RFIDays 2008

Workshop on Emerging Technologies
for Radio-frequency Identification

Book of Proceedings

Gaetano Marrocco Editor

Jointly organized by

University of Roma "Tor Vergata"
CNIPA
IEEE – Italy Section

Internal Report RR-08- 69
Dipartimento di Informatica Sistemi e Produzione
University of Roma "Tor Vergata"

June, 2008

Document edited on June 2008 by

Gaetano Marrocco
University of Roma Tor Vergata
Dipartimento di Informatica Sistemi e Produzione
Via del Politecnico, 1
00133 Roma (Italy)
Tel + 39 06 72597418
e-mail marrocco@disp.uniroma2.it

Please cite as:
"Emerging Technologies for Radio Frequency Identification", G. Marrocco Editor, *Research Report RR-08-69*, Dipartimento di Infomatica Sistemi e Produzione, Università di Roma Tor Vergata, Jun2 2008

RFIDays 2008

Emerging technologies for radio-frequency identification

12-14 May, 2008
University of Roma Tor Vergata,
Roma, Italy
www.disp.uniroma2.it/alab

Organized by

Università di Roma Tor Vergata

CNIPA
Centro Nazionale per l'Informatica nella Pubblica Amministrazione

IEEE
Italy Section
VT-COMM Chapter

Supported By

Europäisches Patentamt
European Patent Office
Office européen des brevets

Symposium Organizers

Gaetano Marrocco, *University of Roma Tor Vergata*
Indra Macrì, Giovanni Rellini, *CNIPA*

Scientific Committee

G.F. Italiano, *University of Roma Tor Vergata*
E. Frezza, *CNIPA*
F. Bardati, *University of Roma Tor Vergata*
M. Orefice, *Politecnico di Torino*
F. Vatalaro, *University of Roma Tor Vergata*

Sponsored by

IBM

Reply
Living network

CAENRFID
THE ART OF IDENTIFICATION

Microsoft

ELSAG DATAMAT

RadioLabs

CONTENTS

Preface

Identification of things and processes by Radio Frequency (RFID) is quickly growing up as the emergent technology in logistics, security and bio-engineering. Different kinds of data may be wireless transmitted from the local query unit (reader) to the remote transponder (tag) that includes the antenna and the microchip transmitter. The best established standards are related to low frequency (HF, 13.56 MHz or less), but the Ultra High Frequency UHF and microwaves (870-960 MHz, 2450 MHz) are the most promising technology and a great interest is originating by Ultra Wide Band (UWB: 2.4–4.8GHz 6-8.5GHz) platforms for detailed localization and high immunity to multi-path.

In addition to commercial applications, where tags replace or complement barcodes, new interesting convergences between RFID technologies, personal telecommunication networks and distributed sensors, are currently envisaged in localization and ubiquitous processing platforms. The most innovative research is related to affordable interrogation within complex environment, to the identification of metal objects, to the efficient scavenging of environmental energy, to the development of miniaturized active system with high autonomy, to low cost sensorial tags and to the biomedical telemetry.

In the Italian context there are both Academic and Industrial organizations performing research and development activities in the field of radio frequency identification. The fast diffusion of this technology, and the related heavy scientific and commercial competition, make really important to encourage meeting opportunities between different skills, with the purpose to share experiences and finalize the most advanced research to applications and products.

On May 12-13 2008 the University of Roma Tor Vergata, Italy, in close collaboration with CNIPA, the Italian Agency for Digital Government, organized the Workshop **RFIDays-2008** on the Emerging Technologies for Radiofrequency Identification. By emphasizing the natural multi-disciplinary nature of RFId context, this two-days event offered a unique review of ideas, algorithms, technology and experimentations, coming from Electronics, Electromagnetics, Telecommunications, Computer science and Logistics and proposed an extraordinary interactions between Universities, research laboratories and companies, stimulating new interests and synergies.

The workshop had an Italian connotation, with about 25 technical presentations and an audience of more than 150 participants. Moreover some foreign speakers illustrated, in an international prospective, the new development trends, the funding and the patent opportunities of the RFID technology.

This proceeding collects many of the most valuable contributes, here organized into two parts respectively concerning new technologies for reader and tag design, and the development and implementation of algorithms for system-level applications with attention to security issues.

The complete collection of the speech slides may be freely downloaded at

www.disp.uniroma2.it/alab

* * *

After a detailed tutorial on RFID by *P. Talone* (U. Bordoni Foundation, Italy), *F. Frederix* (DG INFSO, European Commission) reviewed the European concerted efforts on RFID, the running projects, the RFID thematic networks and the Funding policy for 2009-2010. *M Ricciardi* (European Patent Office) gave a patent-view to RFID technology with attention to still open opportunities for University and Industry.

Reader and tag technology - The most advanced tag design technology brings together low-cost electronics, multi-standard identification with sensing capabilities and ultra-wideband operations. *G. Iannaccone* (University of Pisa) showed how the read range of passive UHF RFID tags may be sensibly improved by a proper optimization of the chip's analog front-end, with the real possibility to keep the chip power consumption below a fraction of microwatts. Compact transponder embedding integrated radiating elements for both HF and UHF services, particularly suited for e-cards were proposed by *A. Toccafondi* (University of Siena). *E. Verona* (CNR, Italy) showed how the electro-acoustic devices (SAW) may be integrated with HF and UHF antennas to achieve passive RFID systems able to send back physical information about temperature, pressure and chemical events. *G. Marrocco* (University of Roma "Tor Vergata") addressed the topic of UHF tags for human body sensing. Ad-hoc antennas, able to host sensors, permits to maximize the read-range when placed onto the body and to achieve multi-band operations with relevant application to the monitoring of human body movements. A new physical approach moreover promises to use the antenna itself as a low-cost sensor of the target's changes. A complete active RFID system was proposed by *G. Biffi Gentili* (University of Siena) with dual-frequency protocol having as main advantage a wide range of activation and very compact size and the possibility to easily host sensors. Preliminary studies on UWB tags, promising high data-rate capabilities and precise localizations, were presented by both A. Toccafondi and G. Biffi Gentili.

Recent advances in reader technology were reviewed by *I. Kipnis* (Intel Corp.) New all on-chip reader electronics is now available from Intel and already implemented into commercial readers, offering a high integration degree within portable low cost devices (*G. Grieco*, Caen, Italy). This technology promises a broad diffusion of the reader-tag paradigm in everyday life. A versatile reader scanning technology were introduced by *M. Orefice* (Polytecnic of Turin) concerning the use of antenna arrays with steerable beams. This system has the capability to dynamically change the radiation patterns, in the UHF and microwave bands, and selectively interrogate particular portions of the environment, or to cover a large area without moving the read station. These features could greatly improve the radio interrogation of crowded environments.

The true performances in term of bit error rate of a reader-tag system within a real scenario including scattering objects were discussed by *C. Piersanti* (University of Bologna) who showed some relevant effects and trade-offs in transponder and reader design. In particular it was demonstrated that the true interrogation region is non-uniform and the communication between tag and reader could fail even at distances lower than the maximum interrogation range due to multipath propagation.

Distributed Systems - The application of pervasive computing to the RFID paradigm, and its integration with existing communication networks was the subject of four contributions. RFID-based indoor localization systems were introduced by *F. Mazzenga* (University of "Roma Tor Vergata") who discussed simple localization algorithms combining both active and passive devices, with particular emphasis to multi-frequency configurations. *M. Mamei* (University of Modena and Reggio Emilia) described how some animal collective intelligence, based on the pheromone multi-agent interactions, may be implemented by an RFID network. Humans and robots could spread/sense pheromones by properly writing/reading RFID tags that are likely to populate our everyday environments. Such pheromones can encode and describe application-specific information useful to achieve object tracking and to provide context-aware information. *A. Sciarappa* (Istituto Superiore Mario Boella, Italy) presented several middleware solutions to implement the "Internet of Things" for dynamic contexts in applications of ubiquitous communications, pervasive computing and ambient intelligence. Finally, *A Moroni* **(**University of Roma "Sapienza") provided an introduction to Near Field Communication, and its integration the GSM/UMTS mobile networks and Wi-Fi towards intuitive, safe and contactless transactions.

The issue of RFID security is a key-concern for a safe and reliable diffusion of this technology. *G. Me* (University of Roma "Tor Vergata") reviewed the possible computer attacks and spamming relevant to RFIDs, and discussed the opportunity of tags with cryptographic codes. Concerning the same issues, *M. Rebaudengo* (Polytechnic of Turin) told how to improve the integrity and the safety of food all along the supply chain by introducing cryptographic procedures within the reader-tag communication protocols, taking advantage of the read/write memory of the tag.

The workshop concluded with the contributes of Industry (*G. Violante*, ElsagDatamat; *L. Di. Pace*, IBM; *M. Caprino*, Reply; G. Zanelotto, *Microsoft*) about system integration upon final users.

Section 1

Reader and tags technology

UHF-HF Integrated Transponder for RFID Applications

Alberto Toccafondi, Cristian Della Giovampaola, and Paolo Braconi

***Abstract*—In this paper an integrated RFID transponder operating at both UHF (868 MHz) and HF (13.56 MHz) bands is presented. It combines in one ISO 7810 ID-1 card both a UHF meander dipole antenna and an ISO 15693 commercial tag arranged in two separate sections. The design of the meander dipole has been optimized taking into account the presence of the coil antenna of the HF tag. In order to allows a better control of the input impedance of the antenna, three EM coupled loading bars have been introduced. It is found that the geometrical parameters of the antenna such as loading bars spacing, width and meander step can be designed to obtain a very good conjugate match between antenna and UHF tag chip. The proposed UHF tag antenna also provides small size, proper bandwidth and omnidirectional pattern in a plane perpendicular to the antenna plane. Numerical simulations and measurements have shown very good overall performances of the integrated transponder.**

***Index Terms*—RFID, UHF Tag antenna, HF Tag, Integrated Transponder.**

I. Introduction

RADIO Frequency Identification (RFID) is a recent outstanding technology which permits the identification and tracking of objects using wireless data exchange between a reader station and small transponders or tags, located on the objects to be identified. Passive RFID systems are typically used in all those applications that require large-scale low-cost tagging.

RFID systems operating at HF band (13.56 MHz) are widely use in the area of ticketing, personnel access control and object tracking. These systems employ the near field inductive coupling to transfer energy and binary data between a reader and the tags. They are characterized by an excellent immunity to environmental noise and electrical interference and exhibits a minimal shielding effects from adjacent objects and the human body. However, the usually small dimensions of the tag and the limits on the maximum permitted magnetic field strength imposed by the european regulations typically limit the maximum attainable reading range to about one meter.

The continuous growing of the market demand has promoted intense research on passive RFID systems, especially at UHF band, where the possibility to obtain middle to long range wireless links is joined with a good reliability of the communication. Passive RFID systems at UHF (866-869 MHz, European band) and microwave (2.45 GHz) bands use the modulated scattered technique to establish a radio link between the reader and the tags. Here, the reflected signal from the tag is modulated by an integrated microchip (IC) directly connected to the antenna. As a consequence, RFID transponder performances are strongly affected by the frequency-dependent impedance match between antenna and IC. However, in these frequency ranges it is possible to resort to antennas with far smaller dimensions and greater efficiency than that employed at frequency ranges below 30 MHz. Depending on the IC sensitivity and on the tag antenna performances, typical ranges of 4-6 m can now be achieved using passive UHF backscatter transponders. Typical fields of application of these systems are in logistic as well as in access control services.

The issue of obtaining a single card with dual standard HF-UHF operation is of considerable interest in all those multi-service applications, where a high level of interoperability between different systems is required. For this purpose, in this paper an integrated RFID transponder operating at both UHF (868 MHz) and HF (13.56 MHz) bands, is presented. It is composed by a ISO 7810 ID-1 card where both a UHF meander dipole antenna and an ISO 15693 commercial tag are arranged in two separate sections. For the HF section a commercial ISO 15693 tag with appropriate dimensions is selected. The UHF antenna is designed and optimized to provide small size, proper matching with the tag IC and omnidirectional pattern in a plane perpendicular to the antenna plane. The design process of the UHF antenna is conducted taking into account the presence of the HF-coil antenna. Moreover, the use of three loading bars, electromagnetically coupled with the meandered dipole, has allowed to improve the impedance matching, without significantly increasing the antenna size. The antenna input impedance is designed for conjugate-matching to high-capacitive input impedance of a commercial IC. UHF antenna characteristics including return loss with respect to the chip impedance and radiation pattern are investigated. A prototype of the integrated transponder printed on FR4 substrate has been developed and measurements has been conducted at UHF frequencies. Experimental results have shown a good overall performance of the transponder, demonstrating its possibility for use in various RFID applications.

II. Transponder structure

The basic idea of this work is to accommodate on a single ISO 7810 ID-1 Card two different passive tags operating at HF and UHF bands, with their own radiating structures. Typical dimensions of a ISO 7810 ID-1 Card are 85.72 x 54.03 mm which represent also the available space to integrate the two tags. In view of this, it is important to conceive the final structure allowing for each tag the necessary space that

Alberto Toccafondi and Cristian Della Giovampaola are with the Department of Information Engineering, University of Siena, Siena, 53100 Italy
Paolo Braconi is with Wavecomm S.r.L., Siena 53100 Italy.

provides the best tradeoff between occupied space and desired operational requirements. Moreover, it is also important that the final layout maintains the typical RFID requirements of flexibility and low cost for mass production and, at the same time, let the performance of each single tag may not be significatively influenced by the presence of the other. The above considerations lead us to select a transponder layout in which the UHF and the HF tag are allocated in two separate sections, each one corresponding on one half of the available card space. It is found that the proposed arrangement provides a satisfactory compromise between overall tags performances and mutual decoupling, with respect to other conceivable layouts.

Low cost for mass production can be achieved designing UHF uniplanar single-layered dipole-type antennas with reduced size, so that the complete transponder could be realized taking advantage of common, and low cost PCB techniques. A simple way to reduce the size of a dipole antenna is to meander its arms, thus obtaining a more compact antenna. It is found that this type of antenna allows to decreasing size maintaining reliable radiation characteristics [1]. To maximize the efficiency and power transfer of the tags, a correct conjugate impedance matching between antenna and the tag IC, has to be obtained. Passive ICs are intrinsically highly reactive because of the necessary power to bias the IC front-end, which is delivered through electromagnetic coupling. In this case, due to the low input resistance and high capacitive reactance of the IC, antenna with a conjugate-match to the IC have to be designed with a low resistive and highly inductive input impedance. However, reducing the size of the meander antenna often makes it difficult to match to high-capacitive tag IC input impedance. Recently, it has been found that the presence of a load bar allows to better control the antenna input impedance [2] and to improve the antenna gain. Here, in order to provide a better match to the chip input impedance without increasing too much antenna dimensions and loose meander symmetry, both a loading bar parallel to the vertical branches of the meander line dipole and two more strips near the final branches of the antenna arms has been added [3]. The antenna layout is shown in the left part of Fig. 1 where the geometrical parameters of the antenna are also indicated. The UHF antenna is 40 mm x 52 mm sized, and it has been designed to fit the maximum allowed dimensions of 42 mm x 54 mm. The RFID chip connected to the prototype antenna is a Philips Ucode with a packaged chip input impedance of $Z_c = (20 - j500)\Omega$.

In RFID systems operating at HF band (13.56 MHz), the tag coupling element is typically constituted by a planar coil antenna that may be realized at low cost using etching or common screen printing technique. In this work, since the wide availability of commercial inlays, for the HF section it is chosen to resort to a commercial ISO 15693 Tag constituted by a copper wire deposited antenna coil connected to a Philips I-Code RFID chip. This commercial Tag is 41 mm x 52 mm sized and it can be easily allocated on the planned one-half card space. However, to take into account the presence of the HF antenna coil in designing the UHF tag, a 3-turn rectangular loop coil is considered in the numerical simulations. The final layout of the UHF-HF integrated transponder is shown in Fig. 1.

UHF ANTENNA DESIGN AND ANALYSIS

The UHF antenna has been designed and optimized using a commercial electromagnetic simulation software. Among all the geometrical parameters considered in the design process, the key ones are the vertical d and horizontal l length branches of the meander line and the distances g_1 and g_2 between the antenna trace and the loading bars. The optimal geometrical configuration of the antenna has been mainly determined by analyzing the family of curves associated to these parameters.

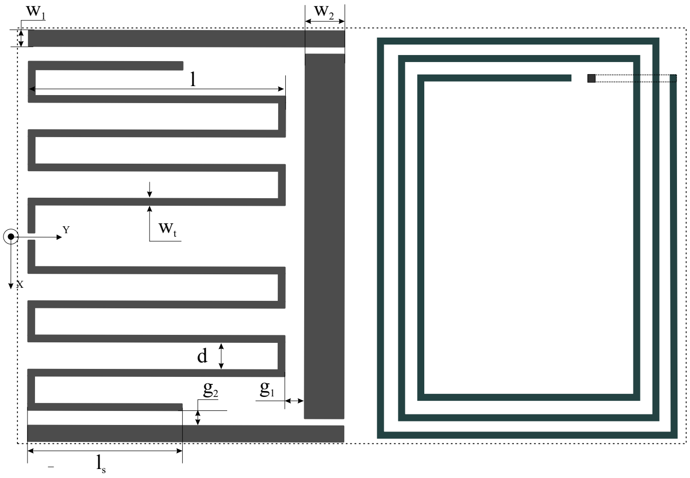


Fig. 1. UHF-HF integrated transponder

A. Return loss

As mentioned above, a proper impedance match between the antenna and the IC is very important in order to maximize tag performances. In RFID tags, the antenna is directly connected to the chip, which typically presents a high-capacitive input impedance. In order to obtain a better conjugate impedance match, it is important to minimize the Kurokawa's power reflection coefficient $|s|^2$ [4]- [5], where s is defined by

$$s = \frac{Z_c - Z_a^*}{Z_c + Z_a} \tag{1}$$

where Z_a is the complex antenna impedance and Z_c is the complex IC input impedance The maximum reading range D_{max} of an RFID system is directly affected by the power reflection coefficient and can be calculated using the Friis free-space formula as

$$D_{max} = \frac{\lambda}{4\pi}\sqrt{\frac{P_{erp}G_r}{P_{th}}p(1 - |s|^2)} \tag{2}$$

where λ is the wavelength, P_{eirp} is the equivalent isotropic radiated power transmitted by the reader, G_r is the tag antenna gain, P_{th} is the minimum power required to activate the chip, and p is the polarization loss factor.

In fig. 2 the return loss as a function of the maximum achievable reading range, for a given antenna gain G_r, is reported. Calculations are made for a chip sensitivity $P_{th} =$

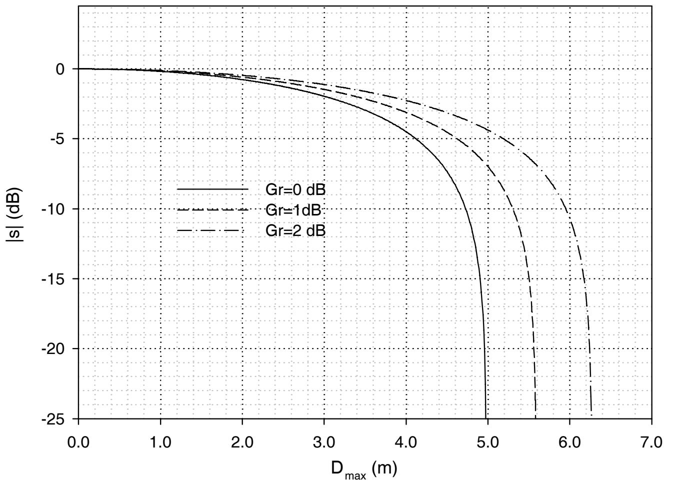


Fig. 2. Return loss as a function of the maximum achievable reading range for different tag antenna gain ($P_{th} = -13$ dBm, $p = 0.5$ $f = 868$ MHz, $P_{eirp} = 35.16$ dBm

-13 dBm, frequency $f = 868$ MHz, polarization loss factor $p = 0.5$, and for the maximum allowed P_{eirp} by the European regulations ($P_{eirp} = 35.16$ dBm). It is shown that for a given tag antenna gain, the reading range may not be significantly increased by increasing the return loss over about $-15 \div -17$ dB. These return loss values can be considered the impedance match requirements in order to obtain good tag performances.

The design process involves as a first step a major tuning of the antenna input impedance obtained by modifying the length of both vertical d and horizontal l length branches of the meander line. In Fig. 3 it is shown the return loss of the conjugate-matching to the high-capacitive IC input impedance as a function of the length d for a copper antenna printed on a FR4 substrate. It may observe that increasing d corresponds to both increase the tag resonant frequency and improve the return loss minimum.

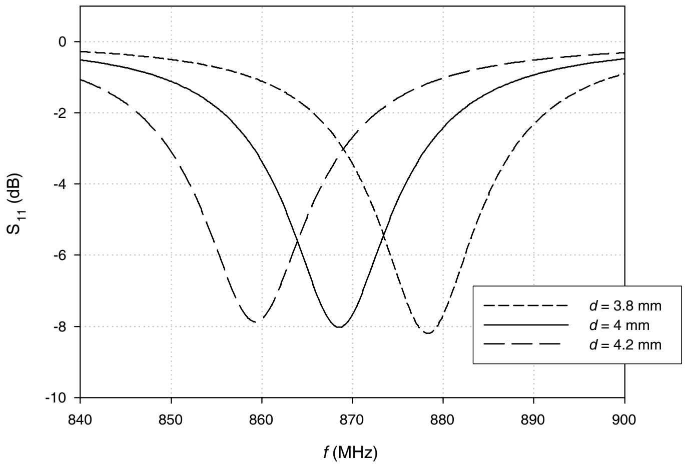


Fig. 3. Effect of d variations on the return loss

A similar behavior is obtained by the analysis performed on the length l of the meander trace, and illustrated in Fig. 4; increasing l corresponds to decrease the tag resonant frequency of the antenna. However, in this case, the minimum of the return loss is more emphasized for lower frequencies and hence for higher lengths of the vertical branch.

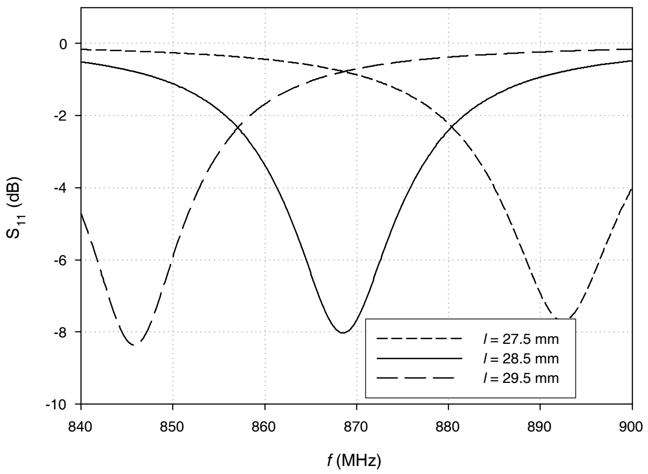


Fig. 4. Effect of l variations on the return loss

At the working frequency, a good conjugate-match between the antenna and the IC input impedance is first obtained as a compromise between the contrasting effects of l and d on the return loss. In Fig. 5 the dash-dotted line shows the behavior of the return loss for the optimized meander line antenna, without loading bars. In the same figure, the dashed line corresponds to the behavior of the return loss for the previous antenna, when a simple vertical loading bar is inserted as suggested in [2]. This result has been obtained after a tuning analysis on g_1 and w_2. It is worth noting that the presence of this loading bar allows to obtain a better return loss at the working frequency. Finally, a solid line is also shown that corresponds to the

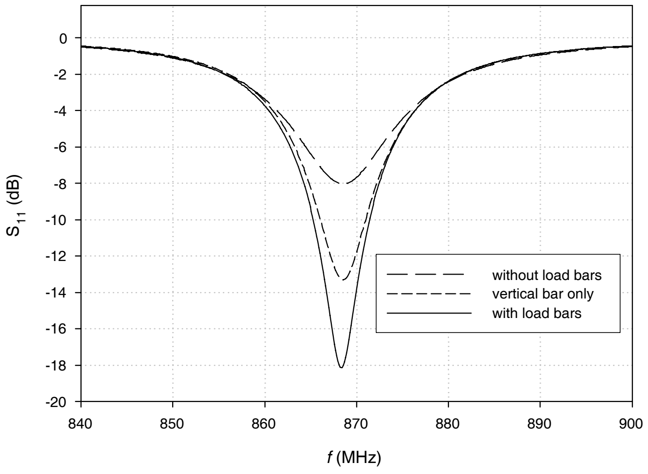


Fig. 5. Effect of loading bars on return loss

behavior of the return loss when also two horizontal loading bars are inserted near the final arms of the meandered antenna. These latter are not connected to the previous one otherwise currents that flow all around the meander line antenna may degrade its radiation characteristics. Of course, this final result has been obtained after a tuning analysis on the g_1 and g_2 distances. It is found that the use of three loading bars allows to significantly improve the quality of the impedance match, without significantly affecting the bandwidth of the matching. The proposed antenna exhibits useful tuning properties acting on the length l_s of the final branch. In Fig. 6 the behavior of the return loss as a function of the l_s parameter is reported. The simulation has been carried out taking into account the presence of the HF coil antenna near the proposed UHF antenna. It is shown that the return loss minima remain almost constant while trimming the final meander trace by $\Delta l_s = 1$ mm. In this case the working frequency moves up by about 4 MHz.

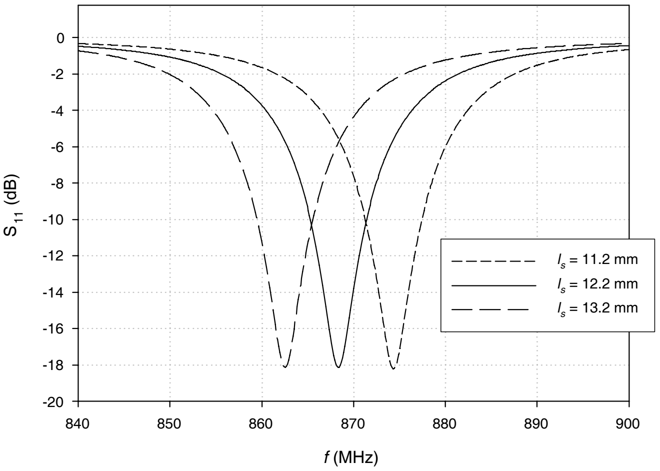


Fig. 6. Return loss for different arms length l_s

B. Radiation pattern

The simulated E- and H-plane radiation patterns at $f = 868$ MHz of the co-polar components for the proposed antenna in its final configuration are shown in Fig. 7. These curves show that the obtained radiation patterns are somewhat similar to that of a typical dipole. It is also observed that the maximum of the radiation pattern in the H-plane plane is slightly tilted towards the side where the horizontal bar is located. However it is expected that this effect does not significantly affects the overall performances of the transponder in the direction perpendicular to its plane. Simulations have also shown a 0.9 dB antenna gain at the working frequency.

III. Prototype

A prototype of the integrated transponder at the European band (868 MHz) with the proposed antenna has been realized using FR4 ($\epsilon_r = 4.4$, thickness = 0.8 mm) as a substrate, copper for the traces and a passive RFID IC by Philips. The final geometrical dimensions of the meander antenna are $l = 24.5$ mm, $d = 4$ mm, $w_t = 0.8$ mm, $l_s = 12.2$ mm, $g_1 = 0.3$ mm, $g_2 = 0.5$ mm, $w_1 = 0.8$ mm, $w_2 = 15$ mm.

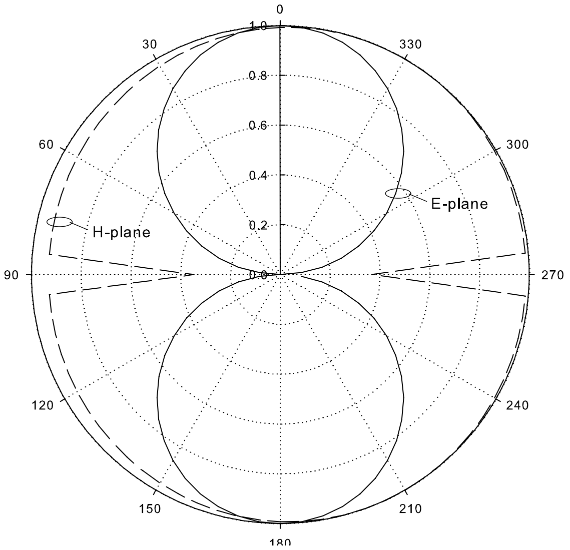


Fig. 7. Fig.4: Radiation patterns at 868 MHz

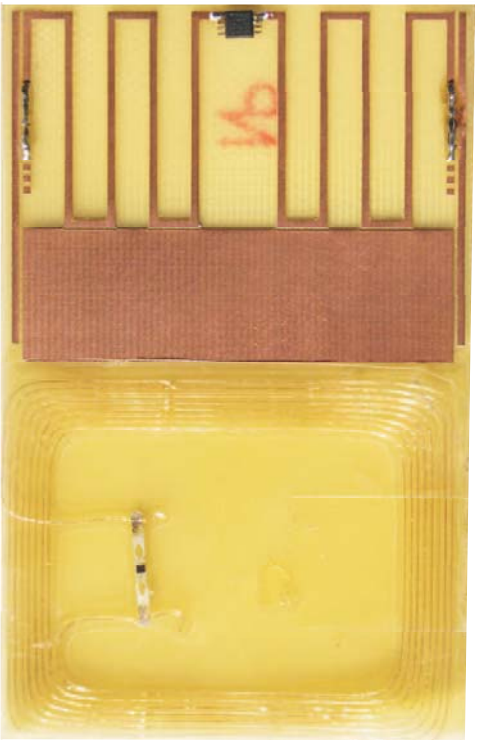


Fig. 8. Prototype of the integrated transponder

A picture of the prototype is shown in figure 8. As mentioned above, the commercial ISO 15693 Tag consists of a copper wire antenna coil deposited on a plastic flexible substrate, that can be easily placed near the UHF tag antenna. The tag read range has been measured using a setup composed by a UHF Reader and circular polarized antenna with gain $G_t = 6$ dBc. In fig 9 is shown the measured maximum read range of the UHF tag of the integrated transponder when it is equipped with the HF tag (circle point) and when is present only the UHF tag. It is found that the presence of the HF coil do not significantly degrades the overall performance of the UHF tag.

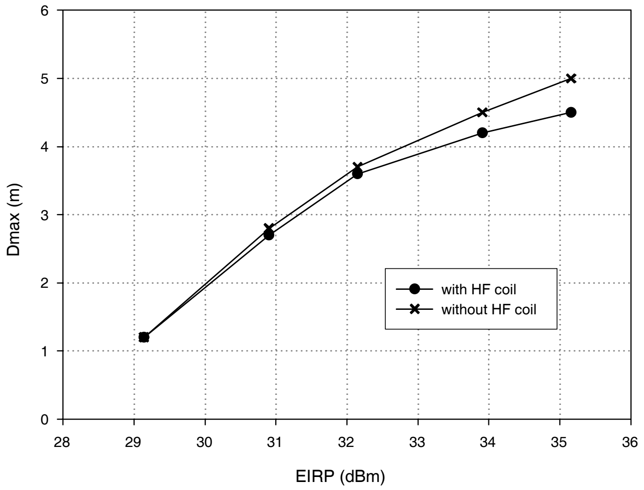


Fig. 9. Measured maximum read range of the integrated transponder for different values of EIRP

It is worth noting that the measured read range (5 m) obtained for the single UHF tag when P_{eirp} is equal to the maximum allowed value, is in a good agreement with that reported in fig. 2, where a read range of about 5.2 m is estimated for a tag antenna gain $G_r = 1$ dB and a return loss $|s| = -10$ dB.

IV. Conclusions

An integrated RFID transponder operating at both UHF (868 MHz) and HF (13.56 MHz) bands has been presented. It is composed by a ISO 7810 ID-1 card where both a UHF passive tag and an ISO 15693 commercial tag are arranged in two separate sections. A novel UHF tag antenna composed by a compact meander line dipole with three loading bars electromagnetically coupled to it has been presented. It is found that the presence of the loading bars allows to obtain a reduced sized meander antenna with both a satisfactory return loss and useful tuning characteristics. Simulated radiation patterns have also been presented which are found in a good agreement with that of a typical planar dipole antenna. The results of the tests conducted on the final prototype obtained using a UHF readers, have demonstrated a good overall performances of the integrated transponder.

References

[1] G. Marrocco. Gain-optimized self-resonant meandered line antennas for RFID applications. *IEEE Antennas and Wireless propagation letters*, 2:302–305, 2003.

[2] K.V. Seshagiri Rao, Pavel V. Nikitin, and Sander F. Lam. Antenna design for UHF RFID tags: a review and a practical application. *IEEE Transactions on antennas and propagation*, 53(12):3870–3876, Dec. 2005.

[3] Alberto Toccafondi and Paolo Braconi. Compact load-bars meander line antenna for uhf rfid transponder. In *Proc. of the 2006 European Conference on Antennas and Propagation*, Sept. 2006.

[4] Pavel V. Nikitin, K.V. Seshagiri Rao, Sander F. Lam, Vijay Pillai, Rene Martinez, and Harely Heinrich. Power reflection coefficent analysis for complex impedance in rfid tsg design. *IEEE Transactions on Microwave Theory and Techniques*, 53(9):2721–2725, Sept. 2005.

[5] K. Kurokawa. Power waves and the scattering matrix. *IEEE Transactions on Microwave Theory and Techniques*, 13(3):194–202, Mar. 1965.

A review of RFID and Wireless Sensors based on Surface Acoustic Wave Devices

M. Benetti, D. Cannatà, F. Di Pietrantonio, E. Verona

Istituto di Acustica *O.M.Corbino*, CNR, via del Fosso del Cavaliere, 100, I-00133, Rome, Italy
Phone:+39 06 4993 4481; Fax:+ 39 06 20660061; e-mail: verona@idac.rm.cnr.it

Abstract – **Wireless systems based on surface acoustic wave (SAW) devices show great potentialities for both identification and sensor applications. SAW-ID-Tags can be easily designed to operate in the ISM frequency band and in harsh environments and, moreover, the technologies for the fabrication are well-know and low-cost for mass production. In this paper, we report some examples of SAW ID-Tags systems and the principle of operation is described. Our attention is focused on design of interdigital transducers (IDTs) and reflectors, analyzing standard bidirectional structures and introducing newer structures such as unidirectional transducers or programmable reflectors. Some mentions on coding strategies and the *P*-matrix formalism are given. Finally, a SAW ID-Tag, combining identification and sensor capability, and developed in our laboratories, is presented.**

I. INTRODUCTION

Surface acoustic wave (SAW) devices have been introduced almost forty years ago [1] and nowadays are largely used in many application areas of both linear and non-linear signal processing (filters, resonators, delay lines, correlators, etc). Every year, manufacturers produce more than 300 million of SAW filters for mobile phones and color TV sets and the market is continuously growing due to the very good trade-off between costs and performances of these devices.

More recently, the increasing demand for high performance sensors, have brought to a wide research activity for the development of different kinds of sensing devices based on electro-acoustic transduction and exploiting the propagation of both bulk and surface acoustic waves. SAW sensors have shown to be a powerful tool to measure physical parameters such as force, acceleration, pressure, electric and magnetic fields, potentials, etc., or chemical and biochemical values, such as gas, vapour or ion concentrations in both gaseous and liquid environments.

For wireless systems, SAW devices have two distinguishing features that suggest their use: the high achievable operational frequency, in the GHz range, and the low phase velocity (3500-4000 m/s), that allows to obtain electric delays up to several tens of µs in a small analog device and hence to easily separate the data signals from the echoes due to multipath radio effects in the VHF/UHF range. Moreover SAW ID-Tags combine sensor and identification capability in low-cost passive devices that can be used in any harsh environments.

Several examples of wireless systems based on SAW devices are reported in the literature and also commercial operation already started for both identification or sensor purposes. For example, in toll booths in Oslo, SAW ID-Tags at 2.45 GHz with 33 bits (reflectors) divided in four tracks, are used [2], [3].

SAW Tags can be useful as clinical thermometers or as temperature sensors on rotating turbine blades, train brakes, in centrifuges, tires, or electrical motors [4], [5], [6], [7], in high voltage plants [8], and in dangerous or inaccessible process chambers. In the Munich subway net, the SAW ID-Tags were used to monitor the brakes temperature of the railway cars. The system showed interference immunity even under extremely harsh conditions [9].

Wireless pressure sensors can also be implemented by SAW devices pattering a reflective delay line or a resonator on a quartz diaphragm that bends under hydrostatic pressure. A continuous monitoring of the tire pressure in road cars by SAW delay lines is reported in [10], where a resolution of 1000 Pa is obtained. Resolution of a few Pascal can, however, be achieved in other applications [11].

Magnetic fields, as well the current producing the fields, can be detected by SAW devices implementing an impedance-type sensor with a magnetoresistor acting as a variable load Z_L [12].

Finally, chemical and biochemical SAW sensors are fabricated by coating the free surface of the devices with a proper chemically interactive material (CIM). Adsorption and desorption of the analyte from the membrane give rise to changes in the CIM's properties (mass density, but also elastic, viscoelastic, electric properties), that are detected as shifts in the resonant frequency of the SAW component [13].

In this paper we will describe the principle of operation of a SAW ID-Tags system focusing our attention on the main components of an RFID Tag: IDT transducers and reflectors.

II. PRINCIPLE OF OPERATION

The structure of an RFID Tag system is shown in Fig. 1. It consists of a piezoelectric substrate with an interdigital transducer (IDT), connected to a proper integrated antenna and a coded array of SAW reflectors. The interrogation signal, consisting of a RF pulse, or of a compressible signal, like in the radar technique, is picked up by the antenna and converted by the IDT into a SAW that propagates towards the reflectors, distributed in a characteristic barcode-like pattern, and then is partially reflected back at each reflector. The reflected acoustic bursts are collected by the IDT and, through the antenna, converted into electromagnetic waves, that are finally collected by the reader. The received signal, after processing, contains information about the number and location of the reflectors.

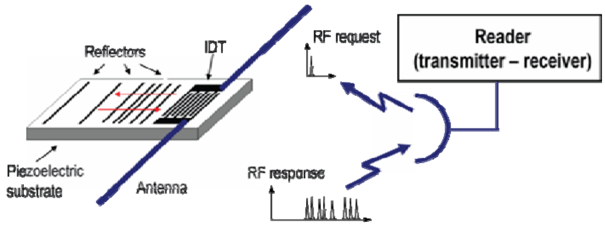


Figure 1. Structure of a SAW ID-Tags system.

On considering the European ISM bands, SAW-based RFID Tags could, in principle, operate at each of the frequencies between 13.56 MHz and 5.8 GHz. At the frequency of 13.56 GHz, however, the size of both Tag and antenna would be too large, with an increase in the volume and cost of the device, while, at 5.8 GHz, the line-width resolution required to implement IDTs and reflectors, would be of the order of 170 nm, requiring more sophisticated fabrication technologies. The most suitable frequencies for SAW RFID Tags operation are, therefore, 434 MHz or 2.45 GHz.

The request units of wireless SAW ID-Tags are similar to those used in radar applications [4], [5]. Systems based on pulse radar, pulse compression radar and frequency-modulated continuous-wave (FM CW) radar architectures can be implemented. In Fig. 2 are reported the distance interrogation range for unlicensed low power devices in the ISM frequency bands.

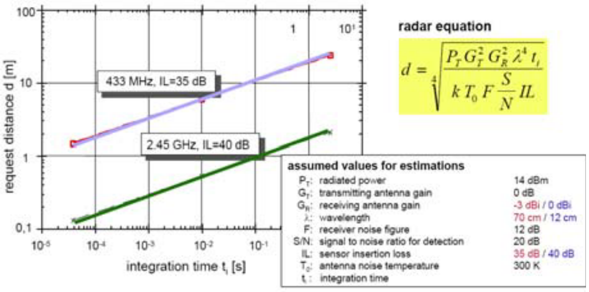
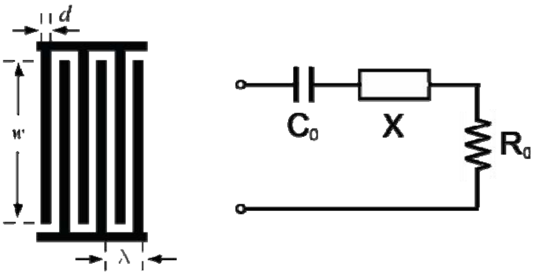


Figure 2. Distance interrogation range in ISM bands for unlicensed low power devices.

III. IDT TRANSDUCERS AND REFLECTORS

Fig. 3a shows the configuration of a simple IDT transducer. It consists of two interdigitated metal comb electrodes, deposited in the form of thin films on the free surface of a proper piezoelectric layer. The spatial periodicity λ of the IDT corresponds to the SAW wavelength and is related to the SAW velocity v and to the frequency of operation f_0, by: $\lambda = v/f_0$, being the width of electrodes and spaces usually equal to λ/4. Other IDT project variables are the number N of finger pairs, affecting the frequency bandwidth of operation ($\Delta f/f_0 \approx 1/N$) and the finger overlap w, that fix the acoustic near field distance (A·w^2/λ), being A a coefficient taking into account the velocity anisotropy in the propagation plane; in the isotropic case A=1.



(a) (b)

Figure 3. Simple IDT transducers (a) and its equivalent circuit (b).

The equivalent circuit, at the electrical port of the IDT, is reported in Fig. 3b. It consists of a static capacitance, C_0, taking into account the capacitance of the interdigital electrodes, a motional reactance X, equal to zero at the center frequency and, usually, not taken into account since negligible with respect to the static reactance, and, finally, a radiation resistance R_0. The expressions for the radiation resistance and the static reactance, at center frequency are:

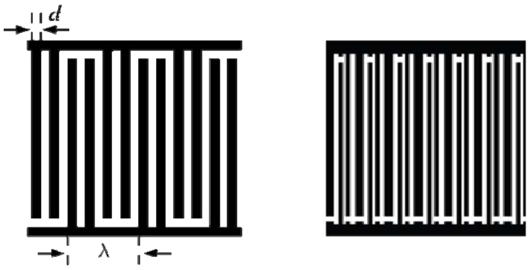
$$R_0 = \frac{1}{v \cdot C_s \cdot d \dfrac{\pi^2}{2K^2}} \qquad X_0 = \frac{1}{2\pi \cdot v \cdot C_s N \cdot d}$$

being C_s [F/m] the surface capacitance of the piezoelectric substrate, K^2 its SAW electromechanical coupling coefficient and d = w/λ the IDT's directivity. At frequencies different from f_0, the radiation resistance is given by:

$$R(f) = R_0 \frac{\sin\left(N\pi \dfrac{f - f_0}{f_0}\right)}{\left(N\pi \dfrac{f - f_0}{f_0}\right)}$$

Maximum power transfer conditions require the IDT and antenna to be designed with equal radiation resistance and opposite reactance.

The IDT reported in Fig. 3a shows a constructive interference of the SAWs reflected at the finger edges, thus limiting its use to IDTs with a low number of finger pairs and/or to low electromechanical coupling coefficient materials. To overcame this effect, the split finger configuration (Fig. 4a) can be preferred, in spite of an higher line-width resolution required at high frequency operation. The two transducers are both bidirectional; unidirectional operation is possible using the single phase unidirectional transducer (SPUDT), shown in Fig. 4b, and exploiting both λ/4 and λ/8 electrodes [14], [15], [16].



(a) (b)

Figure 4. Split finger transducer (a) and SPUDT transducer (b).

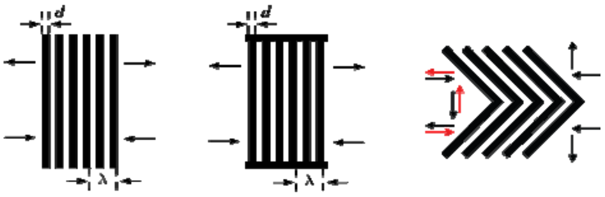
Fig. 5 shows examples of reflecting structures, consisting of either open (a) or electrically shorted (b) metal strips. The SAW is partially reflected at each strip edge; constructive interference between reflected signals takes place when the edge distance is λ/4. Larger is the number of reflectors, higher is the reflected signal and lower is its frequency bandwidth. The reflection, at each strip edge, is due to the different SAW velocity for free surface condition (v_f) or for electrically short-circuited surface (v_s); we have:

$$K^2 = 2\frac{\left(v_f - v_s\right)}{v_f}$$

and thus:

$$v_s = v_f\left(1 - \frac{K^2}{2}\right)$$

The reflectivity of each strip depend on the electrical conditions, in particular for high piezoelectric substrate, such as lithium niobate ($LiNbO_3$), open and shorted strips exhibit significant reflectivity values [17].



(a) (b) (c)

Figure 5. Reflecting structure: open (a), electrically shorted (b) and chevron type (c).

The bidirectional operation of these reflecting structures can produce signal degradation due to multiple reflection (triple) interferences (see Fig. 7). To this purpose the chevron type structure, shown in Fig. 5c, allows a better rejection of multiple reflection paths, in spite of a lower reflection efficiency, since the main reflected signal is produced by a double reflection.

A different type of reflector, shown in Fig. 6, is the multistrip one, consisting of a number of thin strips (<λ/4). It can operates with an high efficiency, over a wide-band of frequencies, showing cut-off when the strips width correspond to λ/4.

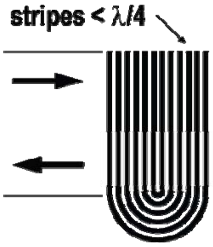


Figure 6. Multistrip reflector.

Multiple reflection mechanism, whose signals can interfere with other symbols and, consequently, affect information is shown in Fig. 7. A possible way to reduce multipath responses and thus prevent information degradation is shown in Fig. 8; it exploits both the IDT bidirectionality and the distribution of the reflector array over partial beams on the transducer aperture. The advantage is a better uniformity, when the number of reflectors increases [3], at expense of an higher signal insertion loss.

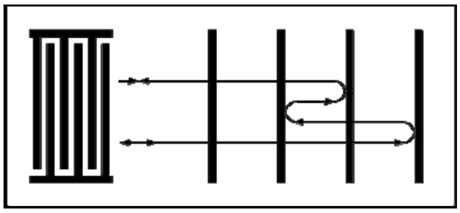


Figure 7. Multiple reflection interferences.

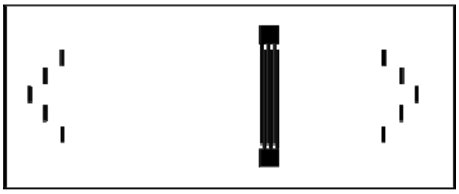


Figure 8. Distribution of reflectors on multiple tracks.

A different approach to implement SAW RFID Tags is based on use of bus-bar connected arrays of IDTs, as shown in Fig. 9, very similar to tapered delay-line transducers. When excited by a RF pulse, each IDT element generates a train of waves that propagates along opposite directions and is collected by all the other IDTs. A proper choice of the number, polarization and location of the single IDTs, allows to produce a given code. Unlike the reflecting structures, where high K^2 substrates, such as $LiNbO_3$, are preferable, here small dielectric and piezoelectric constant materials, such as quartz, can be exploited.

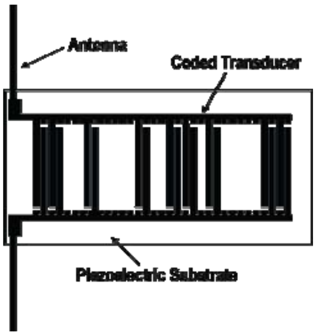


Figure 9. SAW ID-Tag based on bus-bar connected array of IDTs.

IV. MIXED MATRIX FOR SAW ID-TAGS

The mixed matrix or *P*-matrix formalism is a useful description for IDTs or reflectors in ID-Tags. In fact, the values of this matrix are not dependent, as for the scattering matrix notation, on external components connected at the transducer [18].

For the generic network (Fig. 10), the amplitudes of outgoing elastic waves and the current intensity are expressed in terms of incoming wave amplitudes and voltage:

$$\begin{pmatrix} b_1 \\ b_2 \\ I \end{pmatrix} = \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ V \end{pmatrix}$$

where: P_{11} and P_{22} are the reflections coefficients and P_{12} and P_{21} the transmission ones. P_{13} and P_{23} describe the efficiency for SAW generation, while P_{31} and P_{32} the reverse efficiency for the current intensity. Finally, P_{33} is the input admittance at the electric port.

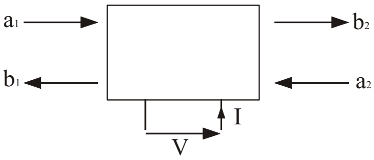


Figure 10. *P*-matrix formalism for a generic network.

