Wired and wireless sensor networks for industrial applications

Alessandra Flammini a,*, Paolo Ferrari a, Daniele Marioli a, Emiliano Sisinni a, Andrea Taroni b

a Department of Electronics for the Automation, University of Brescia, Via Branze 38, 25123 Brescia, Italy
b Università Carlo Cattaneo, LIUC, C.so Matteotti 22, 21053 Castellanza, Italy

ARTICLE INFO

Article history:
Received 28 March 2008
Received in revised form 7 July 2008
Accepted 21 August 2008

Keywords:
Sensor network
Smart sensor
Wireless communication
Real-time system

ABSTRACT

Distributed architectures for industrial applications are a new opportunity to realize cost-effective, flexible, scalable and reliable systems. Direct interfacing of sensors and actuators to the industrial communication network improves the system performance, because process data and diagnostics can be simultaneously available to many systems and also shared on the Web.

However, sensors, especially low-cost ones, cannot use standard communication protocols suitable for computers and PLCs. In fact, sensors typically require a cyclic, isochronous and hard real-time exchange of few data, whereas PCs and PLCs exchange a large amount of data with soft real-time constrains. Looking at the industrial communication systems, this separation is clearly visible: several fieldbus have been designed for specific sensor application areas, whereas high-level industrial equipments use wired/wireless Ethernet and Internet technologies.

Recently, traditional fieldbus were replaced by Real-Time Ethernet protocols, which are “extended” versions of Ethernet that meet real-time operation requirements. Besides, real-time wireless sensor networking seems promising, as demonstrated by the growing research activities.

In this paper, an overview of the state-of-the-art of real-time sensor networks for industrial applications is presented. Particular attention has been paid to the description of methods and instrumentation for performance measurement in this kind of architectures.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

The low cost of microcontrollers facilitated the diffusion of new sensors, often called “smart sensors” or “smart transducers” [1]. This new generation of devices is replacing traditional sensors with a standard analog interface (e.g. 0–5V, 4–20mA, and so on). In fact, smart sensors provide improvements in terms of linearity, signal-to-noise ratio and diagnostic features; in many cases, they also support network connectivity. A formal definition can be found in the standard IEEE 1451.2 [2]: “a Smart Transducer provides functions beyond those necessary for generating a correct representation of a sensed or controlled quantity. This functionality typically simplifies the integration of the transducer into applications in a networked environment”. Generally, smart sensors can be seamlessly managed regardless of peculiarities due to its vendor or the adopted interfacing protocol, since they are PLUG & PLAY, eliminating errors due to manual configuration and data entering. They can be installed, upgraded, replaced or moved with minimum effort. Moreover, a smart transducer implements a general model for data, control, timing, configuration and calibration. It can contain a standardized structure (called Transducer Electronic Data Sheets (TEDS) in [2]) with manufacture-related data.

The simplified block diagram of a smart transducer is given in Fig. 1. The key aspect is the network capability: analog point-to-point interfaces can be substituted by a single, digital, low-cost, and reliable field area network. Unfortunately, there is no unique communication standard for sensor networking. If communication among computers is dominated by few technologies (e.g., Ethernet [3], Internet protocols, etc.), several incompatible sensor networking solutions are competing with each other for market leadership in a particular application. Sensor networks, in fact, are used in many application areas: military, agriculture, environment monitoring, home automation, health and welfare, automotive, industrial applications. Each field has its own requirements: for instance, health and welfare need sensor compactness and Wireless [4], while low cost is imperative for automotive and home automation [5]. Even focusing only on industrial applications, factory automation (motion control, discrete automation) and process control (power plants, refineries) have very different network requirements. More generally, a sensor network is evaluated with respect to some characteristics such as the following: extension of the network, generally affected by the physical mean; compactness; mobility, that implies wireless.
sensors with autonomous power source [6]; cost; performance, that is roughly represented by bit rate but that could depend on general timing requirements, like latency and jitter; and, last but not least, robustness, that is the safety and security [7] of the system.

1.1. The industrial scenario

In industrial applications robustness is a key factor, because strong electromagnetic power sources (welders, smelting furnaces, motors and so on) [8] can reduce the transmission quality, causing errors. On the other hand, sensor compactness and mobility are not so critical requirements, since sensors are usually installed in a fixed place and main power supply is generally available (especially in factory automation).

Performance is one of the most important issues: the ability of sensors to transfer information within a small, fixed and known time is a crucial point. It should be said that the term usage in industrial communications could be quite confusing, and expressions like “real-time” or “determinism” are often misused. According to IEC 61784-2 [9], the “real-time” is the ability of a system to provide a required result in a bounded time. Consequently, a real-time communication system is able to transfer data in real-time. In some cases it is used as a distinction between “soft real-time”, with a statistical real-time behavior (i.e. some deadline violations are tolerated), and “hard real-time”, where there is a maximum delay, i.e. latency, that is respected in all cases. Determinism is related to the ability to set an imposed and invariable latency, that is the required result is provided in a fixed, known and repeatable time. Determinism is a more stringent constraint than “hard real-time”, but sometimes the terms are used as synonymous. The term isochrony refers to the ability of a system to be strictly repetitive in time. In an isochronous communication system, the data transfer is strictly cyclic and has a very low jitter. The term jitter indicates the difference between the maximum and the minimum time that elapses between cyclic events.

There are some industrial applications, such as packaging, manufacturing, wood machining or plastic extrusion, that require high performance systems to achieve a cost reduction [10]. Data exchange must be fast, reliable and deterministic; latency times must be in the order of hundreds of microseconds to correctly close control loops between twin drives, and jitter must be one order of magnitude lower than latency.

The best way to respect deadline in data transfer is a centralized architecture, where transducers are read/written, when needed, by a central system through point-to-point connections. In this case, the event reaction time, that is the delay between an input event and the related output actuation, is minimal and well known. Use of smart sensors in a centralized architecture is uneconomic and reductive, since a few of their qualities can be exploited (e.g. self-calibration, reduced diagnostic features).

The distributed architecture is more suitable for smart sensors, thanks to their network interface. Advantages of distributed architectures are countless and include increased flexibility, improved performances, predictive maintenance, simple installation, and cabling cost reduction. Unfortunately, the use of distributed architectures implies transmission delays that could heavily affect performance. Particular care must be taken in order to assure correct operations.

For the sake of clarity, an example is analyzed. A simple program such as “if A>B then immediately actuate C”, which properly works in a centralized architecture, could present some problems if we suppose a simple distributed architecture, as depicted in Fig. 2. In fact, A and B quantities could be sampled in different instants, that is A = f(t0) and B = f(t1), and therefore can be only roughly compared. Even supposing T0 ≈ t1, it could be difficult to estimate the transmission time tA,A and tB,B of A and B to the controller; typically they can be approximated by known limits (Tmin ≤ tA,A ≤ TMAX). Controller elaboration should start as soon as sensor messages arrive and the actuator should actuate C as soon as the controller message arrives. In this case elaboration takes time tact, and the controller message takes time tA,A to reach the actuator and, consequently, there is a delay time T0. If sampling and actuating time is ignored, Td = max(tA,A + tA,B + tB,B + tact, Td) could be significant and variable, because tA,A, tB,B and tact could depend on network traffic. This simple example shows how performance of a distributed architecture could be affected by network and application behavior. It should be highlighted that the considered strong hypothesis leads to a simplified Td expression, which, in real applications, is much more complex.

