application purpose, and must be decided after a thoughtful cost-benefit analysis.

There are three types of tags:

- Passive (battery-less);
- Semi-passive (battery-assisted) - allow improved reception;
- Active (battery-powered) - can have additional processing power.

Passive tags are the least expensive. Active tags are the most expensive.

Regarding low-cost RFID, there is a constant struggle between: cost, range and functionality. A rule of thumb is that you can choose, at best, two of them [21].

This state of matters won’t change in the foreseeable future due to physical limitations that are hard to circum-
2.2. Automatic sensing

Automatic sensing extends automatic identification with additional data attributes that are measurements of the ID’s surrounding physical environment, e.g. temperature, air pressure, etc.

The sensor data set is represented in figure 3.

Sensors enable response to changes in physical environment [1].

2.3. Edgeware

Auto-ID and Sensors extend the information system’s reach to the real world. However, to take advantage of their capabilities, they have to be deployed in disperse and remote locations, connected through a network of readers and other devices. The software required to operate this network at the edge of the enterprise is called edgeware. Its main functions are data filtering (to avoid overloading), interoperability of heterogeneous devices (that need to work together) and integration with the enterprise’s information systems.

Beyond the edge of one enterprise, lie other enterprises. Edgeware will be an important part of Business-to-Business (B2B) systems. For instance, in an automatic supply chain, there are interactions with business partners triggered by significant business events, e.g. "a rack has arrived so request invoice to partner". Systems will have to react to calls triggered by the partner systems, human operators and by changes in the physical world.

The experiences with EDI have demonstrated that even with a relatively rigid format standard for B2B message management, trading partners must coordinate their individual implementation of those standards [17].

Although individual parts of Auto-ID and Sensors systems can look simple, the overall system can be quite complex [21].

2.3.1. Standards Open standards play a central role in the adoption of any technology and that is also the case with Auto-ID and Sensors. Currently the most developed standard for the edge is the Electronic Product Code (EPC) and the EPC Architecture Framework [25], represented in figure 4.

The EPC framework covers tags, readers, data repositories, internal and B2B applications. A key component is the Object Naming Service (ONS), a global look-up system similar to the Domain Name Service (DNS) of the Internet. Using ONS it’s possible to relate any tagged object to the issuer of its EPC code.
The EPC code is a globally unique number associated with an RFID tag, typically 96 bits long. It’s formed by a header, a company identifier, an object class identifier and a serial number.

EPC standards govern interfaces, not implementations, so there will still be a lot of competition room for implementers.

3. Information systems model

All information systems are built with a set of assumptions. Some of them are documented explicitly in a model, but many others are implicit in context. The IoT introduces enough disruption that justifies looking again at the fundamentals and see how and if they change in this new context.

3.1. Data, information and knowledge

The semantics of the words data, information and knowledge is often overloaded with overlapping meanings. In this paper, we use the following definitions. Data are sequences of symbols that represent facts or events. Information is a set of data with a human interpretation context. Knowledge is demonstrated by people when they use information to produce new information or to perform specific tasks, entailing learning, previous experience and sometimes creativity.

3.2. Definition of information systems

An information system is composed of several parts that work together to collect, process, store and give access to information [14].

Information systems are used by people and organizations, to support decision-making, coordination, control, analysis and visualization of work processes.

In a certain sense, people can be seen as part of the information system, as they also consume and produce information.

3.3. Information Systems in the World

Humans have long relied on information keeping systems to support their everyday activities. With computer technologies, bigger and more sophisticated systems are possible, with automated processing and communication.

3.3.1. Types of information systems

Figure 5 shows how different types of information systems can be used across an organization.

Information systems are categorized by Laudon [14]:

- Operational-level systems

- Knowledge-level systems

- Management-level systems

- Strategic-level systems

When properly aligned in an Enterprise Architecture, all these types of systems play a role in the organization, assisting its people and their business processes and integrations with partners and clients.
3.3.2. Perspectives of an information system

Information systems are complex entities. To cope with complexity, architects, like John Zachman [28], look at them in perspectives: conceptual, logical and technological; as represented in figure 6.

Figure 6. Information system’s perspectives

We start from a conceptual view where the business requirements are defined i.e. what the system is expected to achieve. Then, someone intelligent has to analyze the problem and propose a solution i.e. translate from problem space to solution space. The resulting specification is represented using formalisms like objects, relational databases, etc. Finally, someone intelligent has to detail and adapt the specifications to implementation artifacts, that can be executed automatically by a computer. This is usually called programming. This process is constrained by weaknesses and leverages the strengths of existing technologies.

This whole process is iterative in nature, as requirements, specifications and implementation artifacts evolve over time and as people’s perception of the problem and solution changes.

3.4. The World in the Information System

The implementation artifacts of an information system include a representation of the world where the system operates. This data domain is constructed and maintained through a stream of inputs. The system has a control loop [27] where the "sensed" data is continuously used to update the world representation and used to make decisions and to "control" other parts of the system. This sensing loop connects physical objects to virtual objects, as represented in figure 6.