In a SAW ID-Tag, transducers and reflectors are connected in parallel on the electric ports and are cascaded on acoustic ones. The elements of the resulting *P*-matrix can be easily derived [19]:

$$P_{11} = P_{11}^A + P_{11}^B \left(\frac{P_{21}^A P_{12}^A}{1 - P_{11}^B P_{22}^A} \right)$$

$$P_{12} = \frac{P_{12}^A P_{12}^B}{1 - P_{11}^B P_{22}^A}$$

$$P_{13} = P_{13}^A + P_{12}^A \left(\frac{P_{13}^B + P_{11}^B P_{23}^A}{1 - P_{11}^B P_{22}^A} \right)$$

$$P_{22} = P_{22}^B + P_{22}^A \left(\frac{P_{12}^B P_{21}^B}{1 - P_{11}^B P_{22}^A} \right)$$

$$P_{23} = P_{23}^B + P_{21}^B \left(\frac{P_{23}^A + P_{22}^A P_{13}^B}{1 - P_{11}^B P_{22}^A} \right)$$

$$P_{33} = P_{33}^A + P_{33}^B + P_{32}^A \left(\frac{P_{13}^B + P_{11}^B P_{23}^A}{1 - P_{11}^B P_{22}^A} \right) + P_{31}^B \left(\frac{P_{23}^A + P_{22}^A P_{13}^B}{1 - P_{11}^B P_{22}^A} \right)$$

where P_{ii}^A and P_{ii}^B are the mixed matrix of the two connected elements.

For example, if we consider a reflector connected to a external load (Fig. 11), we obtain:

$$P_{11}(Y) = P_{11,sc} + \frac{2P_{13}^2}{P_{33} + Y}$$

For a short circuit condition ($1/Y = 0$), we get $P_{11,sc}$, that is very small for splitfinger reflector. Otherwise, for the open circuit condition ($Y = 0$), the reflection coefficient becomes: $2P_{13}^2 / P_{33}$. A maximum in the reflection is obtained if the imaginary part of P_{33} (mainly capacitive) is compensated with an inductive load. In this way, a programmable reflector is implemented [20].

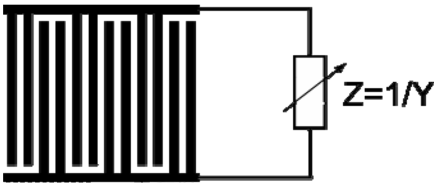


Figure 11. Programmable reflector.

V. CODING

The ID information of a SAW Tag is contained in the reflectors arrangement: each Tag must be coded according to one of a number of different possible codes. Different modulation techniques can be used for SAW ID-Tag applications, including binary amplitude (on/off) keying, where each predetermined possible symbol position is either occupied by a reflector (1-bit) or not (0-bit), second or higher-order phase shift keying (PSK), where the phase of each time response pulse is considered, or, finally, pulse position modulation of the reflectors.

A number of SAW ID-Tags can be read at the same time by a single reading pulse, providing anti-collision codes. The number of not-interfering codes is lower than that corresponding to the number of bits used [21].

VI. SAW ID-TAG AS CHEMICAL SENSOR

We have developed a SAW ID-Tag that combines identification and sensor capability. The IDT was designed to operate at the frequency of 433.92 MHz on a 128° YX $LiNbO_3$ substrate. The finger overlap is 1250 µm with an aluminium film thickness of 100 nm, so to ensure a radiation resistance $R_0 \cong 50\ \Omega$ and, at the same time, an high directivity. The reflectors are located at both the sides of the bidirectional transducer and allocated on two tracks. The sensor and identification functionalities are divided, respectively, on the left and the right side of the IDT as shows in Fig. 12.

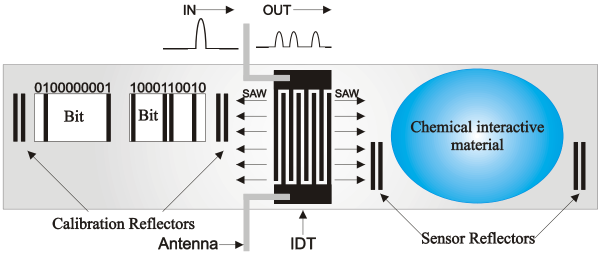


Figure 12. SAW ID-Tag for sensor and identification applications.

Due to the high temperature sensitivity of $LiNbO_3$, $S_v^T = -72\,ppm$, calibration reflectors are used. For the identification, we implemented two words of ten bits, while two reflectors, placed 3.5 mm apart, were designed in the sensor area. Each bit-reflector consists only of two open metal strips in order to ensure a good uniformity of echoes, while the sensor and farther calibration reflectors are made by four strips. A photo of the ID-Tag is reported in Fig. 13.

The time-domain response of two different-coded ID-Tags are shown in Fig. 14, while their main features are reported in Table I.

Figure 13. Photo of the SAW ID-Tag operating at 433.92 MHz.

TABLE I. SAW ID-TAG MAINS FEATURES

Feature	Typical Value	Obtained Value
Insertion Losses	< -40 dB / -60 dB	-48 dB
Uniformity	< 3 dB / 5 dB	1.6 dB
Dynamic	>20 dB	28 dB

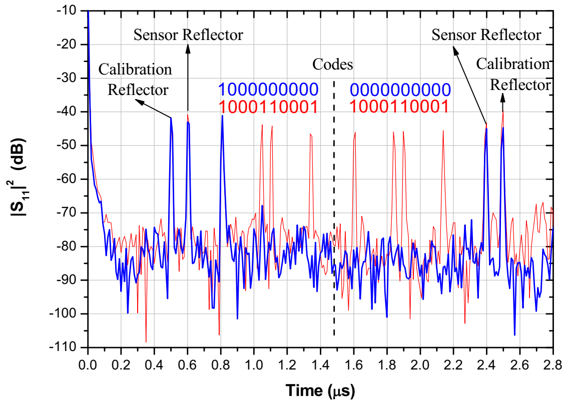


Figure 14. Time-domain responses of two different-code ID-Tags.

VII. CONCLUSIONS

SAW ID-Tag are very versatile systems for many applications. The reported examples show the advantage of this technology to monitor, especially in harsh environments or in inaccessible locations, physical or chemical quantities. The presented SAW ID-Tag shows good features respect to insertion loss, uniformity and dynamic, and can be used as chemical sensor if coated with a proper CIM.

REFERENCES

[1] R. M. White and F. W. Voltmer, "Direct piezoelectric coupling to surface elastic waves" *Appl. Phys. Letters,* vol. 7, pp. 314-316, 1965.

[2] A. Wolfe, "SAW Filters Replace Bar Codes in ID Systems" *Electronics Week,* May 1985.

[3] L. Reindl, G. Scholl, T. Ostertag, A. Pohl, and R. Weigel, "Wireless Remote Identification and Sensing with SAW Devices" in *IEEE 1998 MMT/AP International Workshop on Commercial Radio Sensor and Communication Techniques*, 1998, pp. 83-96.

[4] F. Schmidt, O. Sczesny, L. Reindl, and V. Magori, "Remote sensing of physical parameters by means of passive surface acoustic wave devices (‘ID

TAG')" in *IEEE Ultrasonics Symp.*, Cannes (France), 1994, pp. 589-592.

[5] A. Pohl, F. Seifert, L. Reindl, G. Scholl, T. Ostertag, and W. Pietsch, "Radio signals for ID Tags and sensors in strong electromagnetic interference" in *IEEE Ultrasonics Symp.*, Cannes (France), 1994, pp. 195-198.

[6] A. Pohl and F. Seifert, "Wireless interrogable SAW sensors for vehicular applications" in *IEEE Instrumentation and Measurement Conference*, 1996, pp. 1465-1468.

[7] A. Pohl, G. Ostermayer, L. Reindl, and F. Seifert, "Spread spectrum techniques for wirelessly interrogable passive SAW sensors" in *IEEE International Symp. on Spread Spectrum Techn. and Appl.*, 1996, pp. 730-734.

[8] V. Hinrichsen and G. Scholl, "Online-Temperaturmessung an Metalloxid-Überspannungsableitern mit Hilfe von funkabfragbaren Oberflächenwellensensoren - Ein neues Verfahren der Ableiterüberwachung" *Elektrizitätswirtschaft Jhg. 97,* 1998.

[9] http://www.siemens.de/vt.d/sofis/inhalt.htm

[10] A. Pohl, G. Ostermayer, L. Reindl, and F. Seifert, "Monitoring the Tire Pressure at Cars Using Passive SAW Sensors" in *IEEE Ultrasonics Symp.*, Toronto (Canada), 1997, pp. 471-474.

[11] M. Benetti, D. Cannatà, F. Di Pietrantonio, C. Marchiori, P. Persichetti, and E. Verona, "Pressure sensor based on SAW resonators" in *Sensors and Microsystems Proceedings of the 13th Italian Conference (AISEM)* Rome (Italy), 2008, p. in press.

[12] L. Reindl, G. Scholl, T. Ostertag, H. Scherr, U. Wolff, and F. Schmidt, "Theory and application of passive SAW radio transponders as sensors" *IEEE Trans. Ultrason. Ferroelectr. Freq. Contr.,* vol. 45, pp. 1281-1292, 1998.

[13] M. Benetti, D. Cannatà, A. D'Amico, F. Di Pietrantonio, A. Macagnano, and E. Verona, "SAW sensors on Aln/Diamond/Si structures" in *IEEE Sensors Proc.*, Vienna, 2004, pp. 753-756.

[14] C. S. Hartmann and B. P. Abbott, "Overview pf design challenges for single phase unidirectional SAW filters" in *IEEE Ultrasonics Symp.*, Montreal (Canada), 1989, pp. 79-89.

[15] S. Lehtonen, V. Plessky, C. S. Hartmann, and M. M. Salomaa, "Unidirectional SAW transducer for Gigahertz frequency" *IEEE Trans. Ultrason. Ferroelect. Freq. Contr.,* vol. 50, 2003.

[16] C. Kappacher, O. Maenner, W. Ruile, and R. Dill, "Design and analysis of single phase unidirectional transducers" in *IEEE Ultrasonics Symp.*, Lake Buena Vista Florida (U.S), 1991.

[17] S. Lehtonen, V. Plessky, and M. Salomaa, "Short reflectors operating at the fundamental and second harmonics on 120° LiNbO3" in *IEEE Ultrasonics Symp.*, Munich (Germany), 2002, pp. 333-337.

[18] G. Tobolka, "Mixed matrix representation of SAW transducers" *Trans. Son. and Ultrason.,* vol. SU-26, pp. 426-428, 1979.

[19] D. P. Morgan, "Cascading formulas for identical transducer P-matrices" *IEEE Trans. Ultrason. Ferroelect. Freq. Contr.,* vol. 43, pp. 985-987, 1996.

[20] L. Reindl and W. Ruile, "Programmable reflectors for SAW-ID-Tag" in *IEEE Ultrasonics Symp.*, Baltimore (U.S), 1993, pp. 125-130.

[21] C. S. Hartmann, P. Hartmann, P. Brown, J. Bellamy, L. Claiborne, and W. Bonner, "Anti-collision methods for Global SAW RFID Tag Systems" in *IEEE Ultrasonics Symp.*, Montreal (Canada), 2004, pp. 805-808.

Sensing-oriented UHF-RFID Tags

Gaetano Marrocco

***Abstract*-- This paper addresses the design of new UHF tag antennas for sensing applications, e.g. able to host sensors and additional electronics but also to act as passive sensor itself for some features of the target. The new theories and the proposed solutions are preliminary demonstrated by computer simulations and by measurements on fabricated prototypes.**

***Index terms*—RFID, tag, planar antennas, sensor**

I. Introduction

The recent advances in Wireless Sensor Networks (WSNs) [1] for applications in mobile and time-varying environments, are generating a growing attention to low-cost and low-power wireless nodes, equipped with radio/sensing ability, which are spatially distributed to ensure a cooperative monitoring of physical or application-specific conditions and parameters. Typical fields of applications for WSNs include environmental and habitat monitoring, disaster relief [2], health care, inventory tracking and industrial processing monitoring, security and military surveillance, smart spaces applications. A novel technological trend is the integration among wireless sensor networks and Radio Frequency IDentification (RFID) technologies. Such a convergence of sensing and identification technologies may enable a wide range of heterogeneous applications which demand a tight synergy between detection and tagging.

A new frontier is the wireless monitoring of people within Mobile Healthcare Services [3] with the purpose to reduce the hospitalization of patients, to support disaster relief or to get an epidemic under control. An RFID system could provide real-time bio-monitoring and localization of patients inside hospitals or domestic environments, as well as in extreme conditions like a Space Capsule (Fig.1). In these cases the tag should be placed on the human body and equipped with bio-sensors (temperature, blood pressure, glucose content, motion) and, when activated by the reader, tag ID and bio-signals could be transferred to a remote units and then stored and processed.

Up to date, several approaches have been proposed to provide RFID devices with enhanced sensing and detection capabilities. The main solutions make use of active or passive RFID transponders and Surface Acoustic Wave (SAW) devices [4]. A significant example of enhanced passive RFID system is given by the Wireless Identification Sensing Platform (WISP) project [5] which introduced the concept of ID modulation making use of inertial switches and enhanced power harvesting units to mechanically toggle between two commercial RFID integrated circuits.

These devices could be *passive*, harvesting energy from the interrogating system, *semi-active* when a battery is included only to feed the sensors, or fully *active* where a local source directly feeds a microcontroller as well as the transmitting radio. However, the large battery packs required for active techniques, in addition to the use of protruding antennas, could be suboptimal for some applications and additional issues have to be considered such as the compromise between a long battery-life and a miniaturized design.

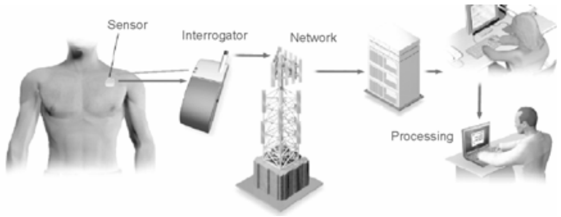


Fig. 1. Typical scenarios for a Mobile Healthcare Network.

In passive RFIDs, together with the microchip sensitivity, the tag antenna plays a key role in the overall system performance, such as the reading range and the compatibility with the tagged object. In case of RFID with sensing capability, the antenna should be additionally suited to electrical and physical integration with sensing electronics. Moreover, it is well experienced that the RFID system performances are greatly dependent on the kind of object where the tag is attached on. When a same tag is placed onto different targets, the tag antenna's input impedance may in some case undergo a mismatch and may produce a change of the read distance. Conventional general-purpose tags are designed in free space, but when they are required to be applied over objects having high values of the permittivity, as in the case of liquids and humans, the strong pattern distortion and the efficiency loss caused by energy dissipation and scattering, need to be taken into account in the first stage of the design.

The author is with the Dipartimento of Informatica Sistemi e Produzione, Università di Roma Tor Vergata, Via del Politecnico, 1, 00133 Roma (Italy). Marrocco@disp.uniroma2.it

Within this scenario, this contribution has a twofold objective: the description of a new class of UHF tag layouts suited to host sensors and to be attached onto high dielectric and lossy targets, such as the human body, and to introduce the novel concept of *self-sensing* tags wherein the antenna itself acts both as a data transmitter and as a sensing device of some tagged body's features. These two arguments are both related to the modelling and handling of *dense* objects, and the *self-sensing* idea originates just from one of the main conventional drawbacks of UHF RFID framework, e.g. the dependence of reading performances on the dielectric value of the tagged object. The self-sensing tags are multi-chip antennas (multi-port system) exploiting the dependence of the tag's input impedance and radar cross-section on the physical and geometrical features of a real target.

The two subjects are here described both theoretically and corroborated by several preliminary prototypes and experimentations.

II. Basic Definitions for RFID Systems

At the beginning of the reader-to-tag communication protocol [6], the reader first *activates* the tag, placed over a target object, by sending a continuous wave which, on charging an internal capacitor, provides the required energy to perform actions. During this *listening mode*, the microchip exhibits an input impedance $Z_{chip} = R_{chip} + jX_{chip}$, with X_{chip} capacitive, and the antenna impedance $Z_A = R_A + jX_A$ should be matched to Z_{chip} ($Z_A = Z^*_{chip}$) for maximum power transfer. The maximum fraction $P_{R\to T}$ of the reader input power that is absorbed by the tag is

$$P_{R\to T} = \left(\frac{\lambda_0}{4\pi d}\right)^2 G_R \tau G_T P_{in} \qquad (1)$$

$$\tau = \frac{4R_{chip}R_A}{\left|Z_{chip} + Z_A\right|^2} \qquad (2)$$

where λ_0 is the free-space wavelength, d is the reader-tag distance, G_R is the gain of the reader antenna and G_T is the gain of the tag over the target, having assumed polarization-matched antennas aligned for maximum directional radiation. τ is the power transmission coefficient. The tag is activated when the absorbed power exceeds the tag's microchip sensitivity threshold: $P_{R\to T} > p_T$. Having fixed the effective power ($EIRP_{R=}G_RP_{in}$) transmitted by the reader, the tag antenna gain (G_{tag}) and the sensitivity (P_{chip}) of the tag microchip, e.g. the RF power required to the microchip electronics to turn on and complete its tasks, the maximum activation distance of the tag along the (θ,ϕ) direction is therefore given [2] by

$$d_{\max}(\theta,\phi) = \frac{c}{4\pi f}\sqrt{\frac{EIRP_R}{P_{chip}}\tau G_{tag}(\theta,\phi)} \qquad (3)$$

During the next steps of the communication, the tag receives the command coming from the reader and it finally sends back the data through a back-scattered modulation of the continuous wave coming from the reader itself. The tag's IC acts as a programmable switching device which toggles between a low or high impedance Z_{mod}. During the data transfer, the RFID system can be considered as a monostatic radar and therefore it can be characterized by the radar range equation which, for the case of typical RFID tags, can be expressed [6] in the form

$$\frac{P_{R\leftarrow T}(d)}{P_{in}} = \left(\frac{\lambda_0}{4\pi d}\right)^4 G_R^2 G_T^2 \rho \qquad (4)$$

$$\rho = \frac{4R_A^2}{\left|Z_{\mathrm{mod}} + Z_A\right|^2} \qquad (5)$$

where $P_{R\leftarrow T}$ is the power received back by the reader and ρ is a modulation parameter related to the tag's radar cross section.

The presence of the tagged object, and in particular of the human body, with its high permittivity and conductivity, will favour the antenna miniaturization but nevertheless will induce a strong power absorption. The antenna gain, and hence the link distance, will be sensibly reduced with respect to the free space. The maximum transmitted power allowed to the reader is constrained to local regulations. In Europe the relevant standards for UHF RFID applications are the ETSI EN330-220 and Draft TESI EN302 208-2. In particular within the 865.6-867.6MHz the maximum EIRP is 3.2W, which overcomes the previous limit 0.8W. In the U.S.A. the FCC allowed band is 902-928MHz with maximum transmitted EIRP=4W.

Microchip power activation threshold is continuously improving, reducing from 1mW in the year 2001 to some microwatts in today current products or even less in the state of the art ASICS [7].

It is easy to show from equation (3) that antennas with averaged realized gain ($G_{tag}\tau$) not less than -10dB (when placed over the human body) could be in principle compatible with reading distances of the order of 5m if the microchip sensitivity is less than 10μW.

III. Tag antennas for Human-body Applications

Three tag geometries are here described together with the related design methodologies. These layouts are slot-type antennas suited to be easily integrated with sensors and additional electronics and useful for placement on high dielectric target such as human-body districts. All the presented tags are numerically modelled by a Finite-Difference Time-Domain solver [8], having considered the antenna placed onto realistic models (box, layered cylinders) of the tagged body.

A. The Nested-Slot Suspended Patch (NSSP) Antenna

The first tag antenna family is a nested-slot suspended-patch (NSSP), [9]. Small size slot antennas are naturally inductive and therefore appear more suited than dipoles to achieve conjugate impedance matching. The basic geometry is visible in Fig.2. Since the slot sizes are comparable with the patch surface, the radiation features are related to both the objects. In particular, the maximum antenna gain is mainly fixed by the patch side *L*, while the impedance tuning can be changed by acting on the slot size *a* and *b* (Fig.3). Depending on the shape of the internal slot, the antenna mainly radiates either as a *dumbbell H*-slot or as a pair of rectangular loops sharing the sourced conductor (Fig.4).

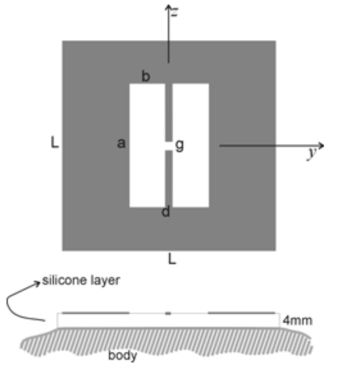


Fig. 2: NSSP tag. Parameters of the proposed planar slot antennas. The microchip transmitter should be placed in the central gap of size *g* x *g*.

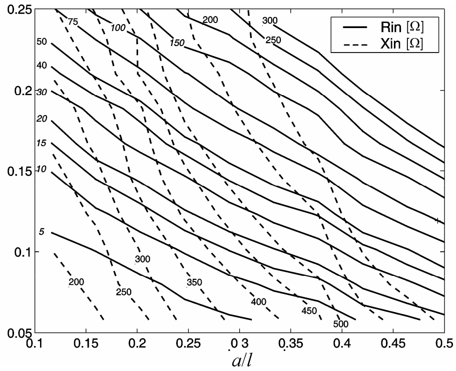


Fig.3: NSSP tag. Matching chart to design the H shape factor to match the particular microchip's impedance.

Fig.5 show a fabricated prototype of the body-matched NSSP tag and the measurement set-up where the tag is attached onto a Perspex box (ε_r=2.7, σ=0) having 5 mm thickness and 20 cm width, filled with a muscle-type solution, and the achieved antenna input impedance and power matching, compared with simulations, are shown in Fig.6a. The experiment demonstrated a relevant impedance tuning agility and a read distance suited to small or even average room (Fig.6b).

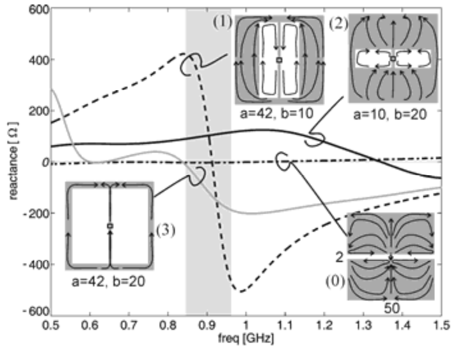


Fig.4: Typical antenna input impedance for some choice of the H-slot parameters (in mm). In all the case the patch size is L=50mm.

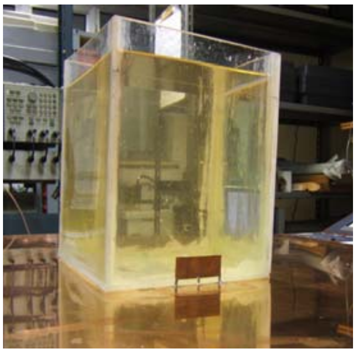
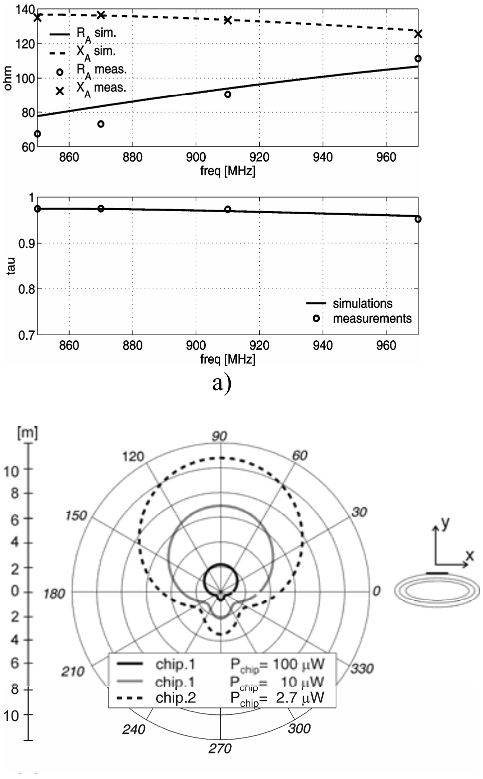


Fig.5: NSSP tag. Fabricated Half-plane NSSP antenna in front of a Perspex cubic phantom filled with tissue-equivalent solution made of deionised water, saccharose and sodium chloride. The antenna and the box are placed over a 1m x 1m copper image plane.

B. The Meandered Slot Antenna (MSA)

The previously considered NSSP antennas are symmetric with respect to both the *x* and *z* axis. However this geometry offers additional degrees of freedom in the position of the slot and in the connection to the microchip provided that a larger number of slot discontinuities (Fig.7a) are considered. This new layout is similar to a meandered slot and, when properly optimized, could permit to fulfill several electrical and geometrical constraints, such as the impedance matching to a particular microchip, dual-frequency operations, the embedding of a sensor stage of given size, and a stable response over a large variety of tagged dielectrics. The slot profile can be seen as a slot-line impedance transformer [10], where each discontinuity (tooth) provides energy storage and radiation. A Genetic Algorithm [11] optimization problem is hence formulated to shape the transformer layout, within input impedance and size requirements. As an example, Fig.7 shows the shape and the power transmission coefficient τ for some

870MHz slot-line antennas optimized to occupy only a fraction of the overall metallization, and preliminary experimental prototypes on FR4. When compared with the NSSP tags, this layout permits to achieve a better gain and more space for the electronic payload.



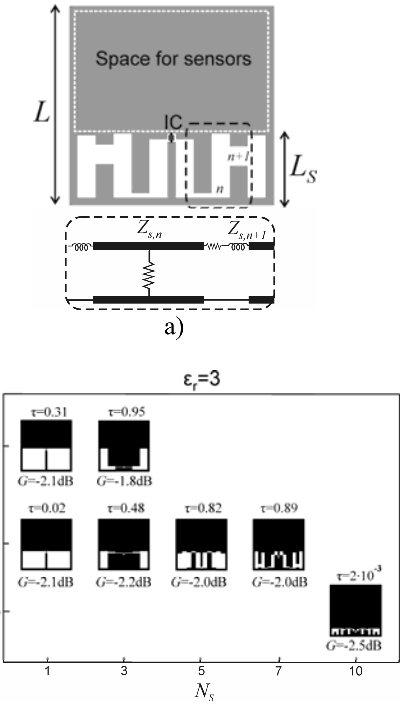
a)

b)

Fig.6: NSSP tag. *a)* Measured and computer-estimated input impedance. *b*) Estimated read distance for different kinds of microchip and 3.2 W EIRP emitted power.

C. *The Slot Inverted L antenna (SILA)*

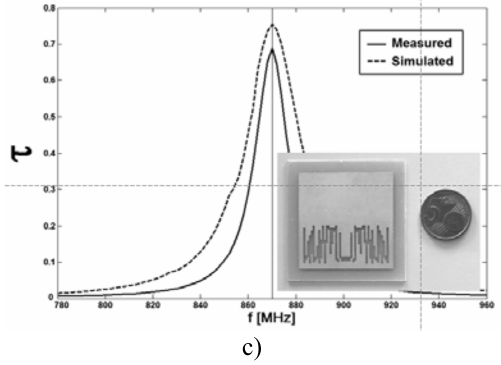
A further evolution of the slot-driven patch comprises an L-type patch folding (Fig.9) with the purpose to increase the antenna radiation and in particular to reduce the power dissipation into the body district where the tag is placed. The folded region acts as a ground plane which partly isolates the antenna from the body. The radiation is now due to the H-slot itself, as in the previous layouts, but also to the current discontinuity in the folding and especially to the patch truncation. When attached onto a leg-like layered cylinder, this layout produces a larger gain than the NSSP and MSA tags, with maximum value of the order of 0dB with back radiation ranging within -5÷-10dB.. Fig.10 shows a fabricated prototype and the measured power transmission coefficient. The resulting read distance is sensibly improved. This antenna is intended to be integrated with inertial switcehs for the monitoring of legs' movement in neuroscience applications.



a)

b)

Fig.7: MSA tag. *a*) Layout of the meandered slot family and slot-line model. *b*) Examples of antennas with *L*=5cm, placed over a ε_r=3 dielectric half-space, which have been optimized for an IC with Z_{chip}=15-j450Ω, for different sizes L_s of the antenna region and for different number N_s of slot-line sections. It is assumed a symmetric layout and therefore N_s represents half the overall slot transitions. *G* is the maximum gain in the air half-space.



c)

Fig.8: MSA tag. Fabricated 5cm x 5cm prototype and in-air measurement of the power transmission coefficient.

IV. Self-Sensing RFID tags

Like any antenna immersed or located close to a real object, the input and radiation characteristics of a passive RFID transponder placed on a target, as well as the strength of the back-scattered power, are closely related to the physical properties of the tagged object itself, e.g. on its constitutive material, shape, temperature, humidity or other. We denote with $\boldsymbol{\Psi}$ the set of the relevant target's features which could undergo changes along with the time, or have to be monitored in someway. If the tag antenna has been designed for optimal performances when placed on a target with nominal set of features Ψ_T, e.g. such that the antenna impedance $Z_A(\boldsymbol{\Psi}_T)$ equals in this condition Z^*_{chip}, a change of one or more target's parameters with respect to Ψ_T may produce a variation of the input impedance and hence the mismatch $Z_A \neq Z^*_{chip}$. Accordingly, also the back-scattered power collected at the reader port will be modified (Fig.12). In the limiting case, the tag may be completely mismatched so that $P_{R\to T} < p_T$ and the tag is therefore inactive. For the sake of clarity, let us focus on the simplified case for which a single target's feature is subjected to change, and such a parameter be the relative dielectric permittivity (simply permittivity ε in the following). It is now useful to define the tag's Activation Set $A_\varepsilon(d)$ for a link length d, as the set of target's permittivity values for which the power harvested by the tag is enough to activate it: $A(d) = \{\varepsilon \mid P_{R\to T}(d,\varepsilon) \geq p_T\}$.

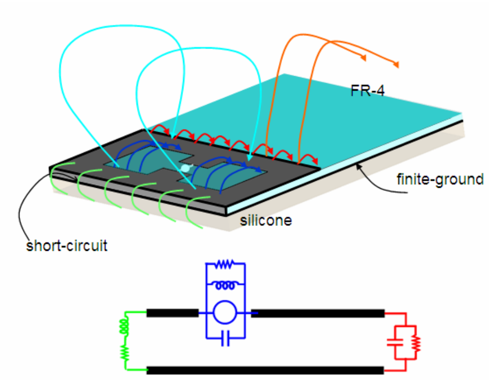


Fig.9: SILA Tag. Layout of inverted slot antenna, with principal radiating mechanisms and its transmission line equivalent

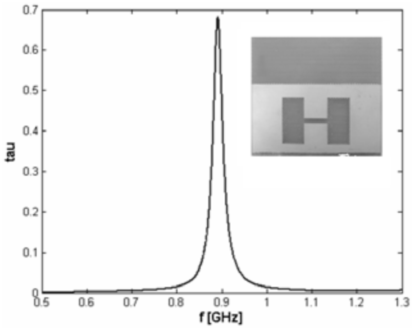
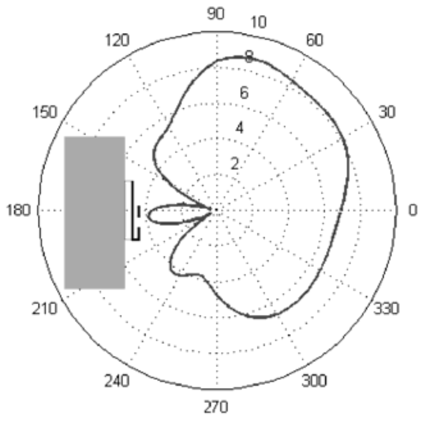


Fig.10: SILA Tag. Experimental prototype of the SILA antenna and measured power transmission coefficient for a microchip impedance Z_{chip}=10-j90 ohm.



b)

Fig.11: SILA Tag. Estimated activation distance for on-leg application when the reader emits 3.2 EIRP and the microchip has a sensitivity P_{chip}=10mW.

As suggested by equation (3) if the reader-tag distance were known, the change in the target permittivity could be theoretically detected by monitoring the power back-scattered by the transponder. Nevertheless, a single received data is not adequate to retrieve permittivity information in case of moving objects, or in applications in which the distance and the orientation of the tag with respect to the reader are not a-priori known (non-cooperative targets). To overcome these uncertainties, multiple independent back-scattered signals have to be collected by the reader**.** In the proposed platform, these signals are originated (Fig.13) from either a *cluster* of N tags co-located onto a same target, or from a single tag provided with N input ports under the condition that each port or antenna has a different input impedance. In particular, we denote with $G_{T,n}$ the radiation gain when only the *n*th port is fed and the others are closed to a reference load, and with $Z_{A,n}$ the input impedance at the n th port in the same conditions**.**

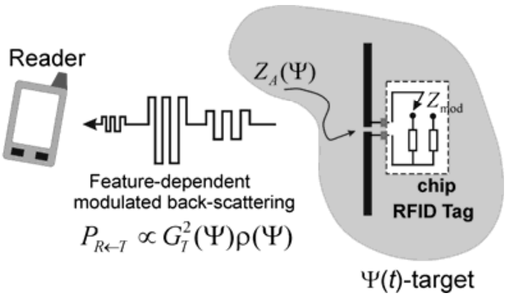

Reader
$Z_A(\Psi)$
Z_{mod}
Feature-dependent modulated back-scattering
$P_{R\leftarrow T} \propto G_T^2(\Psi)\rho(\Psi)$
chip
RFID Tag
$\Psi(t)$-target

Fig.12: Reader-tag scenario wherein the change of the target's features may produce a modulation of the backscattered power signal.

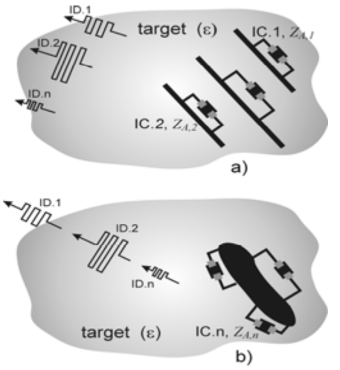


Fig.13: Multi-port tag systems: a) a cluster of co-located single-port tags; b) a single multi-port tag provided with multiple chips.

The multi-port system has to be designed so that, having fixed a target geometry and having chosen N different *reference permittivities* $\{\varepsilon_1,\varepsilon_2, \;..\;,\varepsilon_N\}$, the nth port impedance is matched to the microchip if the target's permittivity value is ε_n (e.g. $Z_{A,n}(\varepsilon_n)=Z^*_{chip}$). It means that, when the multi-chip system is placed on a real target, the ports will be differently mismatched ($Z_{A,n}(\varepsilon)\neq Z^*_{chip}$) and therefore they will originate independent back-scattered power signals, all of them carrying information about the target's permittivity. The resulting overall object is a *multi-port Sensor* RFID (S-RFID) *tag* that employs the same fabrication technology as the conventional RFID tags but, as shown later on, adds specific sensing capabilities to the typical identification features.

A. Sensing the target's permittivity

Depending on the link length d and on the particular design of the multi-port S-RFID tag, there will exist ranges of the target's permittivity for which either multiple ICs respond (overlapping of Active Sets) and hence the reader is able to collect multiple backscattered signals, or only a port is at most activated and the reader may receive a single ID. Two different sensing modes can be correspondingly achieved: *analog sensing* (multiple responding ICs) and *discrete sensing* or *classification* (single responding IC). For both the cases, it is useful to introduce the *Sensing range* $\mathbf{S}(d)$ of the multi-port S-RFID tag, as *the set of all the possible values of the target's permittivity which could be detected, in some way, at a distance d*. Only the analog sensing capability is here described, while a the complete theory could be find in [13].

If the tag has been designed for close reference permittivities { ε_n }, the port impedances will have similar (but not identical) power transmission coefficients τ_n so that multiple microchips will be turned on. In this case the multi-port system will have overlapped Activation Sets $A^{(n)}$. For any couplet of back-scattered signals received by the reader, each with a different modulation parameter $\rho_n(\varepsilon)$, it is possible to drop out the unknown reader-tag distance by calculating the *backscattered power ratio* $p_{i,j}$ between the received powers in equation (4),

$$p_{i,j}(\theta,\phi,\varepsilon)=\frac{P_{R\leftarrow T,i}(d,\varepsilon)}{P_{R\leftarrow T,j}(d,\varepsilon)}=\left[\frac{G_{T,i}(\theta,\phi,\varepsilon)}{G_{T,j}(\theta,\phi,\varepsilon)}\right]^2\frac{\rho_i(\varepsilon)}{\rho_j(\varepsilon)} \qquad (6)$$

However, $p_{i,j}$ is still affected by the uncertainty on the tag orientation (θ,ϕ) with respect to the reader. It is therefore required that the multi-port tag design satisfies the condition of proportional gain patterns, e.g. such that $G_{T,i}(\theta,\phi,\varepsilon)/G_{T,j}(\theta,\phi,\varepsilon)=f(\varepsilon)$. This condition could be roughly satisfied considering a cluster of two antennas having a similar geometry.

The retrieval procedure is now described by means of an example involving a two-port system, e.g. able to backscatter two different IDs toward the reader. An overlapping configuration between the activation ranges is illustrated in Fig.14. When both the ID_1 and the ID_2 are received by the reader, the unknown target dielectric permittivity ε_T will belong to the intersection of the two activation regions, e.g. $\varepsilon_T\in[A_1\cap A_2]$, and therefore the p_{12} ratio can be calculated as in (6). The value of the target's permittivity is hence retrieved by using a *calibration curve* $\varepsilon(p_{12})$ which associates to the actual backscattered power ratio, measured by the reader, a permittivity value for the target (Fig.15). Such a $p_{12}\rightarrow\varepsilon_T$ relationship is specific for the particular application, e.g. for a particular class (geometry) of targets and needs to be produced off-line through measurements or numerical simulations on simplified, or well representative, target models by operating a sweep of the parameter under observation and calculating the resulting backscattered power ratio. The application of such a technique therefore requires preliminary electromagnetic processing to produce calibration curves for the specific class of objects and the so obtained database, together with the retrieval procedure, have to be embedded in the reader's (post)processing unit. The S-RFID range $\mathbf{S}(d)$ is given by the merging of the Activation Sets shared by couplets of ports: $S(d)=A_m(d)\cap A_n(d)$.

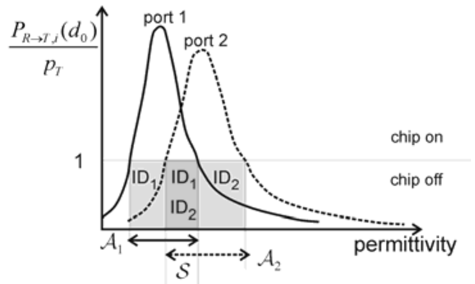
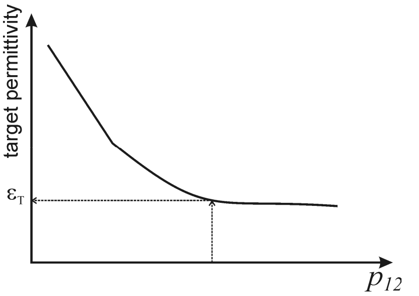


Fig.14: Typical Activation Sets, and Sensing Range, of a two-ports RFID tag, designed to work in analog-sensing mode.



b)

Fig.15: Example of calibration curves $\varepsilon(p_{ij})$ relating the measured backscattered power ratio to a target's permittivity value.

B. An experiment: sensing the filling percentage of a container

A very preliminary laboratory experiment is here discussed. The purpose is to demonstrate the validity of the basic principle concerning the possibility to govern the variation of the two-port tag antenna features with respect to the change of a real tagged body. In particular, we have designed a two-MLA (Meander Line Antennas) tag [13] for the sensing of the filling level, h, of a box. The variation of the shape of the target modifies the apparent permittivity sensed by the antennas and hence all their relevant parameters. With the aim to isolate and characterize the response of the antennas themselves, the experiment does not consider the RFID chip mounted on, nor the interrogation from a real reader. Moreover, readers able to measure the strength of the backscattered signals are up to now not so common on the market.

The target is again the perspex cubic box already used in Fig.5, now filled up to a height *h* (changed during the experiment) by sugar powder (ε_r=3,σ=0). The box is placed over a large copper sheet (1m × 1m) acting as an image plane.

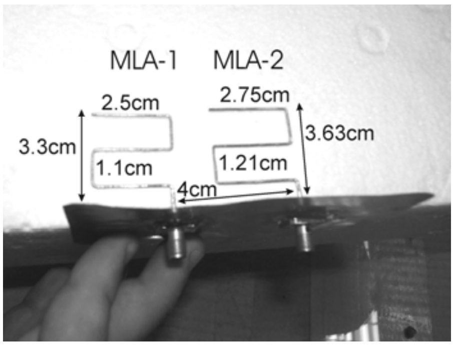
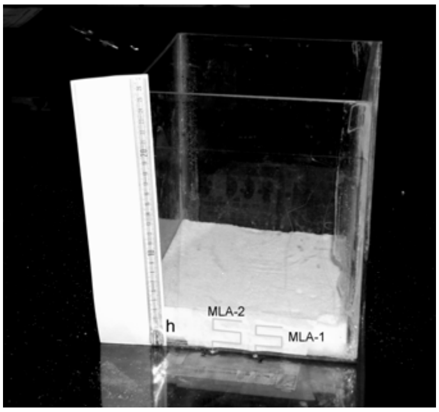


Fig.16: Powder level: meander-line-antennas prototype. Only half the structure is considered since the copper ground plane acts as an image plane.

The two MLAs are intended to be placed vertically on the external side of the box. Due to the presence of the ground plane, monopole configurations have been considered. Consequently, the impedance measurement results greatly simplified since no balun device is required. The antennas have regular turns and they have been optimized for the best τ such that the MLA_1 and MLA_2 are matched, at 870MHz, to the microchip (Z_{chip}=50-j200Ω) when the sugar level is *h*=10cm and *h*=0cm (empty box), respectively. The two MLAs are scaled replicas. The overall antenna heights are 3.3cm and 3.63cm, respectively. The distance between the MLAs' gaps is 4cm. The tag prototype has been fabricated by 1mm-radius copper wire, and is shown in Fig.16. The MLA monopoles are terminated on SMA connectors soldered on a 10cmcm copper sheet which is then placed in front of the perspex box as indicated in Fig.17. At this purpose, the large ground plane was properly drilled to accommodate the SMA connectors for the connection to the HP 8753C Vector Network Analyzer by means of flexible coaxial cables.



b)

Fig.17: Powder level: experimental set-up comprising the two-MLA tags and a perspex cubic box of 20cm by 20cm cross-section partially filled with sugar up to a level *h*. The antennas are fixed to the box's vertical side by means of an adesive ribbon.

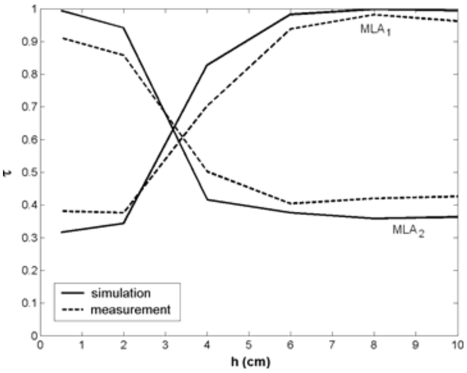


Fig. 18. Powder level: a) theoretical (simulated) and measured power transmission coefficients for the MLA_1 optimized for h=10cm, and MLA_2, optimized for an empty box (h=0). b) theoretical (simulated) and measured ρ_2/ρ_1 coefficients. the sugar level *h* inside the box.

The Z_{11} and Z_{22} impedance of the tag are measured, having de-embedded the SMA connectors, when the filling level is increased in the range $0<h<10$cm with steps of 2cm. The measurements are repeated in the reverse order (by emptying the box) and the two resulting sets of data finally averaged.

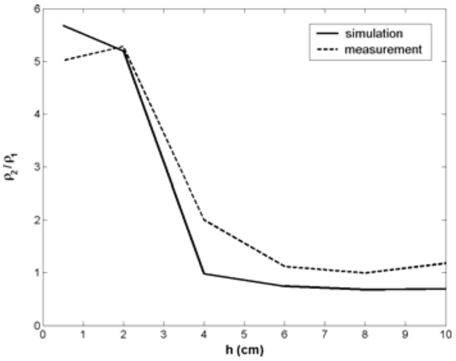


Fig. 19. Powder level: theoretical (simulated) and measured ρ_2/ρ_1 coefficients versus the sugar level h inside the box.

The resulting matching diagram of the power transmission factor, estimated by FDTD and measured, is shown in Fig.18. It is possible to appreciate, beside the nice agreement between simulations and measurements, that the τ-curves are monotonic with the change of *h* and that each port is rather mismatched in the condition for which the other one exhibits τ.

The sensing ability of the two-port tag depends on the calibration curve $p_{21}\leftrightarrow h$, and on the ratio in (6). Our equipment only permits impedance measurements and hence we have only considered the function $\rho_2/\rho_1[h]$, shown in Fig.19 However, we preliminary evaluated, by numerical simulations, that the gain ratio $(G_2/G_1)^2$ is nearly unitary for the whole considered variation of the level *h*, at least for observation in front of the tag.

The behavior of the calibration curve is monotonic, except for a very early short part with a good dynamic ($1< p_{21}<5$) when $2<h<8$. A saturation effect is clearly visible for levels higher than h=8cm, e.g. when the powder level greatly exceeds the vertical height of the antennas. In this condition, further increases in *h* do not produce additional variation of the antenna responses and such a change of the target could not be therefore sensed since the sugar powder acts as an infinite medium for the two antennas. The sensed dynamic of the powder level could be increased by designing longer tags or using a vertical arrays of properly tuned tags.

V. CONCLUSIONS

The proposed antenna configurations seem to be attractive for on-body applications or even for different usage involving highly dense dielectrics. The relevant number of degrees of freedom beside enabling a great matching agility over different microchips and conditions, gives the additional feature to allocate space to host sensors and electronics.

Preliminary simulations have demonstrated that, provided the microchip transmitter is sensitive enough (RF activation power equal or less than 10μW), the tag may be activated within a regular room so enabling a continuous monitoring of a moving human subject.

The proposed self-sensing platform, based on the multi-port concept, could find application in many security and industrial contexts, for instance i) to monitor (non metallic) container filled with low-loss liquids which could undergo changes along with the time, ii) to sense the filling percentage of a container, iii) to monitor the opening or the tampering of a case also in consideration that the target history could be stored by the reader in the rewritable memory of the tag's microchip

VI. ACKNOWLEDGMENT

The author wishes to thank the many persons who gave relevant contributions to this research in the last years, and in particular C. Stifano, C. Calabrese, D. Scarana, C. Occhiuzzi, L. Mattioni, G. Lovisolo, R. Pinto, S. Mancini, P. Tognolatti, L. Scucchia and S. Ricci.

REFERENCES

1. M. Aboelaze, F. Aloul, "Current and future trends in sensor networks: a survey", *2nd IFIP International Conference on Wireless and Optical Communications Networks*, pp. 551-555, Mar 2005. *Automation*, Vol. 1, pp. 537-544, 16-19 Sept 2003.
2. K. Lorincz, D.J. Malan, T.R.F. Fulford-Jones, A. Nawoj, A. Clavel, V. Shnayder, G. Mainland, M. Welsh, S. Moulton, "Sensor Networks for Emergency Response: Challenges and Opportunities", *IEEE Pervasive Computing*, Vol. 3, No. 4, pp. 16-23, Oct 2004.
3. Healthcare : L. Cheng-Ju, L. Li, C. Shi-Zong, W. Chi Chen, H. Chun-Huang, C. Xin-Mei, "Mobile healthcare service system using RFID", *IEEE Int. Conf. Networking Sensing and Control 2004*, Vol.2 pp.1014-1019, 2004
4. L.M. Reindl, A. Pohl, G. Scholl, R. Weigel, "SAW-Based Radio Sensor Systems", *IEEE Sensors Journal*, Vol. 1, No. 1, pp. 69-77, Jun 2001
5. A.P. Sample, D.J. Yeager, P.S. Powledge, J.R. Smith, "Design of a Passively-Powered, Programmable Sensing Platform for UHF RFID Systems", *IEEE International Conference on RFID*, Grapevine, pp. 149-156, 26-28 Mar, 2007
6. P.V. Nikitin, K.V.S. Rao, Theory and Measurement of Backscattering from RFID Tags, *IEEE Antennas and Propagation Magazine*, Vol. 48, No. 6, pp. 212-218, Dec 2006
7. J. Curty, N. Joehl, C. Dehollain, M. J. Delercq, "Remotely powered addressable UHF RFID integrated system", *IEEE J. Solid-State Circuits*, Vol.40, N.11, pp. 2193-2202, Nov. 2005
8. G. Marrocco, F. Bardati, "BEST: a finite-difference solver fo time electromagnetics", *Simulation Practice Theory*, N.7, pp. 279-293, 1999
9. G. Marrocco, "Rfid antennas for the UHF remote monitoring of Human subjects", *IEEE Transaction on. Antennas and Propagation*, N.55, N. 6, pp. 1862-1870, June 2007

10. C. Calabrese, G. Marrocco, “Meandered-Slot Antennas for Sensor-RFID Tags”, to appear on *IEEE Antennas and Wireless Propagation Letters*, 2008
11. D. S. Weile and E. Michielssen, “Genetic algorithm optimization applied to electromagnetics: A review,” *IEEE Trans. Antennas Propag.*, vol. 45, no. 3, pp. 343–353, Mar. 1997
12. G. Marrocco, L. Mattioni, C. Calabrese, “Multi-port sensor RFIDs for wireless passive sensing of objects – basic theory and early results”, to appear *on IEEE Trans. Antennas Propagat*., 2008
13. G. Marrocco, “Gain-optimized self-resonant meander line antennas for RFID applications,” *IEEE Antennas Wireless Propag. Lett.*, vol. 2, pp. 302–305, 2003
14. G. Marrocco, “The art of UHF RFID antenna design: impedance matching and size-reduction techniques”, IEEE Antennas and Propagation Magaz., 2008

A New Versatile Full Active RFID System

Guido Biffi Gentili, Claudio Salvador

***Abstract*— A novel approach to implement a full active, highly versatile RFID system is described. The new system is based on the concept that individual activation fields are established at natural control points for the purpose of awaking incoming Tags. In the new system a compact microwave Illuminator operating in the 2.45 GHz ISM frequency band generates a well defined wake-up footprint. The Tag replay is done with a robust long-range signal in the 433 or 868 ISM frequency bands assuring high immunity from the interferences and high real-time location accuracy.**

***Index Terms*— RFID, Tag, active, asset, RTLS, SALs**

I. Introduction

PAssive and active RFID systems are conceptually similar but very different regarding implementation, perspectives and market opportunities [1]. Much of the opportunities surrounding active RFID is the result of the added capability it has over passive RFID, which is somewhat constrained to dock door, conveyor or shelf applications. Active alternatives bring visibility to the depot, lot, port, healthcare centers and entire world. Additionally, the increasing ability for active Tags to sense and report their environments leads to an enormous range of new and exciting applications. At last active RFIDs are suitable to incorporate many technologies including Real Time Locating Systems (RTLS), Ubiquitous Sensor Networks (USN), Smart Active Labels (SALs), WiFi, Zig Bee and Ultra Wide Band (UWB).