Let us now consider that traffic between the controller and transducers can be organized in a cyclic way, as shown in Fig. 3. The controller periodically (every cycle time) exchanges information organized in time slots with field devices. If the chosen physical and medium access layers ensure the respect of time slot bounds, the communication is real-time, deterministic and isochronous (i.e. frames exhibit a constant inter-arrival time). If this communication approach is used for the application shown in Fig. 2, the whole system at the application level, characterized by the delay time Td, could still show a considerable jitter. In fact, time between event (A>B) and reaction (C actuated) depends on sensor sampling time and, more generally, on synchronization among application tasks in the distributed nodes.

In order to fix this problem, industrial communication protocols often provide some synchronization services, as input/
output synchronization commands (e.g. global read, global write) or application synchronization utilities in order to achieve determinism at the application level. In fact, if all the nodes share a common sense of time (i.e. they have synchronized clocks) and the isochronous scheme of Fig. 3 is adopted, then determinism could be obtained modifying the control program to:

- samples $A$ and $B$ at time $t_0$ (i.e. the start of the cycle $k$), and transmit data in the same cycle
- if $A(t_0) > B(t_0)$ then actuate $C$ at time $t_1$

where the time interval between $t_0$ and $t_1$ must be greater than three times the cycle time. Supposing the elaboration is synchronized with the start of the $(k+1)$th cycle and the elaboration time is less than the cycle time, the message to $C$ can be delivered and actuated in the $(k+2)$th cycle.

A good synchronization among nodes could be useful even if a sensor network does not imply actions to be taken in real-time, but is only used to collect data from sensors. In this case, the only need is to "accurately" reconstruct the temporal sequence of events, alarms or data samples (data timestamping in every node), because time accuracy/resolution affects data value accuracy/resolution.

In conclusion, due to hard constrains in terms of performance, robustness and cost, communication technologies (also known as fieldbuses) for the realization of sensor networks for industrial applications are often "tailored" solutions. Fieldbuses are used in most industrial plants to digitally link subsystems and to transfer few data in real-time, except for close-range applications [11] with few and simple sensors, which still adopt traditional centralized architectures. Resuming, fieldbuses typically have cyclic behavior, synchronization utilities, low/medium data rate (few Mbps), good efficiency (number of data bit with respect to transmitted bit), a good network extension (in the order of 100 m), low cost and pay special attention to safety [12–15]. There are several international (open) standards that describe fieldbus protocols, but they are still very similar to proprietary systems. Actually, they can reach satisfactory performances in a system with only a single fieldbus but cannot coexist with other technologies.

However, fieldbus interfaces, such as DeviceNet or PROFINET, are quite simple and can be easily integrated in low-cost microcontrollers, reducing the need for external components (e.g. some 8-bit microcontrollers provide a CANbus 2.0B interface).

2. From fieldbus to Ethernet and RTE

Ethernet is widely used in industrial plants for communication among Programmable Logic Controllers (PLCs) and Supervisory Control And Data Acquisitions (SCADAs); often, in this case it is called "Industrial Ethernet" [16]. Such high-level communication systems adopt solutions based on TCP/IP [17], as for instance Ole for Process Control (OPC) [18]. Since, as discussed in the previous section, fieldbuses are used at lower levels, an integration problem arises: how do high levels and low levels exchange data? [19]. In addition, fieldbuses are often poor regarding remote-diagnostic features and self-configuring tools, while Ethernet is profiting by more efficient switch-based architectures, increased transmission rate and availability of low-cost devices [20].

These considerations led, in the last years, to the idea to use Ethernet even at the field level. Ethernet seems unsuitable for real-time applications because the a priori estimation of the maximum transmission time of a data packet is impossible [21].

This is mainly due to the Carrier Sense Multiple Access with Collision Detection (CSMA/CD) for the Medium Access Control (MAC) and to unpredictable delays introduced by switches, which depends on network topology, traffic conditions, switch technology ("Store&Forward", "Cut-Through", [22]), and so on. An important incentive to the diffusion of Ethernet in industrial plants comes from IEC61784-1 [23], that describes and acknowledges some commercial solutions of Industrial Ethernet; in some cases they could be used down to the sensor level, as HSE (Fieldbus Foundation for Ethernet), Ethernet/IP, and PROFINET. These protocols do not guarantee performance suitable for most real-time control applications; therefore other solutions are emerging, called Real-Time Ethernet (RTE), such as Powerlink [24], PROFINET IO [25], EtherCAT [26], MODBUS-RTPS [27] and so on, including dedicated solutions [28]. These technologies allow more powerful performances when compared to traditional fieldbuses, taking advantage of the high performance of Ethernet (e.g. 100 Mbit/s or more instead of the typical 1–10 Mbit/s of high-performance fieldbuses).

Real-Time Ethernet (RTE) is defined by IEC61784-2 as the ISO/IEC 8802-3-based network that includes real-time communication [3]. RTE solves the non-determinism problem of Ethernet modifying media access rules by means of software protocols (e.g. master–slave protocols based on Time Division Multiple Access—TDMA) or thanks to ad hoc switches or network interfaces. There is no universally acknowledged single RTE protocol and the above-cited solutions differ in the way they achieve determinism.

Synchronization among nodes, that is all the nodes follow a common clock (master clock), takes a very important role in RTE. Synchronization methods can vary from simple proprietary protocols [29], as broadcast triggering messages, to the use of standard solutions, as Network Time Protocol (NTP) or Precision Time Protocol (PTP) described in standard IEEE1588. Standard IEEE1588 estimates frame propagation delays through the infrastructure (cables, switches, etc.). Currently, it seems the most promising synchronization method because it is independent of technology and it allows full-software realizations but, in that case, it strictly depends on the application level. In addition, in industrial plants, star topologies are considered unsuitable, so if many switches are cascaded [32], propagation delay of a frame can be asymmetric and IEEE1588 could yield to considerable estimation errors. Hardware–software solutions can be used in order to increase the performance of IEEE1588 achieving an accurate timestamping of frames; in this way, a synchronization in the order of 100 ns can be achieved, but RTE protocols are not supported [33]. Anyway, IEEE1588 is not a power-efficient protocol and other new approaches have been proposed to synchronize nodes, e.g. in the field of wireless sensor networks [34]. Moreover, IEEE1588 was originally designed for LAN application and if a larger area has to be covered or mobility must be ensured, the Global Positioning System (GPS) can be used to obtain a Universal Coordinated Time (UTC) reference as described in [35].

Another common aim of an RTE network is to be compatible with TCP/IP traffic. In fact, industrial communications should support at the same time, on the same media, a fast (isochronous) real-time data exchange and, if an event occurs (alarm or diagnostic or configuration activity for a certain node), complex and acyclic communication. Some research activities have been carried out to quantify RTE performance reduction as a function of bandwidth dedicated to TCP/IP traffic [36], but methodologies for the test environment setup, like load generation and profile, are rather rough [37,38].