People are an essential part of the systems, as they are the ones who make sense of it. However, Auto-ID and Sensors tighten the sense-control loop, making the connection between both worlds more up-to-date and more independent of human intervention and interpretation. This poses a series challenges that must be address in IoT information systems. Some of these challenges are presented next.

4. Challenges

This section states the most important and specific challenges in IoT information systems.

4.1. Intrinsically distributed systems

The IoT is intrinsically a distributed system. Its sheer scale makes it impractical to look at it as a centralized system. There isn’t a single shared domain data, but multiple and fragmented views and there are also different points of failure [26].

The IoT can be considered a network, but its addressing will be limited in the sense that although every part will be connected, not all parts will be connected all the time.

The IoT will be a heterogeneous system at the device and network levels.

These characteristics should be recognized as part of the solution and not as part of the problem, so that expectations on requirements for IoT information systems can be reasonable.

4.2. Dealing with uncertain data

The issue of distribution also impacts IoT databases. These databases can be viewed as a cache of reality that will be out-of-date by default. The problem however goes beyond delays; in most cases, the system will have to operate using uncertain data due to sensing and interpretation errors.

To deal with the intrinsic distribution and the uncertain facts, IoT applications will need different data-access modes.

One data-access mode will be the event filtering and collection. Data flows upward, from readers to information
systems, in a hierarchy. Only events considered important are propagated to the level above.

Another data-access mode will be data on-demand. Here the need for data goes downward, from information systems to the readers, because there is something important being decided and the confidence level in data is not sufficient, so it must be double checked. This may entail integration on-demand, because the involved systems may be different and not usually connected. This requires very flexible middleware architecture.

4.3. Real-time requirements

A real-time system has time as a key parameter. The system is expected to meet deadlines with precision (within a limited time range) or else it is considered to fail. Usually there is the distinction between hard real-time (e.g. industrial control systems) and soft real-time (e.g. multimedia streaming). In the latter, missing an occasional deadline can be acceptable [23]. Some IoT applications requiring data on-demand will have soft-real time requirements.

4.4. Recognizing business events

The purpose of event filtering and collection is recognizing relevant business events in the midst of a massive stream of ordinary events.

Auto-ID and Sensors give literal “what”, “when” and “where” data. The problem with these is that “what” is a plain numeric identifier, the “when” is a time value relative to a local clock, and the “where” is a reader identifier or network address. None of the previous is particularly useful for business purposes.

“Identifier 061278 observed at time -1939032000 by reader with identifier 181177 at address 192.168.3.23”

Business events need business “what”, “when” and “where” along with “who” and “why”. Using these, appropriate actions can be triggered.

“Order 2333 from supplier Acme has arrived at 2008-07-22T12:00 at store 22 as part of instance 104 of the Stock Replenishing process”

The challenges here is how this all can be achieved.

Names are used to attach meaning to plain identifiers [24], so a naming service is a mandatory requirement. In IoT applications the naming service will have to be public. This entails political issues about who controls such a service, similar to the ones existing in the Internet’s Domain Name Service (DNS). Probably in operation it will be seldom necessary to look up each identifier on a central name repository. A look-up will only be used for exceptions i.e. unrecognized products and, in most cases, recognizing the object class will be sufficient.

4.5. Security

The security concerns in IoT applications are distinct in “object tagging” and in “people tagging” and solutions will have to be different for each. Figure 8 lists applications fields and makes this distinction [2].

Figure 8. CERFIDs RFID applications reference model

The protections to put in place will depend on the value of the transactions. For instance, a supply chain requires the identification, authentication, and authorization of each participant in the system. The levels of access for different stakeholders will have to change dynamically under changing circumstances [17]. Trust schemas will play a very important role in all of this, as mediated trust will be a requirement in most applications cases [9]. Figure 9 shows a generalized trust schema, with a trust chain of two or more trusted third parties.

Figure 9. Generalized trust schema
When people are being referenced through tagging, personal privacy becomes the most important security issue. The consumers' interactions with RFID labeled products can be a source of data about the consumers' behavior. Due to the invisible reading, it can occur without explicit consent [15]. In this context, other security mechanisms must be limited somehow.

4.6. Smart (enough) things

The way IoT information systems will be designed will greatly depend on the capabilities of the smart things being deployed and, in turn, these depend on cost.

The challenge here is finding the right cost-benefit balance that makes sense for an application. This has mostly to do with the value handled by the system. Intuitively, we can say that things will have to be smart enough to achieve the application’s purpose.

It would be very helpful if there were frameworks for assessing this issue and assisting the decision to opt for smarter or less smart things.

5. Conclusion and research paths

First we present a brief conclusion of the current IoT situation and then present possible research paths to follow.

5.1. Conclusion

Before the IoT can become a reality, there are several requirements to be fulfilled. The following are just some of the most important [4]:

- Commonly agreed methods of communicating and operating with support for international standards (e.g. RFID and network standards);
- Agreement on ownership of data resources;
- A pervasive federated network in which unregulated personal area and local area networks can interoperate with and through regulated communication services;
- An evolutionary strategy for accommodating network, mobile and wireless systems with accommodation of legacy systems and regulations.