In passive tagging small low cost Tags are energized and read by an RF Reader thorough the field produced by an activation antenna.

Passive RFID systems can be classified accordingly to the operating frequencies used, that are assigned by some open standards.

Low frequency RF passive RFIDs typically operate in the 135 KHz or 13.56 MHz bands and because the antenna dimensions are very small compared to the wavelength no matter the frequency used, radiation is negligible and the Tag is energized and communicates through the near inductive field that reduces its amplitude with the cube of the radial distance from the antenna.

Conversely UHF passive RFIDs [2] operate in the 868 MHz band and the reading antenna, typically of the patch type having dimensions comparable to the wavelength, posses high radiation efficiency [3]. Therefore the Tag is energized and communicates through the radiated far field that reduces its amplitude only with the first power of the distance from the transmitting antenna.

Passive RFID systems generally suffers of poor communication reliability and reduced reading range due to the low Tag signal strength and interference susceptibility. In order to increase the reading distance low-frequency Readers make use of a bulky antenna of the loop type and UHF Receivers transmit with high EIRP levels, up to 38 dBm.

Adding a small battery that continuously powers the RF communication circuitry the passive Tag is transformed in a semi-active or full active device.

Active RFIDs allows very low-level signal to be received by the Tag and the Tag can backscatter high-level signals to the Receiver. The robustness of the reading/writing communication channels is therefore increased and smaller Reader antennas along with reduced EIRP levels can be employed.

At present the most commercially available active RFID systems uses low frequency RF signals to awake the Tag because the "near field" technology is well assessed and many customized monolithic integrated circuits are available to fabricate system units. As an example of the current technology we can mention the "dual active" system produced by Axcess Inc., Dallas [4].

In our system the activation is done at microwave frequency, therefore we use radiative far fields instead reactive near fields for the awaking and writing function.

II. System description

The new full active RFID system [5], depicted in Fig. 1, take great advantage of using microwave frequencies for the purpose of awaking incoming Tags normally staying in a latency state.

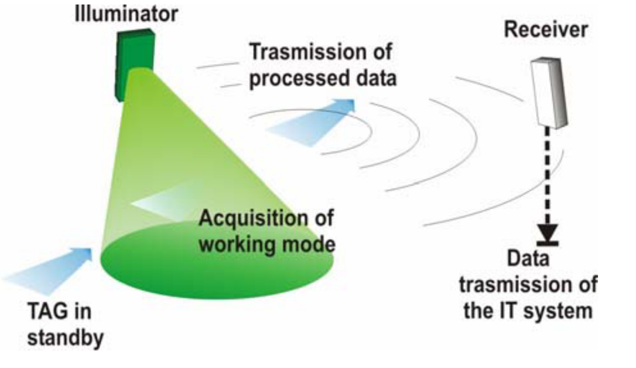

Illuminator
Trasmission of processed data
Receiver
Acquisition of working mode
TAG in standby
Data trasmission of the IT system

Fig. 1. Conceptualization of the new Active RFID System

G. Biffi Gentili is with the Electronic and Telecommunication Dept, University of Florence, via S. Marta 3 50139 Firenze (e-mail: guido.biffigentili @ unifi.it)
C. Salvador is with Advanced Microwave Engineering (AME), via del Monasteraccio 4, 50143 Florence (e-mail claudio.salvador@ameol.it)

The main advantages are:

- greater range of activation;
- very compact Illuminator architecture that integrates directional antenna;
- well defined footprint of the antenna pattern;
- possibility of transmitting data from the Illuminator to the Tag (writing and setting functions) with greater speed and channel robustness;
- smaller dimensions of the Active Tag.

A lower frequency, typically 433 or 868 MHz, is instead used for the long-range mono or bi-directional communication between the Tag and Receiver Units.

The use of two different physical layers, the former essentially dedicated to the awaking function and the latter to the data exchange, allows to greatly increase the operational flexibility of the system and the robustness to the interferences.

The main system units are the Illuminator, the Receiver and the Tag, synthetically described in the following.

A. *Illuminator Unit*

The Illuminator Unit (IU) essentially consists of a 2.45 MHz PLL oscillator cascaded with a OOK modulator and a medium power MMIC amplifier. The Illuminator antenna, shown in Fig. 2 is a low-cost stacked suspended substrate CP (Circularly Polarized) patch directly connected to the output of the amplifier.

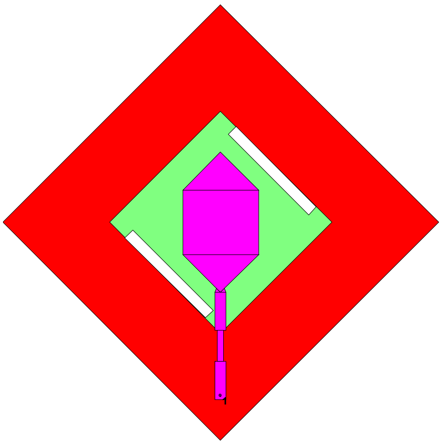


Fig. 2. The stacked 2.4 GHz Illuminator antenna on finite ground plane (80x80 mm). Green layer: patch; violet layer: feeder; reed layer: ground plane.

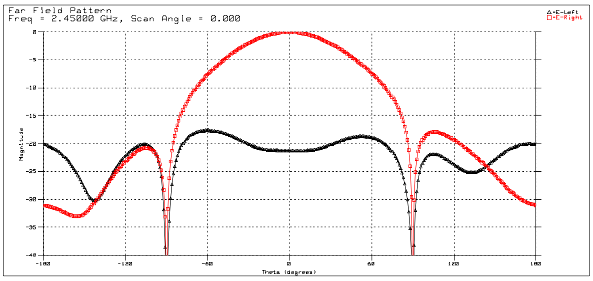


Fig. 3. Far field pattern of the stacked Illuminator antenna. Red curve: RHCP component; black curve: LHCP component

The radiation pattern of the antenna has 120 degrees –8 dB angular aperture as shown in Fig. 3.

Circular polarization is employed because the orientation of the Tag, that uses a linear polarized antenna, is unpredictable in many applications.

The signal transmitted by the IU provides a programmable ID code and few more setting commands that are used for programming the operation mode of the Tag entering in the field pattern of the Illuminator. The RF power output of the IU can be set from 0 dBm to 20 dBm depending on the regulations and the required Tag activation range that can reach up 20 m. By dynamically setting the IU output power the distance of the responding Tag can be estimated with a fairly good accuracy.

B. *Receiver Unit*

The Receiver Unit (RU) essentially consists of a 433 or 868 MHz single chip high sensitivity (-100 dBm) receiver and an omnidirectional monopole antenna. The signal received from the Tag is then put on the net using an Ethernet or 802.11b wireless interface.

C. *Tag Unit*

The Tag Unit (TU) is composed by a 2.45 GHz receiving section and a 433/868 MHz transmitting section, as depicted in Fig. 4. Both sections are controlled by a very low power microprocessor with 4 Kbites of data memory.

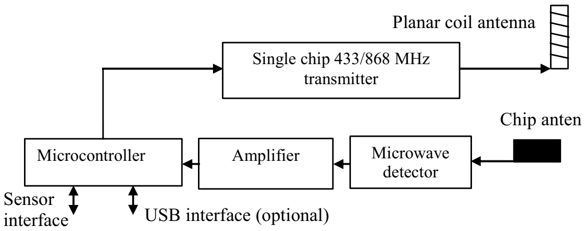


Fig. 4. Block diagram of the Tag Unit

The sensitivity of the receiving section depends on the Tangential Sensitivity (TS) of the microwave detector. If a low cost zero bias Schottky microwave diode is employed, a TS of -55 dBm is obtained. Better sensitivities up to -70 dBm are expected by employing as detector a more expensive tunnel diode. A low noise microwave amplifier stage can not be used due to its excessive current sink and cost.

To maximize the receiving sensitivity and the S/N ratio the diode is connected to the antenna through a narrow band matching section. The antenna is a very small coil wrapped around a high permittivity ceramic core.

The transmitting section, operating in the 433 or 868 MHz ISM frequency band consists of a commercially available mixed signal MMIC transmitter that consumes less than 1 µA in the idle state. The final stage of the transmitter is connected to a planar coil performing as a low efficiency narrowband radiating antenna. Despite this low efficiency the Tag can reliably communicate with the RU at distances more than 60 m outdoor and 20 m indoor.

All the Tag components have been carefully selected in order to greatly reduce dimensions, power consumption and costs. Fig. 5 shows the layout of the Tag board in the more compact version.

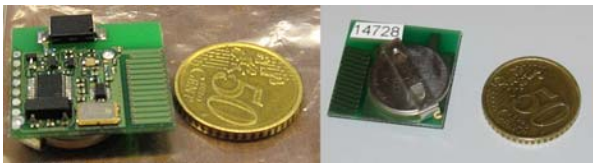


Fig. 5. The Active Tag Board with the 2.65 GHz antenna (upper side) and 433 MHz printed coil antenna (right side). (Courtesy of AME Srl)

The upper side of the board contains circuit components while the opposite side accommodates the battery, of the coin type. With this very small battery the Tag allows more than 200.000 transmissions that means 2 to 4 years life depending on the operating mode.

It is worth noting that the entire board, that integrates both transmitting and receiving antennas, has dimensions no much greater than that of an half Euro coin.

III. Operational modes

When the Tag enters in the activation field of an Illuminator the microwave diode detects it and a signal is generated that is sensed by the sleeping microcontroller. As the Tag wakes up, the microcontroller gets out of the stand-by mode and starts to receive the microwave OOK modulated signal. A proprietary protocol is structured as a data/command string and the information can be addressed to a specific tag or broadcasted.

Using setting commands it is then possible to set one of the operational mode that are pre-programmed in the Tag firmware.

As an example the tag could work as a conventional transponder, transmitting its own code after the interrogation from the illuminator or could act as a beacon, transmitting its code every N seconds. In the transponder mode the Tag associates to its own code the ID code of the Illuminator, previously received.

The flexibility of the tag firmware allows to implement in real time the operational mode(s) to be used to tailor system parameters for a specific function (for example positioning and tracking).

The protocol allows the tag memory to be written by the illuminator thus obtaining full read/write capabilities as is require in logistics and asset control.

IV. Recent research activities

Recent research activity has been focused in the fields of the Ultra Wide Band (UWB) technology [6] and beam-scanning Illuminator antennas. Both technologies are very useful to perform Tag location in indoor and outdoor complex environments and to implement advanced RFID applications.

Although often considered a recent breakthrough in broad-band wireless technology, UWB has actually experienced well over 40 years of technology advancements.

One of the more recent and fascinating application field for UWB technology has been in the area of precise localization [7]. In these applications, one takes advantage of the fact that very short pulse waveforms permit an accurate determination of the Time of Arrival (TOA) of a burst radiated from a transmitter and detected from a receiver. With distances computed from the relative time of flight measurements using a set of sparse receivers, one can then determine two-dimensional (2-D) and three dimensional (3D) position of an UWB transmitter using conventional multilateration algorithms.

Conceptually the narrowband Active Tag can be converted in a UWB Tag simply by substituting the UHF transmitter with a sub-nanosecond pulse generator radiating through a UWB antenna [8], as depicted in Fig. 6.

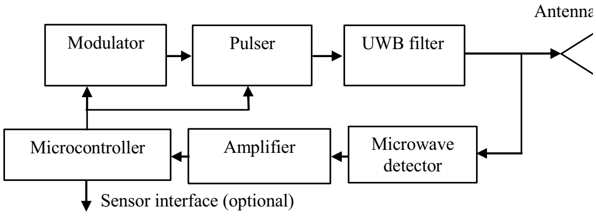


Fig. 6. Block diagram of the UWB Tag Unit

The first research objective was the design of a transmitter capable of generating Gaussian pulses with a duration ranging from 200 to 400 pS, according to the recently approved ETSI standards.

Using the AWR design environment [9] a nearly “all digital” pulser architecture [10] has been simulated in frequency and time domain. Results proved that the proposed pulser circuit is capable of generating 0.3 W peak 300 pS low duty cycle pulses with an average power consumption of only 60 μW @ 3 V supply. It is expected that the number of transmission (pulse burst) could reach million using this UWB transmitter powered by the same coin sized battery. This means that the Tag life could reach then years.

Antennas for UWB Tag operating in the 3.1 – 10.6 GHz band are difficult to be realized in very compact size [11].

In order to maximally reduce the overall dimensions, a dielectric loaded conical antenna [12], shown in Fig. 7, has

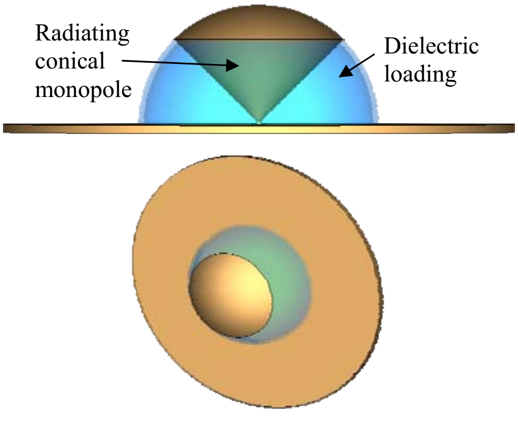


Fig. 6. The optimized UWB hemispherical antenna with 16 mm radius and 90 degree cone aperture.

been devised and analyzed.

EM simulations made with CST Microwave Studio [13] have shown that the designed hemispherical shaped antenna has VSWR<2.5 in the entire UWB band. Planar UWB antennas [14], [15] could also be employed due to their compactness and very low cost.

The second research objective was the development of a planar array for the most advanced RFID applications that require an Interrogator antenna having beam scanning capability in azimuth (horizontal plane) and beam shaping in elevation (vertical plane). The proposed array that derives from that originally designed for radar polarimetry applications [16], consists of two to four circular polarized serially feed linear microstrip subarrays, as depicted in Fig. 8.

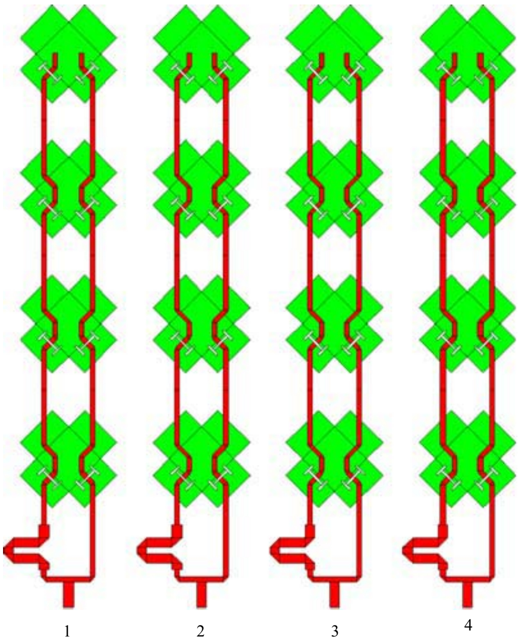


Fig. 8. The multilayer planar array for advanced RFID applications. Green layer: crossed patches; red layer: microstrip feeding lines.

Patches are serially fed through H shaped slots etched on the ground plane separating the patch layer from the feeding layer. Beam shaping in the vertical plane is obtained by a synthesis procedure that varies the slot size to adjust the coupling level to each patch so as to realize the prescribed current distribution.

Beam scanning in the horizontal plane is implemented by using MMIC digital phase shifters at the input ports of each subarray or a coplanar Butler matrix, if only discrete azimuth angles are required.

V. Conclusions

The active RFID system described in this article has been developed in the context of a scientific and technical collaboration between University of Florence (Antenna and Microwave Laboratory) and AME (Advanced Microwave Engineering), a spin-off company founded in 1999.

The new full active RFID system has evolved from a laboratory prototype to a robust and reliable product capable of responding to the impelling demands of system integrators in the fields of security, healthcare, homeland defense, supply chain, asset control, automatic identification and sensor networks. In this last area the Tag has sensor embedded or inputs where sensors can be connected from outside world. Temperature, shock, humidity, moisture, vibration and motion are common sensing devices but any two wire digital sensor output can be connected to trigger an alarm condition or transfer measured data. UWB technology is very promising to implement ad hoc and ubiquitous sensor networks because of the expected very long life of the Tag and the possibility of accurate localization and tracking also in dense and complex environments.

References

[1] K. Finkenzeller."RFID Handbook". *Wiley & Sons*, second edition, 2003.
[2] C. Chen, Majid B. Nejad, Li-Rong Zheng, "Design and Implementation of a High Efficient Power Converter for Self-Powered UHF RFID Applications", *IEEE International Conference on Industrial and Information Systems* 2006, Sri Lanka.
[3] K.V. Seshagiri Rao, Pavel V. Nikitin, and Sander F. Lam. "Antenna design for UHF RFID tags: a review and a practical application". *IEEE Transactions on antennas and propagation*, Vol. 53, Dec. 2005.
[4] Axcess International Inc, Dallas. www.axcessinc. com
[5] Brevetto Italiano n° FI2000A000221 del 06/11/2000; European patent EP01126317.5 (2008 - Granted).
[6] R.J. Fontana, "Recent System Applications of Short-Pulse Ultra-Wideband Technology", *IEEE Trans. on Microwave Theory and Tech.,* Vol. 52, No 9, Sept. 2004.
[7] J. Schroede, S. Galler and K. Kyamakya "A Low-Cost Experimental Ultra-Wideband Positioning System" *IEEE Conf. on Ultra Wideband,* ICU 2005, Sept. 2005.
[8] A.F. Kardo-Sysoev "Generation and radiation of UWB signals", *33rd European Microwave Conference*, Munich 2003.
[9] Microwave Office. Applied Wave Research Inc.
[10] L. Smaini, et Alii, "Single-Chip CMOS Pulse Generator for UWB Systems", *Proceedings of ESSCIRC,* Grenoble, France, 2005.
[11] H.G. Schamtz, "Introduction to Ultra-Wideband Antennas", *IEEE Conference on Ultra Wideband Systems and Technologies,* Reston, Virginia USA, Nov.2003.
[12] G. Biffi Gentili, M. Cerretelli, L. Cecchi, "Coated Conical Antenna for Automotive Application", *Journal of Electromagnetic Waves and Applications*, vol. 18, no. 1, 2004.
[13] MW Studio 5 by CST Computer Simulation Technology, Darmstadt, Germany.
[14] S. Gupta, M. Remesh and A.T. Kalghatgi, "Ultra Wideband Embedded Planar Inverted Conical Antenna", *Microwave and Optical Technology Letters,* Vol. 48, No. 12, Dec. 2006.
[15] C.J. Pan, C. Lee, C.Y. Huang and H.C. Lin "Band Notched Ultra-Wideband Slot Antenna", *Microwave and Optical Technology Letters,* Vol. 48, No. 12, Dec. 2006.
[16] A. Vallecchi and G. Biffi Gentili, "A low-cost shaped-beam hybrid-feed microstrip planar array antenna for X-band polarimetric radar systems," in *2006 European Radar Conference*, Manchester, UK, 13-15 Sept. 2006.

A 900MHz UHF RFID Reader Transceiver IC

Issy Kipnis, Scott Chiu, Marc Loyer
Intel Corporation

Abstract-- **The architecture and design of a single chip transceiver for multi-class worldwide RFID reader applications is described. The transceiver utilizes a direct conversion architecture on the receive and transmit paths. The chip integrates all the RF blocks, synthesizer, sigma-delta converters, digital filtering, and the digital modulation and demodulation functions. The chip delivers an output power of +11dBm in linear mode, and +20dBm under class-C amplitude-modulation operation, and has a sensitivity down to -85dBm in the presence of a 0dBm self-jammer signal. Total power consumption is 1.25W. The die is fabricated on a 0.18μm SiGe BiCMOS process.**

I. INTRODUCTION

Automatic identification procedures have become indispensable in today's service, manufacturing, purchasing and distribution, and material flow industries. The omnipresent barcode labels that revolutionized the identification systems many years ago, are being found to be inadequate because of their low storage capacity, the fact that they can not be reprogrammed, and the line-of-sight requirement to read them out [1]. Alternatively, storing data on a chip has been available for some time now in the form of smart cards (like telephone and bank cards) that operate upon mechanical contact. Clearly a contact-less method of transferring the data is more flexible and enables many new user models. Radio Frequency Identification (RFID) systems provide a more optimal technical solution. Near-field RFID systems using inductive coupling and operating between 100KHz and 30MHz are fully deployed (our Intel badges, for example). But the read range for those systems is limited to distances shorter than 1m. A nascent technology using RFID tags that operate in the UHF 860-960MHz band is presently being deployed worldwide. The benefits of UHF RFID systems are extended range (up to 10m) and fast data rates (up to 640Kbps). Although supply chain is the main application driving the technology, it is believed that item-level tagging will follow once the economies of scale are in place, making these RFID systems as ubiquitous in the near future as barcode technology is today.

A UHF RFID system consists of a reader and passive tags. The communication between reader and tags is half duplex, as indicated in Figure 1 [2]. The reader initiates an inventory round by sending an un-modulated CW signal. The tag, being a passive device, harvest the RF power from the reader to power up. The reader continues to send RF power throughout the end of the inventory round (the time needed to read all tags within a field of view) to keep the tags alive. Once the tags are powered up, the communication is initiated by the reader by modulating its carrier. After the reader completes the first command, it stops modulation and sends a CW RF signal. At this time the tag or tags of interest respond by reflecting, or back-scattering, the reader's CW signal. The tag modulates the back-scattered wave by changing the amplitude or phase of its antenna's impedance.

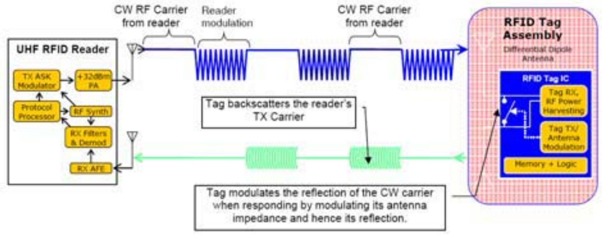
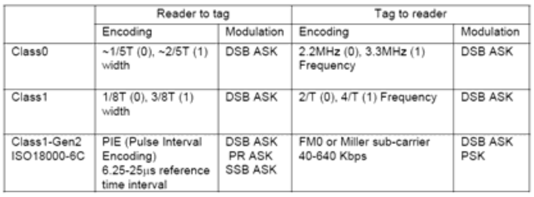

CW RF Carrier from reader
Reader modulation
CW RF Carrier from reader
UHF RFID Reader
TX ASK Modulator
+32dBm PA
Protocol Processor
RF Synth
RX Filters & Demod
RX AFE
Tag backscatters the reader's TX Carrier
Tag modulates the reflection of the CW carrier when responding by modulating its antenna impedance and hence its reflection.
RFID Tag Assembly
Differential Dipole Antenna
RFID Tag IC
Tag RX, RF Power Harvesting
Tag TX/ Antenna Modulation
Memory + Logic

Fig 1: UHF Reader System

UHF RFID systems are presently deployed using three different protocol standards: class0 [3], class1 [4] and class1-gen2, also known as ISO 18000-6C [5]. There is a chronological order to the appearance of the protocols, each being progressively more advanced. With time it is expected that Gen2 tags will dominate the market (or perhaps even newer versions of the protocol), but for the time being all three classes need to be supported. The chip was optimized for Gen2 operation, because of the prevalence of that class of tags in the market.

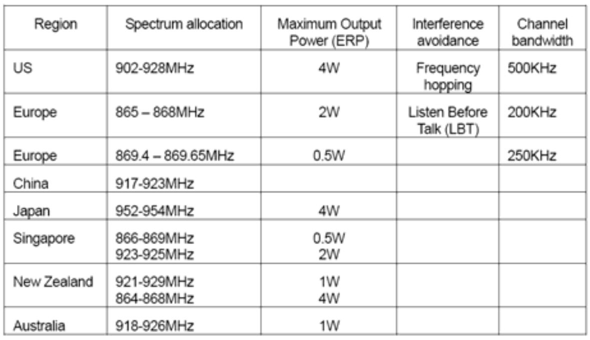
Table 1 indicates the encoding and modulation for the forward and reverse links for all three classes of tags. The modulation is always a variant of Amplitude Shift Keying (ASK): either Double Side-Band (DSB), Single Side-Band (SSB) or Phase-Reversal (PR)



	Reader to tag		Tag to reader	
	Encoding	Modulation	Encoding	Modulation
Class0	~1/5T (0), ~2/5T (1) width	DSB ASK	2.2MHz (0), 3.3MHz (1) Frequency	DSB ASK
Class1	1/8T (0), 3/8T (1) width	DSB ASK	2/T (0), 4/T (1) Frequency	DSB ASK
Class1-Gen2 ISO18000-6C	PIE (Pulse Interval Encoding) 6.25-25μs reference time interval	DSB ASK PR ASK SSB ASK	FM0 or Miller sub-carrier 40-640 Kbps	DSB ASK PSK

Table 1. Encoding and modulation for all three classes of tags

Although the UHF RFID system occupies in general the 860-960MHz IMS band, there are specific regulations by geography that need to be met. The specific air-interface requirements for the US and Europe, for example, are specified in references [6] and [7] respectively. The local regulations specified the spectral mask, maximum output power, interference avoidance mechanism , channel bandwidth, etc. Table 2 shows a summary of some of the air-interface requirements by geography. This paper describes a single chip transceiver IC that contains all the RF, mixed-signal and digital baseband functions of the physical layer of a UHF RFID reader. The protocol microcontroller, the crystal oscillator and the RF power amplifier required to deliver 1W (+30dBm), are the only elements not included on the chip. The body of the paper will describe the architecture and topology of each of the main blocks of the IC.



Region	Spectrum allocation	Maximum Output Power (ERP)	Interference avoidance	Channel bandwidth
US	902-928MHz	4W	Frequency hopping	500KHz
Europe	865 – 868MHz	2W	Listen Before Talk (LBT)	200KHz
Europe	869.4 – 869.65MHz	0.5W		250KHz
China	917-923MHz			
Japan	952-954MHz	4W		
Singapore	866-869MHz 923-925MHz	0.5W 2W		
New Zealand	921-929MHz 864-868MHz	1W 4W		
Australia	918-926MHz	1W		

Table 2. Air-interface requirements by geography

ARCHITECTURE

The functional block diagram for the chip – code named Tilden are shown Figure 2 and Figure 3. The receiver is a direct conversion architecture. After down-conversion the major part of the DC is removed by resetting AC-coupling capacitors. The analog IF filter provides coarse channel selectivity. For class1 operation the poly-phase IF filter is configured as a low-pass filter with a zero-IF, the IF mixers are latched in transparent mode and the signal is processed by a 48MS/s sigma-delta analog-to-digital (ADC) converter. For class0 operation the poly-phase IF filter is configured as a band-pass filter with a low IF=2.75MHz. The subcarrier modulated signal is then down converted to baseband by the mixers and further processed by the ADC. The coarse analog filtering is supplemented by sharp and well controlled digital filtering. The demodulation is also performed digitally.

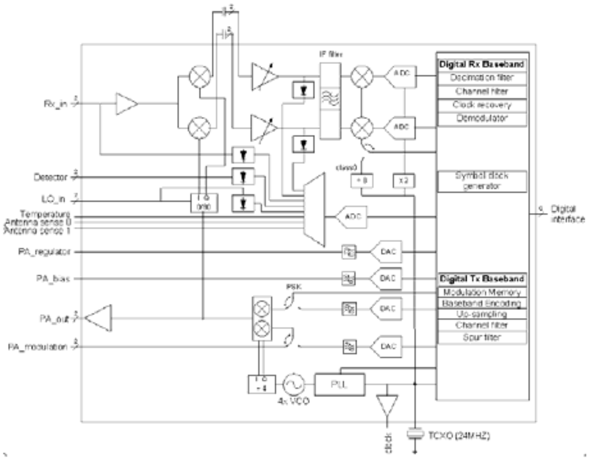


Fig.2: Functional analog block diagram

The transmitter uses a high-efficiency polar-modulation architecture for DSB-ASK modulation and a traditional linear IQ architecture for SSB-ASK and PR-ASK modulation. The baseband encoding and pulse-shaping is done with a look-up table to minimize latency. In the case of SSB-ASK transmission the baseband signal is filtered with a Hilbert filter to create a complex IQ signal with suppressed negative frequencies. The signal is then offset in frequency to center the SSB-ASK spectrum in the channel. The digital I and Q signals are converted into the analog domain by sigma-delta DACs. In DSB-ASK transmission the baseband encoding and pulse shaping is done just as for SSB-ASK, but the shaped signal is pre-distorted to compensate for non-linearity in the AM transfer function. The pre-distorted AM control signal is converted into the analog domain by the Q sigma-delta DAC.

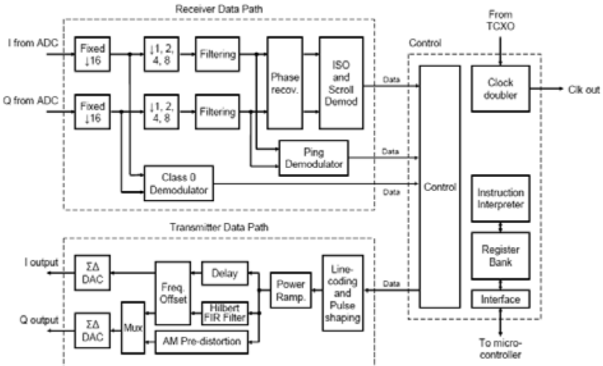


Fig. 3: Functional digital block diagram

The PLL is an integer-N synthesizer with a fully integrated 4x VCO. The loop filter is external. The clocks for the digital blocks are derived from a 24 MHz reference frequency coming from an external TCXO. The sigma-delta DACs are running directly on the 24 MHz signal. The sigma-delta ADCs are running on a 48 MHz clock generated by an integrated frequency doubler.

There is an auxiliary ADC that serves various functions (temperature, antenna sense and power detection). There are two auxiliary DAC for PA-regulator and PA-bias functions. Tilden supports two interfaces, one low speed parallel interface with a data rate of up to 20Mbps and one serial interface with data rate of 150Mbps for reader-to-tag and up to 450Mbps for tag-to-reader communication. The interfaces are multiplexed on the same pins, and the interface is determined during power up. Both interfaces are operated at 3.3V. Low level instructions are written into a FIFO and executed one at a time. All information is transferred via the register bank. The control of the chip is state machine driven.

RECEIVER RF FRONT-END AND ANALOG BASEBAND

The biggest challenge for the receiver front-end is coping with the leakage from the full power CW signal being transmitted to keep the passive tags powered up. As there is no frequency division multiplexing, this leakage falls in the middle of the receive band. The typical transmit output power at the antenna connector is +30dBm, and due to antenna isolation (in the case of separate Rx and Tx antennae) or antenna reflection (in the case of single Rx/Tx antenna) the receiver needs to tolerate a self-jammer signal of 0dBm at the LNA input. The large power of the self-jammer limits the amount of gain of the LNA/mixer. At the output of the mixer the self-jammer is converted to DC, and is easily removed via AC coupling capacitors. The transceiver is expected to achieve a sensitivity level of about -80dBm, which requires a front-end circuit with instantaneous dynamic range of over 100dB.

The LNA uses a common base configuration, as it is best suited to handle the large self-jammer power. The collectors of the LNA feed directly into the mixer quad. The LO signal is tapped at the output of the external PA, it is attenuated and used to drive the poly-phase filter quadrature splitter. There are buffers on the LO path to allow a large range of LO input powers and to optimally drive the quadrature mixers. Simulations indicate that the front end dissipates 300mW. It has a voltage gain of 13dB, a noise figure of 12dB and 18dB without and with self-jammer present, respectively, and an input 1dB compression point of +2dBm.

Because of the nascent nature of the UHF RFID market, with its standards still under development, It was critical to add as much programmability as possible to the receive chain. Therefore most of the signal processing was pushed to the digital domain. The main purpose of the receiver baseband analog section, thus, is to perform the following signal conditioning functions.

1. Sample-and-hold offset correction circuit to remove the DC voltage created by the reflected/coupled transmit carrier that is down converted to DC in a direct-conversion receiver. The sample-and-hold circuit varies the high pass filter bandwidth from wide (sample phase during tag initialization,) to very narrow during receiving tag responses (cancellation phase), and to close to 0 during reader transmit (hold phase). Not being limited by one filter bandwidth enables the reader to have a short turn-around time, while keeping the low frequency received data contents intact. Figure 4 shows a high level representation of the correction circuit – the shunt switch acts as the sampling switch, and the series switch the hold switch. During the tag power up phase, the shunt switch is closed, and the series switch is open; transmit coupled carrier power is sampled onto the external AC coupling capacitor. When the reader talks to the tags, the shunt switch and the series switch are open to hold the DC value in the AC coupling capacitor. The series switch closes after the reader ends talking to the tag, and shortly before the reader starts receiving tag responses. Since there is no settling of the high pass filtering, the turn-around time from transmit to receive is greatly reduced.

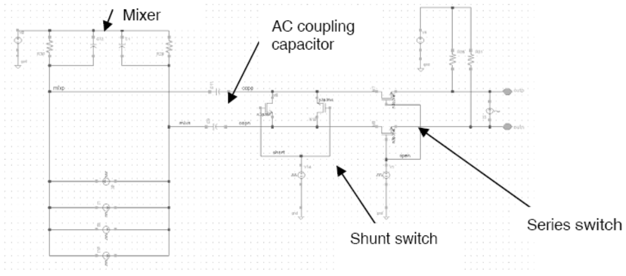


Fig. 4: DC offset correction circuit

2. Low noise baseband amplifier (IF-LNA) to optimize the receiver dynamic range with respect to noise figure and linearity: Due to the high RF power coupled in at the receiver antenna interface, the RF front-end gain is very limited. The IF-LNA's noise performance thus greatly impacts the receiver noise figure – the target input referred noise voltage is 3.6 nV/sqrt(Hz). To achieve this performance down to 10KHz, bipolar devices are used to minimize 1/f noise contribution; a 4-input pseudo-differential 2-stage OPAMP architecture is selected to meet gain, linearity, and noise specifications.

3. A programmable baseband analog filter to attenuates the blockers and interferers to meet various local regulations and standards. A 5th order active RC Chebyshev low pass filter with 0.5 dB pass-band ripple is used, with 64 programmable bandwidths ranging from 120KHz to 1.2MHz, to meet different local regulations. A Chebyshev characteristic was selected for better rejection of adjacent/alternate channel interferers. The filter transition band is placed far above or below the 3rd order harmonic of data modulation so that the filter phase non-linearity does not affect symbol timing recovery. The filter bandwidth is tuned by calibrating the capacitance using a master-save configuration after selecting a target resistance value. After tuning the master with a duplicate

of the resistor and capacitor array, the tuning results are copied to the slave, the filter under tuning. The low pass filters on I and Q channel can be cross coupled to form a complex bandpass filter [8] for Class-0 RFID tags. The purpose is to attenuate the image of the Class-0 tag response at IF before using a baseband mixer to center the response around DC in the following stage.

4. Baseband VGA/IF mixer. This block performs a 2nd gain control stage in the analog signal conditioning chain. The purpose is to maximize the input signal to the ADC without clipping. For Class-0 tag demodulation mode, it also serves as the IF down-converter to shift the tag response around DC. The mixing quad is latched during other modes of operation.

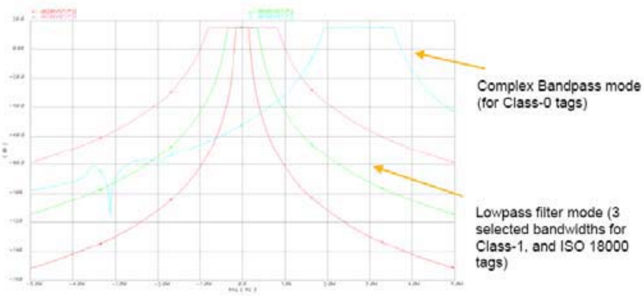


Fig. 5: Analog baseband filter frequency response

5. Receive ADC. The chip uses a continuous-time, 3rd-order 48MHz ΔΣ ADC with 1-bit feedback to interface to its on-chip DSP core. The implementation uses cascaded resonator with feed forward structure (CRFF) to optimize the integrator range. A limiter at the last stage integrator degrades the loop filter to 2nd order in case of overload to guarantee stability in all conditions. The sampling clock is derived from the 24MHz crystal reference using a doubler. To mitigate the sensitivity to deterministic clock jitter from the doubler, a discrete time switched capacitor is used on the 1-bit DAC feedback [9].

RECEIVER DIGITAL BASEBAND

The 48 MHz sigma delta serial output exists for both I/Q channels. This is followed by a fixed decimation chain of 16 to produce 3 MSps (samples per sec) waveforms. ISO 18000-6C standard supports data rates in the range [40,640kbps]. To keep the relationship between signal bandwidth and the filter sample rate relatively constant for efficient use of the main digital channel filter, an additional data rate dependent decimation is employed. The benefit of this is that for the slowest to fastest data rate we can maintain the same channel filter size.

The digital channel filter block provides the lowpass narrowband digital filtering to preserve the information content of the inband signal whilst attenuating out of band noise and interference.

This is achieved using a combination of FIR and IIR filter. The FIR filter is a programmable filter with a minimum tap length of 36 taps scaleable in 6 tap increments to 72 taps. In addition to the DC removal provided by the analog AC coupling, Tilden also has a digital DC removal scheme based on a DC estimator filter. This is implemented as a 2nd order IIR Butterworth filter with a very low cut off frequency which subtracts from the output of the channel filter.

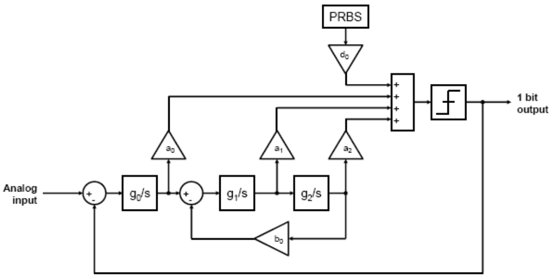


Fig. 6: Receive ADC Block diagram

The phase recovery block follows the digital filter whose function is to re-align the modulated signal with the real axis by applying a phase rotation to I/Q samples. Due to the type of modulation used by the tag all of the signal information is present on the I channel. This is required only for class 1G1 scroll and all ISO modes.

Depending on the RFID protocol selected one of three separate demodulators are used;

- Class-0; demodulation is performed by computing a 16 bin DFT and comparing the peak power in the data-0 bins to the peak power in the data-1 bins. The soft decisions for the data-0/data-1 bins are passed onto the MAC for bit slicing.

- Cclass-1 Ping ID; a power triggered zero crossing detector is used. Ping reply is a special case where the reply contains no preamble which makes timing recovery difficult. Therefore, demodulation is performed by estimating data 0/1 bits using a zero crossing detector followed by matched filters that are matched to 32 combinations of a ping reply for a given ping ID.

- Class-1 Scroll ID; and all modes of ISO 18000-6C; uses a coherent demodulator comprised primarily of a rate estimator/timing recovery and bank of matched filters which attempt to match the received baseband signal against the modulation basis functions.

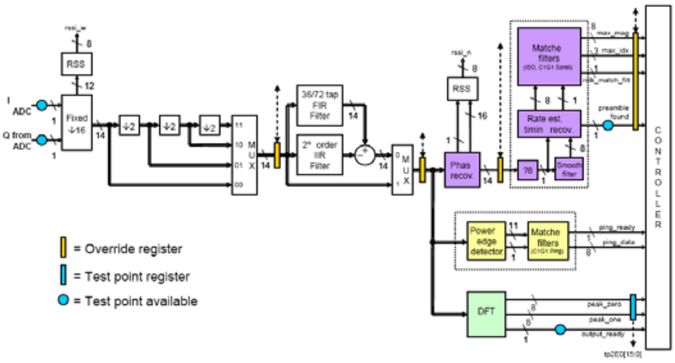


Fig. 7: Tilden receiver baseband

- To reduce the required bit widths in the demodulator the signal is normalized. To increase timing accuracy the incoming signal is upsampled by 8 using a 16-tap polyphase filter. The timing recovery scheme is comprised of closed loop system which is used to adjust the phase of the sampling clock. This uses a Gardner timing detector to measure the sample time error with a first order digital filter as well as a NCO which generates the sample clocks

- Due to the large startup frequency errors produced by the tags a separate data rate estimator is required. This attempts to match the received signal to a preamble pilot tone using a bank of matched filters operating at different data rates. Once the rate is estimated a frequency adjustment is provided to the NCO.

- The preamble correlator is comparing the sign of the signal with the sign of the expected preamble waveform. When a match is found a signal is generated and sent to the controller. At the same time the matched filter bank is enabled. The filter bank consists of eight matched filters running in parallel. Each filter is matched to a three-bit sequence comprising of the coding basis functions for ISO schemes FM0, Miller and class1G1 scroll. The largest estimate of the received bit from this filterbank demodulates the data to either data-0 or data-1.

In addition to the tag demodulation, the baseband receiver also provides information about the signal power of the baseband signal. A wideband/narrowband RSSI measurement is available pre/post channel filter. The integration time for both RSSI measurements is made programmable.

This information is utilized for programming the IF Mixer/AGC to provide the receiver with the maximum dynamic range in presence of interferers. Overall, AGC programming is controlled by the MAC. For the ETSI standard [7], the listen before talk protocol uses the RSSI to determine whether a selected channel is occupied before it attempts a reader to tag interrogation.

TRANSMITTER DIGITAL BASEBAND

To reduce the turn-around time between transmit and receive, pre-calculated waveforms are stored in a look-up table to avoid the group delay that would result if a programmable digital filter was used. The symbol look up methodology is implemented using a programmable custom microcode sequencer as follows. The MAC microcontroller accesses the transmit function through a 16-element deep reader-to-tag (R2T) FIFO. A number of commands are supported in this FIFO, including single bit transmit, byte transmit, random sequence transmit, RF ramp up and ramp down, end-of-transmission flag (which includes a field for delaying the enable of the receiver), and a number of user-programmable transmit sequences.

The basic wave shaping function is implemented as a 24-bit wide, 64 location deep lookup table, containing the wave shapes for a data-1 symbol and a data-0 symbol. In general 10 samples points for each Tari are stored. This table is accessed through another table (the microcode table) containing instructions for stitching portions of the wave-shape table together to form more complex waveforms. Finally, the microcode table is accessed via the commands in the R2T FIFO. Certain commands written into the FIFO identify a range of instructions to interpret in the microcode table, which in turn specify the locations to be accessed in the waveform LUT's. Thus, a single command in the FIFO can generate a complicated waveform, such as a framesync (Fig 8).

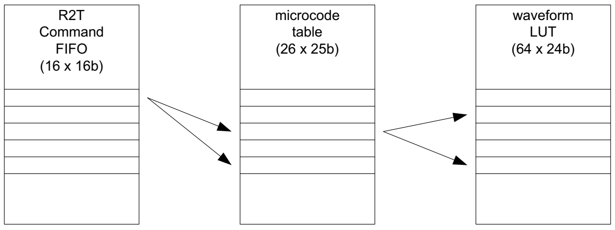


Fig. 8: Transmit FIFO, microcode and look-up table

To support different modulation coding, the baseband data path is reconfigured to provide real or I/Q complex modulation; in addition, a single-sideband mode is devised to include a programmable Hilbert transform filter, and a CORDIC based frequency shifter in the signal conditioning path. A 5th order programmable polynomial pre-distortion block is included to cope with the nonlinearities introduced by the PA in supply modulation mode.

TRANSMITTER ANALOG BASEBAND AND RF

The digital encoded data is converted to analog through a single-bit 24MHz 3rd order ΔΣ DAC with a transfer function similar to receive ADC (Fig.). Depending on the input sampling rate, an up-sampling may be performed to relax the anti-aliasing requirements. The 24MHz crystal reference clock is used in the single-bit output of the sigma delta DAC to avoid transmit phase noise created by sampling clock jitter.

The reconstruction filter is a 6th order filter (5th order Butterworth with one more pole at the output) with programmable and tunable bandwidths. It utilizes the same master-slave tuning controller as the receiver filter does. The high order analog filter is chosen not only to provide anti-alias capability, but also some improvement in spectral shaping for meeting local regulations.

An important differentiation of an RFID reader are the temporal envelope shaping constraints, in addition to the spectral mask requirements. Since the tags behave as an envelope detector, the reader needs to transmit a waveform with well defined temporal characteristics to aid the tags in demodulating the signal. This time domain constraints include envelope rise/fall time, envelope ripple, bit period, and modulation depth on a bit-by-bit basis. Figure 10 shows an envelope of the modulated waveform, and the associated timing limits.

After the modulated Tx signal has been converted to analog and filtered out, it is directly converted to RF by a traditional I/Q upconverter mixer. The LO quadrature signals are obtained from the divide-by-four circuit following the VCO. The Tx analog/RF chain features 16dB of variable gain in 1dB increments, to control the output power; the DAC provides an additional 0.5dB step of power control. After the mixer the signal is applied to a three-stage amplifier with integrated inter-stage matching. The complete Tx chain has a target OIP3= +29dBm and OPsat= +22dBm.

SYNTHESIZER

Tilden carrier generation specification is determined by a combination of the various air interface standards and applicable local regulations. Frequency resolution is defined through local regulations: i.e., 250 KHz in the US, and 25KHz in Europe. Settling time is not explicitly defined by any local regulations/standards. However, it is desirable to keep it to within a few hundred µsec to enable fast monitoring of available spectra. Because of the lax requirements on settling time an integer-N PLL architecture with a frequency resolution of 25KHz was selected.

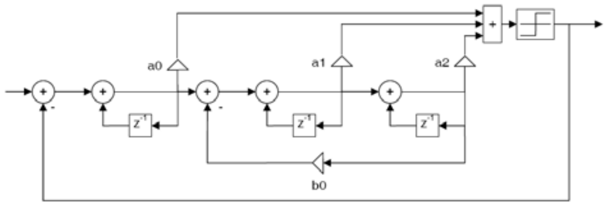


Fig. 9: TX DAC structure

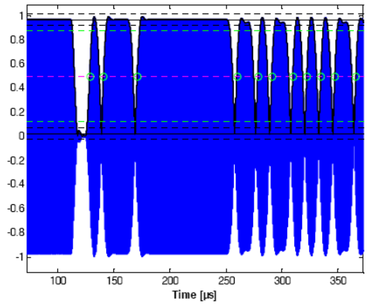


Fig.10: PR-ASK RF Envelope Waveform

Phase noise specification is set by the [7] adjacent channel interferer and out-of-band spurious emission requirements. This leads to a carrier requirement of -116 dBc/Hz at 200KHz offset, and -144 dBc/Hz at 3.6MHz offset. Due to the variations of local regulations, and the stringent phase noise demanded, the PLL loop filter (a 2nd order low-pass filter) is kept off-chip. Extra attention is applied to protect the single-ended control voltage input to shield coupled noise on-chip as well as from the application board. A low phase noise LC VCO was designed which achieves a phase noise of -119 dBc/Hz at 200KHz offset, and -146 dBc/Hz at 3.6MHz offset. This VCO is constructed using cross coupled NPN transistors (for better close-in phase noise) with a high Q symmetric inductor (using the top layer metal with 5.5 mΩ/□ sheet resistance). Coarse frequency tuning is performed using a bank of 8 switched capacitors, and fine frequency tuning is applied using a junction diode varactor to make sure the VCO tank Q is dominated by the inductor. To avoid on-chip noise coupling at the carrier frequency, the VCO is chosen to operate at 4 times the carrier frequency (~ 3.6GHz) after trading off the inductor and varactor performance. Separated/dedicated VCO supply and reference supply are used to further decouple the supply noise from other on-chip blocks. In addition, this block is surrounded by a guard ring to mitigate substrate noise coupling. To keep the VCO in its optimal phase noise operating condition over process and temperature variations without long term reliability concerns, an amplitude detector with programmable VCO current bias works under the control of the MAC processor to avoid going above or below preset voltage thresholds in oscillation amplitude. This optimization procedure is invoked by baseband processor at power up, and periodically during normal operation.