As an example of an RTE, PROFINET IO is briefly described [39]. PROFINET IO performance is described with the class number: RT Class 1 (RT, Real-Time) is used in jitter-tolerant systems requiring cycle time down to tenth of milliseconds; RT Class 2 (also called IRTflex, Isochronous Real-Time with Flexible network...
switches must be used, thanks to a powerful ASIC\cite{41}, forward structure must be PROFINET IO compliant; in fact, special the network infrastructure. This means that the network infra-

several other indicators can be used\cite{48,49}. For instance, "Stack Traversal Time" is the time required by data to pass through the communication stack from top (application layer) to bottom (physical layer). "Event Reaction Time" is the time required by the system to acknowledge an external event.

Furthermore, several other indicators can be used\cite{48,49}. For instance, "Stack Traversal Time" is the time required by data to pass through the communication stack from top (application layer) to bottom (physical layer). "Event Reaction Time" is the time required by the system to acknowledge an external event.

3. Metrics and instruments for RTE

RTE networks are an example of emerging technology where scientific research and industrial interest converge. RTE networks are a new topic and recently some workshops are appearing\cite{42–44}. Besides a lack of widespread knowledge, a general absence of measurement methods and instruments characterizes RTE-based applications.

Particularly, a complete set of suitable parameters to characterize an RTE-based application is not defined; moreover, even if a feature derived from the Information and Communication Technology (ICT) field seems adequate, measurement methodologies and test environments are often not available. For instance, bandwidth and latency are well known and widely used in Ethernet\cite{45} and Internet\cite{46} and they appear correct for RTE too. However, real bandwidth measurement is rather difficult, because it depends on data and on the state of linked nodes actually, the peak value is often considered as a good indicator. As regards latency, it is usually measured in an empirical way, thanks to instruments that measure the normally called "roundtrip delay", which is defined as the time interval between the transmission of special frames and the receipt of the related acknowledgment\cite{47}. The most famous method is the Ping command that is based on Internet Control Message Protocol (ICMP). Obviously, this method does not support the resolution required by RTE networks. The above-cited IEC61784 suggests the following performance indicators:

- **Delivery time**: the time needed to convey application data from one node (source) to another node (destination).
- **Time synchronization accuracy**: the maximum deviation between any two node clocks.
- **Non-time-based synchronization accuracy**: the maximum jitter of the cyclic behavior of any two nodes when such cyclic behavior is established by means of periodic events over the network. For instance, this is the case of some RTE protocols that use a network message to signal the start of a cycle. In such protocols the sharing of a common clock reference is not required.
- **Redundancy recovery time**: "the maximum time from failure to become fully operational again in case of a single permanent failure".
- **Throughput RTE**: the total amount of RTE application data (in octet length) on one link per second.
- **Non-RTE bandwidth**: "the percentage of bandwidth, which can be used for non-RTE communication on one link". The total link bandwidth shall also be specified, since they are related to each other.

Furthermore, several other indicators can be used\cite{48,49}. For instance, "Stack Traversal Time" is the time required by data to pass through the communication stack from top (application layer) to bottom (physical layer). "Event Reaction Time" is the time required by the system to acknowledge an external event.

Fig. 4. PROFINET IO cycle.
A new, low-cost, multi-probe instrument has been recently proposed [58]. The general architecture is shown in Fig. 5. The instrument can be viewed as a network of probes designed to simultaneously log Ethernet traffic in different links of a target RTE network. This “parallel” network, called measurement network has been used. Generally, if 100BaseT high-bandwidth RTE protocols are considered, the measurement network should work with 1000BaseT or more. In the realized implementation, a 1000BaseT measurement network has been used.

One of the objectives during the development of the new instrument architecture was cost limitation. This led to have single-chip FPGA-based probes and a single Monitor Station instrument architecture was cost limitation. This led to have single-chip FPGA-based probes and a single Monitor Station. Probes are requested to associate a reliable timestamp to every frame that transits on the Ethernet link they monitor. This results in a special probe architecture that enables RTE full-duplex logging together with strict time synchronization among probes.

Monitor stations must store and elaborate all the incoming data; thus the only critical point is the system bandwidth, which is the ability to manage all the data without dropping frames. In fact, logged frames and timestamping-related data must be transferred, resulting in quickly growing bandwidth requirements. Generally, if 100BaseT high-bandwidth RTE protocols are considered, the measurement network should work with 1000BaseT or more. In the realized implementation, a 1000BaseT measurement network has been used.

Regarding measurement instrumentation, in the ICT field some instruments are used to associate a time reference to Ethernet frames: from the PC-based instruments like WireShark (formerly Ethereal), a well-known network analyzer software, [53,54] with resolution in the order of 0.1 ms, to the high-performance network analyzers. The latter allows a time resolution in the order of tens of nanoseconds that could be suitable for RTE networks; on the other hand, limits are the compactness, the cost and the robustness typically needed by industrial environments. As an example of new instruments designed for ICT, the WAND group [55,56] of University of Waikato in New Zealand has developed a new instrument based on programmable logic devices that adds timestamps to every Ethernet packet. At present, about 100 of these instruments, synchronized by GPS, are used all over the world to perform statistical analysis of Internet traffic.

As software-based instruments are not adequate and network analyzers allow only a costly and localized measurement, often RTE performance characterization is done looking at input and output signals, for instance measuring the event reaction time with respect to external events and subsequent reactions.

In order to develop instruments tailored to RTE networks, a multi-probe approach must be considered, taking advantage of recent developments in the FPGA technology and in the availability of network processors [57]. In fact, by means of a multi-probe architecture, it is possible to experimentally measure delay times (i.e. through a switch) and verify synchronization among nodes.

### 3.1. A new instrument

A new, low-cost, multi-probe instrument has been recently proposed [58]. The general architecture is shown in Fig. 5. The architecture is focused on providing simulation tools. Obviously network simulators like OPNET [50] or OMNeT++ [51] do not natively support RTE protocols; therefore much effort is spent to develop an effective model of an RTE node [52].

Regarding measurement instrumentation, in the ICT field some instruments are used to associate a time reference to Ethernet frames: from the PC-based instruments like WireShark (formerly Ethereal), a well-known network analyzer software, [53,54] with resolution in the order of 0.1 ms, to the high-performance network analyzers. The latter allows a time resolution in the order of tens of nanoseconds that could be suitable for RTE networks; on the other hand, limits are the compactness, the cost and the robustness typically needed by industrial environments. As an example of new instruments designed for ICT, the WAND group [55,56] of University of Waikato in New Zealand has developed a new instrument based on programmable logic devices that adds timestamps to every Ethernet packet. At present, about 100 of these instruments, synchronized by GPS, are used all over the world to perform statistical analysis of Internet traffic.

As software-based instruments are not adequate and network analyzers allow only a costly and localized measurement, often RTE performance characterization is done looking at input and output signals, for instance measuring the event reaction time with respect to external events and subsequent reactions.