In this paper we presented a set of challenges that will have to be effectively overcome before these requirements can be fulfilled. Next we propose possible research paths.

5.2. Research paths

5.2.1. Composite smart things One RFID tag technology doesn’t suit all uses. The same can be said for Readers and almost other parts in an Auto-ID and Sensors system. There is also a diversity of networks that can be used to communicate. If the IoT is going to be a universal purpose system, then these limitations have to be abstracted.

An interesting research path would be to aggregate a set of physical tags and sensors into one composite smart thing. This way, the system would differentiate between the logical thing and the different bindings to a physical reality. The same principle could be applied to create an overlay model for the networks.

Currently there are some hybrid approaches underway, like bar code printers with RFID module embedded in them that allows them to work as both an RFID label and also as traditional bar code [22].

This kind of approach could be initiated by creating frameworks for the development of smart things applications. First we could create distinct frameworks for distinct technologies and then consolidate them with module reuse and concepts generalization.

5.2.2. Visible computing, invisible computers The main idea in this research path is having information systems show their internal state in ways easily understandable by humans (e.g. traffic signs, audio feedback) while at the same time, the computers would disappear into everyday objects. For example, Readers could “disappear” into container objects like shelves, boxes, tables, etc.

This “disappearance” could also be helped by the decoupling of the tag-reader relationship itself, by separating the power and signal stages, e.g. a light bulb would give power and a single centralized reader would collect all data [21].

5.2.3. Continuous understanding of the world Auto-ID and Sensors enable a continuous sensing of the world. However the understanding of the world to extract business meaning is a different matter.

A research path would be to explore the concept of a blackboard [5] that would keep an up-to-date representation of the state of the world. This board would be continuously and asynchronously updated and annotated by different contributors, each with a unique perspective that could complement the understanding.

The different perspectives could result from different locations, capabilities and available resources. For instance, some contributors could refer to basic models of behavior (e.g. rules in the physical movement of objects) to derive new assertions [20].

Some of the board annotations would have to be probabilistic assertions to provide a means for dealing with uncertainty.

The board implementation will have to be distributed and will need data merging mechanisms akin to long-running.

---

3 This concept is inspired by the continuous understanding of spoken dialogue, by James F. Allen of the University of Rochester.
There is already some research underway concerning distributed data models and information fusion [16].

5.2.4. Society of things Instead of designing IoT systems with a grand vision of how everything should work, perhaps it would be preferable to create a limited set of rules and then leave the smart things behave as agents, with their own sensors, actuators and goals [11]. An analogy for this concept is human society’s laws, where citizens have rights and obligations.

This approach could allow for interesting mash-ups of physical and virtual behavior. For instance, the presence of the object in a physical room could give it access to a specific set of data services.

Another possible advantage of this approach is that it provides a way to achieve a universal purpose system without having to deal with all the complex details up-front.

5.2.5. Business events detection and handling Auto-ID and Sensor systems need to translate low-level events to business events and use them to trigger orchestrations, because different data needs to be forwarded to different applications and data stores [13]. A way to start mapping these data flows are high level collaboration diagrams, where the participants, collaborations, responsibilities are outlined.

As a general rule, only state changes are relevant in Auto-ID and Sensor systems [17].

To do these kinds of processing, a generic, efficient, rule-based engine is required [7].

There are working applications in the Financial sector that also have to handle high throughput, online tasks. Palmer’s strategies for financial applications [17] can be applied to Auto-ID and Sensors:

- Process data close to its source;
- Cluster various pieces of data into logical events;
- Utilize data concentrators;
- Cache contextual information;
- Federate data locations for more efficient distribution of data among sites;
- Continuously filter data events;
- Automate exception handling.

5.2.6. REST approach REST stands for Representational State Transfer [8] and is about the naming and addressing of a information system using URIs (Uniform Resource Identifiers).

The main distinction of a REST approach versus a functional approach is that in REST instead of having a variable set of “verbs” (operations) we have “names” to represent data entities and a fixed verb set, like HTTP’s GET, POST, etc.

REST is promising as a way to address the IoT system. Some of the URI parameters would give access to indexed snapshots of world using for time, location and other meaningful properties.

REST can also be easily adapted to realms [27], a hierarchical layout for the system, to allow better scalability.

5.2.7. SOA approach When creating a logical perspective for IoT information system, services appear to be a good choice as modeling formalism [4].

Service-Oriented Architecture (SOA) core concepts are: autonomy, loose coupling, abstraction and contracts [6]; and exploring them might be the path to achieve integration on-demand for data access.

So far, most SOAs have been built under the implicit assumption that service consumer’s are people or other information systems or people. These can be called “downcalls”, coming up from the presentation layer down to the underlying systems. Extending SOAs to be sensor-enabled means that they will have to handle “upcalls”, coming from events sensed in the edge and rising up to the information system.

This is also a very interesting research path to follow.

References


4 Even though the modeling of the whole system is done with services, this doesn’t entail that all of them will exist as callable Web Services. Probably most of the edge layer services won’t even be practical.