CONTROL SECTION

This portion of the design consists of all programmable registers, the interface to the MAC microprocessor, the R2T and tag-to-reader (T2R) FIFO's, and a number of finite state machines. The microprocessor interface supports two modes of operation, a multiplexed address/data ‘8051-like mode, and a high bandwidth synchronous serial port mode. State machines for auto-tuning the IF filter, generating a dither signal for the receive ADC, and logic for the synthesizer lock detect are implemented in the control block.

The parallel interface is a multiplexed address/data scheme utilizing a 4-bit bus plus 4 control signals. This provides a simple low speed mechanism for interfacing the device to ‘8051-like busses. Access to the full register space is indirect; that is, three 4-bit wide address registers are programmed, which then access the selected programmable register through four 4-bit wide data registers. Some often used registers have been made available in the direct map, including the R2T command FIFO, the T2R data FIFO, the interrupt status register and the interrupt mask register.

The device also supports a synchronous serial interface. This mode of operation uses three signals: clock, data, and frame.. Each access is formatted into a 32-bit frame, and consists of 4-bit control information, 12-bit address, and 16-bit data. The device supports one synchronous serial channel in the R2T direction, and three channels (which share a common clock and frame signal) in the T2R direction. Serial clocks up to 150MHz are supported in order to facilitate bypass mode, as discussed below. The serial port functions are multiplexed on the same pins as are used for the parallel interface. Operational mode is selected through a hardware pin.

Issuing transmit commands is performed through the R2T command FIFO. Commands include a number of data-related instructions, as well as special commands for enabling the receiver and the RSSI function. Data-related commands range from simply transmitting a single bit to sending an entire byte, a long random sequence of bits, or more complicated waveforms such as framesync’s and preambles. After an entire transmit sequence has been queued, the microcontroller must issue an end-of-transmission (EOT) command in order for the device to schedule when the receiver should start searching for a valid preamble. The EOT command includes a delay field for this purpose. A similar command is used to control the integration start for the RSSI measurement.

Once the receiver has successfully detected a preamble, subsequent decoded data is buffered in the T2R data FIFO. This FIFO is 16 elements deep; data is placed in the FIFO eight bits at a time, except in Class 0 operation, where only single bits are placed in each FIFO location. If no preamble is detected, no data is decoded and the device eventually times out and sets a bit in the interrupt status register.

In order to provide flexibility in the receiver function, the device supports “bypass” nodes. These nodes are points along the receiver processing chain that can be configured to automatically generate read frames on the serial interface when a new sample arrives. Supported nodes include the output of the decimation chain, the output of the channel filter, and the output of the phase recovery block. This capability helps “future-proof” the device in the face of developing standards.

Finally, many of the programmable registers implement design-for-debug capabilities, including observability of state machine outputs as well as manual override of those outputs in selected instances. For example, RSSI measurements may be directly initiated, preamble search status may be observed, soft outputs from the matched filters may be observed; all through various dedicated registers. Additional debugging capability is provided through two digital test hardware pins; these pins provide observability of a number of internal signals, and can also be used as GPIO by the MAC microcontroller.

RESULTS

The complete digital portion of the chip was ported to a FPGA platform to emulate all the implementation details and to enable running a large number of packets, in order to get an accurate measurement of the packet error rate (PER). Figure 11 shows the emulation results for the sensitivity - receive power level at the antenna port at which PER=1% and BER~10^{-4} – versus data rate for FM0 and Miller subcarrier modulation. The plots are shown for four different gain setting. The lowest gain setting is used in interference limited scenarios as it provides the reader with the highest tolerance against blockers and interferers.

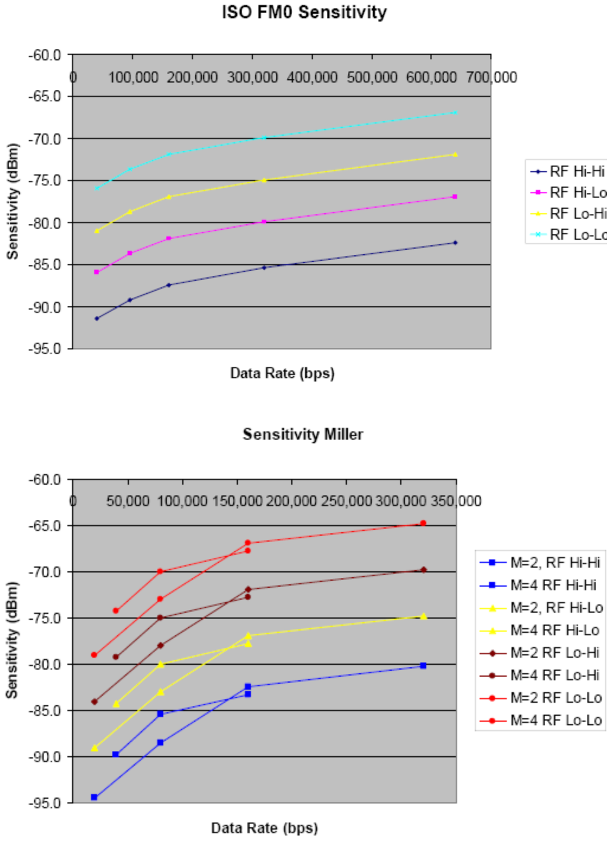


Fig. 11: Sensitivity vs. data rate for FM0 and Miller modulation

Figure 12 shows a measured transmitted spectrum in CW mode and for PR-ASK modulation with a reference time interval Tari=25μs. Also indicated is the spectral mask requirement in Europe. The waveform does not account for the distortion of the external power amplifier which will cause spectral re-growth on the mask.

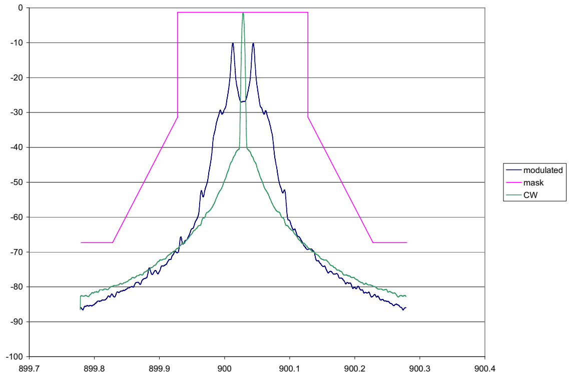


Fig. 12: Measurement trasmitted PR-ASK spectrum, Tari 25 μs

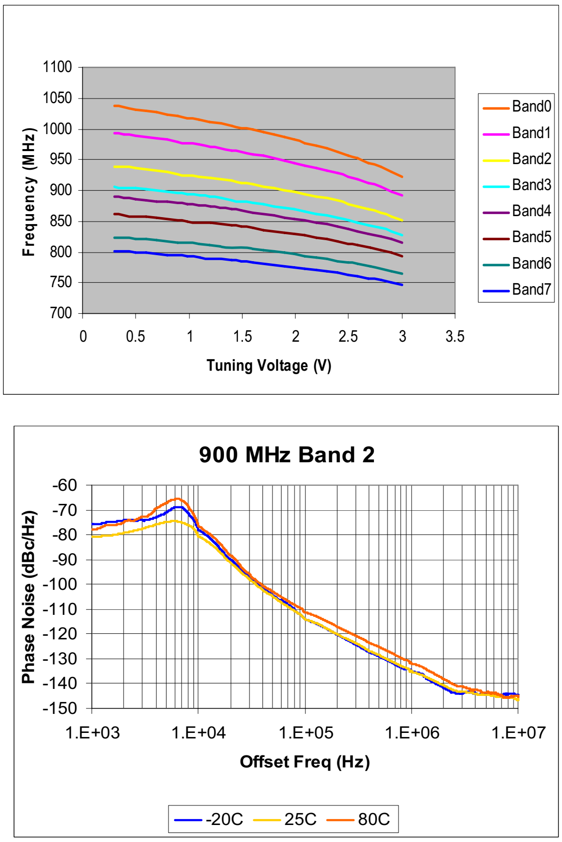


Fig.13: Measured VCO performance: a) Frequency v. tuning, b) phase noise

Figure 13 shows the measured results of the VCO. The up figure shows the VCO frequency v. tuning voltage for all 8 bands. The VCO operates from 750 to 1040 MHz at room temperature; over process and temperature variations the VCO covers the required 860-960 MHz band. The down figure shows the measured phase noise over temperature at 900MHz. The VCO meets the design requirements of -116 dBc/Hz at 200KHz offset and -144dBc/Hz at 3.6MHz offset . Figure 14 shows the measured OIP3 for the on-chip PA. The graph indicates an OIP3= +27dBm. The result is about 2dB lower than simulated, but it was traced back to the loss of the off-chip matching network used. If a narrower band, lower loss matching network is used, the OIP3 should be close to the design value of +29dBm.

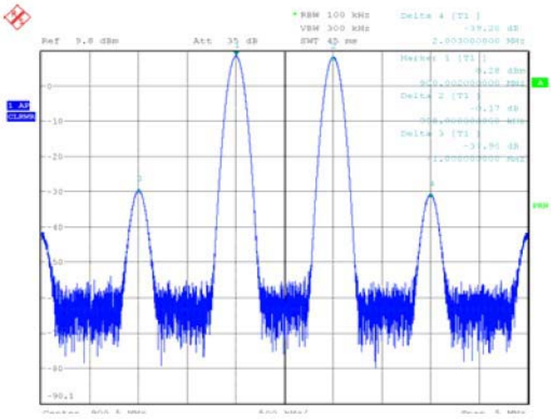


Fig. 14. Measured OIP3 of the on-chip PA

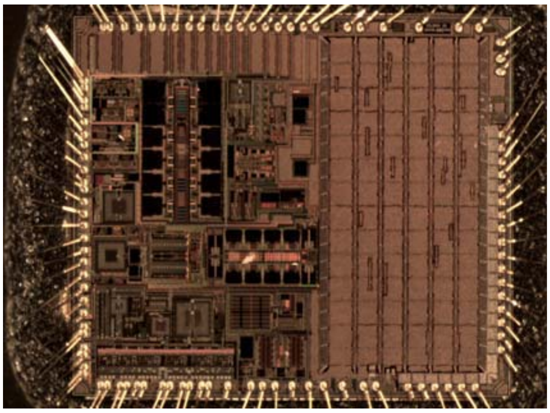


Fig 15: Photograph of the die

SUMMARY

The architecture and design of a single chip transceiver for multi-class worldwide RFID reader applications was described. The transceiver utilizes a direct conversion architecture on the receive and transmit paths. The chip integrates all the RF blocks, synthesizer, sigma-delta converters, digital filtering, and the digital modulation and demodulation functions. Results from the emulation platform and the full functionality A0 chip were presented.

All UHF RFID readers in the market today are based on discrete designs. This is the first single chip transceiver in the market and will provide a reduction of 100's of components and more than $150 off the reader's BoM. This highly integrated transceiver should create an inflection point towards low-cost ubiquitous UHF RFID readers.

Bibliography/References

[1] K. Finkenzeller, *RFID Handbook*. England: John Wiley & Sons, 2003

[2] I. Kipnis, J. Posamentier, T. Barnes, "UHF RFID Systems", Tutorial, DTTC 2005

[3] MIT Auto-ID Center, Class 0 RFID Tag Protocol Specification, February 2003

[4] MIT Auto-ID Center, Class 1 RFID Tag Protocol Specification, Version 1.0.1, November 2002

[5] ISO-IEC_CD 18000-6C, Version 2.1c2, July 2005

[6] FCC 47 CFR Ch. 1, part 15, 10-1-98 Edition

[7] ETSI EN 302 208-1, Version V1.1.1, September 2004

[8] K. Martin, "Complex signal processing is not complex", *IEEE Transactions on Circuits and Systems – I* , vol. 51, no. 9, pp. 1823 – 1836, Sep 2004

[9] R. Veldhoven, "A triple-mode continuous time sigma delta modulator with switched capacitor feedback DAC for a GSM-EDGE/CDMA 2000/UMTS receiver", *IEEE J. Solid-State Circuits*, vol. 38, no. 12, pp. 2069 – 2076, Dec 2003

A UHF READER ANTENNA WITH STEERABLE BEAM

Mario Orefice, Gianluca Dassano

Abstract-- **This paper presents some preliminary results of a study on a low cost steerable beam reader antenna for RFID applications: the prototype is in the UHF ISM frequency range 2.40-2.48 GHz, although it can be easily scaled at other frequencies (e.g. in the lower UHF). The antenna has two configurations: one consisting of a linear array of 4 elements, with a wide beam in the plane orthogonal to the steering plane; the other consists of an array of 4x4 elements on a regular rectangular lattice, with a relatively narrow beam in both planes, fed by a microstrip BFN.**

The beam steering is in one plane, and is obtained by a single analog voltage control of a number of diodes used as voltage controlled capacitors. Experimental results for the phase shifters and for the array well fit with the theoretical analysis.

Index terms-- **RFID antennas, Scanning arrays, Planar arrays, Analog phase shifters.**

I. INTRODUCTION

Technologies for RFID applications are one of the most promising and innovative fields for the solution of many logistics problems, and they are becoming widely diffused in all sectors of commerce and transportation, as well as many other applications in everyday's life. Several frequency bands are in use, ranging from LF (125-134 kHz) to HF (13.56 MHz) to UHF (433 MHz, 868 MHz, 2.45 MHz) and higher. Advantages and disadvantages are related to each of these frequency bands: typically, lower frequencies have a better behaviour near conductive (or partially conductive) bodies, as water, metals, human bodies, while higher frequencies have a larger reading range and a higher transmission speed. In fact, the amount of data that can be stored in a tag and transmitted generally increases with the frequency: at UHF there also the possibility of using, besides to passive transponders (whose typical reading ranges are several meters), also active transponders, that increase the range to tens of meters and have a much larger memory size. However, disadvantages of such higher frequencies are that the propagation characteristics are more sensitive to the characteristics of the object (type of the material where the tag is placed) or of the environment (humidity); typically, the presence of metallic parts may reflect, scatter or obstruct the RF signal.

One of the first applications of the UHF RFID technology was to containers. In 1991 the international standard ISO 10374 was drawn up to provide a basis for use of this technology in the identification procedure. High frequency RFID bands (around 900 MHz in North America, 2.4-2.5 GHz in Europe) are used for this purpose to ensure a high bit rate communication and long range detection, up to 13m [1]. In this case RFID active tag was used supplied by a battery to ensure a lifetime of about 10-15 years.

The positioning of the tags strongly depends on the type of application, and this can imply the design of reader antennas often with non conventional radiation patterns, e.g. it could be shaped to guarantee uniform coverage to reach all tags located in the area. Moreover, the allowed limits of RF power is generally not much high: at present the ETSI EN 300 440-1 defines the power limits for RFID systems in the 2.45GHz ISM band as follows:

Frequency band	Power limit EIRP (1)	Use of equipment	Comments
2.446-2.454 GHz	+27dBm	No restriction	FHSS or unmodulated carrier (CW) only
2.446-2.454 GHz	+36dBm (2)	In-buildings only	FHSS only

Note 1: EIRP including an antenna with the following data:

a) equal or less than ±45° horizontal beamwidth;

b) equal or more than 15dB SL attenuation;

c) physical protection (e.g. antenna dome) which dimension limits a power transfer from the RFID antenna to a quarter wave matched dipole at positioned at an extreme close proximity to ≤+15dBm.

Note 2: the use of power levels above +27dBm (EIRP) shall by technical means be restricted to in build use only and shall have a duty cycle less than or equal to 15% averaged over any 200ms period.

The authors are with the Laboratory of Antennas and EMC (LACE), Dipartimento di Elettronica, Politecnico di Torino, C.so Duca degli Abruzzi 24, 10129 Torino, Italy

It can be observed that the RFID equipment shall have two EIRP levels, 4W and 500mW or less, as default. The equipment should automatically switch via software the two levels to ensure the correct use. In particular for fixed mounted system the controlling software will be destroy in case the system was removed from the installation point to disable the capability to operate with 4W EIRP In case of portable 4W RFID equipments the system shall be configure to ensure that the high power level will be used only in-building situation and shall be commutate at lower power level otherwise.

A possible significant improvement can be obtained by introducing scanning beam antenna arrays: a possible use of such antenna is to ensure the communication between the RFID tag (e.g. on a container) and the RFID read station in case of long distance and of a large number of container stocked inside or outside a building. With a steerable antenna it is possible to have a higher gain antenna that can cover a large area without moving the read station. The system can therefore use a less expensive read station to cover the same area or, whit the same power, increase the coverage area. In particular in case of in-building applications, the system could be easily controlled (e.g. through the varactor polarization voltage) to ensure that the 4W EIRP covers the inside area only, setting the antenna beam scanning angles using a mask defined by software.

Beam steering can be obtained in many ways, e.g. by using digital phase shifters or switching through matrices (e.g. Butler matrix, Blass matrix): however in general beam steering requires complex and expensive control systems and components.

In this paper we will describe the design and construction of an array and its microstrip BFN, where the variation of the output phasing is done through a single analog voltage control of a number of diodes in reverse polarization, used as voltage controlled capacitors. This type of phase shifter has been described in different forms by various authors (see e.g. [2-6]) although with some difference with respect to the one used here.

II. Description of the structure

The antenna and the BFN operate in the frequency range 2.4-2.48 GHz. The BFN, shown schematically in fig.1, is for the power distribution and phase control of the radiating elements in the scanning plane ("phasing section"): the scanning plane is generally horizontal. If the volume to be covered is wide also in the vertical direction (e.g. identification of elements located on shelves, or stacked) the radiating element can be a single antenna (e.g. a dipole or a patch). Conversely, if the space to be explored is narrow in the vertical plane, the radiating element should be a subarray, so that the BFN requires a second section ("subarray section") for the power distribution in the plane orthogonal to the scanning: these two configurations will be called here Wide Vertical Beam (WVB) and Narrow Vertical Beam (NVB), respectively. The BFN phasing section has 1 input port and 4 output ports, across which the output voltage has a uniform distribution in amplitude; the phase shift between adjacent ports is constant and controlled by continuously variable DC voltages V1 and V2.

The basic element of the BFN is the voltage controlled variable phase shifter, consisting in a 3 dB hybrid with 2 ports loaded by two low cost commercially available voltage controlled diodes, used in reverse polarization as capacitors, shown schematically in fig. 2(a), where the polarization circuit has been omitted. A picture of the phase shifter is shown in fig.2(b). Phase shifters of this type have been studied in the past by other authors (see e.g. [2-3]).

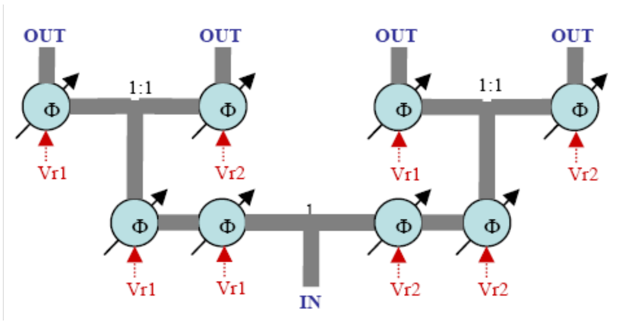


Fig.1: BFN scheme (1 input port, 4 output ports).

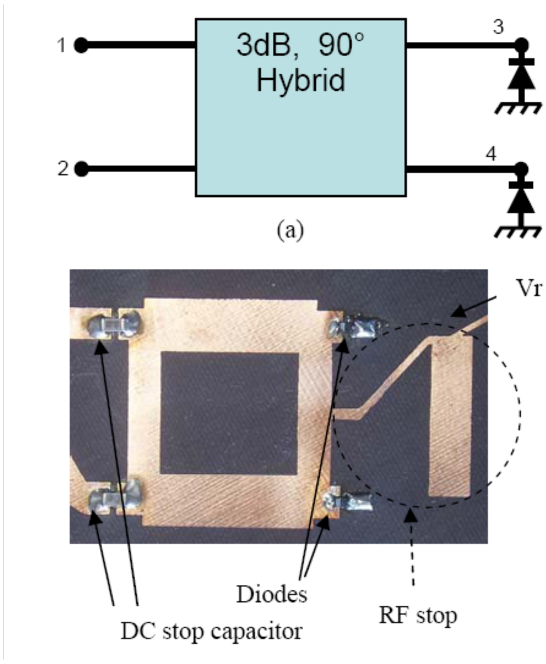


Fig.2: Scheme (a) and photograph (b) of the variable phase shifter.

The matching of the phase shifter is not constant because the hybrid has a variable load. Figure 3 shows a 3-D plot of the measured reflection coefficient at port 1 vs. frequency and polarization voltage. From this plot it can be seen that in the frequency range of interest the reflection coefficient is always below 0.2, and it is near this value, as the worst case, for low polarization voltages (about 0 V) and in the lower part of the band. However, the full operating bandwidth is from 2.4 to 2.55 MHz, larger than the requirement.

The measured phase shift between ports 1 and 2 is almost constant, for each polarization voltage, in the considered frequency range, and it is shown in fig.4. Conversely, the insertion loss depends essentially, in addition to the losses of the lines, on the internal series resistance of the diode: if this is small the losses are very low. For more commercial devices, as those used in this applications, having a resistance around 1.5-2 Ω, losses can be close to 1 dB.

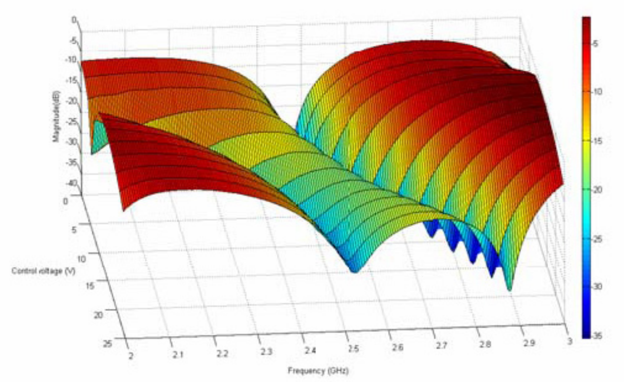


Fig.3: 3-D plot of the measured reflection coefficient at port 1 vs. frequency and polarization voltage.

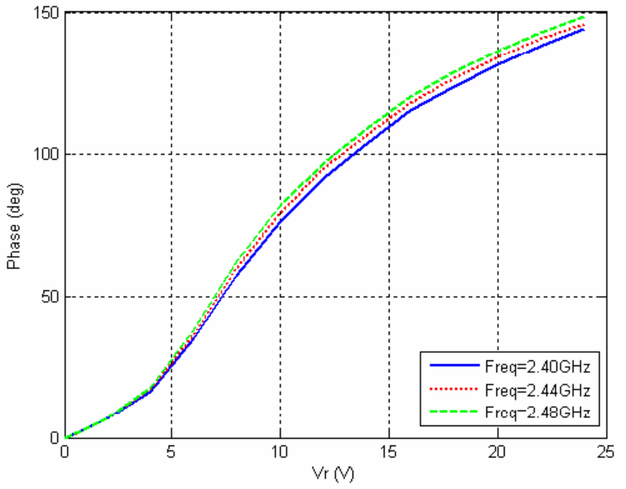


Fig.4: Phase shift vs. polarization voltage for different frequencies.

In order to maximize the phase variation between adjacent ports, for each branch of the main power divider two phase shifters have been used, fed with two voltages V_{r1} and V_{r2} between 0 and V_{max}, such that $V_{r2}= V_{max} -V_{r1}$: in this way we obtain a wider range of phase variation and an increase of the linearity of the phase vs. control voltage.

The phase difference between adjacent elements can reach about ±150°, as shown in fig. 5, where theoretical (solid lines) and experimental (dashed lines) values of the phase difference among the ports are reported. Theoretical and experimental values are very near, but looking at the differences it can be seen that the phase progression between elements, with varying voltage, is not exactly linear. This leads to some increase of the sidelobes for particular angles. To reduce such errors a more efficient but more complex voltage control for the phase shifters should be used: this option was not considered, aiming to develop an antenna with reduced costs, although with non optimum results.

The final configuration of the BFN is shown in fig. 6 for the WVB prototype. Considering separately the phasing section and the sub-array feeding system, it can be seen that the phasing section has a good uniform power splitting.

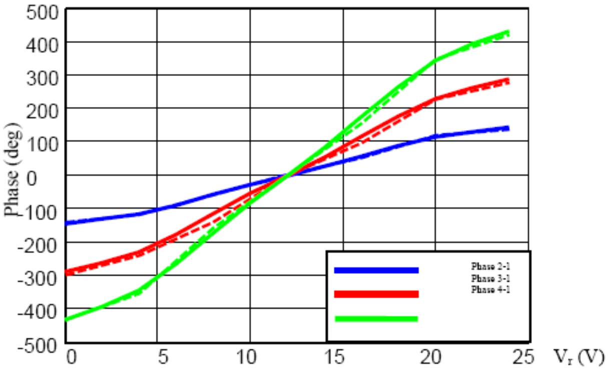


Fig.5: Phase differences for the complete BFN vs. control voltage: experimental (dashed) and theoretical from the phase of the single phase shifter (dotted).

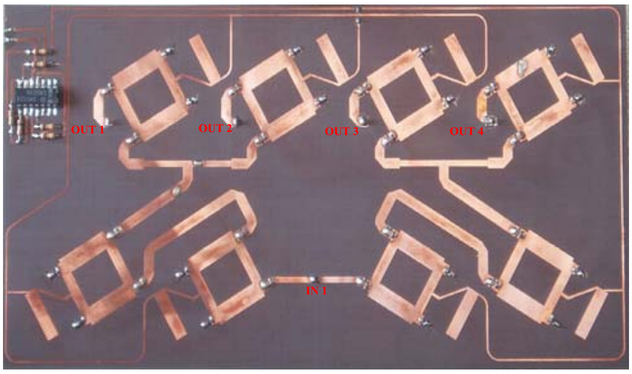


Fig.6: The complete BFN for the WVB configuration.

The amplitude of the transmission coefficient, for the 4 ports, is shown in fig. 7, for the particular polarization voltage of 12V (corresponding to zero phase shift). Without losses, it should be -6dB (uniform amplitude distribution), while the measured results show, in the band of interest, values oscillating between -10 dB and -8 dB. The oscillations may be due in part to discontinuities in the measurement setup, and the average value is around -9 dB. Consequently this section of the BFN introduces approximately 3 dB losses, due in large part to the series resistance of the diodes. Additional losses are then introduced by the subarray power distribution.

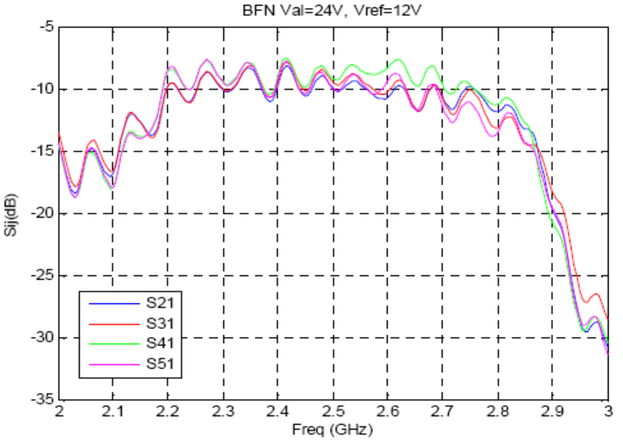


Fig.7: Amplitude of transmission coefficients (Sj1) for the phase shift section of the BFN.

III. The Array

In the VWB configuration each output port of the BFN phasing section feeds the BFN subarray section to feed a subarray of 4 elements: a 4x4 planar array is therefore obtained, and the achievable phase difference allows a wide steering of the array in one plane.

Depending on the distance between subarrays, the scan angle can vary almost from -90° to 90°: however, to obtain the full half plane coverage it is not necessary to reach a pointing to ±90º, because the number of subarrays is small and the beam is wide. In must be also considered that, to have a full ±90º scanning, non directive (i.e. omni-directional or quasi omni in the upper half space) radiating subarrays must be employed. In this application, thick dipoles parallel to the ground plane have been used as subarray element. This type of element has a relatively wide bandwidth (more than 20% for a SWR<1.5) [7]; in particular they have been chosen because this element has a beamwidth in the scan (H) plane larger than of other elements (e.g. a patch) and therefore allows a wider scan. A view of the complete antenna is shown in fig. 8.

In this case, the radiating elements have been designed to have 100 Ω input impedance, in order to have a simpler power distribution network. The outer elements are connected to a 100 Ω transmission line, to keep the impedance constant along the line, so that when they are connected in parallel to the inner elements an input impedance of 50 Ω is obtained for the left and right side, then connected to the phasing section by a power splitter. The 100 Ω line is 1.5λ_g long (the minimum possible multiple of half-wavelength for the given dimensions), so that an exchange of the feed point is necessary.

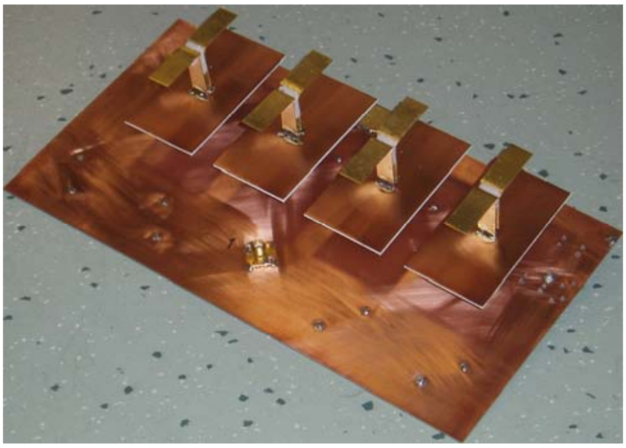


Fig.8: View of the 4 elements array WVB prototype.

When no voltage is applied to both diodes, the phase difference among the output ports is zero, and therefore a broadside array is obtained. In figure 9 the experimental radiation pattern is shown in both horizontal (H-plane) and vertical (E-plane) cuts. A good symmetry of the main beams and of the nulls can be observed; some differences on the sidelobe levels on left and right sides are attributable to slight inaccuracies in the mounting of the element, as well as coupling among the elements, non-uniform spacing due to the feed point reversal, phase shift inaccuracy.

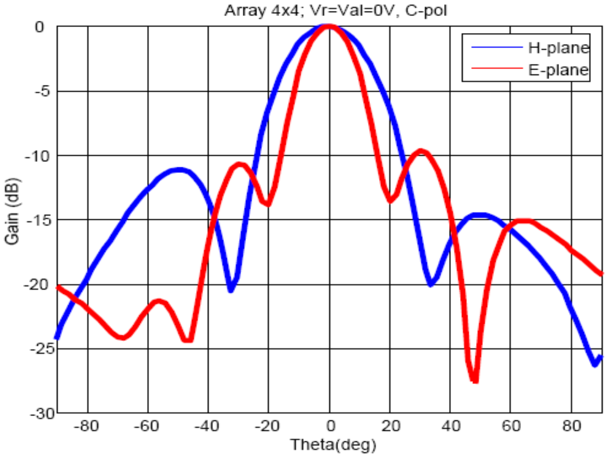


Fig. 9: Measured radiation patterns of NVB without diode polarization (2.44 GHz).

When the diodes are polarized, the main beam is scanned in the horizontal plane. A few examples of radiation patterns are shown in fig. 10, measured at 2.44 GHz. The diode polarization voltage varies from 0 to 24 V and the phase shift between adjacent elements varies from -140º to 140º. For relatively small scan angles (e.g. for 9 V or 15 V) the gain doesn't change, while for larger angles the gain decreases and the sidelobes increase because of the characteristics of the element pattern. As it is known, the 3 dB beamwidth of a dipole over a ground plane is, in the H plane, ±60º: therefore, the scan angle should not go beyond those limits.

In this prototype the spacing among the elements was chosen as about 0.44λ, so that with the maximum phase shift the scan angle is about ±50º. In fact, with this spacing, a minor part of the grating lobe enters in the visible space, and it can be seen as the third sidelobe on the left (or on the right) for the V=0 an V=24 V. In any case, the reduction of sidelobes, if required, is straightforward because it can be easily obtained the power splitting ratio in the power dividers between inner and outer elements.

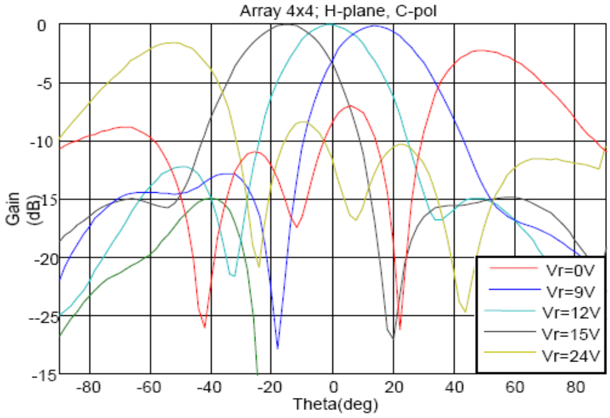


Fig.10: Radiation patterns for various scan angles.

IV. CONCLUSIONS

In this paper some preliminary results about a low cost steerable beam planar antenna for RFID applications in the UHF bands, in particular in the frequency range 2.40-2.48 GHz, have been presented. The antenna consists of an array of 4x4 elements on a regular rectangular lattice, fed by a microstrip BFN. The beam steering is in one plane, and is obtained by a single analog voltage control of a number of voltage controlled capacitors. Experimental results for the phase shifters and for the array well fit with the theoretical analysis.

Further refinement are in progress to improve the efficiency and to meet the sidelobe specifications, that require a -15 dB level and therefore a non uniform feeding of the radiating elements and of the subarrays:

From the results obtained, it can be seen that this array can be easily controlled and steered, without significant degradation of the main beam, for a very wide angular range.

V. ACKNOWLEDGMENT

This work was partially supported by EC Network of Excellence ACE (Antenna Center of Excellence).

REFERENCES

[1] K. Finkenzeller, *RFID Handbook*, Wiley, 1999.

[2] E.C. Niehenke, V.V. DiMarco, A. Friedberg: “Linear analog hyperabrupt varactor diode phase shifters”, *1985 IEEE MTT-S Digest*, pp.657-660.

[3] C.-L. Chen; W.E. Courtney, L.J. Mahoney, M.J. Manfra, A. Chu, H.A. Atwater,” A Low-Loss Ku-Band Monolithic Analog Phase Shifter”, *IEEE Transactions onMicrowave Theory and Techniques*, Vol. 35, No. 3, Mar 1987 pp.315-320.

[4] D.M. Krafcsik, S.A. Imhoff, D.E. Dawson, A.L. Conti, “A dual-varactor analog phase shifter operating at 6 to 18 GHz”, *IEEE Trans. on Microwave Theory and Techniques,* Vol. 36, No. 12, Dec 1988, pp.1938-1941.

[5] T.-W. Yoo, J.-H. Song, M.-S. Park, “360° reflection-type analogue phase shifter implemented with a single 90° branch-line coupler”, *Electronics Letters,* Volume 33, No. 3, 30 Jan. 1997, pp.224-226.

[6] N. Gupta, R. Tomar, P. Bhartia, “A Low-Loss Voltage-Controlled Analog Phase-Shifter Using Branchline Coupler and Varactor Diodes”, *Proc. ICMMT '07*, April 2007, pp. 1-2.

[7] M. Orefice, G.L. Dassano; L. Matekovits; P.Pirinoli; G. Vecchi. B. Shurvinton, “A wide coverage scanning array for smart antennas applications”, *Proc. 31st EuMC,* London, September 2001, vol. 3 pp. 477-480.

Electromagnetic Analysis of the RFID Link

Carmine Piersanti[†], Franco Fuschini*, Francesco Paolazzi[‡], Vittorio Degli-Esposti[†], Gabriele Falciasecca[†]

Abstract-- **The recent widespread diffusion of radio-frequency identification (RFID) applications operating in the UHF band has been supported by both the request for greater interrogation ranges and greater and faster data exchange.**
UHF-RFID systems, exploiting a physical interaction based on Electromagnetic propagation, introduce many problems that have not been explored (or only partially) for the previous generations of RFID systems (e.g. HF). Therefore, the availability of reliable tools for modeling and evaluating the radio-communication between Reader and Tag within an RFID radio-link are needed. The paper discuss the impact of real environment (e.g. Electromagnetic Coupling, Multipath propagation) on system performance in terms of error-probability and received powers, showing some relevant effects and trade-offs in transponder and reader design.

Index terms-- **RFID, Bit Error Rate (BER), Multipath, Electromagnetic Coupling, Modulation Index.**

I. INTRODUCTION

Radio-Frequency IDentification (RFID) technologies have been used for many years for identification of objects or people, electronic toll collection, asset identification, retail item management and vehicle security. The widely diffused technology based on inductive coupling at 13.56 MHz is useful for penetrating water or metals, but also offers low reading capacity and the interrogation (read) range (maximum distance between the Tag and the Reader) is approximately equal to the size of the reader antenna, and therefore usually limited to less than 1 meter (often to only a few centimeters). The request of longer read-ranges and faster data-transmission has provided to the recent widespread diffusion of applications operating in the UHF band. The increasing operating frequency (from HF to UHF bands) involves a radical change in the physical interaction mechanism between Reader and Tag, which is no more based on the inductive coupling but rather on the transmission and reception of electromagnetic (EM) waves propagating between the terminals [1].

In passive UHF-RFID systems the RF energy radiated by the Reader is used both to supply the digital section of the transponder and to allow data transmission.

After the Tag demodulates the Reader interrogation message (Downlink), most of the passive UHF-RFID systems exploit modulation of the backscattered radiation to transmit data from Tag to Reader (Uplink): the Reader sends an unmodulated carrier, which partly provides the Tag power supply, partly is reflected and modulated by means of a proper variation of the load (Z_L) of the Transponder antenna (back-scatter modulation); typically with amplitude-shift keying (ASK, or in particular OOK) or phase-shift keying (PSK) [2]. Assuming that the Tag activation threshold is enough for a correct Downlink communication, a reliable and complete evaluation of system performances requires the capability to analyze the two main cases of the system mis-functioning: Power Limitation, if the power at the Tag logic control is too small to supply it, and BER Limitation, in case of a missed modulation operation. Since power supply and modulation efficiency are conflicting necessities, in order to maximize the operating range, it is also important to achieve a nontrivial trade-off between the desired error probability (BER) at the reader and the power available for supplying the transponder logic control (P_{TAG_IN}).

RFID systems are often studied and planned in ideal, Free Space conditions, but they are then required to operate in real environment, and therefore they may be requested to cope with several possible impairments, such as Multipath propagation, Non Line of Sight (NLOS) propagation and unexpected changes in the Tag antenna impedance (Z_{AT}) due to the EM coupling with the tagged object (many common materials, including metals and aqueous liquids, have strong effects on the performance of UHF tag antenna) and/or other surrounding Tags ([3, 4]).

II. TRANSPONDER ACTIVATION ANALYSIS

The power absorbed by the tag (P_{TAG_IN}) should be enough to turn on the device, i.e. to properly feed the logic control (containing the memory) and the back-scatter modulator.

If ideal, free space (FS) conditions are assumed, the tag received power may be expressed by the Friis formula:

$$P_{TAG\_IN} = EIRP \cdot g_T \cdot \left(\frac{\lambda}{4\pi r}\right)^2 \cdot \tau \cdot \rho_T \qquad (1)$$

being EIRP the Reader equivalent isotropically radiated power, g_T the tag antenna gain, r the distance between terminals, λ the wavelength, $\tau \in [0:1]$ the polarization factor (related to the polarization (mis)matching between the receiving antenna and the incident EM field) and ρ_T the power absorption coefficient [5]:

$$\rho_T = \frac{4 \cdot \Re(Z_{AT}) \cdot \Re(Z_L)}{|Z_{AT} + Z_L|^2} \qquad (2)$$

Z_{AT} represents the Transponder antenna impedance and Z_L the load impedance (seen at the Tag chip terminals) which is switched over the modulation values Z_{L1} and Z_{L2}, according to the usual back-scatter modulation procedure [1].

III. BIT ERROR RATE EVALUATION

III.a. Analysis of the tag back-scattered field

It is well known that an object illuminated by an incoming wave scatters EM energy as the result of currents induced by the incident field [6,7,8]. The total amount and the radiation pattern of the back-scattered power depend on the geometrical and EM properties of the object (size, shape, constitutive materials, etc.).

In particular, objects designed as antennas or constituted of an antenna as main part can be associated with two different modes of scattering the incident energy, one independent of the antenna load impedance Z_L (structural mode scattering), the other strongly load-dependent (antenna mode scattering) [7,8].

According to the consideration described in [9], backscattered field by a tag loaded by the impedance Z_L can be easily factored into the following form:

$$\vec{E}_{scat}(P\,|\,Z_L) = j\frac{\lambda\cdot\left|\vec{\mathrm{E}}_{\mathrm{inc}}\right|}{4\pi r}\cdot\sqrt{g_T(\theta,\phi)\cdot g_T(\theta,\phi)_{inc}}\cdot(1-\rho)\cdot \\ \cdot\left[\hat{p}(\theta,\phi)_{inc}\cdot\hat{p}_{inc}\right]\cdot e^{j2\arg(Z_{AT})}\cdot e^{-j\beta(d+r)}\cdot\hat{p}(\theta,\phi) \qquad (3)$$

being:

- $(\theta,\phi)_{\mathrm{inc}}$ the direction of arrival of the incoming field $\vec{\mathrm{E}}_{\mathrm{inc}}$;
- $\hat{\mathrm{p}}$ the polarization vector of the Tag antenna;
- $\hat{\mathrm{p}}_{\mathrm{inc}}$ the polarization vector of the incident field $\vec{\mathrm{E}}_{\mathrm{inc}}$;
- ρ is the reflection coefficient at the Tag antenna terminals [4]: $\rho = \frac{Z_L - Z_{AT}^*}{Z_L + Z_{AT}}$;
- A: a complex, load-independent coefficient related to the current induced on the antenna conducting surface by the incident wave [9,10]; its exact value is then dependent on the geometrical layout of the Tag antenna and the EM properties of the materials it is constituted of [10].

The back-scattered field to the Reader can be immediately achieved by imposing $g(\theta,\phi) = g(\theta_{inc},\phi_{inc})$ and $\hat{\mathrm{p}}(\theta,\phi) = \hat{\mathrm{p}}(\theta_{inc},\phi_{inc})$ in (10).

The variation of the load impedance Z_L between the values Z_{L1} and Z_{L2} (load modulation) produces two different values of the power reflection coefficients (ρ_1 and ρ_2) and therefore the modulation of the field scattered by the Tag which switches over the corresponding values $\vec{\mathrm{E}}^1_{\mathrm{scat}}$ and $\vec{\mathrm{E}}^2_{\mathrm{scat}}$.

Since the whole scattered field doesn't generally depend only on the current induced at the antenna terminal, then A ≠ 1. Nevertheless, it is known that A ≈ 1 for thin dipole [11], and therefore A = 1 is often assumed as useful, rather reasonable approximation for the performance evaluation of general RFID system, since UHF RFID Transponder are usually equipped with thin, small, dipole-like antennas.

III.b. Bit Error Rate

Depending on the value of $\vec{\mathrm{E}}^1_{\mathrm{scat}}$ and $\vec{\mathrm{E}}^2_{\mathrm{scat}}$ the voltage signal (V_{mod}) received by the Reader varies between the values V_1 ($\mathbf{E}_{\mathbf{scat}} = \mathbf{E}_{\mathbf{scat1}}$) and V_2 ($\mathbf{E}_{\mathbf{scat}} = \mathbf{E}_{\mathbf{scat2}}$) [9,12]. The signal V_{mod} is then demodulated and the information stored into the Tag are therefore acquired by the Reader. In order to achieve a satisfactory reception, the BER due to the noise of the receiving device is requested to be lower than a proper maximum threshold (BER_{th}).

Assuming ideal, matched-filter demodulation and additive white gaussian noise (AWGN) with standard deviation σ at the input of the detector, the BER at the Reader side for both PSK and ASK (OOK) modulation can be simply expressed as [12]:

$$BER = \frac{1}{2}erfc\left(\frac{|V_0|\cdot m}{2\cdot\sqrt{2}\cdot\sigma}\right) = \frac{1}{2}erfc\left(\frac{m\cdot\sqrt{P_{READER\_IN\_TPM}}}{2\cdot\sigma\cdot\sqrt{\Re(Y_{LR})}}\right) \quad (4)$$

$m = |\rho_1-\rho_2|/2$ is the modulation index[1] [13], Y_{LR} the load admittance at the Reader antenna terminals, V_0 and $P_{READER\_IN\_TPM}$ the voltage at the input of the receiver and the received power in case of Tag Perfectly Matched ($Z_L = Z_{AT}^*$).

If ideal, free space conditions are considered, the Reader received power may be expressed as (more details in the Appendix):

$$P_{READER\_IN} = EIRP\cdot g_R\cdot g_T^2\cdot\left(\frac{\lambda}{4\pi r}\right)^4\cdot\tau^2\cdot\rho_S \qquad (5)$$

where g_R represents the Reader antenna gain and ρ_S is the Tag scattering coefficient [5]:

$$\rho_S = \frac{4\cdot\left[\Re(Z_{AT})\right]^2}{|Z_{AT}+Z_L|^2} \qquad (6)$$

III.c. Environmental effects

Each RFID system is always required to operate in real environment, and therefore it may be afflicted by several possible environmental impairments which should be carefully taken into account since they may strongly reduce the real performances with respect to the ideal ones.

Multipath propagation –because of the usual presence of objects around and/or between the RFID devices, multipath

[1] ρ_i i=1,2 are the complex reflection coefficients at the Tag antenna [4]

propagation may arise [12]. Multipath interference may have a strong impact on P_{TAG_IN} and $P_{READER\_IN\_TPM}$ (and therefore V_0) ,whose value can be no longer estimated by means of simple free space formulas () () which must be corrected by a proper fading margin (according to a suitable statistical description of the RFID channel [13]) or completely discarded in favour of more adequate deterministic models (such as ray tracing tools). On the contrary, if flat frequency fading is assumed (since the RFID signals bandwidth can be usually supposed less than the channel coherence bandwidth), then no changes are suffered by the modulation index, which is independent of the presence of multipath [13].

EM coupling – the presence of objects in the proximity of the Transponder may produce electromagnetic coupling and therefore unwanted variations in the Tag antenna impedance Z_{AT} [3, 14]; this effect involves not only changes in all the related coefficients (such as the modulation index) but also both phase and amplitude distortion and therefore a mixed (phase-and-amplitude) modulation. It turns out that the value of the modulation index in (1) should take into account the impact of the coupling effects on the Z_{AT} actual value.

IV. REAL ENVIRONMENT BEHAVIOR

According to the remarks carried out in the previous sections, if real environmental conditions are considered, RFID systems are exposed to many possibilities of interference and degradation that the simple, analytical Free Space model previously described cannot take into account; therefore, a different analysis approach is needed for a more reliable and complete estimate of the system performance.

IV.a. Scattering Matrix and System Simulator Modeling

The RFID radio channel can be modelled as a two-port (four-terminal) network (Figure 1) and is described through its scattering matrix S, which is then embedded into a system level simulator (Figure 2) in order to estimate both Received Powers (P_{TAG_IN}, P_{READER_IN}) and Bit Error Rate (BER).

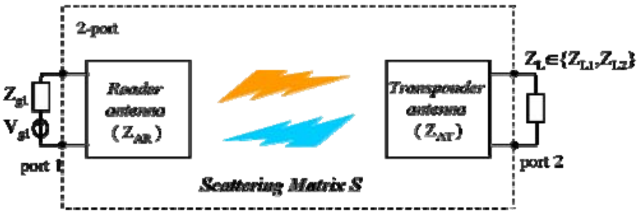

2-port
Z_{g1}
V_{g1}
port 1
Reader antenna (Z_{AR})
Transponder antenna (Z_{AT})
$Z_L \in \{Z_{L1}, Z_{L2}\}$
port 2
Scattering Matrix S

Figure 1 - RFID radio channel as a 2-port

In particular, the scattering coefficients S_{ij} (i, j ∈ {1,2}) are achieved by means of complete EM simulation [13].

Thanks to this approach, the impact of real environment on system performance is carefully and properly modelled and both far-field and near-field effects (Multipath and NLOS propagation, EM coupling with near objects) are automatically taken into account.

With reference to the flow diagram shown in Figure 2, the random-generated bit stream firstly undergoes the coding scheme compliant with the ISO 18000-6c standard [14] (FM0 or Miller) and then drives the Transponder antenna load, which is switched over the chip modulation impedances Z_{L1} and Z_{L2}.