In order to develop instruments tailored to RTE networks, a multi-probe approach must be considered, taking advantage of recent developments in the FPGA technology and in the availability of network processors [57]. In fact, by means of a multi-probe architecture, it is possible to experimentally measure delay times (i.e. through a switch) and verify synchronization among nodes.

### 3.1. A new instrument

A new, low-cost, multi-probe instrument has been recently proposed [58]. The general architecture is shown in Fig. 5. The instrument can be viewed as a network of probes designed to simultaneously log Ethernet traffic in different links of a target RTE network. This “parallel” network, called measurement network, conveys data (logged by probes) towards a supervision equipment, called monitor station. Probes are requested to associate a reliable timestamp to every frame that transits on the Ethernet link they monitor. This results in a special probe architecture that enables RTE full-duplex logging together with strict time synchronization among probes.

Monitor stations must store and elaborate all the incoming data; thus the only critical point is the system bandwidth, which is the ability to manage all the data without dropping frames. In fact, logged frames and timestamping-related data must be transferred, resulting in quickly growing bandwidth requirements. Generally, if 100BaseT high-bandwidth RTE protocols are considered, the measurement network should work with 1000BaseT or more. In the realized implementation, a 1000BaseT measurement network has been used.

One of the objectives during the development of the new instrument architecture was cost limitation. This led to have single-chip FPGA-based probes and a single Monitor Station implemented using a PC. Moreover, the monitor station can use open-source user interface programs like the above-cited WireShark. The probe local time is constantly synchronized with a reference clock despite local crystal oscillator variations (temperature, aging, etc.). Synchronization Unit can operate using multiple synchronism sources: 1-PPS signal from an external source or IEEE1588 Sync Message coming from the measurement network. The local time is synthesized with an adder structure [59]. Briefly, an increment step is summed to the time register at every clock period of the local oscillator. Drift and offset can be compensated adjusting the increment step with a suitable control algorithm. The increment step is refreshed each time a synchronism event happens; it means 1 s with 1-PPS and typically 1 or 2 s with IEEE1588 PTP.

A two-probe prototype has been experimentally characterized comparing performance to a powerful single-probe network analyzer: Endace NinjaCapture 1500 [60]. The test network is the PROFINET IO Class 1 system shown in Fig. 6.

Two Ethernet TAPs have been inserted in the network in order to capture traffic before and after the switch. A TAP can duplicate full-duplex traffic, so it has two monitoring ports (A and B) one for...
Table 1
Switch propagation delay

<table>
<thead>
<tr>
<th>Direction</th>
<th>Node propagation delay (ns)</th>
<th>Ave.</th>
<th>Std. dev.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ninja Capture</td>
<td>IOD → IOD</td>
<td>6283</td>
<td>1670</td>
<td>9487</td>
</tr>
<tr>
<td></td>
<td>IOD → IOC</td>
<td>6403</td>
<td>1650</td>
<td>9492</td>
</tr>
<tr>
<td>Single probe</td>
<td>IOD → IOD</td>
<td>6324</td>
<td>1653</td>
<td>9483</td>
</tr>
<tr>
<td></td>
<td>IOD → IOC</td>
<td>6416</td>
<td>1658</td>
<td>9512</td>
</tr>
<tr>
<td>Two probes</td>
<td>IOD → IOD</td>
<td>6309</td>
<td>1696</td>
<td>9589</td>
</tr>
<tr>
<td></td>
<td>IOD → IOC</td>
<td>6385</td>
<td>1683</td>
<td>9567</td>
</tr>
</tbody>
</table>

IOD: IO controller, IOD: IO device.

As previously stated, traditional networking offers many advantages but requires cables to interconnect devices. This leads to high installation and maintenance costs, e.g. due to low scalability and high failure rate of connectors. For this reason, wireless technologies gained an enormous success in the consumer goods industry in the last few years. In addition, the adoption of wireless solutions at the sensor level offers other advantages: continuous, high-resolution, ubiquitous sensing; support for mobility; redundancy; and compactness.

In particular, wireless sensor networking uses Micro Electro Mechanical System (MEMS) and nano-technology. Even if miniaturization is not the primary driver for such kind of research in sensors application (the size of the sensor depends on measurement itself, and the sensors need to be packaged in an appropriate way to be correctly handled), the real advantage of nano-scale structures will be the development of new materials and new sensing elements. It will become possible to create huge arrays of similar sensing elements and thus develop novel sensing principles, for example for chemical and biochemical sensors [61].

Besides high power consumptions, high area coverage and high cost solution such as the well-known and mature mobile phone technologies (GSM, GPRS, and UMTS just to cite few of them), two standards have monopolized the market of the Local/Personal Area Networks: IEEE802.11 [62] and IEEE802.15.1 [63]. The former is the wireless counterpart of the Ethernet standard and implements lower levels of WiFi [64], while the latter constitutes lower levels of Bluetooth (BT) [65]. The main attraction of both of them is that they do not require any sort of frequency licensing because they operate in the Industrial, Scientific and Medical (ISM), radio frequency region. However, WiFi and BT have been designed to address requirements of office/personal communication, and cannot efficiently be used to realize Wireless Sensor Networks (WSNs), as better explained in the next sections.

Obviously, advantages due to the absence of cables could be usefully exploited in several fields and many efforts have been carried out in this direction. For example, in the past, novel trends [66] have emerged in the agricultural sector converged in the so-called "precision agriculture", which concentrates on providing the means for observing, assessing and controlling agricultural practices. In this way, it would be possible to detect parasites on the field and automatically choose the right type and amount of pesticide. Another field where wireless technologies have been widely used is "environmental monitoring"; just to mention some applications, it is possible to monitor air quality in real-time by means of unattended stations or collect data in places discourage human presence. Another interesting application is in the field of "smart structures", which comprises home and building automation; in this case, a wireless sensor and actuator network is integrated within a building to improve living conditions and reduce overall energy consumption. Also, "medical and health care" are fields where WSNs have been successfully employed, for example, it is possible to ensure patients continuous monitoring without limiting their mobility.

4.1 Wireless sensor networks

As stated in previous section, wireless communications are an effective and reliable solution in home and office automation. Generally speaking, several media can be exploited, including light and ultrasound, but considerations regarding data size, rates and area coverage make radio frequency (RF) links more attractive. Many standards have been proposed to satisfy requirements of the consumer world, as proved by the IEEE802 subgroups that cope with these topics (refer to Fig. 7), but the most interesting for WSN applications are probably those comprised in the IEEE802.15 [67] working group, whose effort focuses on the development of Personal Area Networks or short-distance wireless networks (≈ 10 m).

In particular, here is defined the concept of Personal Operating Space, a spherical region that surrounds a wireless device with a radius of 10 m. Even if originally designed for portable and mobile computing devices such as PCs, personal digital assistants (PDAs), cell phones, pagers, and consumer electronics, it may be successfully applied to WSNs. However, it must be underlined that large-scale applications in the sensor networking area are yet in the development stage.

First of all, it is important to distinguish between the idea behind the WSN concept and implications related to an industrial scenario, better described further. From a general point of view, a WSN is made up of a large number of tiny devices (sensors), which
are densely deployed and collaborate to monitor and analyze a phenomenon of interest [66].