By means of a proper ri-normalization process of the scattering matrix with respect to the actual value of Z_L [15], both voltage and current at ports 1 and 2 can be achieved:

$$V_2 = \frac{p_2}{\sqrt{\left|\Re(Z_L)\right|}} \cdot Z_L \cdot S_{21} \cdot a_1 \qquad (7)$$

$$I_2 = -\frac{p_2}{\sqrt{\left|\Re(Z_L)\right|}} \cdot S_{21} \cdot a_1 \qquad (8)$$

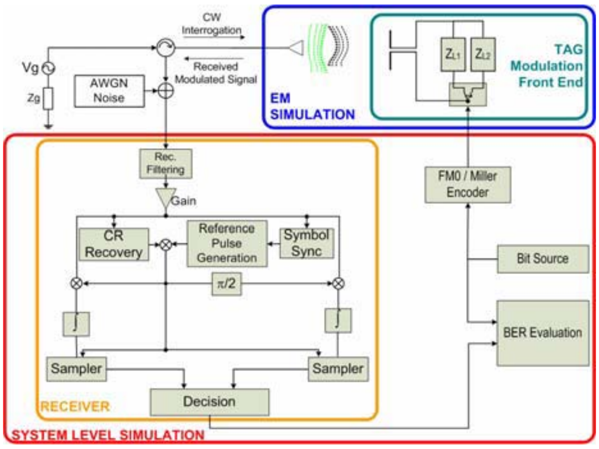


Figure 2 - Simulator Flow Diagram

$$V_1 = V_R = -\frac{(Z_{AT} + Z_L) \cdot Z_{g1}}{2 \cdot \Re(Z_L) \cdot \sqrt{\Re(Z_{g1})}} \cdot p_2 \cdot p_1 \cdot S_{21} \cdot S_{12} \cdot a_1 \qquad (9)$$

$$I_1 = I_R = \frac{(Z_{AT} + Z_L)}{2 \cdot \Re(Z_L) \cdot \sqrt{\Re(Z_{g1})}} \cdot p_2 \cdot p_1 \cdot S_{21} \cdot S_{12} \cdot a_1 \qquad (10)$$

being a_1 the incident power wave at port 1 [15], Z_{g1} the generator internal impedance **loading** the port 1, and the coefficients p_1, p_2 defined as in [15].

Moreover, Equation (10) easily provide the following expressions for the power received by Tag and Reader:

$$P_{TAG\_IN} = \frac{1}{2}\Re(Z_L) \cdot \left|I_{TAG}\right|^2 = \frac{1}{2}\Re(Z_2) \cdot \left|S_{21}\right|^2 \cdot \frac{\left|a_1\right|^2}{\Re(Z_2)} = \frac{1}{2} \cdot \left|a_1\right|^2 \cdot \left|S_{21}\right|^2 \, (11)$$

$$P_{READER\_IN} = \frac{1}{2}\Re(Z_{g1}) \cdot \left|I_{\mathrm{mod}}\right|^2 = \frac{1}{2}\Re(Z_{g1}) \cdot \frac{\left|Z_{AT} + Z_L\right|^2}{4 \cdot \left[\Re(Z_L)\right]^2 \cdot \Re(Z_{g1})} \cdot \left|S_{21}\right|^2 \cdot \left|S_{12}\right|^2 \cdot \left|a_1\right|^2 \, (12)$$

It is worth noticing that the interrogation signal feeding the Reader antenna is not included in Equations (9) and (10), which only account of the received signal related to the Tag backscattering modulation. Therefore, the receiving block can be fed by the modulated voltage V_R so that the received bit-stream can be evaluated and the BER analysis can be performed.

ERP	2 W
g_R	3 dB
g_T	1 dB
τ	1
f	868 MHz (λ ~ 34.5 cm)
Z_{AT}	10+j245
$Z_{L1,2}$	OOK modulation: $Z_{L1} = (Z_{AT})^*$; $Z_{L2} = +\infty$
	PSK modulation: according to what suggested in [4] for any possible m value.

Table 1 - Free Space case radiation and modulation parameters

In more detail a simple linear base-band model of the receiver architecture is assumed, with an ideal behaviour of the circulator back to the Reader antenna and perfect Carrier Phase and Symbol Timing recovery. Moreover, standard correlation demodulation technique with euclidean distance metrics decision is performed to achieve BER evaluation. It is then supposed that the noise at the input of the receiver is Additive White Gaussian Noise due to the thermal noise of the apparatus (prevalently antenna and LNA). In order to compare results from system simulation with those from Free Space formulas, it must be reminded that the value σ^2 in Equation (4) represents the variance of the noise at the input of the decision block.

V. FURTHER CONSIDERATIONS

V.a. . Free Space

According to the remarks carried out in the previous sections, an RFID system may be "power limited" or "BER limited"; in the former case the "activation range" r_1 is lower than the "demodulation range" r_2, whereas it is $r_1 > r_2$ in the latter one. Obviously, the real (interrogation) range of the system is $r_I = \min\{r_1, r_2\}$.

The particular case corresponding to $r_1 = r_2 = r_I$ (i.e. the Tag turns off at the same time the BER at the Reader exceeds the threshold BER_{th}) may be referred as a sort of "optimal, reference case", since in such a case each device would be properly dimensioned and designed with respect to the other.

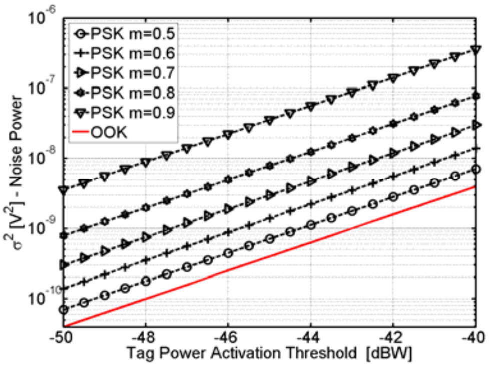


Figure 3 - - Noise Power vs. depending on Tag Power and modulation index m

READER MAIN PARAMETERS	
ERP	2 W
g_R	3.7 dB
Interrogation Frequency	868 MHz
Noise Power (σ^2)	$2 \cdot 10^{-9}$ V^2
BER_{th}	10^{-3}
TAG MAIN PARAMETERS	
g_T	3.6 dB (0.5 dB in the link direction)
Modulation	OOK with Z_{L1} = 10-j245 ; $Z_{L2} = +\infty$
Z_{AT}	17+j255
P_{TAG_MIN}	-48 dBW

Table 2 - Reader – Tag Main parameters

If Free Space propagation is supposed, this "reference case" may be easily investigated by means of Equations (1) and (5), in particular, for any given Tag Activation Threshold P_{TAG_MIN} (i.e. for any given r_1), the Noise Power value (σ^2) corresponding to $r_2 = r_1$ is evaluated. Results of Figure 3 have been achieved according to the radiation and modulation parameter briefly described in Table 1.

According to the usual range of the activation threshold ([-20÷-10] dBm), it is clear that higher quality Reader (sensitivity of about -80 dBm, Noise Power < 10^{-12}) are typically power limited [16]. Nevertheless, Readers sensitivity may usually range from about -80 dBm to about -60 dBm; therefore, in case rougher, higher noise Readers are used (such as simpler handheld device could be), then the Noise Power can raise to about 10^{-9} and the system range can be easily limited by BER impairment.

In order to show a useful example, the free space communication between the RFID devices briefly described in Table 2 may be considered.

In particular, the value Z_{L1} may be referred to the chip XRAG2 produced by STMicroelectronics.

Moreover, the antenna parameters values in Table 2 are the outcome of proper EM simulation [17] of both the Tag and the Reader. Actually, full and accurate antenna design would not be strictly requested for FS evaluations, but it is necessary in order to properly compare FS results with those from system level simulation (see section *V.b*).

In particular, the designed Tag antenna is an usual meander dipole whose layout is represented in Figure 3.

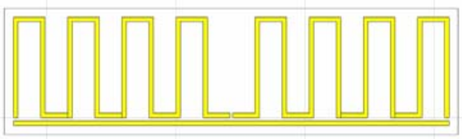


Figure 4 - Meander Tag Layout

It is worth noticing that optimal OOK modulation would require $Z_{AT} = (Z_{L1})^*$ and this is not exactly the value of Table 2.

However, the non-optimal value of Table 2 could be considered more realistic than the ideal target one, since it may account for possible, unpredictable imperfections in the Tag antenna designing and manufacturing processes.

As far as the Reader antenna is concerned, a rather common shorted patch in free air substrate has been considered (3.73 dBi gain, 50 Ω input impedance at 868 MHz).
Then, assuming Tag and Reader oriented so that perfect polarization matching is achieved ($\tau = 1$), P_{TAG_IN} and BER vs. distance may be computed and plotted (Figure 4).

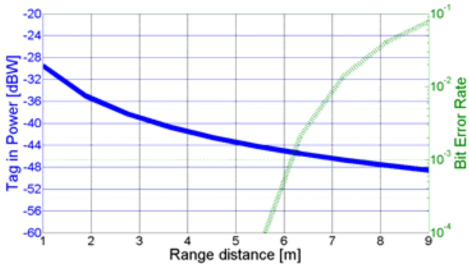


Figure 5 - Performance in case of OOK modulation

Reminding that the Tag activation threshold is assumed to be -48dBW and BER_{th} 10^{-3}, it is well evident that the Tag turns off at the distance of about 8 m, whereas the BER threshold is crossed at about 6.2 m. Therefore, the system is BER-limited.

Since OOK modulation forces m = 0.5, it is not possible to tune the system parameters in order to get $r_1 = r_2$.
On the contrary, if PSK modulation is considered (assuming $Z_{L1,2}$ according to [4] instead of those of Table 2), then m = 0.8 allows to achieve the "reference case" previously defined (Figure 3). This is clearly confirmed also by the Figure 6.

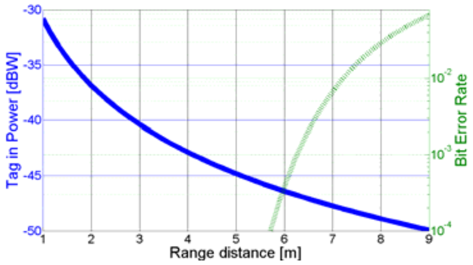


Figure 6 - Performance in case of PSK modulation

V.b. Environmental effects

Slight modifications to the FS propagation scenario are usually sufficient to highlight the possible, strong impact of real environment on the performance of UHF RFID systems.
Starting from the FS RFID systems considered in the previous section, a simple, interesting case study can be achieved by adding a single object close to the Tag, thus accounting for the obvious fact that the Transponder is always associated with something to be tagged.
In particular, a square, watery slab has been considered, having side and thickness equal to $\lambda/2$ (~ 17 cm) and 5 mm, respectively (Figure 7); $\varepsilon_r = 81$, $\gamma = 0.01$ S/m are the adopted values for the electromagnetic parameters. The distance between the Tag and the slab is set to 2 mm.

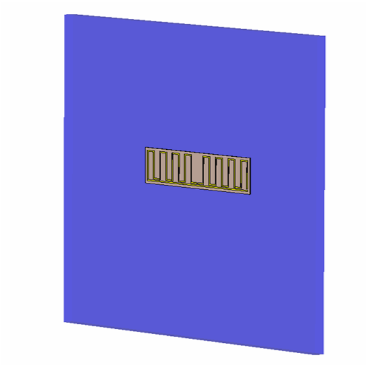


Figure 7 - Tag-watery slab coupling

The presence of the slab in the vicinity of the Tag antenna may involve variations in its antenna impedance Z_{AT} (and therefore in m, ρ_S and ρ_T values), radiation pattern and efficiency [2,3]. Obviously, these effects cannot be taken into account by the Free Space model, but they are automatically considered by the electromagnetic , system-level simulation approach described in section IV.
Results are shown in Figure 8 and **Figure 9** for the OOK and the PSK modulation, respectively. It is evident that the interrogation range strongly decreases with respect to the FS in both cases.

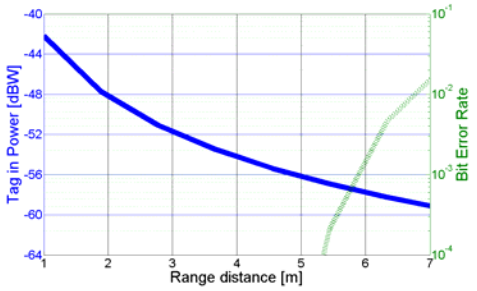


Figure 8 Environmental effect in case of OOK modulation

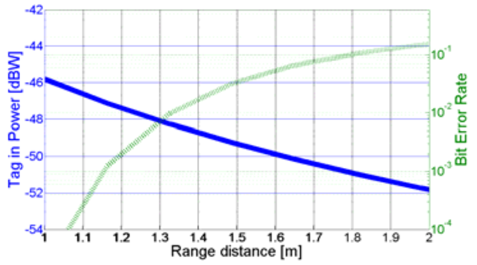


Figure 9 - Environmental effect in case of PSK modulation

Obviously, further objects can be considered within the operating environment (in addition to the tagged one) for a more realistic modelling of the propagation scenario.
In presence of objects around and/or between the RFID devices multipath propagation arises and therefore Tag and Reader are requested to cope with possible signals impairment due to interference of the multiple received echoes.
As well as EM coupling, multipath phenomena are completely neglected by the Free Space model, but they are

automatically accounted by the suggested electromagnetic system-level simulation.

By way of useful trial, the multipath scenario represented in Figure 10 and constituted of two metallic object in addition to the watery slab can be considered.

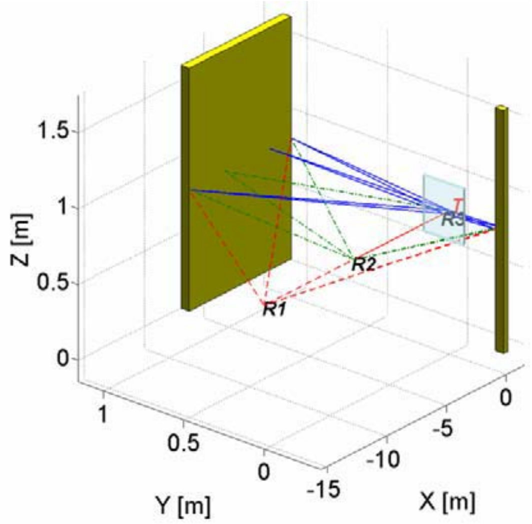


Figure 10 - Multipath scenario and main multipath components (plotted according to the Geometrical Optics Theory)

Depending on the Reader position, multipath contributions properties change (numbers of received paths, time delays, angles of arrival/departure, amplitudes, etc.) thus generating the evident spatial fading shown in Figure 11.

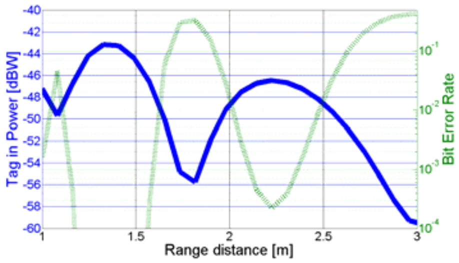


Figure 11 - Impact of spatial fading on RFID system performance (PSK modulation)

If compared to Figure 9, because of multipath the maximum interrogation range is slightly increased to about 2.4 m, but this result is obviously related to the particular considered scenario (reduction in range could be observed in different cases).

Moreover, the "interrogation region" appears non-uniform and the communication between Tag and Reader can fail even at distances lower than the maximum interrogation range. These kind of "coverage holes" represent a general problem typically related to the multipath propagation effects and can strongly limit the reliability of an RFID system.

Figure 8, Figure **9** and Figure **11** have obtained by interpolation of few simulation points; for each point, EM simulation has required ca. 8 hours computing time on a common 1processor 2GHz RAM Personal Computer.

VI. CONCLUSIONS

Under the usual assumption of Free Space conditions UHF RFID systems can be described and designed by means of simple analytical formulas. Unfortunately, environmental effects (such as electromagnetic coupling and multipath propagation) may afflict the system functioning and the real performance may be sensibly lower than the expected ones. Therefore the analytical Free Space models must be rejected in favour of different analysis approach for a more reliable and complete evaluation of the system performance.

In this work, a system level simulator for RFID radio link modeling is described; thanks to the electromagnetic simulation approach, the possible impact of real environment is automatically taken into account, as shown by some case studies presented in the paper.

The present work is mainly addressed to RFID systems operating through radiowave propagation (electromagnetic backscatter coupling [1]), but the system level simulator is also suitable for the evaluation of RFID systems based on inductive coupling.

VII. ACKNOWLEDGMENT

The Authors would like to thank DATALOGIC S.p.a. for having supported some topics of the research activity on RFID systems.

REFERENCES

[1] K. Finkenzeller, *RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification*, 2nd ed., John Wiley & Sons Inc., 1999

[2] U. Karthaus, M. Fischer, "Fully integrated passive UHF RFID transponder IC with 16.7-μm minimum RF input power" *IEEE J. Solid-State Circuits*, vol. 38, NO. 10, pp. 1602–1608, Oct. 2003

[**3**] P. Raumonen et. al., "Folded Dipole Antenna Near Metal Plate" *Antennas and Propagation Society International Symposium 2003,* vol. 1, pp. 848-851

[4] D. Dobkin, S. Weigand, "Environmental effects on RFID tag antennas" 2005 *IEEE MTT-S International Microwave Symposium*, paper TU3A-2

[5] K.V.S. Rao, K.V. Nikitin, S.F. Lam, "Antenna Design for UHF RFID Tags: A Review and a Practical Application", *IEEE Transactions on Antennas and Propagation*, vol. 53, NO. 12, December 2005

[6] E.F. Knott, *Radar Cross Section*, Artech House, 1993;

[7] K. Penttila, M. Keskilammi, L. Sydanheimo, M. Kivikoski, 'Radar cross-section analysis for passive RFID systems', *IEE Proceedings Microwave, Antennas and Propagation,* vol. 153, no. 1, pp. 103-109, February 2006;

[8] R.C. Hansen, "Relationships between antennas as scatterers and radiators*", Proceedings of the IEEE,* vol. 77, No. 5, May 1969;

[9] F. Fuschini, C. Piersanti, F. Paolazzi, G. Falciasecca, "Analytical approach to the back-scattering from UHF RFID Transponder", in press on *Antennas and Wireless Propagation Letters*;

[10] C. C. Yen, A. E. Gutierrez, D. Veeramani, "Radar Cross-Section Analysis of Backscattering RFID Tags", IEEE Antennas and Wireless Propagation Letters, vol. 6, 2007;

[11] C. A. Balanis, *Antenna Theory*, Wiley & Sons, 1997;

[12] F. Fuschini, C. Piersanti, F. Paolazzi, G. Falciasecca, "On the efficiency of load modulation in RFID systems operating in real environment", in press on *Antennas and Wireless Propagation Letters*;

[13] F. Fuschini, C. Piersanti, F. Paolazzi, G. Falciasecca, "Electromagnetic and system level co-simulation for RFID radio link modelling in real environment", *European Conference on Antennas and Propagation"*, Edinburgh (UK), November 2007;

[14] International Standard ISO/IEC 18000-6:2004/FDAM 1:2006(E), July 2006;

[15] K. Kurokawa, "Power Waves and Scattering Matrix", *IEEE Transactions on Microwave Theory and Techniques*, vol. 13, NO. 2, pp 194-202, March 1964

[16] P.V. Nikitin, K.V.S. Rao, "Performance limitations of passive UHF RFID systems", *IEEE Antennas and Propagation Symposium*, Albuquerque, NM, July 2006

[17] Computer Simulation Technology®, www.cst.com

Section 2

Distributed Systems

Indoor Localization Techniques based on Active and Passive Devices

Franco Mazzenga
University of Rome Tor Vergata
email: mazzenga@ing.uniroma2.it

Marco Vari
RadioLabs
email: marco.vari@radiolabs.it

***Abstract*—A brief survey and a comparison among the most important wireless technologies for indoor localization are presented. The performances of indoor localization systems based on RF power measurements with active or passive devices are discussed in terms of the accuracy, complexity and costs.**

The operating principle of the considered localization methods is simple: the terminal to be located measures the RF received power and/or the identity of each retransmitting device belonging to the localization system. This information are then sent by the terminal to a central server using a wireless communication technology. The server uses these information to determine the position of the terminal inside the area and then to retransmit-it to the terminal. A localization system based on passive RFIDs is presented and its performance analyzed. Analysis is carried out through simulation in a typical office-like environment. Parameters of the localization system such as the density of the RFID devices in the area are varied. It is shown that passive RFIDs based networks offer adequate performance in terms of the achievable localization accuracy and costs.

I. Introduction

At the present time several techniques for indoor localization have been proposed or are still under investigation. These techniques envisage the use of active (e.g. WiFi, HiperLAN, UWB and/or acoustic devices such as Crickets, [1]) or passive devices (typically RFID or surface acoustic wave devices (SAW) [2]). Often the former devices are integrated in the user terminal and can be adapted for localization purposes by using simple software to be installed or downloaded in the terminal. However, as shown in the following, such techniques often lack of adequate precision and reliability especially when compared with RFIDs based solutions such as that presented in this paper. In addition their setup in a real environment can be problematic because in most cases the exact characterization of the RF environment is required for localization purposes.

Ultra-wideband (UWB) devices used for setting up communication networks offer localization at, practically, no additional costs. This is due to the very large bandwidth allocated to UWB signals which allows to extract position information based on the time of arrival of the first path. Localization techniques based on UWB can be very accurate even for indoor environments. Even though UWB technology seems to be very promising, UWB-based localization systems and algorithms are still under study [3] and their achievable performances can be compared with those based on RFID devices considered in this paper.

The paper is organized as follows: in Section II we briefly review the most important techniques that can be used for setting up a localization system for indoor applications. Techniques are first classified as RF-power based and TOA based and then compared in terms of achievable localization accuracy. An expanded analysis on the Wi-Fi localization based technique is presented. In the same section we detail the characteristics of the localization system based on passive RFID devices. Its operating principle are detailed in Section III and results obtained from simulation are described in the same Section. Finally Conclusions are drawn.

II. Localization techniques for indoor environment

In this Section a short review is provided on the main features of the actual localization techniques for indoor environments.

A. Time based techniques

Time based localization techniques exploit the measurements of the time of arrival (TOA) and/or on time difference of arrival (TDOA) of signals transmitted by fixed points in the area whose position is known. TOA/TDOA techniques can be used to perform localization for outdoor as well as for indoor environments even if in the latter only some kind of signals can be successfully used to achieve accuracy below one meter.

1) UWB based techniques: For positioning systems employing UWB radios, time-based schemes provide very good accuracy due to the high time resolution which is achievable thanks to the very large bandwidth occupied by UWB signals, [4]. Detection and estimation problems associated with a signal traveling between nodes have been well studied in radar and other applications. An optimal estimate of the arrival time is obtained using a matched filter, or equivalently, a bank of correlation receivers. In the former approach the instant at which the filter output attains its peak provides the arrival time estimate, whereas in the latter, the time shift of the template signal that yields the largest cross correlation with the received signal gives the desired estimate.

In indoor environments, because the presence of multipath, multiple (delayed) replicas of the transmitted signal partially overlap and shift the position of the correlation peak and this value may not return the true TOA estimate. In this case the multipath channel creates mismatch between the received signal of interest and the transmitted template used. As a result, instead of auto-correlation, cross-correlation is obtained which does not necessarily peak at the correct timing. To prevent this effect (complex) estimation techniques, have been proposed, however, multiple correlation peaks can still be present and it is important to consider algorithms to detect the first arriving signal path.

Moreover, in NLOS cases only signals generated by multiple reflections of the UWB pulse reach the receiving node. Therefore, the delay of the first arriving pulse does not represent the true TOA. Since the pulse travels an extra distance, a positive bias called the NLOS error is present in the measured time delay. In addition NLOS propagation may also cause a situation where the first arriving pulse is usually not the strongest pulse. Therefore, conventional TOA estimation methods based on the strongest path would introduce another positive bias into the estimated TOA.

UWB signals can also be used to determine location of one or more UWB nodes forming a network in an area. If two nodes have a common clock, the node receiving the signal can determine the TOA of the incoming signal that is time-stamped by the reference node. Since the achievable accuracy under ideal conditions is very high, clock synchronization between the nodes becomes an important factor affecting TOA. If there is no synchronization between a given node and the reference nodes, but there is synchronization among some reference nodes (Access Point, AP), it is possible to deploy a localization system which exploits the TDOA of different beacon signals. In the absence of a common clock between the nodes, round-trip time between two transceiver nodes can be measured to estimate the distance between two nodes.

2) Acoustic devices: Acoustic devices use a combination of RF and (ultra)sound technologies to provide location information to attached host devices, [5]. Wall and ceiling-mounted beacons placed through a building publish information on an RF channel. With each RF advertisement, the beacon transmits a concurrent ultrasonic pulse. Listeners attached to devices and mobiles listen for RF signals, and upon receipt of the first few bits, listen for the corresponding ultrasonic pulse. When this pulse arrives, the listener obtains a distance estimate for the corresponding beacon by taking advantage of the difference in propagation speeds between RF and ultrasound. The listener runs algorithms that correlate RF and ultrasound samples (the latter are simple pulses with no data encoded on them) and to pick the best correlation. Even in the presence of several competing beacon transmissions, those systems achieves good precision and accuracy quickly.

Due to the small propagation speed of the sound in the air as compared to the speed of light, it is possible to perform accurate ranging measurements. Assuming that the radio beacon reaches the receiver instantaneously, the distance between receivers and transmitter is obtained by solving the simple equation

$$d = \Delta t_f \times v_s \quad (m), \tag{1}$$

where d (m) is the distance, Δt_f (s) is the time difference between the radio beacon arrival and the sounding beacon arrival and v_s (m/s) is the speed of sound in the air. Combining such measurements it is possible to triangulate and then to localize with very deep accuracy.

In some products (such as Cricket) beacon and listener are identical hardware devices. A unit can function as either beacon or listener, or can be used in a 'mixed' mode in a symmetric location architecture, all under software control, however those techniques requires an expensive dissemination of sensors in the environment. Moreover active sensors are penalized by the requirements of continuous power supplying and this could make the entire infrastructure for localization too much expensive because cabling costs and maintenance.

B. Techniques based on RF power measurements

Localization procedures based on RF power measurements operates in accordance to two sequential phases:

1) *measurement phase*: the terminal measures the received power(s) of the signals transmitted by devices (specifically) used for localization;
2) *processing phase*: power samples are first recorded in the terminal and subsequently processed to estimate the position of the terminal.

Data processing phase can be centralized or distributed. In the first case, power data are retransmitted by the terminal to a local server while in the second case terminal owns all the necessary side information required for position calculation.

1) WiFi based techniques using RF maps: Typical indoor localization based on WLAN-WiFi apparatus, [6], performs comparison among the received powers of the signals transmitted by the access points (APs) in the area with those stored in a-pre built RF-map of the area. The RF-map can be located in a central server and/or it can be downloaded and periodically updated in the user terminal by means of the WiFi communication technology itself. Successful localization is achieved when it is possible to re-map the set of measured powers on the available RF-map without ambiguities. The point in space such that the measured powers are the closest to the powers stored in the RF-map is selected as the estimate of the terminal position.

In general, the accuracy of position estimation is strictly related to the propagation characteristics of the environment, on the number of transmitting devices (e.g. the APs) and on the spatial resolution of the RF-map. It should be observed that the accuracy of localization information greatly suffers the instability of the measurements of the environment. As an example, in an office with moving people, doors (closed or opened), the characteristics of the RF environment rapidly change and this can lead to a significant departure of the powers measured by the terminal from those stored in the RF map. Another relevant factor influencing the position estimation is the power measurement accuracy guaranteed by the hardware inside the terminal to be located.

Although the performance of these techniques can be slightly improved with the use of motion prediction and estimation based on available algorithms such as Viterbi-like or Kalman filters etc.[7]-[10], in all cases the precision/accuracy of the estimated position is always strictly related to the resolution of the RF-map and in its "persistence". Furthermore, the positioning of the active devices in the area (i.e. the APs) is a critical issue for localization performance. In an IEEE 802.11 based system, the access points (APs) can be positioned to achieve "best" coverage with a minimum number of APS, thus leading to a reduction of the overlap among the APs' coverage areas. In this case the terminal to be located could receive one or two APs at maximum and, as shown in the following, this can impair the performance of the localization algorithm. Obviously this could not be optimal for localization where it is necessary to increase the number of APs simultaneously seen by the terminal.

The effectiveness of the Wi-Fi-based localization techniques has been evaluated by the authors in [11]. Some results are now repeated. The topology of the considered office scenario is depicted in Fig.1. The environment is characterized by small rooms aligned along two parallel corridors. Offices are accessed through fire proof doors. Small/medium size walls are dominant in this kind of environment. A maximum number of 21 APs have been considered for performance evaluation and the data in the RF-map have been obtained neglecting any noise measurement effect. This assumption is representative of a realistic situation since the measured powers, used to create the RF-map, are commonly obtained by averaging them over a long time and using accurate instrumentation.

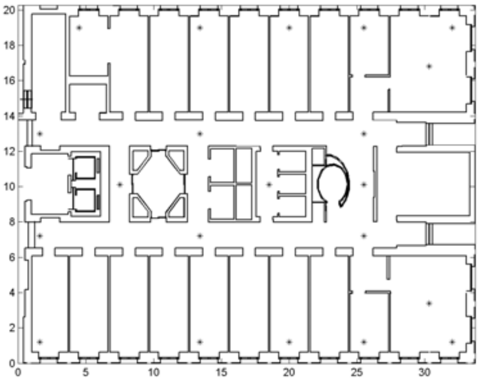


Fig. 1. Schematic representation of the considered scenario; the dots indicate the position of the active devices; Area $33.7 \times 20\ m^2$.

For performance evaluation purposes the position error defined as the distance between the estimated and the true terminal position in the area is introduced. If the position error is above a threshold a localization failure is occurred. The performance of the WiFi based localization system has been expressed in terms of the localization error probability, P_p i.e. the probability that localization failure is occurred. The calculation of P_p in a closed analytical form seems to be a very difficult task since it depends on several parameters such as: the number of active devices turned on in the area, their positions, the instantaneous propagation conditions (fast fading due to obstacles in the area), the accuracy of the power measurement in the terminal, the accuracy of the RF-map and on the topology of the area.

In Fig.2 we plot the maximum, the mean and the minimum values of the P_p as a function of the number of available APs in the area. The grid step of the RF-map was set to $d = 2$ m. The positions and the numbers of the APs participating in the localization were randomly selected at each iteration. The large variations in the P_p are mainly due to the geometric arrangement of the APs used for localization. In particular, since the APs participating in the localization were randomly selected in each iteration, it was observed that the largest values of P_p can be obtained when the APs used for position measurement result to be located along a straight line and almost LOS conditions exists with the terminal. In this case due to the symmetric configuration the same power vector may indicate different points in the area. Another case corresponding to large values for P_p occurs when the APs are (randomly) concentrated in a relatively small area as compared to the service area. In this case for several points in the area the differences among the power vectors are not so marked and, due to measurement errors in the RF power, localization errors are more frequent. It was also observed that better performance (corresponding to the minimum values of P_p in Fig.2), were obtained when no symmetries exist in the layout of the APs and/or when the APs are sparsed in the area. Finally when the number of APs participating to the localization is maximum (e.q. 21) the three curves intersect since the layout of the active devices is unique.

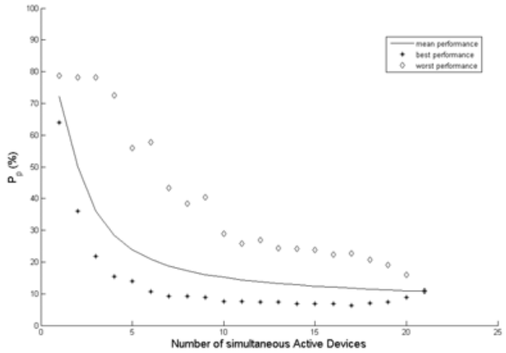


Fig. 2. P_p as a function of the number of active devices in the area.

From Fig.2 it is also interesting to observe that even increasing the number of active sensors in the area the P_p cannot decrease below a threshold even in the best cases. This is due to noise and quantization error in the power measurements influencing the result of the comparison with the data in the RF-map.

Finally, it should be noted that P_p can be decreased by improving the receiver sensitivity, [11]. This fact is not surprising since improvement in the receiver sensitivity allows to increase the (average) number of active devices simultaneously seen by the terminals thus providing better localization performance (see Fig.2). However when active devices are also used to provide communication services the simultaneous visibility of more than one active device from the terminal to be located could lead to interference situations that impair the normal operation of random access schemes such as the carrier sense multiple access with collision avoidance (CSMA/CA).

2) RFID-based localization techniques: RFID devices could ensure simple, economic and efficient setup of a (indoor) localization infrastructure. The basic idea on a localization system based on passive RFID devices consists in the revealing of proximity among user and tags (sensors) disseminated in the environment. User's device obtains information about the identifiers of its closest RFID tags. Then it send the RFIDs' IDs via a traditional communication infrastructure (such as WiFi, GPRS or WiMAX) to a central server (CS). The CS computes the localization information using an indoor map where the position of each tag in the area has been recorded for example during the installation phase.

The architecture of the considered RFID system is depicted in Fig.3. Thanks to the knowledge of the tags' positions, the CS can intersect the estimated tags' coverage areas and then estimates the position of the terminal with very high accuracy. As an alternative to processing in the CS, RFID maps can be downloaded on the user terminal and RFID data could be processed inside the terminal itself thus enabling self-localization. The performance of the proposed passive localization technique was compared in [11] with that achievable by the Wi-Fi based technique.

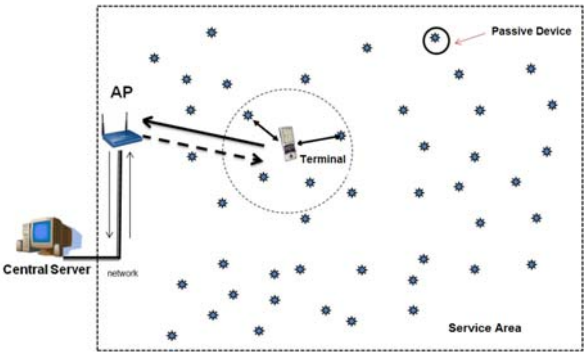


Fig. 3. Architecture of the considered passive localization system

Before concluding this Section, in Table I the typical performance figures of each one of the considered localization techniques have been summarized. UWB and RFID techniques seems to be the most promising in terms of achievable accuracy and costs. However, UWB technology is still in its infancy and localization algorithms are still not well established and validated. For what concerns the adoption of RFID technology some problems still need to be solved as indicated in the following Section where the performance of the RFID technique is analyzed in more detail.

Technology	**coverage radius**	**Achieved accuracy**	**Note**
Acoustic Sensors (e.g. Cricket)	$> 6m$ (LOS)	About $15cm$ with $6-7$ sensors $50-60cm$ with $4-5$ sensors	Localization possible in open environment; accuracy suffers the presence of obstacles
RadioMap-based (WiFi@2.4 GHz)	$>$ $15/20m$ (indoor)	About $2.5m$ with more than 4 sensors	Criticism in RadioMap creation /maintenance
UWB	$\cong 10m$	from few cm to tens of cm with minimum 4 sensors	precision related to the multi-path characteristics
RFID	$< 2m$	$0.5m$ with a density of 0.5 devices for m^2	Very high number of required sensors

TABLE I
COMPARISON AMONG DIFFERENT TECHNOLOGIES IN TERMS OF ACHIEVABLE LOCALIZATION ACCURACY.

III. PERFORMANCE ASSESSMENT OF PASSIVE RFID LOCALIZATION

In this Section we illustrate the algorithms used to obtain localization using RFID devices. We neglect the effects of the interference due to simultaneous answers coming from the interrogated passive devices i.e. the user terminal is always able to assess the identities of the interrogated RFIDs. Procedures to avoid (or reduce) collisions at the terminal's receiver, are illustrated in [2].

In order to improve the performance of the proposed technique, we assume that the terminal can measure the received power w_i of the i-th responding device. The value of w_i can be related to the i-th RFID-terminal distance, d_i, as:

$$w_i = \frac{w_T(\lambda/4\pi)^4 G_{tx}^2 G_{rx}^2}{d_i^4 I_L} \tag{2}$$

where $\lambda = 0.125\ m$ is the wavelength associated to the selected operating frequency (2.4 GHz), I_L is the passive device insertion loss, G_{tx} and G_{rx} are the transmitting and receiving antenna gains and w_T (W) is the transmitted power of the interrogating signal. Equation (2) is valid provided that the terminal transmitted power investing the single RFID device is above the threshold required to enable it.

In general, the CS can estimate the position of the terminal using the identities of the responding RFIDs devices and/or the measured w_i, $i = 1, 2, ..., N_a$ where N_a is the number of answering RFID devices (see Fig.3).

1) The identities of the responding devices can be used to rapidly restrict the area where the terminal is located. This can be achieved by determining the uncertainty area

of the terminal which is obtained as the intersection of the coverage areas of the responding RFIDs devices. The barycenter of the uncertainty area is taken as position estimate.

2) The position estimate achieved within the uncertainty area can be further refined using the values of w_i in (2). Inverting (2) with respect to d_i a position estimate (x, y) of the terminal can be obtained solving the (over-determined) nonlinear system of equations:

$$d_i^2 = (x - x_i)^2 + (y - y_i)^2, \quad i = 1, 2, \ldots, N_r, \quad (3)$$

where (x_i, y_i) are the coordinates of the N_r responding devices. The solution of the system in (3) was obtained using standard algorithms implemented in the *fsolve* routine of Matlab.

A. RFID system analysis

The passive devices used for localization have been positioned as depicted in Fig.4. The arrangement of RFID devices in Fig.4 is only for illustrative purposes and the total number of RFID devices considered for simulation is higher than that shown in Fig.4. It is further assumed that devices cannot re-radiate through walls. To simulate different densities, the number of RFID devices participating in the localization was varied during simulation. In particular, at each iteration, some of the devices in the area have been randomly turned off in accordance to a uniform distribution. Furthermore, we assume for the RFIDs an irregular radiation pattern having two notches randomly located in the angular interval $[0, 180^o]$. The angular extension of each notch was 3^o and the losses due to the two notches were set to 5.5 and 28 dB.

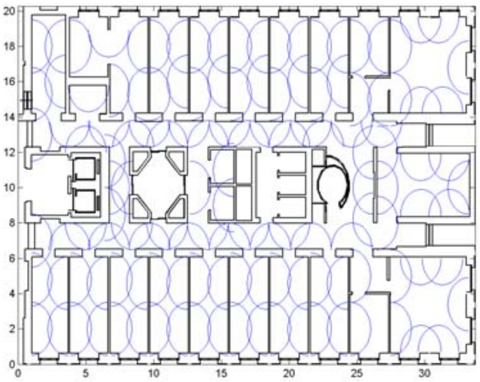


Fig. 4. Layout of a subset of the passive devices in the area and illustration of their coverage area.

The presence of one notch in the RFID radiation diagram can impair the measurement of the RFID retransmitted power e.g. w_i; however, depending on the format of the signal re-radiated by the RFID, identity could be yet decoded and used by the CS to restrict the area where the terminal could be located. It is assumed that the CS doesn't know the exact shape of the radiation pattern of each RFID instead, for calculation purposes, the CS assumes it as circular (see Fig.4). Thus the uncertainty area as calculated by the CS is given by the intersection of N_a half-circles.

The coverage radius of the RFIDs was varied during simulation from 1 m up to 2 m[1]. The sensitivity of the terminal receiver was set to -90 dBm, $G_{tx} = G_{rx} = 0$ dB (omnidirectional antennas on the horizontal plane) and insertion loss $I_L = 20$ dB. When the maximum value of d_i is below 2 m we assumed that the free space propagation model applies. The terminal to be located was randomly positioned in the area in accordance to a uniform spatial distribution. Finally, the power measured by the terminal receiver for each responding device was affected by a zero mean Gaussian random noise with standard deviation $\sigma = 2.5$ dBm (this value was obtained from the measurements reported in [11].

B. Results

In Fig.5 we plot the average P_p as a function of the density of passive devices. Relatively small densities have been considered. Different values for the error estimation threshold (above which a location failure is declared) have been considered $(0.5, 1.0, 1.5$ and 2 m). The RFID coverage radius is assumed to be fixed and equal to 1.5 m. The average of P_p is obtained with respect to the positions of the terminal to be located. As expected, the P_p decreases with the density of devices and is practically negligible when the error threshold is above 1.5 m. This means that the system can guarantee an estimation accuracy lower than 1.5 m for most of the time. System performance can be greatly improved by increasing the

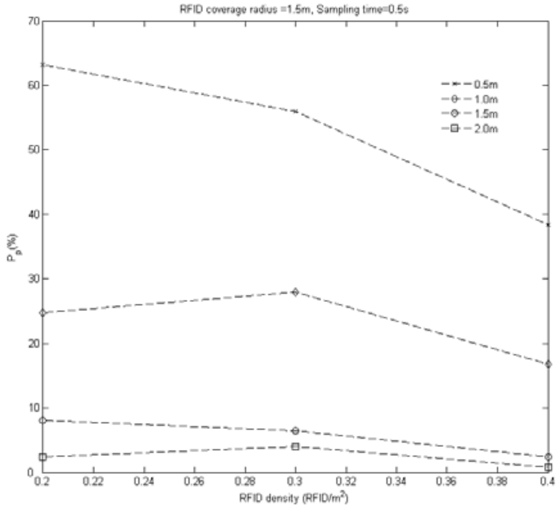


Fig. 5. P_p as a function of the density of the passive devices - variable threshold of the estimation error.

[1] In practice this means that the interrogating terminal could vary its transmitted power to increase/decrease the maximum tag response distance (see (2)).

density of RFID devices in the area as shown in the following results.

In Fig.6 we plot the probability that the number of answering devices is equal to $0, 1$ or 2 or 3 or above 3 as a function of the density of the passive devices in the area. As expected the percentages of having 0 or 1 answering device decreases with the density thus leading to a significative reduction of P_p.

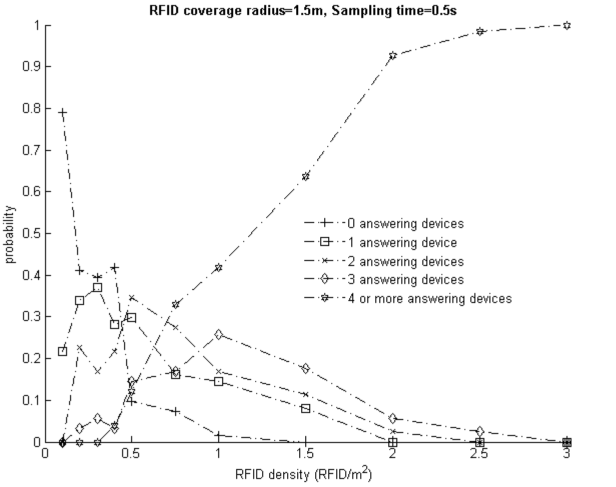


Fig. 6. Probability that the number of responding passive devices is equal to $0, 1, 2, 3, > 4$ as a function of the density of RFIDs.

In Fig.7 we plot the average position estimation error as a function of the density of the passive devices. Data corresponding to 0 and 1 responding devices have not included in Fig.7. It can be observed that when the number of responding devices is lower than 3 the position error increases. This fact is shown in Fig.6 where it can be observed that, for small densities, the percentage of times we have two responding devices is higher. In particular, a threshold value for the RFID density (e.g. critical density) has been evidenced in Fig.7. If the density of the RFID devices is below the critical density value the system performance rapidly degrades; in addition the achievable improvement for densities above the critical value rapidly saturates.

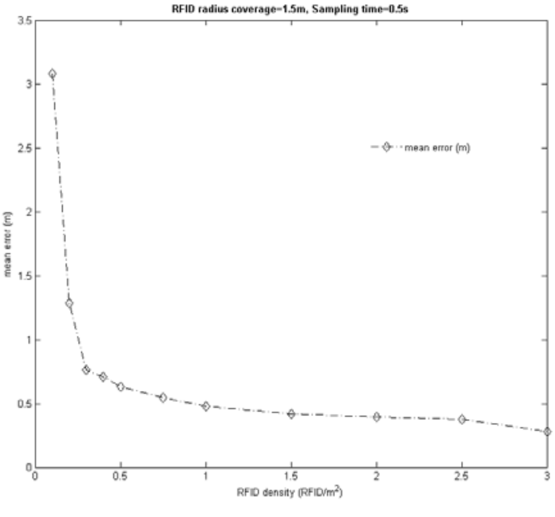


Fig. 7. Mean position error as a function of the density of passive devices; power data are available at the CS.

From the results in Fig.7 it can be further observed that even when the number of responding devices is greater than 1 the position estimation error remains within tolerable limits even for relatively small densities of the devices in the area. This is due to the small coverage area that allows to restrict the area where the terminal can be located.

IV. Issues on power consumption

In this Section we briefly discuss on the power-energy required by the interrogator to ping the RFID devices. For simplicity we assume that the interrogator inside the terminal is build using the technology adopted in IEEE802.11b products. As an illustrative example we consider the power consumptions of the Cisco Aironet PCMCIA card's indicated in [12]. In order to transmit an RF power of 100 mW, the overall power consumption is 2.25W for a transmission speed of 1Mb/s. During reception the power consumed by the device is 1.35W for receiver processing. Finally Cisco also declares a consumption of 0.075 W in sleep mode. Using the previous data it is possible to obtain the average energy required to transmit one bit at 1Mb/s i.e.: $E_b = 2.25/10^6 = 2.25\mu$J/bit. If the energy packet required to activate the RFID has an equivalent duration of 40 bits, the energy to be transmitted is $E = l \cdot E_b = 90\mu$J. Indicating with l the number of bits retransmitted by the RFID tag the energy required in the receiver for processing is $l \cdot 1.35/10^6$. Assuming for example $l = 40$ we obtain an energy consumption of 54 μJ that should be added to the required transmitted energy[2]. To calculate the total energy consumption required to process data obtained from tags we need to consider the number of responding tags that can range from 1 up to 4. In this case the energy for the interrogation can vary from $90 + 54 = 144$ μJ up to $90 + 4 \cdot 54 = 306$ μJ. Furthermore, from our simulations, the average number of responding tags in the area was 2.45 so that the average energy consumption is $90 + 2.45 \cdot 54 = 222.3$ μJ. Note that previous energy calculations assumed that RFID passive devices had a low sensitivity level i.e. they can respond even when the power at their input is very small (e.g. -24 dBm in our case). This corresponds to a realistic future technological objective since semiconductor techniques are rapidly advancing to reduce the RFID sensitivity toward tens of μW, [13]. If we assume as a realistic value of the RFID sensitivity -10 dBm [14], applying the link budget formula in (2) for a interrogator-RFID maximum distance of 1.5 m, we obtain a required transmitter RF power of about 2.3 W (in line with the data in the current literature [15]) which correspond to an overall power consumption of about 11.5 W.

[2] We implicitly assumed that the processing of the l-bits returned from the RFID should follow the same processing of a WLAN packet. This could be not true for the interrogator.

A final observation should concern the operating frequency of the interrogator. We assumed that the interrogator operates in WLAN frequency band (e.g. 2.4 GHz) which is also used to convey data to the CS. In this case the WLAN packets transmitted by the terminals or by the access point can activate the RFID devices. RFID responses can create background interference noise on the received WLAN packet. This could be easily avoided if the frequency of the RFID devices is different from that of WLAN. Many RFID devices exist on the market having operating frequency well below the 2.4 GHz. The adoption of double frequency RFID devices that can be activated on the WLAN band should not be discarded a priori especially if RFIDs could respond to interrogation on frequency outside the WLAN band. In this case interrogation would be at no additional energy costs since it is stimulated by normal packet transmission and by the transmissions of APs' beacons.

V. Conclusions

We presented a short review of the main wireless technologies that can be effective for setting up indoor localization systems. The performance of localization systems based on RF power measurements using active and/or passive devices have been discussed. An effective localization technique based on passive RFID devices has been presented and its performance evaluated considering a realistic office environment and by varying the system parameters such as the density of the RFIDs. The position estimate obtained by the uncertainty area obtained using only the RFIDs identities can be further refined using the measurements of the powers received by the responding RFID devices thus obtaining acceptable performance with localization accuracies well below 1 m.

References

[1] http://cricket.csail.mit.edu

[2] K. Finkenzeller, "RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification", Second Edition, 2003 John Wiley & Sons, Ltd.

[3] PULSERS, "Pervasive Ultra-wideband Low Spectral Energy Radio Systems", EU-IST Programme (FP6), http://www.pulsers.net.

[4] R. J. Fontana,"*Recent Applications of Ultra Wideband Radar and Communications Systems*", http://www.multispectral.com

[5] N.B. Priyantha, A Chakraborty, H. Baladrishnan, "*The Cricket Location-Support System*', in Proceedings of MOBICOM, 00, Aug 2000

[6] ISO/IEC 802-11: 1999(E), "*Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*" ANSI/IEEE Std 802.11, 1999 Edition.