Due to cost and dimension constrains, sensors have limited computational resources; power consumption must be as low as possible to ensure a true autonomous activity. In addition, sensors could be randomly positioned, thus requiring localization and self-organizing capability. Besides issues considered in this paper, there are other questions that must be considered in a wireless system. In particular, security could be a key aspect; air is an open medium and it is easy for an attacker to maliciously alter transmissions or make the link unreliable injecting jam sequences.

The block diagram of a wireless sensor node is represented in Fig. 8.

As every smart sensor, a wireless transducer consists of three main parts: a sensing unit, a processing unit, and a transceiver unit. In addition, a power manager is present to handle on board power sources, such as electrochemical batteries or more exotic power scavenging units [68]. Moreover, most of the performed tasks require also the knowledge of positions and time, furnished by proper localization and synchronization units. Many researchers are currently engaged in the design of proprietary schemes that fulfill such requirements, each one with its pros and cons. However, according to authors, the most promising solution is to adapt standard devices that are already available on the market and can exploit a huge volume of production and mature technologies, to the industrial application under investigation.

In the following, aspects regarding power consumption, localization and communication architecture will be detailed, while in the next section the industrial scenario will be considered.

4.2. Power consumption

Power consumption of a wireless sensor node can be divided into three different domains: sensing, processing, and communicating. The first one is strictly related to the application and in most cases can be neglected. Regarding data processing, usually processor consumption during the active phase decreases by an order of magnitude or less if compared with that needed in the communication phase. As explained in [69], supposing a Rayleigh fading and a fourth-order loss law, the energy required to transmit 1 KB over a distance of 100 m is approximately the same as that for executing 3 million operations by a 100 MIPS/W processor. From another point of view, it is convenient to implement complex algorithms if this results in shorter data packets and/or in a more robust data link that requires less retransmissions. It is well known that consumption is proportional to the voltage supply (Vdd), and to the operating frequency (f), i.e. P ∝ Vdd f [70]. This relationship suggests two strategies to lower consumption: dynamic scale voltage, i.e. reducing the supply voltage Vdd as low as possible, and changing the CPU clock frequency f according to the computational load (usually, microprocessor uses a low- and a high-frequency oscillator in the idle and active phase, respectively). The most demanding unit is thus the transceiver. If we consider short-range (∼10 m) systems operating in the GHz range with low radiation power (∼0 dBm), the energy required to transmit is almost the same as that required in data reception. Obviously, devices spend most of their time doing nothing; this means that low-duty-cycle strategies, probably the most diffused solution in battery-supplied nodes, can be applied. What really matters is the average current consumption Icc, mean of the four-fifth loss law, the energy required to transmit E = 10 dBm), the energy required to transmit E = 10 dBm)

- wakes up every T seconds (it depends on the Medium Access Protocol implemented),

Please cite this article as: A. Flammini et al., Wired and wireless sensor networks for industrial applications, Microelectron. J (2008), doi:10.1016/j.mejo.2008.08.012
Propagation Models have been established [74,75]; RSSI is used to node or on a base station [71–73]. A node that wants to estimate possible to adopt other techniques such as GPS, whose cost is map, which is very time consuming. As an alternative, the Signal movements of devices and obstacles enforce a recreation of the node communication. If higher performances are needed, it is extra constant needs to be subtracted from Eq. (4) to account for model obstructions like walls are not considered; otherwise an versatile small power sources. Neglecting the power unit self- L 

\[ L \approx \frac{\eta C}{K_{cc, mean}} \]  

where C (Ah) is battery capacity, \( \eta \) the power supply efficiency and \( K_{cc, mean} \) the power supply output voltage gain. Batteries are usually divided into primary or not rechargeable cells and secondary cells; according to the adopted electrolyte (NiCd, NiMH, LiION, etc.), they offer nominal voltage in the order of 1.2–3.6 V and capacity up to 3 Ah for the AA format.

### 4.3. Localization

WSN are severely constrained for energy and cost of deployment and operation. Therefore, localization is usually performed employing the same radio transceiver that is also used for inter-node communication. If higher performances are needed, it is possible to adopt other techniques such as GPS, whose cost is justified only in some applications. The basic idea is to exploit the Radio Signal Strength Indicator (RSSI), a standard feature in most networks. Two approaches are commonly accepted in literature-RSSI-maps and Signal Propagation Models. Many RSSI measurements at different locations form a so-called RSSI-map that is stored on a node or on a base station [71–73]. A node that wants to estimate its position compares the measured RSSI-values with the entries in the RSSI map. The position with the most equal entry is then chosen. Although the precision of this technique is relatively high, movements of devices and obstacles enforce a recreation of the map, which is very time consuming. As an alternative, the Signal Propagation Models have been established [74,75]; RSSI is used to evaluate power attenuation, correlating it to the distance with respect to some anchorages points. In fact, if PT is the transmit power, \( PL(do) \) is path loss for a reference distance, \( \eta \in [2,4] \) is the path loss exponent and the random variation in the power measured at the antenna connector is expressed as a gaussian random variable of zero mean and \( \sigma^2 \) variance, \( X \sim N(0,\sigma^2) \), it is well known that relation (4) can effectively describe losses as a function of distance:

\[ RSSI(d) = PT - PL(do) - 10\eta \log\left( \frac{d}{do} \right) + X \sigma \]  

(4)

All powers are in dBm and all distances are in meters. In this model obstructions like walls are not considered; otherwise an extra constant needs to be subtracted from Eq. (4) to account for the attenuation in them (this constant depends on the type and number of obstructions). However, experimental campaigns [76,77] pointed out that RSSI fluctuates for both intrinsic and extrinsic causes, such as:

- **Transmitter variability**: different transmitters behave differently even when they are configured exactly in the same way.
- **Receiver variability**: different receivers behave differently even when all environmental parameters are the same.
- **Antenna orientation**: different antennas have their own radiation patterns.
- **Multi-path fading and shadowing in the RF channel**: channel behavior greatly depends on environmental characteristics.

As shown in [78], RSSI-based ranging and localization can be a cheap and effective alternative only when applied in the right environment.

As a concluding remark, localization is a hot topic in WSNs, especially when mobility is required, but it is still an expensive (computationally and monetary) task. Several researchers are involved in this field and a clear solution has not yet emerged, although it is not a primary topic in industrial application.

### 4.4. Network communication architecture

A preliminary architecture classification of WSNs can be done distinguishing among infrastructure and ad-hoc networks [79], as shown in Fig. 9.

#### Infrastructure

Wireless networks often extend, rather than replace, wired networks, and are referred to as infrastructure networks. A hierarchy of wide area and local area wired networks is used as the backbone network. The wired backbone connects to special switching nodes called Base Stations. Therefore, within infrastructure networks, wireless access to and from the wired host occurs in the last hop, between base stations and nodes that share the bandwidth of the wireless channel. This approach is used in most of industrial applications, because it offers best performance and reliability.

#### Ad hoc

Ad hoc networks, on the other hand, are multi-hop wireless networks in which a set of nodes cooperatively maintains network connectivity. This on-demand network architecture is completely un-tethered from physical wires. They are characterized by dynamic, unpredictable, random, multi-hop topologies with typically no infrastructure support. Mobile nodes must periodically exchange topology information that is used for routing updates.

Referring to the protocol stack, the traditional ISO/OSI model is modified as shown in Fig. 10. The main difference is the presence of “vertical” planes whose aim is to manage power, localization and synchronization units as shown in Fig. 8.

The PHYsical (PHY) layer is responsible for frequency selection, modulation and data encryption. It is well known that long-distance communications are not efficient in terms of power consumption and implementation complexity, suggesting the adoption of short-range transceivers. In addition, this approach can overcome shadowing and path-loss effects if multi-hop

![Fig. 9. (a) Infrastructure and (b) ad-hoc WSN architecture.](image-url)
networks are implemented. Most diffused commercial available solutions implement spread spectrum modulation and offer data rates in the order of 0.1–1 Mbps. The occupied band is the free ISM near the 2.4 GHz portion of the spectrum. Power consumption is in the order of 10 mA for the transmitting/receiving phase down to less than 1 µA in the standby mode.

The data-link layer is responsible for multiplexing of data streams, data frame detection, medium access control (MAC) and error control. Since the MAC controls the radio, it has a large impact on the overall energy consumption, and hence, in the lifetime of a node. The air is a shared medium and it must be fairly assigned to nodes; the MAC decides when competing nodes may access the radio channel and tries to ensure that nodes do not interfere with each other's transmissions. The two major approaches are contention and schedule-based one. The former allows collisions, letting nodes to contend for the resource; the latter regulates accesses scheduling (usually there is a particular node or access point that broadcasts this information) when and how long each controlled node can access the shared medium. Just to give an example, IEEE802.15.4 [80] implements carrier sense multiple access with collision avoidance (CSMA/CA) that belongs to contention methods, while IEEE802.15.1 adopts TDMA, a scheduling approach. The MAC layer operates on a local scale and lacks the global information to optimize network lifetime; it should ensure that the energy it spends is proportional to the amount of traffic that it handles. Schedule-based protocols are the most efficient, at the cost of reduced flexibility; on the contrary, contention-based ones tend to collapse when the load approaches the channel capacity and wastes energy in idle listening, overhearing, and protocol overheads. A rough classification can be done according to the following:

- the number of used RF channels,
- the degree of organization among nodes, and
- the way a node is notified of an incoming packet.

It is not possible to state what is the best solution, that must satisfy trade-offs dictated by the application [81]. Some considerations regarding the industrial environment are carried out in the next section.

The NetWork Layer (NWK) routes the data supplied by the upper layers from the source(s) to the sink(s). In a more formal way, sources are entities that provide data/measurements while sinks are those nodes where information is required. In a single-hop architecture, sources and sinks are directly interconnected by a radio link, whereas in a multi-hop nodes can forward data frame detection, medium access control (MAC) and localization, radio communications and it should also describe the behavior of the media and the environment. Moreover, it should be modular, allowing one to verify the consequences of using requirements, and mobility of the sensor node. A star topology is a single-hop system in which a particular node, called coordinator, manages communications and all remaining nodes communicate only with it. It is a sort of master–slave structure where the coordinator acts also as a bridge towards other networks. Moreover, star topology is a power-efficient solution that ensures a long network lifetime even if a node collapses, but it can only handle a small number of nodes. This is not a real limit since in many cases communications among coordinators use wired links. Mesh topologies are multi-hopping systems in which all nodes are identical and communicate with each other, so that a coordinator or base station is not strictly needed. The multi-hop system allows for a much longer range than a star topology at the cost of higher power consumptions rate and higher latency. In fact, nodes have a high duty cycle since they need to “listen to” messages and network changes and latency is related to the number of “hops” between the source and the sink. The aim of a star–mesh hybrid architecture (also known as cluster tree) is to take advantage of the low power and simplicity of the star topology, as well as the extended range and self-healing nature of a mesh one. Nodes are organized in a star topology around routers or repeaters, which, in turn, organize themselves in a mesh network. However, latency may be a problem.

The transport layer is usually implemented only if end-users access the WSN through the Internet. Upper layers are usually summed up in a generic application layer that makes the hardware and software of the lower layers transparent to the end-user.

5. Available tools

In order to fully understand the complexity of designing wireless protocols that work in real life, it is necessary to model, simulate and also to implement and test on real-world systems. Abreast of traditional instruments as spectrum analyzers, which are out of the scope of this digression, a plethora of simulators and “sniffers” appeared on the market in order to predict and verify the WSNs behavior. Simulation is the most common approach to develop and test new solutions for WSNs. Obviously, simulators must be provided with highly accurate models in order to be effective. Taking into account the fact that the wireless medium poses new challenging troubles, new model components (sensor hardware, batteries, CPU model) and a tight crosslayer coupling, not considered in classical tools, must be included. In order to facilitate the sensor network designer in this task, a useful instrument could be a simulation framework. It is a software object that defines structures representing networks, sensor nodes within networks and their inter-connections. The framework should model the major aspects of sensor networks such as battery lifetime, node localization, radio communications and it should also describe the behavior of the media and the environment. Moreover, it should be modular, allowing one to verify the consequences of using...
various techniques and architectures. There are a number of such frameworks nowadays available, both with commercial and with open-source licenses.

In addition, even if simulations at the packet level are useful in verifying protocols performance, designing an effective WSN involves extensive and accurate knowledge of the environment and radio behavior. In fact, assumptions made in most propagation models do not necessarily reflect the real-world conditions.

Traditional mistakes can be summarized as follows:

- the world is flat,
- radio's transmission area is circular and all radios have equal range,
- if I can hear you, you can hear me,
- if I can hear you at all, I can hear you perfectly, and
- signal strength is a simple function of distance.

Finite element simulations of the environment could overcome these limits but are time consuming and very difficult to implement. For all these reasons a new approach must be pursued.

In [82] it has been shown that traditional stochastic models fail in forecast parameters as packet error rate, especially when applied in industrial scenarios. In addition, classical tools do not consider node components as sensor hardware, batteries, CPU that must be correctly modeled to ensure an accurate estimation of network life. A survey on these topics is furnished by Egea-Lopez et al. [83]. For all these reasons, simulation of WSNs is still an open research field that must be flanked by experimental validations. This task can be partially accomplished by low-cost instruments known as “sniffers”.

These devices are realized using the same hardware of sensor nodes (i.e. they are low cost but purposely designed for a particular physical layer) and detect and capture wireless radio signal packets in the ambience presenting them in convenient graphic displays, furnishing a useful insight on what occurs in air.

6. Wireless sensor networks in the industrial scenario

Although WSNs are generally characterized by fast deployment and cost effectiveness, their applicability in industrial environments is still in the development stage. The vast majority of conventional wireless protocols emphasize bit rate over versatility and reliability, which is unsuitable for industrial control and monitoring applications. For this reason, even if wireless is largely used at the Enterprise and Factory control level, applications at the field level are still very few. It must be remembered that fault assumptions are very different in wireless communication than in wired ones. Even if transmission errors are more frequent than on wired links, they are bursty in nature. On the contrary, errors on wired channels are often of permanent nature due to connector or cable failures. In addition, as stated in the RUNES report [84], “the industrial automation sector, in general, is characterized by conservatism. Companies do not want to take chances with large investments in new installations and require demonstration of practicality”. For all these reasons, new tools must be developed, with particular attention to simulators that should allow one to accurately predict the network behavior in a real scenario.