[7] P. Bahl and V. N. Padmanabhan, A. Balachandran, "*A Software System for Location Mobile Users: Design, Evaluation, and Lessons*", Microsoft Research and University of California at San Diego.

[8] P. Prasithsangareel, P. Krishnamurty, P.K. Chrysantis, "*On indoor Position Location with Wireless LANS*", PIMRC 2002.

[9] P. Bahl, V. N. Padmanabhan, "*RADAR: An In-Building RF-based User Location and Tracking System*", Microsoft Research

[10] P. Bahl, N. Venkata N. Padmanabhan, A. Balachandran, "*Enhancements to the RADAR user location and tracking systems*", Microsoft Research

[11] D. De Luca, F. Mazzenga, C. Monti and M. Vari, "Performance evaluation of indoor localization techniques based on RF power measurements from active or passive devices", EURASIP Journal on Applied Signal Processing archive, Volume 2006 , Issue 1, pp.Pages: 160- 170

[12] http://www.cisco.com

[13] http://www.trolleyscan.com/paper/ecolect.html

[14] Philips EPC 1.19 G2 RFID ASIC, http://www.semiconductors.philips.com

[15] http://www.alientechnology.com/

SWARM INTELLIGENCE VIA RFID TAGS

Marco Mamei, Franco Zambonelli

DISMI – Università di Modena e Reggio Emilia – Via Allegri 13 – Reggio Emilia – ITALY

{mamei.marco, franco.zambonelli}@unimore.it

ABSTRACT

Pheromone-based multiagent interaction has received a growing attention in the past few years. Still, so far, the number of deployed systems exploiting pheromones for coordinating activities of distributed agents/robots situated in physical environments has been very limited. In this context, this paper presents a real-world, low- cost and general-purpose, implementation of pheromone interaction, realized by making use of RFID tags technology. Humans and robots can spread/sense pheromones by properly writing/reading RFID tags that are likely to populate our everyday environments. The proposed solution is tested and evaluated via an application for object-tracking, allowing robots and humans to find "forgot-somewhere" objects. The application works by letting objects spread digital pheromones trails that can be tracked afterwards. The paper presents several experiments to assess the effectiveness of our approach, outlines its limitations, and sketches further potential application scenarios.

1. INTRODUCTION

Pheromone-based interaction, exhibited by social insects to coordinate their activities, has recently attracted a vast number of researches in multi-agent systems [2, 3, 8, 11, 14, 19]. In these works, agents interact by leaving and sensing artificial pheromones (i.e. markers) in a virtual environment. Such pheromones can encode and describe application-specific information to be used to achieve specific tasks.

From a general perspective, the strength of pheromone interaction is rooted on two key points:

1. it completely decouples agent interactions (that are mediated by pheromones) making it suitable in open and dynamic scenarios where agents do not know each other in advance and can come and go at any time.
2. it naturally supports application specific context awareness, in that pheromones provide agents with an application-specific representation of their operational environment (e.g. in ant-foraging, pheromones provide a representation of the environment in terms of paths leading to food sources [3]).

Inspired by this mechanism, a lot of agent applications have been proposed. Still, so far, the number of actually implemented systems exploiting pheromones for coordinating activities of distributed agents situated in a physical environment has been very limited. The great majority of the proposals have been implemented in simulated environments [2, 3, 8], only few of them have been concretely implemented by deploying pheromones in shared virtual data spaces [11], other few realize pheromones by means of ad-hoc physical markers (special ink, metal dust, water on brown-kraft-paper, etc.) [19]. All these approaches will be better described in the related work section, however, in our opinion none of them propose valid solutions to actually spread pheromones in real-world everyday environments (who really wants pheromone-graffiti covering his/her home?!).

The reason for these missing implementations is rather natural: discarding centralized – not scalable- solutions, and still-to-come and costly sensor networks [4], it is not easy to find a suitable distributed infrastructure on which to store digital pheromones.

Inspired by these challenges, in this paper we propose an ubiquitous computing version of pheromone deployment based on RFID tags [16]. The key idea of our approach is to exploit the fact that RFID tags (in stark contrast with other tagging technologies like barcodes) can be written on-the-fly by suitable wireless devices, called RFID readers (which are also writers despite the name). On this basis, RFID readers, carried by a human or embedded in a robotic agent, could deploy pheromone trails across the environment, by storing the pheromone values in the RFID tags located there, as the user roams across the environment. The main point in favor of our approach is its extremely low price since it uses technologies (RFID) that are likely to be soon embedded in the scenario independently of this application.

Relying on such an implementation, a wide range of application scenarios based on pheromone interaction can be realized ranging from multi-robot coordination

[14], to monitoring of human activities [15]. In this paper, for the sake of illustrating our approach, we describe an application consisting in an agent-based system to easily find everyday objects (glasses, keys, etc.) forgot somewhere in our homes. In particular, the application allows everyday objects to leave virtual pheromone trails across our homes to be easily tracked afterwards.

The paper is organized as follows. Section 2 details the idea of RFID pheromone deployment. Section 3 presents the object tracking application, showing in detail the agent-based software implementation. Section 4 presents the actual deployment and experiments. Section 5 discusses further application scenarios and related work. Section 6 concludes.

2. RFID PHEROMONE DEPLOYMENT

In this section, we detail the idea of deploying digital pheromones in RFID tags present in the environment. For the sake of readability, we first describe the RFID technology that is at the core of the proposed mechanism, then we show how RFID tags can be used to store pheromones.

2.1. RFID Technology

At the heart of our approach is a breakthrough in sensing technology. Advances in miniaturization and manufacturing have yielded postage-stamp sized radio transceivers called Radio Frequency Identification (RFID) tags that can be attached unobtrusively to objects as small as a toothbrush. The tags are wireless and battery free. Each tag is marked with an unique identifier and provided with a tiny memory, up to some KB for advanced models, allowing to store data (our test-bed implementation comprise tags with a storage capacity of 512bit). Tags can be purchased off the shelf, cost roughly €0.20 each and can withstand day-to-day use for years (being battery-free, they do not have power-exhaustion problems).

Suitable devices, called RFID readers, can access RFID tags by radio, either for read and write operations. The tags respond or store data accordingly using power scavenged from the signal coming from the RFID reader. RFID readers divide into short- and long-range depending on the distance within which they can access RFID tags. Such distance may vary from few centimeters up to several meters.

Given this technology, our scenario requires that a number of places in the environment (e.g. doors, corridors, etc.) or unlikely-to-be-moved objects (e.g. beds, washing machines, etc.) are tagged with RFID tags. Tagging a place or an object involves sticking an RFID tag on it, and making a database entry mapping the tag ID to a name. In the following of this paper, we will refer to these tags as *location-tags*.

RFID readers accessing one of these tags can lookup the tag ID into the database and infer to be close to the corresponding place. It is worth noting that, other than the pheromone-deployment mechanisms described in the following, such basic technology could implement a primitive localization mechanism [13].

As a final note, it is worth emphasizing that current trends indicate that within a few years, many household objects and furniture may be RFID-tagged before purchase, thus eliminating the overhead of tagging [18]. Moreover, some handheld devices start to be provided with RFID read and write capabilities (the Nokia 5140 phone can be already equipped with a RFID reader [9]).

2.2. Pheromone Deployment

As anticipated in the introduction, pheromones are created by means of data-structures stored in RFID tags. The basic scenario consists of human users and robots carrying (embedding) handheld computing devices, provided with a RFID reader, and running an agent-based application.

The agent, unobtrusively from the user, continuously detects in range *location-tags* as the user roams across the environment. Moreover, the agent controls the RFID reader to write pheromone data structures (consisting at least in a pheromone ID) in all the tags encountered. This process creates a digital pheromone trail distributed across the *location-tags*.

More formally, let us call *L(t)* the set of *location-tags* being sensed at time *t*. It is easy to see that the agent can infer that the user is moving if *L(t)* ≠ *L(t-1)* [1]. If instructed to spread pheromone *O*, the agent will write *O* in all the *L(t)-L(t-1)* tags as it moves across the environment.

For the majority of applications a pheromone trail, consisting of only an ID, is not very useful. Like in ant foraging, most applications involve agents to follow each other pheromone trails to reach the location where the agents that originally laid down the trail were directed (or, on the contrary, to reach the location where they came from). Unfortunately, an agent crossing an-only-ID-trail would not be able to choose in which direction the agent that laid down that trail was directed. From the agent point of view, this situation is like crossing a road without knowing whether to turn left or right.

To overcome this problem, the agent stores in the *location-tags* also an ever increasing hop-counter associated with *O* – we will call this counter *C(O)*. In particular, if an agent decides to spread pheromone *O* at

[1] RFID tag read is not always accurate. So the agent actually verifies that *L(t)* and *L(t-1)* differs for at least 10% of the tags.

time *t*, the agent reads also the counter *C*(*O)* in the *L(t)* set (if *C(O)* is not present, the agent sets *C(O)* to a fixed value zero). Upon a movement, the agent will store *O* and *C(O)+1* in the tags belonging to *L(t+1)* that do not have *O* or have a lower *C(O)*.

In addition, the basic pheromone idea requires a pheromone evaporation mechanism to discard old – possibly corrupted – trails. To this end we store in the tag also a value *T(O)* representing the time where the pheromone *O* has been stored[2].

To better understand how pheromone data structures are stored in RFID tags, it is fundamental to describe how the tag memory has been organized (see figure 1).

RFID tags, other than an unchangeable unique identifier, are typically provided with an array of memory cells, each consisting of few bits.

This is where pheromone data structures will be held. Our proposal is to organize such memory by allocating 3 slots for each pheromone. The first slot will hold the pheromone identifier. The second slot will hold the associated counter. The third will hold the above mentioned timestamp. The very first slot in the tag will be used to hold an index pointing to the first slot available for pheromone storage.

Since RFID tags are completely passive devices, upon a pheromone insertion, the RFID reader must: *(i)* read the tag *index* from *tag[0]*; *(ii)* store in *tag[index]* the pheromone id; *(iii)* store in *tag[index+1]* the pheromone counter; *(iv)* store in the *tag[index+2]* the timestamps; *(v)* store the new *index+3* in *tag[0]*.

2.3. Pheromone Reading and Evaporation

To read pheromones, an agent trivially accesses neighbor RFID *location-tags* reading their memories. Since RFID read operations are quite unreliable, the agent actually performs a reading cycle merging the results obtained at each iteration. Given the result, the agent will decide how to act on the basis of the perceived pheromone configuration.

To realize pheromone evaporation, after reading a tag, an agent checks, for each pheromone, whether the associated timestamp is, accordingly to the agent local time, older than a certain threshold *T*. If it is so, the agent deletes that pheromone from the tag. This kind of pheromone evaporation leads to two key advantages:

1. Since the data space in RFID tags is severely limited, it would be most useful to store only those pheromone trails that are important for the application at a given time; old, unused pheromones can be removed.
2. If an agent does not carry its personal digital assistant or if it has been switched off, it is possible that some actions will be undertaken without leaving the corresponding pheromone trails. This cause old-pheromone trails to be possibly out-of-date, and eventually corrupted.

In this context, it is of course fundamental to design a mechanism to reinforce relevant pheromones not to let them evaporate. With this regard, an agent spreading pheromone *O*, will overwrite *O*-pheromones having an older *T(O)*. From these considerations, it should be clear that the threshold *T* has to be tuned for each application, to represent the time-frame after which the pheromone is considered useless or possibly corrupted.

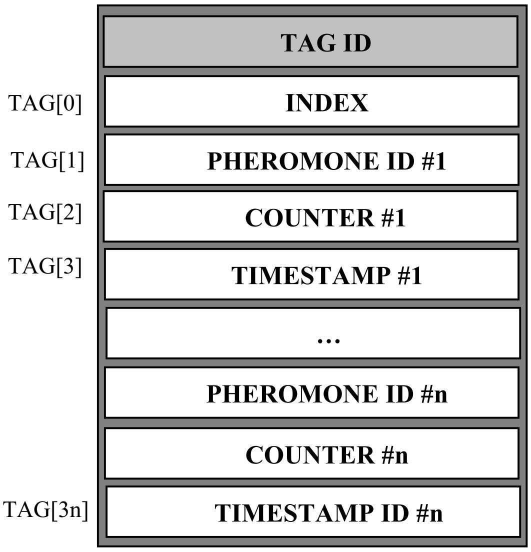


Figure 1. Memory organization of RFID tags.

3. PHEROMONE-BASED OBJECTS TRACKING

In this section, we present a concrete application to test our approach. It consists in an agent-based application to easily find everyday objects (glasses, keys, etc.) forgot somewhere in our homes. The application allows everyday objects to leave virtual pheromone trails across our homes to be easily tracked afterwards.

3.1. Overview

Overall, the object tracking application work as follows (details in the following subsection):

1. The objects to be tracked need to be tagged. For sake of clarity, we will refer to these tags as *object-tags* to distinguish them from the *location-tags* identifying places in the environment.
2. Agents (either robotic or humans) are provided with

[2] *T(O)* is meaningful in a distributed scenario only if the clocks of the agents are synchronized. However, for our purposes, a rough synchronization in the order of the second in more than enough. In this paper we suppose that all the agent are synchronized accordingly.

a handheld computing device, connected to a RFID reader, and running an agent-based application.

3. The agent-based application can detect, via the RFID reader, *object-tags* carried on by the user. Exploiting the mechanism described in the previous section, it can spread a pheromone identifying such objects into the available memory of near *location-tags* (see section 3.2 for details).
4. This allows the object to leave a pheromone trail across the *location-tags* in the environment.
5. When looking for an object, a user can instruct the agent to read in-range *location-tags* searching the object's pheromone in their memory. If such pheromone is found, the user can follow it to reach the object current location (see section 3.3 for details).
6. Once the object has been reached, if it moves with the user (i.e. the user grabbed it), the agent automatically starts spreading again the object associated pheromone, to keep consistency with the new object location.
7. This application naturally suits a multi-user scenario where an user (or a robot), looking for an object moved by another user, can suddenly cross the pheromone trail the object left while moved by the other user.

3.2. Spreading Object Pheromones

To spread pheromones, the agent needs first to understand which objects are currently being carried (i.e. moved around) by the user. To perform this task unobtrusively, it accesses the RFID reader to detect in-range RFID tags once a second.

Let us call *O(t)* the set of *object-tags* being sensed at time *t*, *L(t)* the set of *location-tags* being sensed at time *t*. If the agent senses an *object-tag O* such that $O \in O(t)$, $O \in O(t-1)$, but $L(t) \neq L(t-1)$, then the agent can infer that the user picked-up the object *O* and the object is moving around. In this situation, the agent has to spread *O* pheromone in the new location. To this end, the agent writes *O* in the available memory space of all the *L(t) location*-tags that do not already contain *O*. This operation is performed, for every object *O*, upon every subsequent movement. Similarly, if the agent senses that an object-tag $O \in O(t-1)$, but $O \notin O(t)$ [3], then the agent infers that the user left object *O*. When this situation is detected the agent stops spreading *O* pheromone.

These operations create pheromone trails of the object being moved around.

3.3. Tracking Objects

Once requested to track an object *O* the agent will start reading, once per second, nearby location-tags looking for an *O*-pheromone within the sensed *location-tags L(t)*. If such a pheromone is found, this implies that the user crossed a suitable pheromone trail.

There are two alternatives: either in *L(t)* there are two *location-tags* having *O*-pheromones with different *C(O)*, or *L(t)* contains only one *location-tag*.

In the former (lucky) case, the agent notifies the user about the fact it has crossed a pheromone trail and it suggest to move towards those location-tag having the higher *C(O)* [4]. In the following, we will refer to this as *grad-search*, since it is like following a gradient uphill.

In the latter (unlucky) case, the agent notifies the user about the fact it has crossed a pheromone trail, but nothing else. In such situation, the user has to move in the neighborhood, trying to find higher *C(O)* indicating the right direction to be followed (this is like dowsing - i.e. finding underground water with a forked stick – but it works!). In the following, we will refer to this as *local-search*. Following the agent advices, the user gets closer and closer to the object by following its pheromone trail, until reaching it.

4. EXPERIMENTS

To assess the validity of our approach and the effectiveness of the object tracking application, we developed a number of experiments, both adopting the real implementation and an ad-hoc simulation (to test on the large scale). Basically, our approach consists in developing a simulation matching the real implementation data, and then to use the simulation to extrapolate the results in large-scale scenarios.

4.1 Real Implementation Set-Up

The real implementation consisted in tagging places and objects within our department. Overall, we tagged 100 locations within the building (doors, hallways, corridors, desks, etc.) and 50 objects (books, laptops, cd-cases, etc.). Locations have been tagged with ISO15693 RFID tags, each with a storage capacity of 512 bits (each tag contains 30 slots, 1 byte each, thus it is able to store 10 pheromones). Objects have been tagged with ISO14443B RFID tags, each with a storage capacity of 176 bits (each tag contains only the object ID) [1].

In addition, we set up three HP IPAQ 36xx running Familiar Linux 0.72 and J2ME (CVM – Personal

[3] RFID tag read is not always accurate. If an agent senses than $O \in O(t-1)$, but $O \notin O(t)$, then it performs 3 read operations to assess whether the user actually left the object.

[4] Since we do not require the presence of localization devices, the agent suggests the user to get closet the object having higher *C(O)*, by naming the object – e.g. walk to the front door. The user has to know how to get to that location without further help.

Profile). Each IPAQ is provided with a WLAN card and a M21xH RFID reader [1].

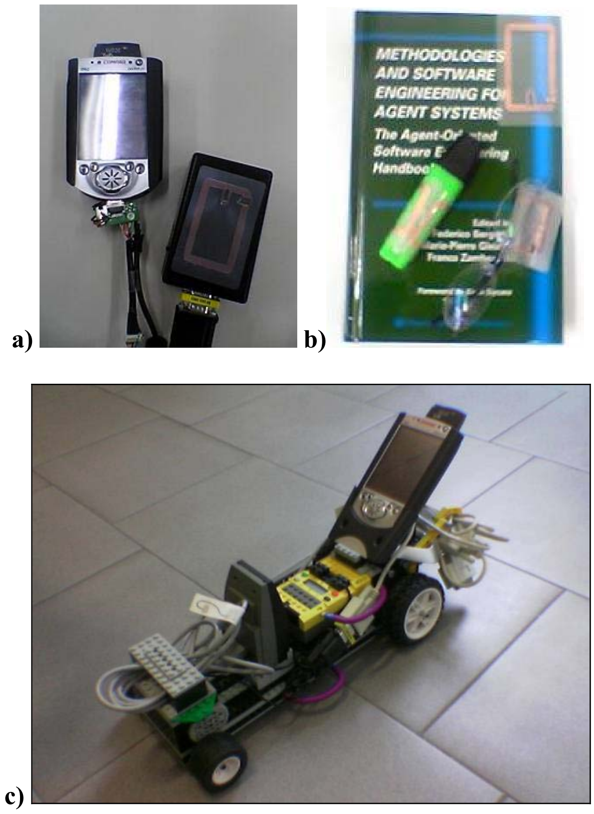


Figure 2. (a) Our test-bed hardware implementation. **(b)** Some tagged objects. **(c)** The Lego Mindstorms robots performing pheromone search.

Moreover, a wirelessly accessible server holds a database with the associations between tag IDs and places' and objects' description (i.e. ID 001 = Prof. Smith's office door). The IPAQ can connect, via WLAN, to the database server to resolve the tag ID into the associated description. Each IPAQ runs the described agent-based application. Finally, a mobile robot has been realized by installing one of our wireless IPAQ (connected to a RFID reader) on a Lego Mindstorms robot [5]. The IPAQ runs an agent controlling the RFID reader and the robot movement. This latter point has been realized by connecting, via IR, the IPAQ to the robot CPU (the RCX Lego brick) enabling the IPAQ to access robot's sensors and actuators. Robots are not provided with a localization device or map of the department. They wander randomly, avoiding collisions, looking for pheromones.

4.2 Simulation Set-Up

To test on the large scale, we realized a JAVA-based simulation of the above scenario. The simulation is based on a random graph of places (each associated to a *location-tag*), and on a number of objects (each associated to an *object-tag*) randomly deployed in the locations-graph. Each tag has been simply realized by an array of integer values.

A number of agents wanders randomly across the locations-graph collecting objects, releasing objects, and spreading pheromones accordingly.

At the same time, other agents look for objects in the environment eventually exploiting pheromone trails previously laid down by other agents. For the sake of comparison, we implemented 3 search algorithms: in *blind-search*, an agent explores the locations-graph disregarding pheromones. In *local-search*, the agent perceives the pheromones in its current node, but it cannot see the direction in which the pheromones increase. In *grad-search*, the agent perceives pheromones together with the directions in which they increase.

The simulator, allows to perform a number of experiments changing a number of parameters such as the graph size, the number of objects, the number of agents involved, the storage capacity of the tags, etc.

Both the real implementation and the simulation have been employed to realize the experiments described in the next subsections and to draw the conclusions described in Subsection 4.4.

4.3 Results of the Experiments

A first group of experiments (reported in Figure 3) aims at verifying the effectiveness of the application. Specifically, we set up two environments: one consisting of 100 tagged places with 100 objects (Figure 3-a) and another consisting of 2500 tagged places with 500 objects (Figure 3-b). 10 agents populate these environments wandering around moving objects and spreading pheromones and, at the same time, looking for specific objects. In the experiments, we report the number of places visited (i.e. number of *location-tags* perceived) before finding specific objects, for different search methods, plotted over time.

These results are the average of a number (over 300) of simulated experiments and verified – on a smaller scale – on the real implementation.

The more time passes the more pheromone trails get deployed. It is easy to see that *blind-search* does not take advantage of pheromone trails and in fact objects are found after visiting on-average half of the places. *Grad-search* takes a great advantage of pheromones, in fact, after several pheromone trails have been deployed, less than 10% of the places need to be visited before finding the object. *Local-search* is useful only in large scenarios: the time taken wandering randomly in a neighborhood, looking for the direction where a pheromone increases, hides pheromone benefits in small environments.

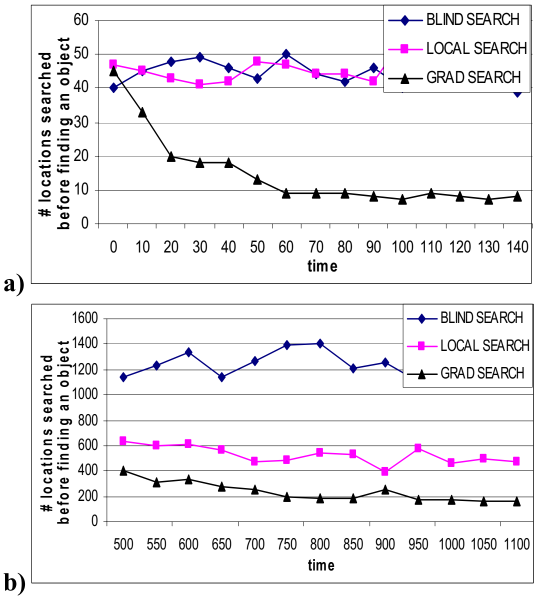


Figure 3. Number of places visited before finding a specific object plotted over time. **(a)** 100 tagged places. **(b)** 2500 tagged places.

A second group of experiments aims at exploring the effects of RFID tag storage saturation upon pheromone spread. This of course represents a big problem, in fact, it can happen that pheromone trails can be interrupted, because there is not available space left on neighbor location-tags, while the object to be tracked moves away. This create a broken pheromone trail leading to a place that is not the actual location of the object.

In Figure 4a, we report an experiment conducted in the 100-tagged-places-environment described before. We plot the number of places visited before finding specific objects for different search methods, over a shrinking tag storage capacity. In this experiments, agents spread pheromones for 150 time steps, then they start looking for objects without spreading pheromones anymore. For sake of comparison, it is worth noting that 150 time steps is exactly the number of steps that concludes the experiment on Figure 3a.

Let us focus on the *grad-search* method that is the most interesting in this context. It can be noticed that, when the tag storage capacity is high, we have good performance (less than 10% of the places need to be visited before finding the object). However, when the capacity fall below 85 pheromones (that is - recalling figure 1 – that the tag has a capacity of less than 85 * 3 = 255 slots = 255 bytes), performance starts decaying really fast and when the capacity is lower than 25, *grad-search* works equal to *blind-search*. It is rather easy to explain this phenomenon: when the tag capacity is low, there are a lot of broken pheromone paths degrading the performances. An agent, reaching the end of a broken pheromone trail, has no choice but starting the search from the beginning.

Figure 4b shows the same problem from another perspective. This time the tag capacity has been fixed to 50 pheromones (150 bytes), and we plot the number of places visited before finding specific objects, for different search methods, over time. Let us focus again on the *grad-search* behavior. It is easy to see that, when time is close to zero, *grad-search* works equal to *blind-search*, since no pheromone trails have been already laid down. After some time, *grad-search* works considerably better than *blind-search*, since pheromone trails drive agents. However, as time passes, tags capacity tend to saturate, the objects are moved, but no pheromone trails can be deployed. This situation rapidly trashes performance leading back to *blind-search* performance.

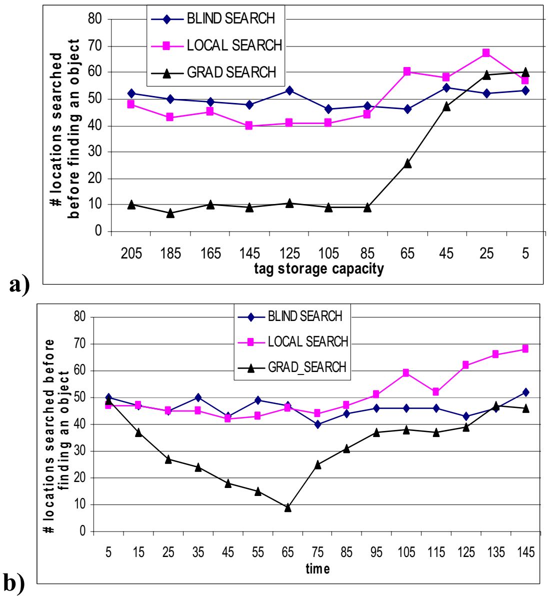


Figure 4. (a) Number of places visited before finding a specific object plotted over a shrinking tag storage. **(b)** Number of visited places before finding a specific object plotted over time, when tags tend to saturate.

Finally, in the experiments depicted in Figure 5, we tried to assess whether the pheromone evaporation mechanism can help in such situation. Figure 5 plots the number of places visited before finding specific objects over a shrinking tag storage capacity (the same of Figure 4a). This time, however, only *grad-searches* are depicted and each plot is associated to a different threshold *T* of pheromone evaporation. Unfortunately, it is rather easy to see that the pheromone evaporation is rather ineffective.

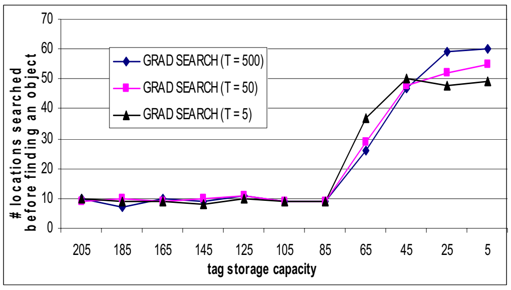


Figure 5. Number of places visited before finding a specific object plotted over a shrinking tag storage space, for various evaporation threshold.

4.4 Lessons Learnt

We get two main lessons from the experiments described in the previous section.

First Lesson: in small environments *grad-searches* work considerably better than *local-searches*. However, this is not longer true in large environments, where the two methods have almost the same performance. This is clearly because the cost of "orienting" in a local neighborhood becomes negligible when the environment is large. Moreover, the drawback of *grad-searches* is the need for *longer-range* (more costly) RFID reader: the reader, in fact, must be capable of reading tags in a "one-hop" neighborhood. On the contrary *local-searches* can work with *shorter-range* (cheaper) RFID reader as well. Overall, the experiments conducted show that in near-future environments (with thousands of objects and places being tagged) *local-search* is a promising approach.

Second Lesson: the limited storage capacity of the RFID tags is a big problem. Basically, if the number of objects to be tracked is greater than the available slots on the RFID tag, in the long run the problem is unavoidable. Sooner or later, a new object will cross to an already full tag, breaking the pheromone trail. We still do not have a solution for this problem. Our research with regard to this topic is leading in two main directions: *(i)* we are currently researching more advanced pheromone evaporation mechanisms. *(ii)* We are considering the idea of spreading pheromone trails not only in *location-tags* but also on *object-tags*. The advantage would be that the more objects are in the system, the more storage space is available for pheromones, letting the system to scale naturally. The problem is how to manage the fact that *object-tags* containing pheromones can be moved around, breaking the pheromone trail structure. As a partial relief from this problem, it is worth reporting that, recent RFID tags have a storage capacity in the order of the KB, making possible to track thousands of objects without changing our application.

5 APPLICATIONS AND RELATED WORK

In the last few years, pheromone interaction and stigmergy have attracted more and more researches due to their power in supporting agent coordination. It is not surprising then, that both there are a number of applications – beside object tracking – that could possibly take advantage or our approach, and there are a number of related work proposing approaches to realize stigmergy either in virtual – simulated – or physical environment. In this section, we discuss these topics.

5.1 Other Application Scenarios

Other than the proposed object tracking application, our approach to actually spread virtual pheromones in the physical environment, well suits a wide range of application scenarios. Here we report two particularly broad and important examples.

Context-awareness: pheromones (i.e. data structures) spread in the environment could provide context information to users. As introduced previously in the paper, localization is a simple –but important - example of this application. Our RFID implementation naturally suits this scenario, in that: *(i)* RFID tags can store – or link to –semantically rich contextual information, *(ii)* context-data can be actually spread in the environment where it will be most useful. Moreover, RFID tags stuck on objects or person could hold information of such objects and person. Reading such tags could be very valuable to asses the application context (e.g. reading the tag associated to the user boss and of a video projector can let infer that the user is in a sort of important meeting with his/her boss) [15].

Motion coordination: pheromones spread in the environment could enable a group of users (both humans and robotics) to coordinate their respective movements. An exemplary application would be distributed environment exploration. Users could decide to explore a specific area if there are not pheromones pointing in that direction (the area is truly unexplored).

In this context, it is important to remark that our approach clearly requires the presence of RFID tags before pheromones can be spread. If the environment does not contain tags at all, our approach could not be used. However, on the one hand, RFID tags are likely to be soon densely present in everywhere (embedded in tiles, bricks, furniture, etc.). On the other hand, it is possible to conceive solutions where agents physically deploy RFID tags while exploring the environment to be used for subsequent coordination. For instance, future development in plastic (and printable) RFID technology [17] let us envision the possibility of enriching an agent with a simple RFID printer to dynamically print in pavements, walls, or any type of surface, RFID tags.

5.2 Related Work

In the last few years, a lot of agent applications inspired by pheromone-interaction and stigmergy have been proposed. In this section, we will report projects trying to realize real implementation of such mechanisms.

In [11] a pheromone-based approach to coordinate Unmanned Airspace Vehicles (UAVs) is presented. This approach has been concretely implemented in a real-world scenario by spreading pheromones in a virtual data-space shared among the UAV agents. In our opinion, spreading pheromones in virtual data-space presents some drawbacks: if the data-space is centralized and globally-accessible, it creates a bottleneck for the application. If it is completely distributed among the locally interacting agents, then consistency may be hard to be maintained.

In [6] pheromones to coordinate robots' movements are spread in a sensor network deployed over the environment. This approach is similar to our RFID implementations: agents connect to nearby sensors and store pheromones in there. In the long-term this would be a really powerful solution: active sensors could implement pheromone evaporation autonomously, but presently it is also very costly. Also, the sensor network solution exhibit battery-exhaustions problems (and, thus, limited life) which our solution prevents.

In [4] and [14] a swarm of mobile robots connect with each other in an ad-hoc network to coordinate their movements. Robots can create distributed data structures (e.g., pheromone trails) over the ad-hoc network defined by the robot themselves. In our opinion, such solution presents problems related to the cost of individual robots, and on the number of robots required to provide a good coverage of the environment and a dense enough network. Also, should the ad-hoc network of robots get partitioned, pheromone trails would be broken.

A terrain-covering robot that exploits sort of pheromones is described in [19]. Here, a prototype robot is presented which is equipped with a pen to leave special ink trails in the pavement. Also, the robot is equipped with proper light sensors to sense the ink trails. In this way, robots can enforce a simple form of pheromone-based coordination (e.g., if an ink trail is sensed, it means that another robot has already covered that part of the terrain. In our opinion, the RFID tag solution is much more flexible, in that it enables using more semantic information for a wider range of applications other than terrain covering.

6 CONCLUSION AND FUTURE WORK

While a preliminary prototype implementation shows the feasibility of our approach, a number of research directions are still open to asses and improve the system practical applicability. In particular, more experiments are required to verify the scalability of the proposed architecture to hundreds of objects being possibly tracked. Moreover, effective solutions to the problem related to broken pheromone trails must be found.

Acknowledgements

Work supported by the project CASCADAS (IST-027807) funded by the FET Program of the European Commission.

7 REFERENCES

[1] Autentiweb, http://www.autentiweb.com

[2] O. Babaoglu, H. Meling, A. Montresor, "A Framework for the Development of Agent-Based Peer-to-Peer Systems", IEEE ICDCS, Vienna (A), 2002.

[3] E. Bonabeau, M. Dorigo, G. Theraulaz, "Swarm Intelligence", Oxford University Press, UK, 1999.

[4] D. Estrin, D. Culler, K. Pister, G. Sukjatme, "Connecting the Physical World with Pervasive Networks", IEEE Pervasive Computing, 1(1):59-69, 2002.

[5] Lego Mindstorms, http://www.legomindstorms.com.

[6] Q. Li, M. De Rosa, D. Rus, "Distributed algorithms for guiding navigation across a sensor network", ACM MOBICOM, San Diego (CA), USA 2003.

[7] J. McLurkin, J. Smith, "Distributed Algorithms for Dispersion in Indoor Environments using a Swarm of Autonomous Mobile Robots", 7th International Symposium on Distributed Autonomous Robotic Systems, Toulouse (F), 2004.

[8] R. Menezes, R. Tolksdorf, "A New Approach to Scalable Linda-systems Based on Swarms", ACM SAC, Orlando (FL), USA, 2003.

[9] Nokia Mobile RFID Kit, http://www.nokia.com/nokia/0,,55738,00.html

[10] V. Parunak, "Go to the Ant: Engineering Principles from Natural Agent Systems", Annals of Operations Research, 75:69-101, 1997.

[11] V. Parunak, S. Brueckner, J. Sauter, "Digital Pheromones for Coordination of Unmanned Vehicles", Workshop on Environments for Multi-agent Systems (E4MAS), LNAI 3374, Springer Verlag, 2004.

[12] V. Parunak, S. Bruekner, J. Sauter, "ERIM's Approach to Fine-Grained Agents", NASA/JPL Workshop on Radical Agent Concepts, Greenbelt (MD), 2001.

[13] D. Patterson, L. Liao, D. Fox, H. Kautz, "Inferring high-level behavior from low-level sensors",UBICOMP, Seattle, Washington, USA, 2003.

[14] D. Payton, M. Daily, R. Estowski, M. Howard, C. Lee, "Pheromone Robotics", Autonoumous Robots, Kluwer Academic Publishers, 11(3):319-324, 2001.

[15] M. Philipose, K. Fishkin, M. Perkowitz, D. Patterson, D. Fox, H. Kautz, D. Hahnel, "Inferring Activities from Interactions with Objects", IEEE Pervasive Computing, 3(4):50-57, 2004

[16] R. Want, “Enabling Ubiquitous Sensing with RFID”, IEEE Computer, 37(4):84-86, April, 2004.
[17] G. Collins, “Next Stretch for Plastic Electronics”, Scientific American 291:74-81, August 2004.
[18] Smart Mobs, http://www.smartmobs.com.
[19] J. Svennebring, S. Koenig, "Building Terrain Covering Ant Robots: a Feasibility Study", Autonomous Robots,16 (3): 313-332, May 2004.

RFID in the Internet of Things: from Static to the Real-Time

Fabio Forno, Antonio Sciarappa

Abstract-- The RFID application scenario is moving from a static context with passive tags to a dynamic one, supported by new technologies that will allow ubiquitous communications, pervasive computing, and ambient intelligence. These concepts will make possible, in the middle-long term view, trend towards always networked devices, where a large population of intelligent objects are connected together introducing the possibility of having concretely the Internet of the Things (IoT).
The paper makes a twofold contribution. It first describes the limitations of the current middleware approach when dealing with a heterogeneous and dynamic set of technologies, explaining how our middleware tackles with them. Then it introduces a novel approach for integrating and federating networks of objects, based on overlay networks built at the top of near real time messaging protocols.

I. Introduction

Current technological development in the area of RFID applications has the main goal to achieve the following advancements:

- Improve current processes with the integration of wireless identification and tracking technologies
- Extend applications with identification facilities working also in difficult physical and environmental context
- Reduce deployment and integration costs exploiting standard hardware and software infrastructure

However this static scenario will change in a dynamic one taking into account the advanced research in the last decade such as ubiquitous communications, pervasive computing, and ambient intelligence
These concepts will make possible, in the middle-long term view, trend towards always networked devices, where a large population of connected intelligent objects will introduce the possibility of having concretely an "Internet of Things" (IoT)[1] at the edges of the current Internet. In the prospective this network will be composed by a countless number of devices, and most of them will be mobile, roaming nodes, without any pre-defined network address, or very limited processing capabilities. Such networks need communication capabilities between individual nodes, and between nodes and access points or gateways that provide connectivity with the outside world. Many of these nodes could be integrated into every day devices; they could be found inside cars, at home, on the human body etc. The application areas based on such networks are varied and numerous, including, for example, intelligent homes, car safety, object-service-process tracking, etc.

This new network is very different from the traditional one and in many ways more complex even than the Internet itself. This fact is due primarily to the number of nodes that could exist in the expanded model of a worldwide RFID-smart objects network, which will be several orders of magnitude larger than the number of nodes on the current Internet. This simply means that traditional computing architectures and infrastructures will not be adequate to handle the relevant higher data volumes expected in a network of RFID tags and some important research activities have to be carried out. In the following a shortlist is reported of technical research challenges that have to be addressed at different levels of the ICT infrastructure with reference more in to the following topics:

1. *Edge technologies*, such as sensors and actuators, passive and active identification tags, or embedded systems that are attached to real objects and make objects "smart" enough to participate in IoT application scenarios.
2. *Networking technologies*, such as fixed, mobile, wired and wireless networks allowing the highly available bi-directional communication on different levels between real objects, applications and services that offer specific functionality.
3. *Middleware systems* that must be scalable, secure in order to put real data into the context of various IoT applications. As the IoT implies that enormous numbers of data sources need to be connected and related to each other, flexible and dynamic middleware has to support the heterogeneity of available devices, sensor networks and other technologies.
4. *Service platforms* that run in the background have to support a relevant management of all involved technical components in an integrated way ensuring scalability, high availability, and the safe and secure execution of the requested functionalities.
5. *Web service technologies* must be improved to provide a new way of making information and services available while reducing interoperability issues and enhancing extensibility, platform independence and standardised exchange of messages.

The authors are with the Istituto Superiore Mario Boella, Torino, Italy, fabio@bluendo.com, sciarappa@ismb.com

In the following paragraphs the paper will focus on the enabling technologies that will make the IoT happen. In particular, in paragraph 2, we will analyze the current middleware approach and the EPC (Electronic Product Code) specifications, and we will identify the limitations towards the integration of heterogeneous pervasive technologies, while in paragraph 3 we will introduce how our RFID Middleware deals with these issues. In Paragraph 4 we will move a step forward in the IoT, identifying the general communication requirements for smart objects and, finally, in paragraph 5 and 6 we will show how most of these requirements can be satisfied by adopting the XMPP protocol and its extensions.

II. THE MIDDLEWARE APPROACH: LIMITATIONS OF THE EPC INFRASTRUCTURE

In recent years middleware has been the common approach for dealing with the increasing complexity of distributed infrastructures such as RFID systems, sensor networks, smart environment[2]. Middleware is a computer software that connects devices, software components, services and applications, offering a common set of abstractions to developers, in order to simplify the production cycle of distributed applications. However, though a middleware dramatically reduces integration costs and time to market, it may present some drawbacks. In fact it usually offers a fixed scheme of abstractions, optimized for tackling the problems of a given application field or hardware technology and, though different applications may have similar solution patterns, the infrastructure is usually difficult to open and extend to different areas. Therefore this approach may lead to isolated silos and to pillars of communication protocols that often replicate efforts for solving similar problems, adding complexity that can easily become unmanageable. Moreover the communication between different middleware solutions (e.g., a sensor network may need to be integrated with a real time positioning system and with RFID tracking) may be hindered by abstractions which are incompatible with a given application, or lack of desired control or simplicity.

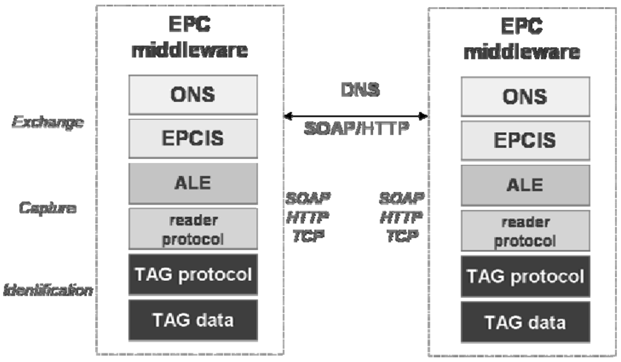


Figure 1. EPC Middleware protocol stack. *Identification* and *Capture* layers are isolated inside the pile of protocols in each middleware instance; the only communication gate is at the highest level, the *Exchange* layer, using standard Internet protocols.

An example is the EPC infrastructure, which defines a complete stack of protocols for collecting, processing, storing and accessing RFID data or events across organization boundaries. The true purpose of the EPC infrastructure is sharing information among all the actors of supply chains with the help of RFID. Therefore the EPC middleware covers these macro areas:

- Tag reading, filtering a collection, with a stack of protocols standardizing the communication with readers (Reader Protocol) and the rules for filtering and reporting events, named Application Level Eventing (ALE).
- Discovery of the owner of tag data, with the help of standard identification schemes and a global directory service named Object Name Server (ONS) built on the model of DNS.
- Information sharing across organization boundaries, leveraging on standard protocols like SOAP and HTTP for transport, and common data storage formats such as the EPC Information Server (EPCIS)

However, though the EPC infrastructure is a successful enabler for a first step towards the IoT, i.e. object identification and tracking, it lacks some of features that could bring effective internetworking to smart objects. In particular it does not deal with:

- Mobile readers; in some systems readers need an autonomous logic since they may temporary work offline (e.g., handheld devices, smart roaming objects with WiFi connections).
- Customizable business logic near the objects; ALE only allows filtering and reporting events, however most the actions taken as answer to input events (e.g., switching on lights, opening gates, etc.) should be managed by customizable logic placed inside the middleware, not by fixed rules.
- Propagation of events across organizations; present EPC infrastructure allows only querying for tag data, there is no support for automatically push events to objects once they are outside the network of their owner (e.g. for software updates), or for receiving events from them (e.g. alarm for critical situations)
- Different identification technologies; NFC and other short range wireless technologies (e.g., Bluetooth or Zigbee) could be used for identification as well.

III. A FIRST STEP TOWARDS THE INTERNET OF THINGS: THE ISMB MIDDLEWARE

ISMB (Istituto Superiore Mario Boella) has developed a novel RFID middleware architecture which is compliant with EPC specifications and which tries to address some of the issues of traditional middleware. In particular our architecture deals with the following issues:

- seamless handling of wired and wireless mobile readers that may be temporary offline

- customization of the business logic inside the middleware, as a dynamic and reconfigurable extension of filtering stage (ALE)

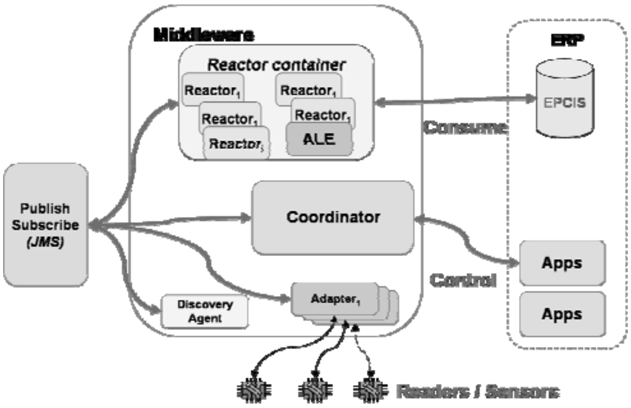


Figure 2. General architecture of the ISMB Middleware. Though compliant with EPC specifications, the middleware offers higher control on lower levels, with discovery services and optimized adapters for mobile readers, and with customisable business logic (reactors).

In order to achieve these goals we have based all the messaging across middleware components on a JMS (Java Message Service) system, which is a high performance communication bus supporting different message exchange patterns. In particular we exploit its publish/subscribe functionalities in order to have a loose coupling between event producers and consumers and the capability to compose arbitrarily event chains (e.g., we may consistently handle events types such as the discovery of a new reader or a tag reading with the same event bus and a processing logic).

Mobile readers are managed by a software component named "Discovery Agent", which is responsible of finding new readers and notifying all system components about their status, while the reader adapters can work as buffers for the temporarily offline readers, both for commands and collected data. All tag observations are asynchronously dispatched to any component via JMS as soon as they are available (mobile readers store them locally until they become connected)..

More interesting is the "Reactor Container" which is responsible for consistently hosting all the business logic of the system. Any part of the business logic is modeled on the concept of "Reactor", which is an abstract class with the following properties:

- it can subscribe and asynchronously react to external events of any type;
- it produces new events that can be consumed by other reactors, thus implementing reusable building blocks of logic that can be composed for assembling complex operations
- its life cycle is manageable by an external module (installation, configuration, start and stop), the Coordinator, which allows applications to hot plug business logic into the system

The ALE module, i.e. rule based filtering, is therefore implemented using a set a "reactors", which receive RFID events through JMS, process them and dispatch aggregated results to other modules via JMS again. The advantage of this approach is that the same engine can be easily extended for handling different types of events, such as device management, sensor inputs and enforcing more complex rules than those allowed by ALE.

IV. BEYOND RFID: INTERCONNECTED OBJECTS AND NEAR REAL TIME MESSAGING

From the perspective of the IoT a RFID middleware is just the instance of a specific application within a specific technology. In fact, as the best definition of the Internet is the "network of networks", the IoT could simply be defined as "a network of networked objects". In both definitions the stress is not on a specific application, but on the basic capacity of communicating between peers without obstacles, thus enabling real innovation at the edges of the network. Therefore the EPC Middleware is a first enabler, but in order to make real the concept of the IoT a more general infrastructure is required.

More specifically, we need general purpose messaging framework able of delivering data and events between objects, and objects and remote controllers. The traditional Internet protocol stack, in fact, can effectively deal only with a network topology which is almost static and always connected. On the contrary connections between objects are often transient and the topology of the network may significantly change in the time (e.g. roaming devices have different addresses). Using a messaging approach we introduce a further abstraction level that can cope with several complexities of the underlying transport:

- use of virtual addresses which are independent from network location and technology, thus allowing correct routing of messages to roaming node;
- offer a uniform transport and quality of service independent from the actual transport (it may be either TCP, UDP or some ad hoc protocol over Bluetooth, Zigbee or other short range technology);
- define gateways between the IP network and non IP networks;
- store messages for nodes that are temporary offline and deliver them as soon as nodes become available.

In particular, for the last requirement, we speak of "near realtime messaging", since the infrastructure should be able to buffer messages and deliver them in almost realtime when the objects become connected, without the need of polling from the destination nodes.

Another important requirement of the messaging framework is the capability of working without a central coordination or authority, but with the cooperation of the actors taking part to the network, following the same exact model of the classical Internet. Within messaging frameworks this concept is known as “federation”, i.e. as the capability of routing messaging between independent networks of messaging servers and their user bases. The best know example of federation is the email system, based on SMTP, but also some instant messaging (IM) systems provide the concept of federation as a built-in feature. In particular open IM systems, such as XMPP[4], SIP[5] or IMPS[6] provide scalable and federated infrastructures that can be used for interconnecting objects.

From a different point of view, messaging systems may be seen also as overlay networks at the top of the network infrastructure enabling uniform routing and services between nodes in heterogeneous network.

V. Overlay networks based on the eXtensible Messaging and Presence Protocol

The eXtensible Messaging and Presence Protocol (XMPP), formerly known as Jabber, is a XML based near realtime protocol for messaging, presence and request-response services. XMPP defines both the format of the exchanged messages (generally named “stanzas”) and a network architecture enabling open federation of services. The basic suggested architecture is based on the client server model in which:

- XMPP entities (end nodes) link up to a server using a TCP socket, and can be addressed using a unique identifier, named Jabber ID (JID), having the form user@server.org/resource. The “/resource” part is an opaque string that allows multiple connections associated to the same identity; they can be seen as specialized mailboxes associated to the same user.
- Federation is enabled by direct server to server communication, in which XMPP servers act as router of stanzas between their user bases. This is made possible since all XMPP server must have a public IP address and a fully qualified domain name associated to them, requirement which is very common for any Internet service

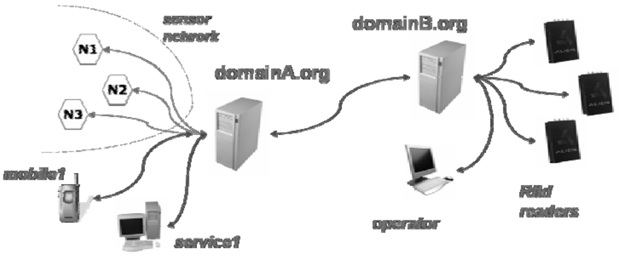


Figure 3. Federation of services and devices in the Internet of Things

This network model builds an overlay network that enables event routing and delivery between any pair of nodes in the Internet, also when nodes are behind firewalls or have private addresses. It becomes therefore possible to address any entity, and any object, independently from its network location. As a result applications and objects in the IoT can exchange events in realtime without having to deal with the complexities of the underlying network. The XMPP infrastructure can deal also with another critical problem in the IoT: end nodes that can be connected with ad hoc non IP protocols, such as sensors nodes which use ZigBee or other protocols optimized for battery consumption. Usually these protocols have incompatible addressing schemes and very limited connections not allowing to contact them. However with XMPP it is possible to build gateways having both the IP and the ad hoc interface. These gateways map foreign addressing schemes to valid JIDs, and they can translate and buffer stanzas, so that any entity can see sensor nodes as if they where regular XMPP peers.

XMPP allows three basic types of stanzas, which can be used for covering all message exchange patterns:

- <message/>: asynchronous one way message between two entities, no reply is required and the delivery is granted (if the target node is offline, the message is stored in the destination server and delivered when the recipient becomes available).
- <iq/>: request-response message for making queries to remote nodes; the behavior is similar to HTTP requests, with the advantage that any entity can asynchronously initiate the request.
- <presence/>: originally meant for distributing presence information in IM systems, presence stanzas are a simple way for broadcasting status information to a list of preauthorized nodes, feature which is extremely useful for controlling the status of sets of smart objects in the environment.

Another significant feature of the XMPP is its extensibility, which allows to transfer arbitrary structured data between applications and objects, making it ideal for machine-to-machine communication (M2M) or for remote controlling of intelligent devices. Extensions can be formally defined by the XMPP Standards Foundation with a peer review process and vote of the board of the Foundation. Several extension are of primary importance as enablers for M2M communication. Examples are the transport of web services over XMPP (SOAP[7] , XML-RPC and lighter forms of web services such as ad-hoc commands[8]), and most important Publish/Subscribe, which we will analyze in depth in the following paragraph.

One more interesting feature of XMPP from the perspective of remote control of intelligent devices is the fact that it is impossible to forge the addresses of message senders, since servers make strict controls on the “from” field, ensuring that it matches the JID used for authentication. Therefore, if the server is trusted, it becomes easy to implement authorization schemes based on access lists, which are reasonably lightweight even for nodes with very limited resources, besides minimizing the number of required messages.

VI. Advanced messagge distribution patterns: Publish/Subscribe

XMPP provides a general framework for end-to-end communication between controlling applications and remote devices, however, in order to scale, more advanced message exchange patterns are needed. Typical IoT applications, in fact, require to address groups of end nodes for sending commands for setting configuration options, and they require to aggregate events or observations from different classes of devices. Moreover it is often required the capability of handling hierarchies or classes of objects and devices. For example a car manufacturer could need to set a configuration option of all his cars, or just of a model type or a of all the cars sold in a given market. Sending a one-to-one message to all the involved cars would be terribly inefficient and error prone, and a sort of broadcast or a one to many message would be preferred.

One more requirement is the capability of loose coupling event producers and consumers. When deploying sensor networks, for example, it is often impossible to forecast all the possible consumers of observed data during the whole life cycle of the system, and reconfiguring the sensors for adding new applications is not cost effective. A better solution is defining fixed and known endpoints where the observations are published, and then attach consumers to those endpoints just changing the configuration at server side.

Generalizing we can abstract most of the communications in the IoT with events, where, accordingly to their function, nodes can be both:

- event producers, such as nodes in sensor networks (for produced data), tag observations of RFID readers, GPS receivers, software coordinators and business elements (when sending commands or alarms), etc;
- event consumers: nodes in sensors networks and RFID readers (for receiving configuration), actuators, mobile devices (e.g. for receiving instructions), coordinators, business elements (processor of observations).

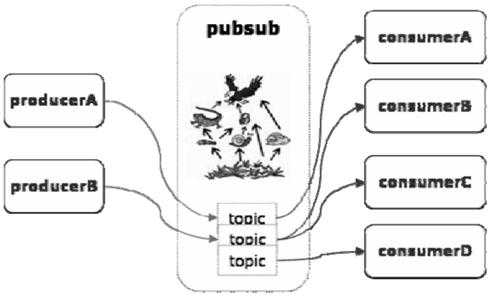


Figure 4. General Publish/Subscribe system

Publish/Subscribe (pubsub) is a pattern for supporting general purpose event distribution, where the producer cannot know in advance the possible consumers. In a *pubsub* system producers publish the events in a well known “topic” or “node”, i.e. a mailbox where events are collected; consumers subscribe to the mailboxes they are interested in and start receiving the related events.

XMPP provides a sophisticated pubsub extension [9] for distributing events to any XMPP entity, and therefore it allows to build complex event distribution systems to a from a wide set of smart devices. The main advantages over other pubsub systems like JMS are:

- seamless event delivery to roaming nodes, nodes in private networks, or nodes that are temporary offline;
- federation of different pubsub systems, since each pubsub servers can be distributed across different domains (more control for the owners of data and better scalability of the whole system, without a single point of failure)
- hierarchical system with the concept of “leaf nodes” (nodes where events are actually published) and “collection nodes” (containers of other nodes, for aggregating events)
- fine grained control of node subscriptions and roles of XMPP entities (e.g. the node creator – the owner – can delegate configuration or publishing to other XMPP entities)

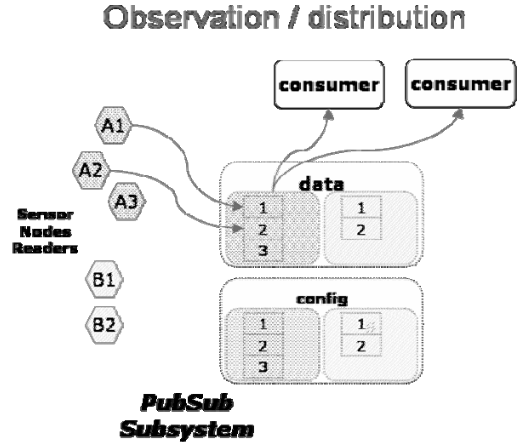


Figure 5. Sensor network observation/event distribution through a pubsub system

The above figure shows a general use case of a pubsub system based on XMPP, which exploits node hierarchy for managing two sensor networks. Two root topics “data” and “config” host respectively events from and to the sensors. All the sensors have a one-to-one association with a topic in the “data” collection, where they publish observations, and each sensor network is associated to a collection node gathering all its sensors. Consumers can therefore receive events from a single node, network or the whole system, by simply subscribing to the relevant topic.

For sensor configuration and control the scheme is similar, since sensors are subscribed to the relevant topics in the “config” subtree, where the coordinators may publish events for sensing commands with the desired granularity

VII. CONCLUSIONS

The Internet as we know is changing radically; in the last few years it became a mass-market, consumer oriented network. Now, it is set to become fully pervasive, interactive and intelligent. Real-time communications will be possible not only among humans but also among things at anytime and from anywhere. The advent of the IoT will be the enabler for the creation of innovative applications and services, which will enhance quality of life and provide new revenue opportunities .

However the introduction of the IoT requires a relevant effort in terms of research and technologies in the area of RFID, short-range wireless communications, and infrastructure for real-time localization and sensor networks that make the task not easy. In particular the next challenge will be building a technologically neutral communication platform for smart objects, as the Internet is for computers, and therefore allowing to add value at the edges of the IoT.

The paper has described studies and development about middleware functionalities that are a first step towards improved and more integrated control of heterogeneous identification and pervasive sensors. Our approach, in fact, allows controlling a wider set of devices than simply RFID readers, with support for dynamic reconfiguring wireless networks. Moreover we have analyzed the requirements for opening such network to a global scale by means of real time messaging protocols, and we have presented a possible solution based on the XMPP protocol, which allows federation of services, a high degree of extensibility, offering support for complex message distribution patterns such as pubsub. That will support main requirements of IoT.

VIII. References

[1] EPoSS Experts, “The Internet of Things - Vision”, EPoSS Expert Workshop “Beyond RFID – The Internet of Things”, Brussels, 11/12 March 2008

[2] S. Hadim and N. Mohamed, “Middleware: Middleware challenges and approaches for wireless sensor networks,” IEEE Distributed Systems Online, vol. 7, no. 3, pp. 1, 2006.

[3] Armenio et al., “The EPC Gblobal Architecture Framework”, Version 1.2, EPC Global, 10 Sep 2007, Status: Final

[4] P. Saint-Andre, “RFC 3920: Extensible Messaging and Presence Protocol (XMPP): Core”, IETF, Oct. 2004, Status: DRAFT

[5] J. Rosenberg et Al., “RFC 2543: SIP: Session Initiation Protocol”, IETF, Jun.2002, Status. DRAFT

[6] “IMPS Architecture”, Ver 1.3, Open Mobile Alliance, Jan.2007, Status: DRAFT

[7] F. Forno and P. Saint-Andre, “XEP-0072: SOAP Over XMPP,” XMPP Standards Foundation, Dec.2005, Status: DRAFT.

[8] M. Miller, “XEP-0050: Ad-hoc Commands,”, XMPP Standards Foundation, Jun.2005, Status: DRAFT.

[9] P. Millard and P. Saint-Andre and R. Meijer, “XEP-0060: Publish- Subscribe,” , XMPP Standards Foundation Sept. 2006, Status: DRAFT.

NFC: Integration between RFID and Mobile, state of the art and future developments

U. Biader Ceipidor, C. M. Medaglia, A. Moroni, G. Orlandi, S. Sposato

***Abstract* - This paper provides a full description of Near Field Communication, a short-range, standards-based wireless connectivity technology, derived from RFID technology that uses magnetic field induction to enable communication between electronic devices in close proximity (few centimeters). NFC enables the so called "touch paradigm" in which users can perform intuitive, safe, contactless transactions, connect electronic devices and access digital contents simply by "touching" or bringing devices into close proximity.**

Moreover, the paper points out the importance of creating an NFC ecosystem based on a standardized environment, which is the main purpose of the European research project called StoLPaN (Store Logistics and Payment with NFC).

***Index terms* - NFC, Mobile, standardization, StoLPaN project**

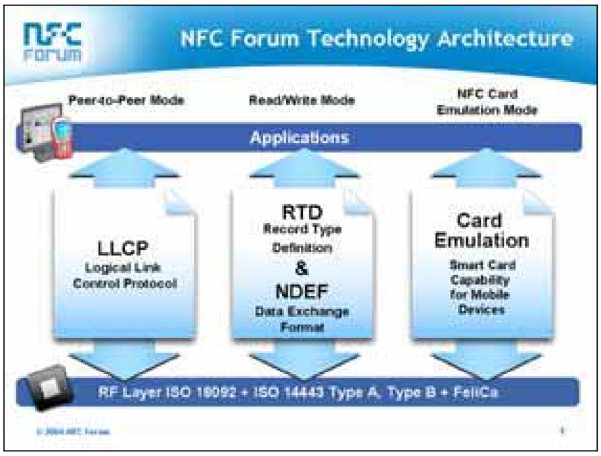
I. INTRODUCTION

What is NFC - Near Field Communication (NFC) [1] is a short-range wireless connectivity technology based on radio frequency communication and designed for intuitive, simple and safe communication between electronic devices. NFC is standardized in ISO/IEC 18092, ECMA-340 and ETSI 102.190. It is also backward compatible with ISO/IEC 14443 type A and type B, which are the standards for Proximity Cards (operating at a maximum of 20 cm).

NFC can be looked at as the merging of two different technologies: the Mobile one, based on GSM standard and the contactless technology, based on the already mentioned ISO 14443 standard. NFC inherits advantages from both, adding benefits for users and other involved subjects (MNO, handset manufacturer, POS vendors, banks, and so on).

About cell phones, these are not only largely used by people all over the world, they are personal objects always brought in pockets or bags by the owner. Moreover, there's a constant ask for new value-added services, over which the operators build most of their earnings.

On the other side, contactless cards are very easy to use. To make a contactless transaction, it's enough to put the card close to the reader (centimeters) and data's exchange takes place in milliseconds, saving time comparing with the same operation made with a "contact" card (like a magstripe card for example, or any card with a "contact" chip).



NFC Forum Technology Architecture

Peer-to-Peer Mode

Read/Write Mode

NFC Card Emulation Mode

Applications

LLCP
Logical Link Control Protocol

RTD
Record Type Definition
&
NDEF
Data Exchange Format

Card Emulation
Smart Card Capability for Mobile Devices

RF Layer ISO 18092 + ISO 14443 Type A, Type B + FeliCa

Figure 1 - NFC Forum Technology Architecture

Contactless cards can rely on a constantly growing infrastructure too. Examples are electronic ticketing in use in London, Milan, Rome, all based on contactless cards.

Many advantages can be taken from merging contactless and mobile technology: first of all, the presence of devices such as a monitor and a keyboard, allow the user to a deeper interaction with the service. Moreover, a cell phone can be easily connected to a net via GPRS, UMTS or Wi-fi, and it opens the doors to new scenarios. For example, a user with a NFC mobile phone could purchase a ticket through the net, via SMS, and validate it on the public transport using NFC technology, using an only device.

II. STATE OF THE ART

Near Field Communication technology was jointly developed by Philips and Sony in 2002. The first trials "on the field" started two years later, with a number of leading handset manufacturers, such as Nokia, Motorola and Samsung. The first commercial implementation of NFC products came in 2005 in Hanau (Germany), after a successful ten-month field trial. With their NFC-equipped Nokia 3220 mobile phone, citizens of Hanau are now able to purchase and validate public transport tickets.

Moreover, in 2004, leading mobile communications, semiconductor and consumer electronics companies formed the non-profit industry association, the NFC Forum [2], to advance the use of NFC technology through standard specifications that ensure interoperability. The Forum now

RFID Lab, University of Rome "Sapienza"

has over 120 member organizations worldwide (as at the end of 2007).

In June 2006, the NFC Forum introduced standardized technology architecture (Figure 1), initial specifications and tag formats for NFC-compliant devices.

The initial set of four tag formats that all NFC Forum-compliant devices must support are based on ISO 14443 Types A and B and FeliCa (derived from the ISO 18092). Tags compatible with these mandatory formats are available initially from Innovision, Philips, Sony and other vendors, and more than one billion tags are already deployed globally.

The main challenge for the market deployment of NFC applications and services is represented by the actual diffusion of mobile phones equipped with NFC technology. However, assessments from researcher companies such as ABI research, foreseen a strong growth of NFC mobile phones in the next years. In particular, ABI research stated that in 2012 there will be 303 million units of NFC-enabled cellular terminals, representing nearly 21% of handsets shipped worldwide in that year [3].

III. NFC STANDARDIZATION

Among the issues that slowed down NFC devices' production by manufacturers, the main one is the lack in standardization, most of all about the architecture.

As we already outlined, the NFC communication protocol has been standardized from ISO, and other organisms such as the NFC Forum take care, among others, to cover non standardized parts promoting the realization and the diffusion of some specifications. Notwithstanding this, some problems remain open.

First of all Secure Element's position, which can be embedded in the hardware, becoming linked to the company who produces the mobile phone, or it could be put into the SIM card, hence linked to the mobile network operator chosen by the user.

It's clear how this issue is crucial for the development of NFC technology and its diffusion through the market; without an agreement on this key point, it would have been hard to start a massive production of devices supporting NFC technology.

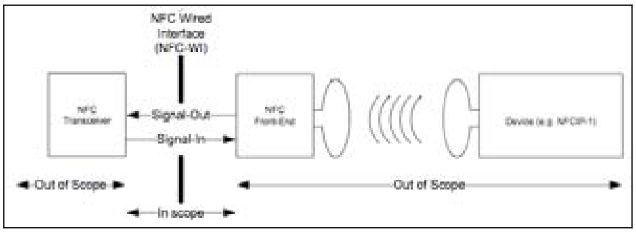


Figure 2 - Context diagram for the NFC wired interface. Source: ECMA

A. Some history

As already mentioned, there basically exist two different architectures with respect to the position of SE. First devices from the manufacturers (Nokia 6131, Samsung x700, Motorola SLVR L7) had the Secure Element embedded into the hardware and connected to the NFC chip through a "full duplex" link: two connections, sign-in and sign-out, between the NFC chip and the security controller (Figure 2).

Using the SIM card as Secure Element it was not possible to keep the double connection with the NFC chip, because seven among the eight contacts of a standard SIM card are busy for other functionalities, as stated from GSM Association in September 2006: five are for GSM/3G nets, the other two are devoted to data download through USB interface.

The only available solution was to connect the SIM card to the NFC chip through a single electric contact. This "half duplex" solution has been introduced by Gemalto in 2006, since data travel on an only wire in both ways, but never at the same time, and has been denominated Single Wire Protocol.

Then it's easy to understand why mobile hardware producers pushed towards this direction, even with concrete initiatives. Thirtyfour of them, which together have 1.3 billion customers, participated in the initiative "Pay-Buy-Mobile" promoted by GSMA to suggest the SWP as the main hardware solution to connect SIM cards to NFC chips [4].

The European Telecommunications Standards Institute (ETSI) adopted the SWP as a standard in October 2007.

Once the hardware part has been standardized, it remained to define a common standard for the software layer, the so-called Host Controller Interface (HCI), which is a key point in order to enable business models to be built upon technical and logical capabilities. On the 22th February of this year, the ETSI Smart Card Platform working group voted in favor of the so-called “Option A”, which counts on a Multi-host interface, based on an open architecture.

This vote was very important for all the companies related to NFC ecosystem, because it opens the last door for a commercial roll-out by giving the possibility to manage different applications through the same HCI.

IV. NFC ECOSYSTEM: THE STOLPAN PROJECT

In order to develop an open architecture for the development, deployment and use of NFC-enabled applications in mobile handsets, the European Commission and Information Society Technologies (IST) program has funded a three-year project called StoLPaN (Store Logistics and Payment with NFC), started in 2006 [5]. The project involves a pan-European consortium of companies, universities and user groups. The aim is to turn NFC enabled mobile handsets into multifunction terminals with bi-directional interaction between the wireless NFC interface and mobile communication channels and to demonstrate the use of this generally applicable new technology in the retail logistical value chain, and also in mobile payment, ticketing and other use cases.

In order to do that, the StoLPaN project proposes a general J2ME Host environment that will enable NFC mobile handsets to run various applications making use of the NFC interface (Figure 3).

After a research phase, new developments will be used to provide a secure, transparent environment in NFC enabled mobile handsets, where the handset's resources will be used for managing the various functions of the NFC chip, enabling it for simultaneously operating as payment card, as secure payment purse and also being the host of multiple other business instruments. Through the new technology secure, independent remote management of the NFC chip will be possible over wireless channels. Secure management of the NFC chip is realized by the extension of existing and development of new APIs, standards.

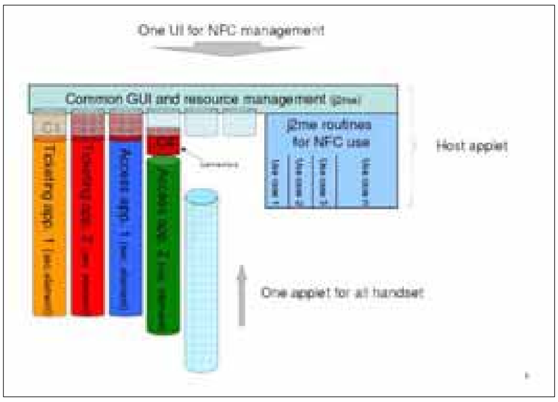

One UI for NFC management
Common GUI and resource management (j2me)
j2me routines for NFC use
Host applet
One applet for all handset

Figure 3 – The StoLPaN host

The described technology developments will be demonstrated in a smart retail environment, where customers will be able to make purchases, pay and check-out using the NFC enabled smart devices without any check-out counters, greatly reducing shopping time and increasing store capacity.

Besides technical research and development the project work includes research activity related to security, legal, privacy and consumer protection issues relevant in case of RFID-NFC use and store operations. Business models supporting the necessary logistical value chain and operation of the NFC services will also be prepared.

Results will be submitted to the relevant trade bodies for adoption by the payment, mobile, transit and ticketing industries to create a standardized NFC ecosystem.

V. RFID LAB'S NFC-BASED APPLICATIONS

RFID Lab of "Sapienza" University of Rome works on designing and creating new applications based on NFC technology. In particular the RFID Lab is focused on analyzing the user interaction model for NFC enabled applications [6, 7]. Many of the today NFC enabled applications are not framed under a general interaction model. This is a significant problem for NFC because it can both be used for simple interactions like touching a secure door with a cell phone to gain access and for more complex scenarios such as buying and validate a bus ticket. Without a stable user model, NFC enabled applications may end up, a mix of poorly thought out interfaces without a unifying interaction model.

Here are some applications realized by the RFID Lab.

A. NFCTicketing

NFCTicketing is a complete system of J2ME applications that allow users to buy and use tickets for public transport. NFCTicketing tries to find a balance between the decreasing of the information necessary to the user and the increasing of the flexibility of the application. In this way, the NFCTicketing service will be not only less complex and then less expensive (keeping the cost of development and of implementation very low) but will also be more usable for a broad community of users. The NFCTicketing service combines latest-generation technologies such as Near Field Communication with well-known and already used technologies such as Short Message Service (SMS). Ticket purchasing and validation procedures are very simple for users, as demonstrated during the usability tests driven on the application into the Usability and Accessibility Lab (LUA) of CATTID (Figure 4).

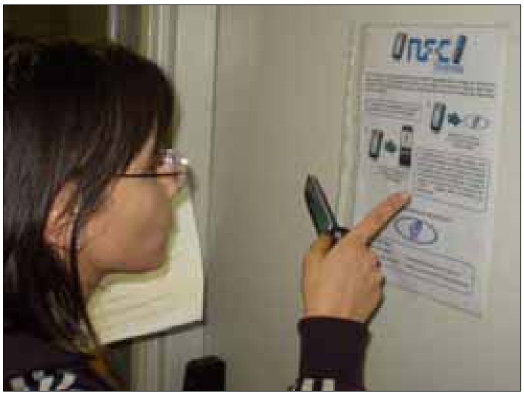


Figure 4 – A user involved in the usability test

Placing Smart Posters with NFC passive tags in different locations such as bus stops, train and metro stations, tobacconists and so on, NFCTicketing gives users the possibility to purchase bus or metro tickets directly by the phone, avoiding having to carry coins and also solving the problem when ticket retailers are closed. Once installed in the phone, the application starts automatically when the user draws up his phone near the Smart Poster. The proximity with the tag wakes up a midlet that guides the user in the shopping process, step by step. After a welcome message, the user has two different possibilities: to purchase tickets or to check the number of tickets already available in the phone. If the user chooses the first possibility, the user has to insert the exact number of tickets required, using the mobile's keyboard. The system waits for the user's confirmation and then sends an SMS containing the ticket request to the service center. The server replies with another SMS, confirming to the user that the tickets have been charged into the Secure Element of the phone in order to avoid external intrusions. The ticket can be then validated via NFC at the metro gate or on the bus, otherwise it will remain on the user's mobile phone until it is used. The conductor equipped with an NFC mobile phone or PDA can check the validity of the ticket simply drawing up his device to the user's one.

B. SIMplyCity

SIMplyCity is an application for NFC mobile phones that helps tourists in their sightseeing around the city. SIMplyCity has not only the capabilities to highlight the most interesting places to visit, but it also permits to orient tourists as a local, giving them information about restaurants, shops, offices and public services (postal offices and so on). SIMplyCity could be also used by local citizens who wish to have information about their city in a very simple and cost-effective way. Placing Smart Posters with NFC passive tags in different locations such as train stations, airports etc., SIMplyCity offers the possibility to obtain information about points of interest in the city simply holding the phone near the pictures highlighted in the poster. A midlet, stored in the phone, is started up by the proximity with the tag placed behind the Smart Poster and it shows to the user several services which he could make use of. For example, it offers the possibility to know how to reach the highlighted places, to buy a bus ticket, or to obtain information about entertainments offered in the city during the day.

C. Touch'n'PAy

Touch'n'PAy will facilitate the communication between citizens and Public Administration, thanks to the integration of mobile-phones with smart cards and the NFC technology. One of the immediate effect will be the reduction of the queue at the front offices. A citizen equipped with an NFC-integrated mobile phone could pay several taxes (for example the tax on the house or the tax on the garbage) straight by his phone, avoiding the queues that overload nowadays Italian public offices. Placing NFC totems or Smart Posters with passive Tags working at 13,56 MHz in public areas like post offices, but also tobacconists', banks and so on, should significantly reduce queues, allowing people to authenticate and use government services at these locations.

In the first release of the application, it is necessary for citizens to register at the Touch'n'PAy service via web through a government server: in particular, citizens have to declare their personal details (name, surname, taxpayer's code) and bank account information, and associate them with their phone number, that will actually be the univocal parameter of authentication.

After the registration process, every time a citizen would make a payment towards government, he only needs to bring the phone near the Smart Poster: the NFC antenna in the phone reads the Tag's information which are the government server web address and the univocal code paired with the Tag (which is necessary to identificate the type of payment). Now, combining registration information and information stored in the Tag, the application running in the phone will open a GPRS/UMTS connection with the government server, identifying citizen through his phone number.

VI. CONCLUSIONS

We saw how NFC technology provides services to facilitate and enhance user's experience. For example, users could buy and validate public transport tickets combining NFC technology with SMS connectivity (NFCTicketing) or with Internet Mobile capabilities (Touch'n'PAy). NFC can also be used to securely exchange small amount of data (personal user data, credit/debit card information).

At the moment, a limit for a wide application of NFC comes from the lack of a complete standardization. This is what the StoLPaN project aim to overpass, building a new Ecosystem, basis for new developments.

In particular, the StoLPaN project's objectives are:

1. to develop a JAVA based mobile host controller interface that provides a transparent environment for the simultaneous operation of various NFC based service applications, by neutralizing specifics of the handset design and taking care of resource, security and communication management;
2. to establish the back-office architecture and necessary communication protocols that ensure the secure, remote management of the various NFC applications hosted in the mobile handset. Porting of selected contactless applications to the StoLPaN specification, and preparation of workflow like design guidelines for the development of new service profiles;
3. development of support devices, and reorganization of traditional business procedures to allow full NFC support for the retail check-out and payment process;
4. business models and standards supporting the management and operation of the various NFC use-cases.

VII. ACKNOWLEDGMENT

The authors would like to thank their partners in the framework of the IST-FP6 project StoLPaN (Store Logistics and Payment with NFC) and Francesco Visconti for the useful discussions and support.

REFERENCES

[1] William Webb, "Wireless Communications: The Future", ISBN 0470033126, John Wiley & Sons, 2007.

[2] NFC Forum, *www.nfc-forum.org*.

[3] ABI Research, "Mobile and Contactless Commerce Forecasts", (Q4, 2007).

[4] GSMA, "Pay-Buy-Mobile, Business Opportunity Analysis", *Public White Paper*, Version 1.0, November 2007.

[5] StoLPaN Project, "State of the Art NFC Technology", *Public Deliverable*, 2007.

[6] Anokwa Y., Borriello G., Pering T. Want R., "A User Interaction Model for NFC Enabled Application", *Fifth Annual International Conference on Pervasive Computing and Communications Workshops (PerComW'07)*, 2007.

[7] Zoe A., Srikant V., "Intuitive Mobile User Interaction in Smart Spaces via NFC-Enhanced Devices", icwmc, p. 86, *Third International Conference on Wireless and Mobile Communications (ICWMC'07)*, 2007.

RFID Security

Gianluigi Me, Giuseppe F. Italiano, *DISP*, Università degli Studi di Roma “Tor Vergata”

Abstract—**The RFID adoption is widely spreading among different application fields. Unfortunately, early stages of this technology posed concerns on security and privacy. This fact resembles some unsecurity practices of the past, known in the literature, as the WLAN early stages in the market: the object lessons seems to have no impact on the RFID manufacturers. In this paper, the authors review the attacks to RFID cryptography, identifying some frequent practices and approaches, i.e. security through obscurity, which represent the ante fact of security and privacy vulnerabilities.**

Index terms-- **RFID, security, Tags, authentication, threat**

I. Introduction

The radio frequency identification (RFID), dubbed as one of the "Five Disruptive Technologies to Watch in 2007" by Information Week [15], is expected to become an important and ubiquitous infrastructure technology: in fact, in 2005, the growing market for RFID technologies reached $1.94 billion, while, by 2015, it is predicted to reach $24.5 billion [14].

The RFID tags main use is tracking objects in supply chains (more than $1845 trillion in transactions annually [19]), and are working their way into the pockets, belongings and even the bodies of consumers. Further typical application domains are

- shipping containers, carry $185,000 worth of cargo on average (up to $2 million to $3 million each [16]);
- Pharmaceuticals; e.g. the Trizivir, a three-in-one HIV drug from GlaxoSmithKline (one of the 32 most commonly counterfeited and diverted drugs [17]), costs $825 per month [18];

Because both opaque and contactless communication, there are several security and privacy risks: e.g., the RFID circuits might be read by their corresponding readers without the authorization of the owner of the tagged item. Communication can also be eavesdropped and interfered by interested and malicious attackers, acting as man in the middle, or simply spoofing messages to RFID.

The main threats include

- physical theft, e.g., where the thieves, in 20 minutes long operations, stole twice the soccer star David Beckham's cars in six months, breaking cryptography via wireless laptop when the car was parked [20];
- identity theft, e.g skimming RFID credit cards. The payment card issuers in the United States have begun mass deployment of radio frequency enabled payment cards. The paper in [21] shows a few examples of this new class of payment card and observe that while the card issuers have implemented some new security features, all of them admit practical attacks to a greater or lesser degree.

An overall RFID security and privacy assessment can be found in [4] and [5].

In order to identify the correct approach to secure the RFID technology, we will briefly overview the most concerning cryptographic flaws pointing out some similarities with known attacks developed in the early stages of the introduction of the 802.11 WEP standard.

II. The object lesson

Discovering a vulnerability in a dual-use technology[1], can lead to two scenarios:

1. alerting the manufacturer to fix the vulnerability, thereby protecting both the attackers and the defenders.

2. undisclosing the vulnerability, thereby leaving the defenders insecure but also leaving the attackers insecure.

This choice, called the *equities issue*, e.g., forced, after 9/11, the NSA to turn back to approach 2 from approach 1, achieved after for very long time debate.

The question we would like to address in this paper is related to a third real side of the vulnerability management approach, not included in the equities issue defined above: *unawareness of the vulnerabilities (with respect to the known attacks in the literature) of the devices on the shelf, thereby leaving the defenders insecure but also leaving the attackers insecure(2).*

As a real-world example, tag cloning (unauthorized copies of legitimate RFID tags) occurred when researchers from Johns Hopkins University and RSA Security cloned a cryptographically protected Texas Instruments (TI) Digital Signature Transponder (DST), which they used to buy

Dipartimento di Informatica, Sistemi e Produzione (DISP),
Università degli Studi di Roma “Tor Vergata”, Via del Politecnico 1, Roma, Italy

[1] A technology used both for attack and defense

gasoline and unlock a DST–based car immobilizer. The paper in [3] presents a reverse engineering attack to the TI proprietary cipher, using an array of 16 FPGAs to crack the 40-bit cryptographic keys on DST tags in just under an hour. These keys were then used to create cloned DST tags, which allowed the group to disable a vehicle immobilizer in a 2005 Ford automobile and to purchase gas at various Exxon–Mobil locations. Moreover, this experience confirms, as well the MiFare Crypto 1 (a lightweight stream cipher used in London's Oyster card, Netherland's OV-Chipcard, US Boston's CharlieCard, and in numerous wireless access control and ticketing systems worldwide) [22], the ineffectiveness of the security through obscurity paradigm. The object lessons could be enriched with the not publicly scrutinized algorithms adopted by Philips, namely HITAG 2 (48-bit secret key, vulnerable to a brute-force attack, not appropriated for long-term security) and SECT (128-bit key).

In fact, these attacks could be easily linked to some security facts in the history. Firstly, the GSM A+ algorithms adoption was under the Security Through Obscurity paradigm, whose effectiveness has been shown to be merely poor, since it was reverse engineered/published on BBS. The STO paradigm can be carefully taken into consideration only when the lifetime of the information to protect is taken into account together with the time-to-break of the encryption algorithm. As a rule of thumb, the "due diligence" Kerchoff principle should rule, stating that the algorithm should be public and the keys private. Furthermore, this experience resembles the 802.11 case: in fact, the inadequate choice of algorithms (or their implementations) leads to vulnerabilities, as for the security suites adopted by manufacturers in the early releases of the 802.11b devices.

In particular, the adoption of :

– RC4 without any specification regarding the Initialization Vector, which value is too short and not protected from reuse;

– The small keys length (40 or 104 bit) and the way they are constructed from the IV;

– The CRC-32 as integrity function, whilst is a mere *linear* parity checking . This led to the bit flipping attack. The CRC-32 is not an integrity checking function: for this reason, the adoption of this function as Message integrity check is not adequate[2].

flawed the early WLAN networks (802.11b), seriously threatening the widespread adoption in that period. The RFID attack cases presented above strictly suffer of the vulnerabilities related to 802.11 (challenge length, key length, not adequate choice of algorithm/function). In the next section, we will present the Keeloq case, the most recently attacked with a side channel attack; as a very long attack experienced algorithm before its adoption by Microchip, we believe it could represent a clever Object Lesson.

III. THE KEELOQ CASE

The Keeloq proprietary block cipher was designed in the mid of 80s and sold to Microchip in the mid of 90s. The encryption schema is depicted in Figure 1. The KeeLoq is a block cipher with a 64-bit key and 32-bit plaintext and ciphertext with a non-linear feedback shift register (NLFSR) where the feedback depends linearly on two register bits, one key bit, and a non-linear function (NLF). The NLF maps five other register bits to a single bit. KeeLoq is used in two protocols, the "Code Hopping" and the "Identify Friend or Foe (IFF)" protocol [6]. In practice, the latter protocol, a simple challenge response protocol, is the most interesting target to acquire the data that is necessary to mount the attack. Because the challenges are not authenticated in any way, an attacker can obtain as many chosen plaintext/ciphertext pairs as needed from a transponder (e.g., a car key) implementing this protocol. The KeeLoq encryption algorithm is widely used for security relevant applications (passive RFID transponders for car immobilizers and in Remote Keyless Entry (RKE)) systems, e.g., for opening car doors and garage doors. Due to the message format, 4 bit are reserved to use the same reader to interact with up to 8 tags: this means that a single attack could open, e.g., both the garage and the car. The KeeLoq cipher is used as RKE by Chrysler, Daewoo, Fiat, General Motors, Honda, Toyota, Volvo, Volkswagen, and Jaguar.

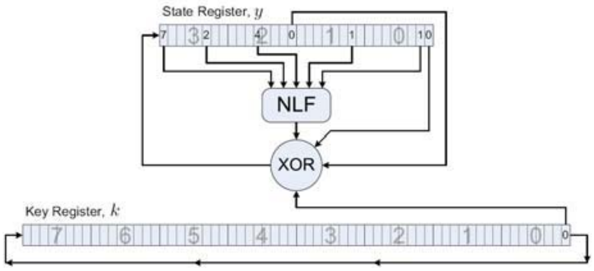
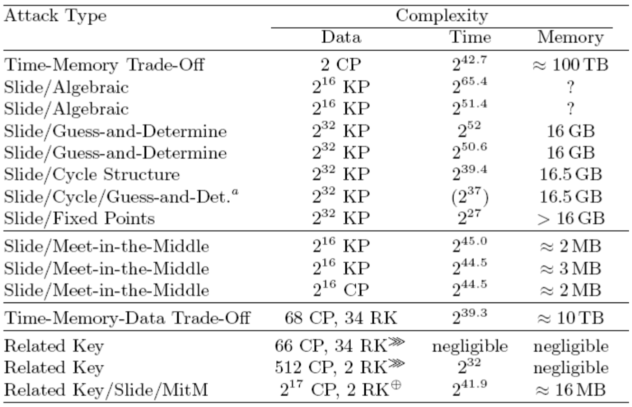


Figure 1:Keeloq encryption

The long Keeloq life collected many attacks. In fact, as shown in Table 1 (details of these attacks can be found in [7], [1], [2], [9] and [23])

[2] In order to check integrity of RC4 is mandatory to use a key generated specifically for integrity checking.



Attack Type	Complexity		
	Data	Time	Memory
Time-Memory Trade-Off	2 CP	$2^{42.7}$	$\approx$ 100 TB
Slide/Algebraic	2^{16} KP	$2^{65.4}$	?
Slide/Algebraic	2^{16} KP	$2^{51.4}$	?
Slide/Guess-and-Determine	2^{32} KP	2^{52}	16 GB
Slide/Guess-and-Determine	2^{32} KP	$2^{50.6}$	16 GB
Slide/Cycle Structure	2^{32} KP	$2^{39.4}$	16.5 GB
Slide/Cycle/Guess-and-Det.^{a}	2^{32} KP	(2^{37})	16.5 GB
Slide/Fixed Points	2^{32} KP	2^{27}	> 16 GB
Slide/Meet-in-the-Middle	2^{16} KP	$2^{45.0}$	$\approx$ 2 MB
Slide/Meet-in-the-Middle	2^{16} KP	$2^{44.5}$	$\approx$ 3 MB
Slide/Meet-in-the-Middle	2^{16} CP	$2^{44.5}$	$\approx$ 2 MB
Time-Memory-Data Trade-Off	68 CP, 34 RK	$2^{39.3}$	$\approx$ 10 TB
Related Key	66 CP, 34 $RK^{\ggg}$	negligible	negligible
Related Key	512 CP, 2 $RK^{\ggg}$	2^{32}	negligible
Related Key/Slide/MitM	2^{17} CP, 2 $RK^{\oplus}$	$2^{41.9}$	$\approx$ 16 MB

Table 1:Keeloq attacks as reported in [1][3]

currently fifteen attacks (mapped on then different techniques) are in the public domain, starting from the Time –Memory Trade Off attack, published in 1980 [10].

In order to provide a straight measure of practicability of the attacks depicted in Table 1, the Slide/Meet in the Middle attack with 2^{16} KP and time complexity $2^{44.5}$, presented in [1], needs a total running time of roughly 500 days. As the attack is fully parallelizable, given *n* CPU cores, the total running time is 500/*n* days. The same attack, requiring 2^{16} CP, needs only 218/*n* days on n CPU cores. For example, as reported by the authors of [1], for 10 000 euro, one can obtain 50 dual core computers, which will take about two days to find the key.

Finally, the most relevant example of the *third approach* of the equities issue is the DPA (Differential Power Analysis) attack, presented in [2]. In fact, in order to successfully performing a DPA attack, some intermediate value of the cipher has to be identified that

a. depends on known data (like the plaintext or the ciphertext);
b. depends on the key bits and
c. is easy to predict.

Considering the generic transformation (Eq. 1) performed by Keeloq block cipher,

$$y_{31}^{(i+1)} = k_0^{(i)} \oplus y_{16}^{(i)} \oplus y_0^{(i)} \oplus NLF(y_{31}^{(i)}, y_{26}^{(i)}, y_{20}^{(i)}, y_9^{(i)}, y_1^{(i)})$$

with trivial manipulations it's possible to write the Eq.2

$$y_0^{(527)} = k_{15}^{(i)} \oplus y_{16}^{(527)} \oplus y_{31}^{(528)} \oplus NLF(y_{31}^{(527)}, y_{26}^{(527)}, y_{20}^{(527)}, y_9^{(527)}, y_1^{(527)})$$

which releases the 15th bit of the key (at round 527), so the whole state depends from k_{15}. This is the case *b*.

Since the DPA attack was presented in 1999 with [8], the underlying techniques are well known between cryptographers by almost ten years, it's trivial to argue that this vulnerability (and the related attack) could be avoided simply checking the Keeloq cipher for the known attacks in the literature.

IV. EXTERNALITIES FOR THE COMPANY INFORMATION SYSTEMS

The vulnerabilities presented in the previous paragraph can lead to theft (e.g. cars) or fraud (e.g. paytoll). The concerns quickly move on Information Systems, where an RFID subsystem can represent one of the gates to enter, as could happen, e.g., for logistics.

Some of the most popular attacks are due to:

– Malware: attackers can use RFID malware, in the shape of RFID exploits, RFID worms, and RFID viruses. RFID exploits can be, e.g., buffer overflows, code insertion, and SQL injection attacks. RFID worms and viruses can be, e.g., RFID exploits that copy the original exploit to newly appearing RFID tags. As shown in [11], attackers could also adopt RFID malware as a tool for identity theft or use RFID tags as a low-risk way to install remote malware, to harvest personal/financial data from the back-end database or for the purposes of extortion or to cripple competition (cryptoviruses, [12]) . The attacker could encrypt data in the victim's database (e.g., using an RFID worm) and then send a ransom note in exchange for the key enabling decryption of the "kidnapped" data.

– "RFID wardriving" (as for WiFi wardriving): attackers wander the streets looking for exploitable RFID readers.

– Spamming (i.e., sending unsolicited emails): RFID tags could be enlisted for "spamming" purposes. For example, an EPC Gen2 tag could have a bogus URI pointing to a banner instead of an Object Name Service server[4] [13], thus earning revenue for the spammer for each tag read.

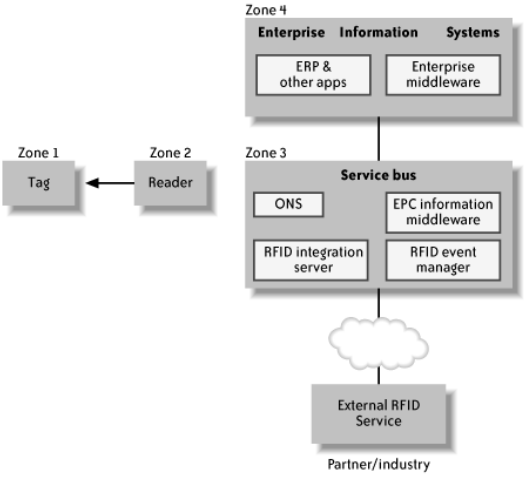


Figure 2:RFID system block system

[3] LEGENDA: Time complexities are expressed in full KeeLoq encryptions (528 rounds).KP: known plaintexts; CP: chosen plaintexts; $RK^{>>>}$: related keys (by rotation); $RK^{\oplus}$: related keys (flip LSB)

[4] The ONS is a subset of the Domain Name System (DNS)

V. CONCLUSIONS

The Return on Security Investment (ROSI) is, beyond doubt, a controversial topic so that defining the value of security investments is definitively challenging.

The cases presented in this paper suggest that hardware technology manufacturers (RFID, currently), as others in the past, consider object lessons weaker than time to market and price. As noted in [3] the causes are mainly related to the substantially increased manufacturing cost (with possible impacts on the overall system architecture due to increased power consumption) and the backwards compatibility. Furthermore, we believe that their approach is more driven by the market demand, which is well known to be based on the prospect theory[5]: in the early stages of the market, *when security is not yet a problem*, a ten-cent crypto-RFID tag is considered less desirable for the market than a five cent non-crypto tag. If security and privacy concerns arise because of losses, or when the potential losses are clear in advance to the customer, due to the vulnerabilities disclosure, the market will demand security, thus accepting to pay more for security enriched devices. It remains not clear for the authors why to adopt the third approach of equity issue.

For these reasons, the time to market and the security scarce consideration in design phase can be considered as the actual adversaries when deploying a technology whose security has not yet reached a maturity level. The novel applications, such as Zero Internet, implantable RFIDs, etc need a strong preliminary security analysis prior to the deployment, in order to avoid security concerns and privacy leakage.

When the RFID technology will absorb the object lesson, releasing secure tags as commodities on the market, and the RFID market target will spread on consumers devices, the security holes will move to user behaviour, arising the problem: ”How to secure the RFID “unaware” usage?”. We hope that we should not witness to new, multiple kind of frauds, mainly based on the user unawareness, promptly warned by the early Schneier’s sentence: ”Only amateurs attack machines, professionals target people”.

REFERENCES

[1]. S. Indesteege, N. Keller, O. Dunkelman, E. Biham, and B. Preneel. A Practical Attack on KeeLoq. In Advances in Cryptology - EUROCRYPT 2008, Lecture Notes in Computer Science. Springer, 2008. to appear;

[2]. T. Eisenbarth, T. Kasper, A. Moradi, C. Paar, M. Salmasizadeh and M. T. Manzuri Shalmani, Physical Cryptanalysis of KeeLoq Code Hopping Applications, eprint.iacr.org/2008/058 ;

[3]. S. C. Bono, M. Green, A. Stubblefield, A. Juels, A. D. Rubin, and M. Szydlo, "Security Analysis of a Cryptographically-Enabled RFID Device," in Proc. 14th USENIX Security Symposium, 2005;

[4]. A. Juels, *RFID Security and Privacy: A Research Survey*, Condensed version to appear in 2006 in the IEEE Journal on Selected Areas in Communication, 2006.

[5]. S. A. Weis, S. E. Sarma, R. L. Rivest, and D. W. Engels, *Security and Privacy Aspects of Low-Cost Radio Frequency Identification Systems*, Security in Pervasive Computing, 2003.

[6]. Microchip. *An Introduction to KeeLoq Code Hopping*. Available in http://ww1.microchip.com/downloads/en/AppNotes/91002a.pdf.

[7]. A. Bogdanov. Attacks on the KeeLoq Block Cipher and Authentication Systems. In 3rd Conference on RFID Security 2007 (RFIDSec 2007), 2007. http://www.crypto.rub.de/imperia/md/content/texte/publications/conferences/keeloq_rfidsec2007.pdf.

[8]. P. C. Kocher, J. Jaffe, and B. Jun. Differential power analysis. In CRYPTO ’99: Proceedings of the 19th Annual International Cryptology Conference on Advances in Cryptology, pages 388–397, London, UK, 1999. Springer-Verlag.

[9]. N. T. Courtois, G. V. Bard, and D. Wagner. Algebraic and Slide Attacks on KeeLoq. In Fast Software Encryption -FSE 2008, Lecture Notes in Computer Science. Springer, 2008. to appear. Also available in http://eprint.iacr.org/ 2007/062.

[10]. Hellman M.E., A Cryptanalitic Time-Memory Trade-Off, IEEE Transactions on. Information Theory IT-26, pp.401-406, July 1980.

[11]. M. Rieback, B. Crispo, and A. Tanenbaum, "Is Your Cat Infected with a Computer Virus?" in Proceedings of the IEEE Pervasive Computing and Communications, 169-179 (Pisa, Italy: Mar. 2006), http://www.rfidvirus.org/papers/percom.06.pdf

[12]. A. Young and M. Yung, Malicious Cryptography: Exposing Cryptovirology (John Wiley & Sons, 2004).

[13]. B. Fabian, S. Spiekermann, Security Analysis of the Object Name Service (ONS) for RFID, http://www.taucis.hu-berlin.de/_download/security_analysis.pdf

[14]. R. Das and P. Harrop, "RFID Forecasts, Players, and Opportunities 2006-2016," IDTechEx, 2006, http://www.idtechex.com/products/en/view.asp?productcategoryid=93

[15]. D. Strom, "5 Disruptive Technologies to Watch in 2007," InformationWeek, Jan. 2007, http://www.informationweek.com/news/showArticle.jhtml?articleID=196800208 ;

[16]. S. Lowe, "Are Container Shippers and Consignees Cutting Cost Corners to Sacrifice the Security and the Safety of the Citizens of USA and Europe?" Directions Magazine, May 2005, http://www.directionsmag.com/press.releases/index.php?duty=Show&id=11727 ;

[17]. M. O'Connor, "Glaxosmithkline Tests RFID on HIV Drug," RFID Journal, 2006, http://www.rfidjournal.com/article/articleview/2219/1/1/ ;

[18]. Anonymous, "FDA Approves Three-in-One HIV Therapy," Drug Store News, 2000, http://www.findarticles.com/p/articles/mi_m3374/is_19_22/ai_68876802 ;

[19]. M. Vargas, "2002 Retail Security Survey Shows U.S. Retails Losing $31 Billion to Theft," About.com, 2002, http://retailindustry.about.com/od/statistics_loss_prevention/l/aa021126a.htm ;

[20]. Anonymous, "Gone in 20 Minutes: Using Laptops to Steal Cars," Left Lane News, 2006,

[5] The prospect theory explains how people approach risk. People tend to be risk averse when it comes to gains, and risk seeking when it comes to losses.

http://www.leftlanenews.com/2006/05/03/gone-in-20-minutes-using-laptops-to-steal-cars/.