For instance, Siemens [85] has just announced a wireless Human Machine Interface (HMI) module that relies on an extension of the IEEE802.11 they call IndustrialWLAN or IWLAN and that supports PROFINET IO protocol. The same manufacturer offers other devices, such as industrial access points and client nodes, which can be used as “cable replacement” of a PROFINET IO Class_1 network. Even if IWLAN sensors are not yet announced, nodes may act as gateways towards standard fieldbuses as PROFIBUS.

A step towards standardization at least in process automation has been reached with the release of WirelessHART specifications [86]. At the application level, backward compatibility with already existing solutions based on HART fieldbus is ensured.

Regarding the wireless link, the physical layer is compliant with the IEEE802.15.4, thanks to its low cost. Upper layers are mostly based on the Time Synchronized Mesh Protocol (TSMP) [87], which improves reliability using frequency, time, and spatial diversity to access the medium. Power consumption is limited, thanks to strict nodes time synchronization; most bandwidth is dedicated in a sort of virtual circuit (link among a source and a destination) in order to minimize the power spent on idle listening and the power wasted on packet collisions. Network topology is the mesh one, i.e. every node can relay incoming data packets towards the final destination (i.e. it acts as a router). The latter choice is dictated by final applications that include monitoring of vast plant like refineries obliging to adopt multiphop protocols. However, as previously stated, this choice greatly affects the latency, on the order of seconds, making this solution useless for factory automation.

An interesting survey has been conducted by the RUNES project, started in September 2004 with the aim to expand and simplify networks of devices and embedded systems. The final conclusion reported in their first meeting [84] is that “Adoption of networked embedded systems (particularly wireless) is slower in the industrial sector than the rest of the sectors examined in this technology roadmapping exercise... While technologies are maturing, wireless should not be used for critical control applications. Monitoring in hazardous and inaccessible areas should be given priority in the short/medium term and in moving towards this some lessons can be learnt from successful telemetry deployments.”

The basic goal of WSN is a reliable data delivery consuming minimum power. Traditionally, sensor networks data delivery is said to be [88]:

- time driven, when sensors communicate their data (continuously) at a pre-specified rate;
- event driven, when sensors report information only if an event of interest occurs; and
- hybrid, when all approaches coexist.

It has been already underlined that Data Delivery Model of industrial communications is time driven for majority of operations (data collection), but, even if infrequent, events as alarms or network management issues must be detected/notified quickly. Real-time, i.e. the respect of temporal deadline, must be ensured with low and predictable delay of data transfer (typically, less than 10 ms); therefore synchronization among nodes must take place. As previously said, radio frequency links are disturbed particularly by low transmission ranges or impenetrable walls and obstacles; therefore, methods for error detection or error correction are crucial. In addition, message lengths are mostly very short; therefore, data efficiency for short telegrams is a very important design guideline. The use of the time-triggered paradigm supports the protocol efficiency because the transmission of message parameters like sender identification, message length, message priority, etc. may be implicitly codified in the communication schedule.

For all these reasons, the best solution seems to adopt small, reliable infrastructured star networks that exploit time division as the medium access policy. In fact, the typical radio link area allows to cover a machinery and several machineries may be

Please cite this article as: A. Flammini, et al., Wired and wireless sensor networks for industrial applications, Microelectron. J (2008), doi:10.1016/j.mejo.2008.08.012
interconnected by traditional wired fieldbuses; in addition, cellular topology allows frequency reuse. Most of the existing wireless systems/standards do not satisfy these requirements, since they all use event-driven data delivery with contention-based MAC protocols. Nevertheless, several efforts have been made to ensure the so-called Quality of Service (QoS), i.e. the ability of a network to deliver predictable results, even in wireless. Just to give an example, the newly released IEEE 802.11e [89] amendment tries to overcome these flaws defining several improvements on the legacy MAC. In particular, the traditional contention-based (CP according to the standard nomenclature) and contention-free periods (CFP according to the standard nomenclature) have been replaced by a hybrid Coordination Function (HCF) that preserves backward compatibility and satisfies QoS requirements. In the CP it is possible to define message priorities using a sort of slotted CSMA ruled by different Arbitrary InterFrame Space (AIFS); in the CFP the hybrid coordinator controls the access to the channel by polling the stations with QoS requirements. The IEEE 802.11e amendment defines also new synchronization mechanisms for the upper layers that dramatically improve the accuracy of the traditional Timing Synchronization Function.

In conclusion, proprietary tailor-made upper layers protocols are usually developed relying over new transceivers compliant with the standard physical level; in this way portability and low cost is ensured without sacrificing performances. As the early adopters of these solutions, it is possible to figure food processing, petrochemical and asset tracking (fast parcel operators) sectors, where monitoring tasks of slow dynamic quantities can be greatly simplified by cables replacement and new scenarios may be envisioned.

Based on the same approach ABB with its acronym of Wireless Interface to Sensors and Actuators (WISA) [90] tries toport wireless also in factory automation. Nodes communication hardware is based on standard IEEE802.15.1 transceiver; the MAC layer in WISA adopts time division multiple access with frequency division duplex (TDMA/FDD), ensuring simultaneous transmission and reception of radio signals. The network level implements the star topology with up to 120 nodes per coordinator. Also a "wireless power supply" has been provided; simple nodes as proximities switches harvest energy by means of inductive coupling with a mains powered supply unit. The downlink, i.e. transmission from the coordinator to nodes, is always active in order to establish cycle and slot synchronization, to send acknowledgments and control data. On the contrary, the uplink, i.e. transmission from node to coordinator, is event driven, in order to lower power consumption. As stated by the manufacturer, for proximity switches that exchange a 1-byte-wide packet, the typical latency between node and coordinator is in the order of 5 ms, with a maximum event rate of 5 Hz.

A proprietary protocol stack, processed by a simple 8-bit microcontroller (HCS08GT60 from Freescale), has been developed to minimize the overhead, to reduce the datagram lengths and to achieve high efficiency decreasing computational effort. An IEEE802.15.4 compliant device has been chosen (formally MC13192 from Freescale). Concerning the MAC, TDMA has been adopted to guarantee the cycle time deadline of sensory data transmission and CSMA/CA for network management purposes.

Regarding the NWK topology, a star architecture has been adopted. In fact, nodes along the barrel are relatively close one from each other and no complex routing strategies are needed; therefore a small firmware footprint can be obtained. With regard to the application layer, it simply encapsulates sensor data within the protocol datagram. No particular attention has been devoted to security, but a simple ciphering method based on a pass-phrase can be adopted.