[21]. Thomas S. Heydt-Benjamin, D. V. Bailey, K. Fu1, A. Juels, and T. O'Hare, RFID Payment Card Vulnerabilities, Technical Report, http://www.nytimes.com/packages/pdf/business/20061023_CARD/techreport.pdf

[22]. K. Nohl, Cryptanalysis of Crypto-1, www.cs.virginia.edu/~kn5f/pdf/Mifare.Cryptanalysis.pdf, 2007;

[23]. Biryukov, A., Mukhopadhyay, S., Sarkar, P.: Improved Time-Memory Tradeoffs with Multiple Data. In: Preneel, B., Tavares, S. (eds.) SAC 2005. LNCS, vol. 3897, pp. 245–260. Springer, Heidelberg (2006)

RFID for Agri-Food Traceability: Methods for Authentication, Integrity and Privacy

C. Demartini F. Gandino B. Montrucchio M. Rebaudengo E.R. Sanchez
Politecnico di Torino, Dipartimento di Automatica e Informatica, Torino, Italy
E-mail: {claudio.demartini, filippo.gandino, bartolomeo.montrucchio, maurizio.rebaudengo, erwing.sanchezsanchez}@polito.it

Abstract[1]

An agri-food traceability system based on public key cryptography and Radio Frequency Identification (RFID) technology is proposed. In order to guarantee safety in food, an efficient tracking and tracing system is required. RFID devices allow recording all useful information for traceability directly on the commodity. The same data used by the traceability system can be used for the supply chain management. The proposed system allows authorized operators to access data recorded on the commodity RFID tag, and it allows competent authorities to directly access all the information on the history of commodities, furthermore a security system avoids industrial espionage, and it protects customer privacy. The security issues are discussed and a method based on public cryptography is proposed and evaluated. The security algorithm uses a nested RSA based structure to improve security. An experimental analysis demonstrated that the proposed system is well suitable on PDAs too.

1. Introduction

Traceability can be considered a key factor in agri-food sector. Improving tracking and tracing without loosing data privacy is requested both by laws and consumer organizations. In several countries laws on traceability have been made during last years:

- in the USA, "Farm Security and Rural Investment Act" requires country of origin labeling for many kinds of food, including perishable agricultural commodities [1];
- in EU, Regulation (EC) No 178/2002 of The European Parliament And of The Council of 28 January 2002 [2] establishes that food business operators shall be able to identify, for the competent authorities, any person who supplied them with alimentary commodities, and any business which takes food from them; they shall, also, label adequately food, in order to facilitate the traceability.

Agri-food companies often apply simple systems, based on paper documents. Some systems exploit barcode to identify commodities: by using the identification number in the barcode, it is possible to find, in the company database, the information about the food. Today, new opportunities for the food traceability come from the Radio Frequency Identification (RFID) technology.

RFID is widely adopted as a contactless identification technology. A typical RFID system is made up of: a reader, which creates an electromagnetic field, and passive tags without an own voltage supply. They can be read only if they are in the interrogation zone of a reader which supplies the power required through a coupling unit. Today, the size of the RFID tag memory allows recording directly on every commodity all useful information for the competent authorities to trace it.

The use of RFID tags both hazards the customer privacy and exposes companies to espionage threats. In the USA, many organizations, such as Consumer Privacy and Civil Liberties Organizations, are requesting attention to privacy threats [3]. In Canada, the Annual Report to Parliament 2005 of the Privacy Commissioner underlines the importance to ensure that RFIDs do not erode informational privacy rights [4]. In EU, in compliance with the Working Document adopted on 2005 by the European Data Protection Working Party [5], the national authorities, set up to protect personal information, established guidelines needed for a safe use of RFID technology [6].

The privacy threads, arose from RFID, involve dangers such as man tracking, personal belongings monitoring and industrial espionage.

Many solutions to the privacy problem have been analyzed, some of them are:

- *killing the tag* [7], a command can stop the tag at the point-of-sale.

[1] This work was partially supported by "Progetto Regionale Ricerca Applicata 2004".

- using *passwords* or *encryption* [8], which try to avoid unauthorized readings of the tag;
- *changing tag ID* [9], the use of different IDs makes difficult to recognize a tag;
- *blocking the anti-collision system of the reader* [10], a special tag stops the correct functioning of the reader.

This paper proposes a system which allows competent authorities to manage alimentary traceability, preventing new privacy problems, and it allows operators of the chain to exploit supply chain management by using RFID tags, preventing industrial espionage. In this system, food business operators shall record on the RFID tag information on their treatments, in compliance with one precise outline. The present size of the tag memory allows using the whole memory for traceability, or leaving a part for other independent aims, such as anti-counterfeit [11] or marketing. Stored data will be protected using the public key cryptography: every operator will record its treatments and only the competent authorities and authorized operators, using the appropriate keys, will be able to decrypt the information. In this way, by means of the resulting ubiquitous data system, authorities could immediately access information on alimentary commodities under examination. The use of encryption allows protecting the memory area of the traceability system, without blocking the memory; it is, moreover, possible to use additional privacy protection systems, in order to ensure the privacy of the whole tag. To improve the security level we propose an algorithms based on the Nested Cryptography Algorithm [12]: Nested Supply Chain Cryptography Algorithm (NSCCA), that uses encapsulated ciphertexts in order to enhance the security optimizing memory occupation.

The remaining of the paper is organized as follows: in Section 2 background about traceability management, industrial espionage, privacy threats, RFID characteristics and cryptography theory are introduced, while in Section 3 the traceability management system and the privacy protection system are detailed. Finally, in Section 4, system abilities and costs are evaluated. In Section 5 some conclusions are drawn.

2. Background

Within this section the description of privacy and traceability management goes into more depth. Tags properties are depicted, spotlighting different nomenclature and current organization. In addition, information theory for cryptosystems is introduced.

2.2. Traceability management

Rules about traceability and food label information change according to the country. According to [2], food business operators shall register the origin and the destination of the alimentary commodities they manage, and they shall label food to facilitate its traceability. In general, alimentary operators shall track the food to allow its tracing. A typical case of food tracking management is shown in Fig. 1:

- a *producer*, yields a commodity;
- a *second operator* buys the commodity, registers the producer data, transforms the commodity or joins it to other commodities and registers its treatments;
- a *distributor* buys the commodity and registers the previous operator data;
- a *retailer* buys the commodity and registers the distributor data.

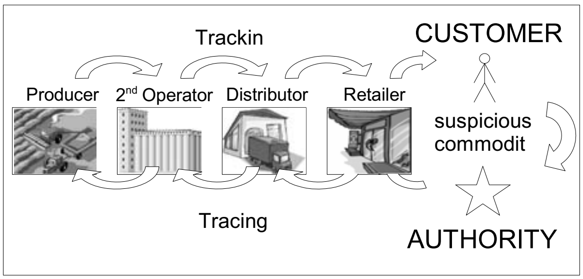


Figure 1. Agri-Food Tracking and Tracing

Whenever there are alimentary sophistications, contamination or infection caused by damaged food, the competent authorities control the retailer which sold them; the operator must search in its own centralized database to make available the information about its treatments and to identify any person who supplied it with food or any other substance included into the commodity. Then, authorities repeat the procedure with the next operator, and so forth. By using RFID tags to label alimentary commodities, every operator could write a copy of its data and of any other useful information directly on the tag, transforming the previous divided databases in only one ubiquitous database, and making the authorities' work easier and faster

2.3. Privacy Threats

Rules about privacy change according to the country, as well. However in many countries there is a great attention on privacy risks. There are many privacy threats connected to RFID [13][14]:

- The serial number of a tag can be associated with the customer's identity, so it is possible to monitor the customer or, knowing the object identified by the serial number, to get information for profiling. Besides knowing which object a person buys, it is possible to know how often a person uses it as well.

- Even without associating a tag number with a person identity, a set of tags can track an unidentified person, violating the “location privacy” [15].
- The transfer of a tag from a set to another set means that an object passes from a person to another one, so it is possible to know that there is a relation between those persons.
- By reading the tag’s memory, it could be possible to know which commodities a person possesses.
- Companies would like to keep private their information, in order to avoid industrial espionage and unauthorized monitoring of their sales.

Privacy threats, due to recording of the tracking information on an RFID tag, are mainly the risk of unauthorized readings of information about the belongings of a person, and the industrial espionage. In this paper a solution to these problems is proposed.

2.4. RFID Tag Properties and Organization

A tag is composed by a radio frequency interface block, a memory component and a logic element. Tags have usually no battery (*passive* ones), so they acquire the power from the external radio frequency communication. Otherwise, a*ctive* tags have their own power supply. Commonly, computational capacities are extremely limited in a tag. The major concern of an RFID reader consists in accessing the tag’s memory. Memory, which plays an important role in the tag architecture, may be a ROM or an EEPROM memory. It contains the unique identification number and may have up to several kilobits of storage capacity. Operational frequency used in an RFID system may vary from low frequencies (several kilohertz) to ultra high frequencies (a couple of gigahertz).

Despite the fact that some RFID tags are able to perform cryptographic operations [16][17][18] because of their internal logic circuitry, the majority of RFID devices have not real capabilities for cryptanalysis functions in part due to their power constraints. Most of RFID tags are passive ones, with limited processor performance and, hence, restricted computational resources.

While first generation tags did not even have memory for an identification number, current versions may have several kilobits for user memory. Our proposed traceability system is aimed for simple-passive tags with user memory.

2.5. Cryptographic Theory

Cryptographic algorithms have been used for decades in order to guarantee communication privacy. The proposed privacy system uses RSA [19] algorithm, that is based on public key cryptography, firstly presented in [20]. Many other applicable algorithms based on public-key cryptography have been proposed in the literature: El Gamal scheme [21], Knapsack scheme [22], Rabin scheme [23].

In a public key cryptosystem, given a pair of families $\{E_K\}_{K\in\{K\}}$ and $\{D_K\}_{K\in\{K\}}$ of algorithms representing inverting transformations, E_K:$\{M\} \rightarrow \{M\}$ and D_K:$\{M\} \rightarrow \{M\}$, on a finite message space {M}, the following must be true:

- for every $K \in \{K\}$, E_K is the inverse of D_K,
- for every $K \in \{K\}$ and $M \in \{M\}$, algorithms E_K and D_K are easy to compute,
- for almost every $K \in \{K\}$, each algorithm equivalent to D_K is computationally infeasible to derive from E_K,
- for every $K \in \{K\}$, it is feasible to compute inverse pairs E_K and D_K from K.

Therefore, by making $K = Ko$, a pair of ciphering functions D_{Ko} and E_{Ko} are fixed. The third property allows making public the key E_{Ko} without compromising the security of the secret key D_{Ko}. In this way, a *plaintext* message $P \in \{M\}$, may be ciphered by means of the public key. The result is a *ciphertext* message $C \in \{M\}$ that can be deciphered using the secret key. Thus, the following relation is true, $C = E_{Ko}(P) = E_{Ko}(D_{Ko}(C))$.

Secret and public keys are generated by means of the RSA algorithm as follows. Two large prime numbers n and p are chosen. The number of elements q in $GF(q)$ is computed by multiplying n and p. A random value E, relatively prime to $(n - 1)(p - 1)$, is picked. Subsequently, the number D is calculated $D=[k(n - 1)(p - 1)+1]/E$, with k chosen in order to make D an integer number. Private algorithm is defined as

$$D_{Ko}(P) = P^D \bmod q = C, \qquad (1)$$

and the public algorithm as

$$E_{Ko}(C) = C^E \bmod q = P. \qquad (2)$$

To avoid risks from chosen plaintext attacks and chosen ciphertext attacks, RSA is normally combined with a padding scheme, such as OAEP [24]. The OAEP processes the plaintext prior to encryption in order to convert the RSA deterministic encryption scheme in a probabilistic scheme, and to prevent partial decryption of the plaintext.

The properties of the public key cryptosystem can be obtained exploiting the apparent difficulty of computing logarithms over a finite *Galois Field* with a number q of elements. Security is measured accordingly with the computational complexity of calculating the logarithmic operation. While it is widely believed that breaking the RSA encryption scheme is as difficult as factoring the modulus q, no such equivalence has proven [25].

While enlarging *q* improves system security, it also places constraints within computational time. The time required to calculate ciphering and deciphering functions is augmented mainly because of the size of the numeric values involved in the computation. Normally, a reasonable value for *q* should be on the order of 2^{1024}. Considering that regular bit length for numerical values is, at most, 64 bits in a computing system, appropriate algorithms should be used to manage 1024 bit or bigger values.

3. Traceability Management System

RFID tags could be defined as an unsecured channel, since they are a means of conveying information that intruders have the ability to read. In our system every operator in the agri-food chain has to write information about its treatments in a specific area of the commodity tag. Unfortunately, unauthorized persons can read tag information to know which kind of commodities an individual owns or to spy a competitor. As to address the privacy needs of a system, unauthorized readings of the tag memory should be forbidden. The traceability management and security system are described in the following.

3.2. General Architecture

At the present time, in order to find the operators that treated a commodity, the authorities have to follow a trail of breadcrumbs. They find the first operator and then they have to trace back, step by step, in order to detect any other.

In order to make easier authorities' work, in [12] was proposed to create a ubiquitous tracking database, by labeling the alimentary commodities with an RFID tag. We propose to create an ubiquitous database, that is used both by authority traceability system and by supply chain management. Every operator of the chain controls a part of the tag memory (*memory slot*) and it has to record its own data and its treatments information on it. Each memory slot is divided in two parts, in the first one the operator writes the data reserved to authorities, in the second one the operator writes data that will be used for supply chain management. In this way all the traceability information are immediately available to the competent authorities, and the operators can access to the information useful for supply chain management.

The tag memory is divided, at logic level, in a sufficient number of areas, to allow a sufficient number of operators to write. On the other hand, the size of a memory slot, that corresponds to the *Maximum Allowed Information Size* (MAIS) of each operator, must be large enough to store all its data. An accurate template is needed to streamline the use of the memory space. In way of employing a smaller memory area than using strings of characters, information must be translated in numerical codes. The use of codes to implement the traceability is under study also by EAN [26]. Codes have to identify operators, their geographic zone, their sector, the kind of commodity and the executed treatment types. The competent authority will fill in a reference table for any kind of code:

- *identification codes (IDC)* reference tables; a group of three tables that identifies the operator:
    - *geographic code (GC)* reference table; the first part of the code identifies the country, the second the region, and the last the municipality; the authorities, by using this code, can immediately identify the origin of a commodity;
    - *sector code (SC)* reference table; the sector code defines the kind of operator, e.g. “farmer” or “distributor”;
    - *operator identification (OID)* reference table; this code identifies the single operator;
- *commodity code (CC)* reference table; this code identifies the kind of commodity; it is useful when a food is made by different elements;
- *treatment code (TC)* reference tables; in every sector a table holds the list of the relevant operations, and their codes.

An operator must also write the IDC of its supplier, in order to enhance system reliability against frauds. Each code written in the memory will start with its identification, so the operator can write the codes in any order, and divide them among the first and the second half of the slot, in accord to its security needs.

In the agri-food chain the commodity follows different steps. Initially the producer stores its data into the first memory slot. Step by step each operator adds its data. The following situation may modify the initial product:

- Simple treatment; the operator adds its data at the bottom of previous information.
- Merge of commodities; if the number of available memory slots is enough, the operator copies the information of all the old tags in the new one. If information regarding the commodities would overfill the memory, it writes only a summary, including a header (the summary special area identifier flag) and the identification codes of suppliers that matched to commodity codes. Then the operator adds its data at the bottom of previous information.
- Partition of a commodity; the operator adds its data at the bottom of previous information, and it tags all the new commodities.

Operators must put in a database the data contained in all tags, in order to be able to prove, in case of an authorities' inspection, their propriety.

3.3. Espionage Protection System

We elaborated a cryptographic algorithm, called the Nested Supply Chain Cryptographic Algorithm (NSCCA), adapted to protect a company from industrial espionage. This algorithm allows reading the data only to the selected members of the production chain and reserves a memory area to protect data available only to competent authorities.

Periodically the competent Authorities establish one Authority Public Key (APuK) and one Authority Private Keys (APrK), and distributes the public key to every operator. At the same time each company establishes a set of Operator Public Keys (OPuKs) and Operator Private Keys (OPrKs) with different lengths, and distributes the public key set to the Authorities and to the operators of the production chain that are supplied by it and that are allowed reading the data.

Every operator encrypts a plaintext that contains the reserved information about the treatments, by using the APuK, and it encrypts a plaintext that contains the information for the other operators by using a public key with a specific length, and it writes the resulting ciphertexts in the appropriate memory areas. The authorities can decrypt all the ciphertexts by using the private keys. The company in the chain can only decrypt the ciphertext of that they have the OPuK. By changing private and public keys periodically, the security increases; in fact, an unauthorized entity which finds some private keys could use them for a short period of time while automatized entities can decrypt old and new ciphertexts.

This system uses pairs of OPuK and OPrK of different length. To understand the benefit of using different key lengths, it is important to remember that enhancing the length of the keys increases security and ciphertext size. Each operator uses a particular OPrK depending on its position in the chronological sequence of the production chain (increasing numbers, e.g., 1 for the farmer, and so on). The MAIS is the same for all operators. The tag memory is, at logical level, divided in slots with this size, the first half of the slot is used for the data reserved to authorities, the second one contains the data for the other operators. The description of the algorithm is shown in the Fig. 2.

The first operator of the chain uses the ApuK to encrypt the reserved part of its data, and it writes the resulting cipthertext (CTA) in the firs half of the first memory slot. Then it uses the first and shortest OPrK, the length of the key is equal to the MAIS; a plain test composed by the remaining information and by its CTA is encrypted by the first OPuK, and the relative ciphertext (CTO) is written in the first memory slot.

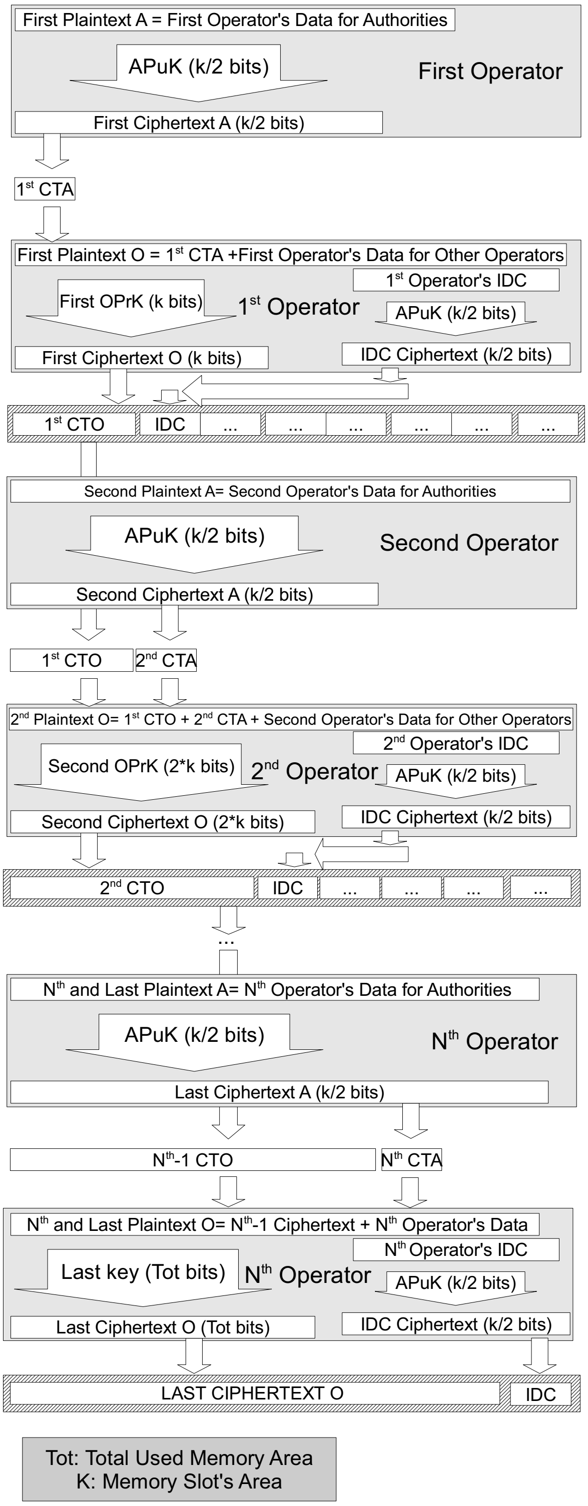

First Plaintext A = First Operator's Data for Authorities
APuK (k/2 bits)
First Operator
First Ciphertext A (k/2 bits)
1st CTA
First Plaintext O = 1st CTA +First Operator's Data for Other Operators
1st Operator's IDC
First OPrK (k bits)
1st Operator
APuK (k/2 bits)
First Ciphertext O (k bits)
IDC Ciphertext (k/2 bits)
1st CTO
IDC
...
Second Plaintext A= Second Operator's Data for Authorities
APuK (k/2 bits)
Second Operator
Second Ciphertext A (k/2 bits)
1st CTO
2nd CTA
2nd Plaintext O= 1st CTO + 2nd CTA + Second Operator's Data for Other Operators
2nd Operator's IDC
Second OPrK (2*k bits)
2nd Operator
APuK (k/2 bits)
Second Ciphertext O (2*k bits)
IDC Ciphertext (k/2 bits)
2nd CTO
IDC
...
Nth and Last Plaintext A= Nth Operator's Data for Authorities
APuK (k/2 bits)
Nth Operator
Last Ciphertext A (k/2 bits)
Nth-1 CTO
Nth CTA
Nth and Last Plaintext O= Nth-1 Ciphertext + Nth Operator's Data
Nth Operator's IDC
Last key (Tot bits)
Nth Operator
APuK (k/2 bits)
Last Ciphertext O (Tot bits)
IDC Ciphertext (k/2 bits)
LAST CIPHERTEXT O
IDC
Tot: Total Used Memory Area
K: Memory Slot's Area

Figure 2. NSCCA Algorithm

The length of any operator OPuK is equal to the MAIS multiplied by the number of the operator position in the chain. All operators, subsequent to the first one, compose the plaintxt for the authorities (PTA) with their reserved information, and they compose their plaintext for the other operators (PTO) adding their CTA and their remainig information to the bottom of the previous CTO. After the encryption, operators write the new CTO in the first part of the memory tag, occupying a number of memory slots equal to the operator position. The last operator, theoretically the retailer, uses always the last and longest key. Its ciphertext occupies all the memory slots, a part the last one.

The last step of every operator is to encrypt its IDC by using the APuK, and to write the resulting text in the first subsequent free memory slot.

Authorities decrypt, one by one, all the ciphertexts. First they decrypt the IDC of the last operator by using the APrK, then by using the correct OPuK they encrypt the last CTO. By using again the APrK they can decrypt the last CTA. Then by using the IDC of the supplier of the last operator, present in the plaintext of the last operator, the authorities can repeat the previous operations for the other chipertext.

At each chain ring the security grows, out of the production chain, the security is to the maximum level. Therefore

This system protects also from frauds by proving the message originality, in fact only the authorized operator can encrypt data by using its OPrK.

4. Experimental Results

We experimentally evaluated the proposed technique implementing a prototype. Initially we filled out part of the code reference tables, sufficient to test the system. The simulation allowed knowing the performance time of the system and the differences among the cryptography algorithms.

To put into operation the system, the authorities need an RFID reader for mobile devices and a PDA with the reading software. The agri-food operators need an RFID reader to write on the tag. A small reader for mobile devices and a PDA with the writing software is enough as well. To increase the efficiency it is possible to use PCs with appropriate readers, instead. We used the following resources:

- RFID tag: SRIX4K from STMicroelectronics, passive tag, compliant with ISO14443, frequency 13.56 MHz, EEPROM with 4 kbits.
- RFID reader: ACG Dual ISO CF Card Reader Module from ACG, compliant with ISO14443, frequency 13.56 Mhz.
- Computing system: PDA with a 624 MHz Intel PXA270 processor.

In the simulation we use the whole memory, of 4096 bits, for the traceability system.

In the proposed algorithm we set the MAIS to 480 bit. There are 8 keys, from 480 bit to 3840.

We implemented the software by using a not optimized implementation of RSA algorithm, so the processing time cannot show the real performance of the system, but it can show the differences when using different key lengths. The authorities' check of a memory slot, encrypted using a 512 bits key, is completed in 3800 ms. Operators employ 500 ms to entirely generate and write their ciphertext. Anyway, by using a PC, with a Pentium 4 at 3.20 GHz processor, the decryption needs 62 ms and the encryption 1 ms; with a 4096 bit key the decryption needs 4125 ms, the encryption 31 ms. The difference between encryption and decryption comes from the use of a very optimized public key. Figures 4 and 5 show the encryption/decryption time. Although, this time table results from the simplicity of the used algorithm implementation; we did not attempt to improve it since its characteristics are not part of this paper objectives.

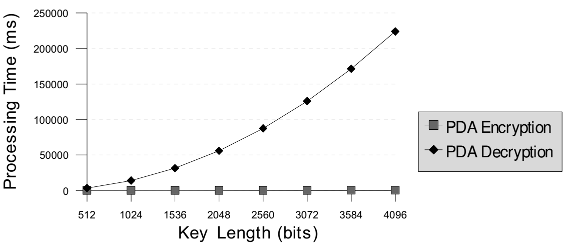


Figure 3. PDA Encryption/Decryption Time

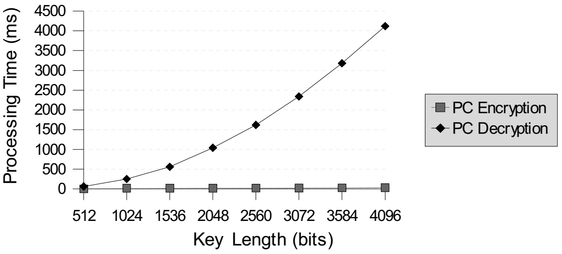


Figure 4. PC Encryption/Decryption Time

5. Conclusion

Today, an efficient management of the traceability is necessary; RFID technology offers the possibility to implement a rapid and effective ubiquitous system. Unfortunately, recording operators and commodity data on a RFID tag involves, in addition to standard RFIDs privacy problems, the risk of unauthorized readings of information about the belongings of a person, and industrial espionage. However, privacy can be protected by using an opportune cryptosystem: the algorithm presented in this paper produces a satisfactory reply to these security and privacy problems.

Even considering the possible optimization of the cryptography algorithm implementation, the decryption time requires the use of a PC, while the encryption can be made simply by a PDA.

In the proposed algorithm it is not possible to lock an area until the subsequent operators have written on the tag. The security algorithm is also an authenticating system, reducing the risk of possible tampering.

Our traceability system, with a suitable RSA implementation, can satisfy security and privacy demands.

References

[1] U.S. Federal Register, "Farm Security and Rural Investment Act of 2002 ", Vol. 68, No. 210, October 30, 2003.

[2] Official Journal of the European Communities, "Regulation (EC) No 178/2002 Of The European Parliament And Of The Council of 28 January 2002", Article 18.

[3] "RFID Position Statement of Consumer Privacy and Civil Liberties Organizations", Privacy Rights Clearinghouse, November 30, 2003.

[4] Privacy Commissioner of Canada, "Annual Report to Parliament 2005 – Report on the Personal Information Protection and Electronic Documents Act", pp. 39-42.

[5] Working Party On The Protection Of Individuals With Regard To The Processing Of Personal Data, "Working document on data protection issues related to RFID technology", *ARTICLE 29 Data Protection Working Party*, January 19, 2005.

[6] Garante per la protezione dei dati personali, ""Smart (RFID) Tags": Safeguards Applying to Their Use", Bollettino del n. 59/March 2005, March 9, 2005.

[7] EPCglobal, *13.56 MHz ISM band class 1 radio frequency (RF) identification tag interface specification*.

[8] Weis, S. A., Sarma, E. S., Rivest, R. L., and Engels, D. W., "Security and Privacy Aspects of Low-Cost Radio Frequency Identification Systems", *Security in Pervasive Computing*, 2003.

[9] A. Juels, "Minimalist Cryptography for RFID Tags," *4th Conf. Security in Comm. Networks* (SCN), C. Blundo and S. Cimato, eds., Springer-Verlag, 2004, pp. 149-164.

[10] A. Juels, R. Rivest, and M. Szydlo, "The Blocker Tag: Selective Blocking of RFID tags for Consumer Privacy," *8th ACM Conf. Computer and Comm. Security*, V. Atluri, ed., ACM Press, 2003, pp. 103–111.

[11] P. Bernardi, et al., "An Anti-Counterfeit Mechanism for the Application Layer in Low-Cost RFID Devices", 3rd IEEE International Conference on Circuits and Systems for Communications, July 2006, pp.207-211.

[12] P. Bernardi, C. Demartini, F. Gandino, B. Montrucchio, M. Rebaudengo, E.R. Sanchez, "Agri-Food Traceability Management using a RFID System with Privacy Protection," *IEEE 21st International Conference on Advanced Information Networking and Applications (AINA-07),* Niagara Falls, Canada, May 21-23, 2007, pp. 68-75

[13] A. Juels, S. Garfinkel, and R. Pappu, "RFID privacy: An overview of problems and proposed solutions," *IEEE Security and Privacy*, vol. 3, no. 3, pp. 34–43, May/Jun. 2005.

[14] A. Juels "RFID security and privacy: A research survey," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 2, pp. 381–394, Feb. 2006.

[15] Alastair Beresford and Frank Stajano. "Location Privacy in Pervasive Computing", *IEEE Pervasive Computing*, 2(1):46–55, 2003.

[16] S. Weis. Security and privacy in radio-frequency identification devices (master thesis), May 2003.

[17] M. Feldhofer, S. Dominikus, and J. Wolkerstorfer, "Strong authentication for RFID systems using the AES algorithm," in *Proc. Workshop on Cryptographic Hardware and Embedded Syst.*, M. Joye and J.-J. Quisquater, Eds. New York: Springer-Verlag, 2004, vol. 3156, Lecture Notes in Computer Science, pp. 357–370.

[18] A. Juels and S.Weis, "Authenticating pervasive devices with human protocols," in *Proc. Advances in Cryptology*. New York: Springer-Verlag, 2005, vol. 3621, Lecture Notes in Computer Science, pp. 293–308.

[19] R. L. Rivest. A. Shamir, and L. Adelman, "A method for obtaining digital signatures and public-key cryptosystems," Commun. ACM, pp. 120-126. Feb. 1978.

[20] W. Diffie and M. E. Hellman, "New directions in cryptography," IEEE Trans. Informat. Theory, vol. IT-22, pp. 644-654, Nov. 1976.

[21] T. El Gamal, "A public-key cryptosystem and a Signature scheme based on Discrete Logarithms," IEEE Trans. on Info. Theory, Vol. IT-31, pp. 469-472, 1985

[22] B. Chor, R. Rivest, "A Knapsack-type public-key cryptosystem based on Arithmetic in Finite Fields," IEEE Trans. on Info. Theory, Vol. IT-34 (5), pp. 901-909, 1988.

[23] M. O. Rabin. "Digitalized signatures and public key functions as intractable as factorization". Techbical Report, MIL/LCS/TR212, MIT Lab. Computer Science, Cambridge, Mass., January 1979.

[24] M. Bellare and P. Rogaway. "Optimal Asymmetric Encryption - How to Encrypt with RSA". In Eurocrypt '94, LNCS 950, pages 92-111. Springer-Verlag, Berlin, 1995.

[25] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone. *Handbook of Applied Cryptography.* CRC Press, 1996.

[26] EAN-UCC *Traceability Implementation*, EAN International, 2003.

Appendix

A Patent View into RFID Technology

Maurizio Ricciardi

***Abstract*— The purpose of this paper is to present an approach based on patents to help understand the technological progress lying behind a given technology, RFId in the present case, and try to derive a trend.**

Patents are defined and the granting procedure before the EPO is outlined. Milestones of RFId technology are then recalled and a selection of applications together with some exemplary patents is presented. The reader is supposed to be familiar with the basic RFId terminology. A technological trend is then extracted with reference to some patents.

A last section including some general remarks about patenting in Italy and a scheme for fostering innovation concludes the paper.

***Index terms*-- RFID, patent, disclosure, technologies, applications, EPO, European Patent Office**

I. What is a Patent

A. *What is a patent*

A patent gives its owner the right to prevent others from commercially using his invention during a period spanning the life of the patent itself, which is 20 years for European patents. The most known legal effect of a patent is the conferment of a monopoly to a person, the patent owner, in the technical field of the invention. The following discussion particularly addresses patents issued by the European Patent Office (henceforth EPO) but most of the considerations described hereafter apply to patents issued by most of the offices worldwide.

From a practical approach, an invention must exhibit a technical character and contribute to the state of the art to be patentable. The object of the invention may fall into different categories, as a product, a product use, a process or an apparatus, or a combination of the previous.

The main criteria for patentability are: novelty, i.e. the invention must not have been previously disclosed; must present an inventive step, i.e. must solve a technical problem in an not obvious way; and must have an industrial application.

B. *Procedure before the EPO*

Life of a patent application before the EPO evolves through four main phases: a filing/reception, a search, an examination and an opposition, the latter being optional depending on observations filed by third parties.

After the filing, the application ends in the receiving section, which carries out a formal check and forwards the application to the search division for performing a prior art search and delivering a search report with a first, non-binding opinion on patentability. The application subsequently enters a substantive examination phase, in which a patent is either granted after all the objections are overcome or refused in case the objections were not set aside. After grant, a patent can be successively be opposed by a third party resulting in a revocation or in the maintenance of the patent in amended form. Every adverse decision can be challenged via an appeal filed before the Board of Appeal of the EPO, which acts as a court of second instance and whose decision is final.

The Legal Board of Appeal and the Technical Boards of Appeal give independent final rulings on appeals against decisions taken during grant and opposition proceedings.

The Enlarged Board of Appeal gives decisions and opinions in order to ensure correct application of the law, or if an important point of law arises.

C. *Benefits of patents*

A comprehensive discussion about the reasons why patents appear to be beneficial for the end user and society as a whole is outside the scope of this paper. Among other aspects, patents help to

1. advance technology from its current state;
2. avoid waste of human and financial resources ;
3. avoid duplication of research in industry and universities;
4. promote creativity and innovation;
5. identify innovative strengths and technological trends, global and regional;
6. identify the lack of advancement in a technology and make improvements.

The author is presently at the European Patent Office, Landsbergerstr. 30, 80339 Munich, Germany.

The previous points can be specialized for a business company or for a university-like research actor. From a company point of view, patents help to

1. retain a market position;
2. increase market share;
3. launch new products and processes;
4. locate new business partners ;
5. provide recognition and motivation for employees;
6. bring together inventors and investors;
7. recover R&D investments and safeguard the results;

whilst for a university entity, a patent consultation[1] helps to

1. avoid duplication of research with industry;
2. identify technological trends and possible developments;
3. foster interaction with industrial partners on common technologies/interests.

In addition to the above-mentioned advantages, EPO offers unitary protection standards in the contracting states allowing the use of a single language to pursue an application in different states, which can be chosen according to the market need.

Another important aspect is that EPO patents offer a strong protection because a thorough search is performed over 60 million patent documents and 5 million non patent documents, followed by an accurate substantive examination, which confers a sound legal protection to every patent granted.

II. A BRIEF HISTORY

The closest parent of a RFID system can be identified as the IFF (identify friend or foe) aircraft transponder used by the allies in World War II. **[quick description]** Transponders are still used by military and commercial aircraft to this day.
Stockman's pioneering work [1] after WWII addressed some relevant themes, the conclusion being somehow sceptical about practical implementation of a back-scattered communication system "…considerable research and development work has to be done before the remaining basic problems in reflected-power communication are solved, and before the field of useful applications is explored."
The user community had to wait almost a quarter of century to see the first patent addressing the problem of radio frequency identification and introduced a passive radio transponder with 16-bit memory, powered by the interrogating signal [2].
On the same year took place a demonstration of a working 12 bit passive and semi-passive RFID tags at 915 MHz at the Los Alamos Scientific Laboratory. The majority of today's UHF and microwave RFID tags still uses this technique.
The first patent to be associated with the abbreviation RFID was granted to Charles Walton in 1983 U.S. Patent 4,384,288 [3].
As of today, almost 12000 patents in EPO database mention RFID in the abstract.

III. APPLICATIONS

Notwithstanding its life spanning more than 30 years, RFID always finds new uses, thus proving itself to be an interesting technology featuring lots of applications. Such proliferation is reflected by the conspicuous number of patents, many of them published in the very last two years, addressing various aspects of an RFID communication system, such as components, tags, or subsystems.

It is outside the scope of this paper to present a full list of applications sustaining RFId technology, which the interested reader can find in [4] or by searching the Internet. The reader is also supposed to be familiar with the basic RFId terminology, a typical system implementation and its technological limitations.

A rather compact selection of prototypical applicative scenarios follows, each scenario presenting a brief list of applications together with some relevant patents.

A. Transportation

Typical applications include:

1. vehicle identification (for granting access to a parking, measure to prevent theft, electronic license plate);
2. automatic toll system;
3. vehicle routing (enable/disable some driving spaces);
4. vehicle performance monitoring;
5. tickets for public transportation [5].

An interesting transportation-related application is presented by a recent patent [6], which presents a secure RFID system processing a real-time transport in which an RFID device receives a password set by a remote operator via SMS.

B. Banking/Commerce

Typical applications include:

1. banking (electronic check book, electronic credit card)
2. banknote authentication

[1] A consultation is a preliminary search in a field

3. commerce (anti theft, ID proof, automatic check-out counter, interactive shelves)

A patent [7] presents an anti tampering RFID device, in which a secure authentication protocol includes a physical uncloneable function (PUF), which verifies if it is being queried by an authorized verifier. Such a system finds application to avoid counterfeiting of banknote via a secure authentication. The verification is based on the bank's unique ability to reveal concealed data, such as data having been created in an enrolment phase at which the RFID tag (or actually the PUF) was registered with the bank. Now, the RFID tag again challenges its PUF to create response data sent to the verifier. The verifier checks whether the response data is correct and, if so, authenticates the device comprising the physical token, since the device is able to produce response data that corresponds to response data concealed and stored in the enrolment phase.

A dressing room is equipped with a mirror reflecting your image and with an interactive display showing images of the apparel item and celebrities wearing it. A webcam also projects an image of the consumer wearing the item on a website for everyone to see, thus creating an interaction between the consumers inside the store and their social network outside the store [8].

[9] presents an intuitive voice authenticated financial transaction system directed to uneducated users via an intuitive voice-driven interface and vocal confirmation. The transaction is validated exclusively if biometric data contained in the RFId chip match those stores on a remote banking or merchant server.

C. *Security*

Typical applications include:

1. Personnel identification [10], automatic gates, surveillance, access control [11]
2. Border control [12], i.e. passports
3. Personnel localization within a facility [13]

[14] discloses a Web-enabled vehicle access control system in which a vehicle associated reader cooperates with a reservation system at which a user places a reservation. At the vehicle, the reader automatically retrieves a unique information code from the user in form of an RFID signal, which is compared with an entry in the reservation system. Access is granted upon validation, most preferably in combination with a further validation step (e.g., via keypad). First-time users are issued a temporary code by the reservation system and receive the unique code upon validation of the temporary code. The reservation database replaces the temporary code with the unique information code. The system allows the real time knowledge of the identity of the driver of each vehicle, thus simplifying the management of a large fleet.

D. *Medical*

Typical applications include:

1. Patient identification [15], prevention of medicine/organ/blood mismatching.
2. Inventory and personnel management.

[16] shows a RFID enabled apparatus for monitoring a medical status of a user, whereby the system compares the medical status of the user to a first threshold value and a second threshold value and automatically detects whether the patient needs a nurse or a doctor. The status can be logged and accessed from a remote terminal. Whenever an alert is issued, a call may be automatically placed to a number in a contact list.

E. *Inventory systems*

Typical applications include:

1. supply chain tracking;
2. promotion tracking (forward buy management, etc.);
3. product tracking (tampering prevention), shipping and handling (item localization);
4. management of libraries and media stores.

[17] presents a supply chain tracking and management system, in which a unique identifier is assigned to a product as well as to each copy of the product and unique identifiers are assigned to all distributors of the product. When a copy of the product is sold from the manufacturer to an end user through the distributors, the unique identifier of that copy is matched with the unique identifiers of all distributors who were involved in supplying the copy to the end user. Sale of that particular item is recorded at the retail point of sale using a RFID identification method, which is then recorded by an accounting server. The distributors in the supply chain that were involved in selling that particular copy of the product are then paid according to a pre-determined schedule.

F. *Miscellaneous*

Some not-so-typical uses of RFId technology are hereby further listed:

1. Some exclusive clubs identify VIP customers via their implanted chips.
2. Animals and pets are tagged for easier identification and vaccination history storage.
3. Off road races often use passive and active RFID systems, the riders having a transponder on their

person, normally on their arm. When they complete a lap they swipe or touch the receiver, which is connected to a computer and log their lap time.

4. RFID tags are now being embedded into playing cards that are used for televised poker tournaments, so commentators know exactly what cards has been dealt to whom, as soon as the deal is complete.
5. Some casinos are embedding RFID tags into their chips. This allows the casinos to track the locations of chips on the casino floor, identify counterfeit chips, and prevent theft. In addition, casinos can use RFID systems to study the betting behaviour of players.
6. Some theme parks (such as Alton Towers in the United Kingdom) use RFID to help them identify users of a ride in order to make a DVD of their time at the park. This is then available for the user to buy at the end of the day. This is voluntary by the user by wearing a wristband given to them at the park.

IV. Technological trend

As the above sections III.A-III.F prove, RFID technology is literally invading our daily life with an ever-growing number of applications in almost every field. This implies that there is a deep interest for a technology that is per se known from the fifties. The continuing process of filing patents thus mirrors the status of a field, which is still researched to cover more and more aspects of applications which just some years ago were not thought to be either beneficial or of commercial interest.

Integrated on chip coils [18] and coexistence of RFId antennae with other radio-frequency communication antennae [19] makes a whole new set of high-speed and low cost devices available, which can bridge and profit from different technologies.

Extensions to new spectral regions, as for instance ultra wideband RFId systems [20] do, allow better resolution of positions and could be used in conjunction with other transmission systems [21] to extend some services or some existing protocols [22] to the end-user.

From the above, it is evident that RFId technology appears to be mature enough. Lots of applications are under way, especially in the field of its integration with different communication systems, as exemplarily witnessed by [23], which discloses the convergence of every possible communication/control system in a typical SOHO (Small Office Home Office) application scheme.

Given such a broad applicative spectrum, it is difficult to foresee which aspects of RFId technology might be more successfully patented, though every improvement could in principle be pursued with an application for a patent. What is probably less known to the technical community is that an improvement on a single component might lead to claims directed to the whole system. If an improved RFId antenna meets the conditions for patentability, an RFId transmitter using that antenna and an RFId system specifically using that transmitter are patentable as well.

V. Considerations on patenting in Italy

It is clear that Italian universities and research centres have neither a sound tradition of filing patent applications nor huge expertise and financial coverage for considering filing a patent application. A quick scan in the EPO database reveals that just 12 Italian companies, out of a grand total of more than 9000, have applied for a patent in this field during the last two years 2006-2008.

On the other hand, a waste of intellectual resources occurs whenever the technical implementation of a commercially valid idea is not pursued. In the long term, such quite careless behaviour leads to a loss of competitive advantage in the technical domain of interest, which might lead to the complete loss of the market for lack of investors. This applies generally to all strategic technical fields in which protection of intellectual property is not undergone.

There is therefore an urge to consider a scheme for fostering a beneficial interaction between the constituted research actors and the industrial exploiters. This scheme is not new per se, but has already proven its benefits abroad and aims at introducing a sound co-operation between universities and companies on equal grounds, thus balancing financial investment and R&D effort. Such a scheme can be summarized as follows:

1. A market analysis is carried out (usually by industrial partner) to establish a need;
2. University performs the required research focusing on the need previously established with the industrial partner and retains right to publish R&D material (under agreement with industrial partner) after a patent application is filed;
3. Company benefits from commercial patent exploitation (marketing, licensing);
4. A third party (commercial exploiter or retailer) might optionally interact with investing company, though university may retain some rights to secure further financing (via binding contracts);
5. In case of commercial success, spin-off companies may be created with staff from university and the investing company to focus on promising/assessed/profitable technologies.

VI. Disclaimer

All views expressed in this article are of the author only and are not to be construed as official views of the European Patent Office.

REFERENCES

[1] Harry Stockman, *"Communication by Means of Reflected Power",* Proceedings of the IRE, pp 1196–1204, October 1948.

[2] US 3713148, 1973. All patents mentioned in this article can be retrieved from Esp@cenet, the free official EPO online database at www.espacenet.com.

[3] US 4384288, 1983.

[4] Quaderno CNIPA 30, pp- 53-69, February 2007, in Italian.

[5] In 1995, Paris issued RFID passes conforming to the Calypso international standard for public transport systems.

[6] KR 2002 0048916, 24 Jun 2002.

[7] WO 2007 116368, 18 Oct 2007.

[8] www.rfidradio.com, 8 Jan 2008.

[9] US 2008 040262, 14 Feb 2008.

[10] In 2004, the Mexican Attorney General's office is reported to have implanted 18 of its staff members with the Verichip to control access to a secure data room.

[11] The majority of large and medium-sized enterprises are replacing traditional swipe cards with RFID contactless based solutions.

[12] Some countries which have already issued RFId-enabled passports: US (2006), Ireland (2006), Japan (2006), Pakistan, Norway (2005), Malaysia (2000), New Zealand (2005), Belgium, The Netherlands (2005), Germany, United Kingdom. Many European Union countries are also planning to add fingerprints and other biometric data, while some have already done so.

[13] Ohio Department of Rehabilitation and Correction (ODRC) approved a contract to evaluate an RFID-based personnel-tracking technology. Inmates wear wristwatch-sized transmitters that can detect attempted removal and alert prison computers. This project is not the first rollout of tracking chips in US prisons. Facilities in Michigan, California and Illinois already employ the technology.

[14] US 2007 285209, 13 Dec 2007.

[15] In July 2004, the Food and Drug Administration started a final review process that will determine whether hospitals can use RFID systems to identify patients or permit relevant hospital staff to access medical records. Since then, a number of U.S. hospitals have begun implanting patients with RFID tags and using RFID systems, more generally, for workflow and inventory management.

[16] US 2008 027288, 31 Jan 2008.

[17] WO 2008 011429, 24 Jan 2008.

[18] US 2008 079587, 3 Apr 2008

[19] US 2008 081631, 3 Apr 2008

[20] WO 03 098528, 27 Nov 2003

[21] WO 2007 017871, 15 Feb 2007

[22] US 2007 205867, 6 Sep 2007

[23] CA 2555122, 28 Jan 2008

Photo Gallery

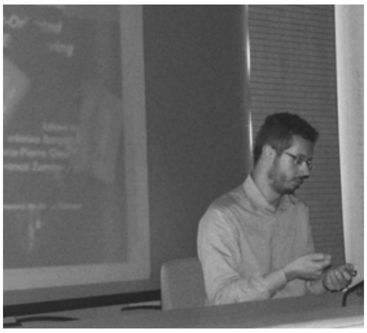
Opening Ceremony: *From the left:* **F. Vatalaro**, *IEEE Italy Section;* **E. Manganelli**, *CNIPA,* **A. La Bella**, *University of Roma Tor Vergata;* **G. Marrocco**, *University of Roma Tor Vergata*

F. Frederix, *European Commission*

F. Frezza, *CNIPA*

M. Ricciardi, *European Patent Offi*

F. Rolleri, *CNIPA*



M . Mamei, *University of Modena Reggio Emilia*

G. Marrocco, *University of Roma "Tor Vergata"*

G. Me, *University of Roma "Tor Vergata"*

G. Zanelotto, *Microsoft*

A.Moroni, *University of Roma "Sapienza"*

M. Orefice, *Turin Polytechnic*

G. Iannaccone, *University of Pisa*

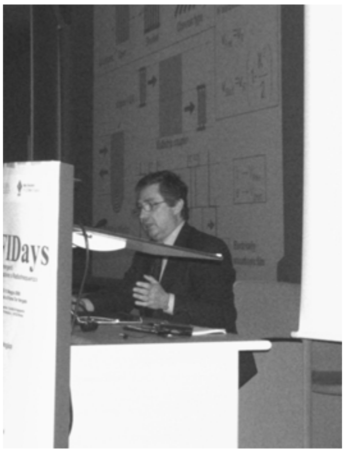
ce

M. Rebaudengo, *Turin Polytechnic*

G. Grieco,*CAEN*

A. Sciarappa, *Istituto Superiore Mario Boella*

P. Talone, *Fondazione Ugo Bordoni*



E. Verona, *CNR*

G. Biffi Gentili, *Univerisy of Florence*

S. Iudicello *&* **C. Occhiuzzi**, *University of Roma "Tor Vergata"*

Document edited on June 2008 by

Gaetano Marrocco
University of Roma Tor Vergata
Dipartimento di Informatica Sistemi e Produzione
Via del Politecnico, 1
00133 Roma (Italy)
Tel + 39 06 72597418
e-mail marrocco@disp.uniroma2.it

Please cite as:
“Emerging Technologies for Radio Frequency Identification”, G. Marrocco Editor, *Research Report RR-08-69*, Dipartimento di Infomatica Sistemi e Produzione, Università di Roma Tor Vergata, Jun2 2008