The network coordinator is mains powered and is always in the on-state and acts as a MODBUS RTU slave. It periodically sends a BEACON packet that delimits the beginning of a new cycle of the WSN. The first part of the cycle is devoted to network constitution (Join Period in Fig. 12); it lasts 32 ms.

A node that wants to join the network waits for the BEACON and sends a JOIN packet with a CSMA/CA approach. If the coordinator accepts, it sends an ACK packet specifying the Network IDentifier (NID) and the node time slot that corresponds to the Device IDentifier (DID). Datagrams are shown in Fig. 13. The PHY level header and the Frame Check Sequence (FCS) fields are imposed by the IEEE802.15.4-PHY. The PROT field is used to distinguish the proposed protocol with respect to IEEE802.15.4 and specifies protocol version. SQN is a sequence number to allow for cycle traceability. SN is the node univocal identifier (factory set). The Time To Wake-up (TTW) indicates the amount of time that must elapse before next wake up and allows for time synchronization, as better explained in the following. BC is the Backup Channel adopted to improve reliability by means of channel diversity. The remaining part of the cycle (Real Time Period in Fig. 12) is devoted to real-time data communication that occurs by means of TDMA; it lasts 96 ms. Once a node is binded with the coordinator, it sleeps for most of the time in the powersaving mode and periodically (every $T_{cycle}$) wakes up and sends its data—DATA packet—to the coordinator that answers with the ACK packet.

**Fig. 12.** CSMA/CA and TDMA hybrid approach. (a) N1 affiliation and (b) N2 affiliation.

Please cite this article as: A. Flammini, et al., Wired and wireless sensor networks for industrial applications, Microelectron. J (2008), doi:10.1016/j.mejo.2008.08.012
ACK packet. The application payload is made up of seven bytes; two bytes are reserved for the temperature information (TCJ), two bytes constitute a progressive sequence number (SQN), two bytes are for diagnostic and identification purposes (DIAG: node status, battery level, cold junction status, etc.); finally, one message integrity byte is computed to check data integrity (CHK) and ensure that the frame belongs to this kind of network.

Obviously, in a distributed time-triggered system, a synchronized timebase is a crucial requirement to enable reliable system behavior. The first requirement is not to overlap two successive 6.6 ms-wide slots. This condition may be complicated by a very low-performance clock that microcontrollers use in the power save mode. The so-called “rate synchronization” has been chosen, i.e. all nodes measure the same time interval lengths. The coordinator detects the time of arrival of each packet sent by nodes and computes the next Time To Wakeup [ms] TTW based on its internal clock. The TTW and $f_{cycle}$ values should coincide, but relative drift between coordinator and node clocks makes them different. A simple P(proportional) I(integral) controller has been implemented to correct the “error” between them. As asserted before, nodes wait for an ACK packet; if this packet gets lost, i.e. a timeout condition is reached, a single retransmission occurs, signaled in DIAG field. No ACK is sent to stay within the time slot duration, thus avoiding collisions.

Measurements on the field showed that the average current of the RF section is $I_{RF,AVG} = 0.5$ mA while other circuits (microcontroller and sensor conditioning) require $I_{OTHER,AVG} = 0.8$ mA. It means that if no retransmissions occur, the node life is about 2 months if a power source of 2.3 Ah is employed (two alkaline AA batteries).

In order to evaluate performances in a real application, some additional measurements have been conducted in a factory building. Four wireless thermocouples were installed on a plastic injection molding machine. In particular, there was no direct line-of-sight between some thermocouple nodes and the coordinator. Traffic “on the air” was sniffed for an hour. Fig. 14 reports a histogram showing the frequency distribution of lost packets. The horizontal axis represents the number of consecutive lost packets (lack of information interval), while the vertical one represents the percentage of the number of occurrences with respect to the overall cycle number NCycle = 28125; in particular, the maximum number of consecutive lost packets is equal to 7, and occurs only one time.

**7. Coexistence issues in the industrial scenario**

When several sensor networks share the same physical medium coexistence problems may arise, in both wired and wireless communication systems (e.g. RTE or WSNs). For this reason, it is easy to foresee a troublesome fight for shared media access among machineries of different manufacturers using different wireless communication solutions.

A typical manufacturing cell, i.e. a stage of an automation line in an industrial plant, has dimensions ranging between 10 and 20 m per side [93], below the area coverage of radio links. Nowadays, available technologies consider interfering signals just as wideband noise, without any knowledge about its origin. Within a densely populated environment, uncoordinated fighting to obtain transmission rights leads to poor performances for the competing devices, since each radio simply “talks louder” than others, increasing overall noise. In addition, as previously stated, industry is a “closed world” and it is not reasonable that a manufacturer borrows some of its bandwidth or shares resources to implement complex routing strategies or solve “communication troubles” of adjacent systems. Thus, it becomes very important to implement a coexistence strategy that should be technology and protocol agnostic. Many efforts have been carried out in this direction, such as those performed by the IEEEP802.19 draft guide [94], whose purpose “is to provide a standardized method for predicting the impact of mutual interference on network performance between dissimilar networks”. The goal is to furnish a document called Coexistence Assurance (CA), which is a sort of report showing how well every new proposed communication solution planned for unlicensed operation coexists with already-developed standards. Anyway, it is still in a draft state. According to the nomenclature adopted in this guide, coexistence is the ability of one system to perform a task in a given shared environment where other systems have an ability to perform their tasks and may or may not be using the same set of rules. On the contrary, standard technologies currently available (like WiFi, Bluetooth, etc.) are intrinsically selfish, i.e. even if they have some channel access mechanisms that allow similar devices to share the same medium, none of them have efficient mechanisms that allow effective discovery and coexistence with devices using different solutions. It is also useful to distinguish coexistence from interoperability, where the last term points out the ability of one or more systems to provide services to and accept services from
one or more other systems and to use these to enable the different systems to operate effectively together. In other words, internetworking can be ensured simply working at the application level, while coexistence involves the lower protocol layers. Researchers are now concentrated in defining the protocol etiquette, i.e. the framework of rules governing medium access, to which all devices using the resource must obey. Collaborative and non-collaborative [95] are the two major categories used to classify coexistence etiquette. They are differentiated by the ability to exchange information about data traffic flows. In the former technique, coexistence algorithm schedules packet traffic in controlled systems to avoid interference while maximizing throughput; in the latter there is no way to exchange information between the controlled systems and they operate independent of each other. For this reason, collaborative techniques usually require an additional component, i.e. the communication arbiter. Examples of both solutions for IEEE802.11 and IEEE802.15.1 devices are described in [96].

Obviously, considering the industrial automation data delivery model, i.e. the real-time periodical exchange of a relatively small amount of data, the most suitable approach is probably the collaborative one. In this way it is possible to efficiently exploit the cyclical nature of communications in order to orthogonalize requirements of competing devices.

Taking as an example the case study of the previous section, the easiest solution in order to ensure coexistence is the frequency diversity, i.e. collocating different networks on different frequency channels. In this way, however, no advantages are gained from the cyclical nature of data exchange. For this reason a collaborative schema that interleaves time slots of interfering networks allowing their survival has been also developed [97]. Obviously, the original communication protocol has been slightly modified in order to avoid ACK exchange in the real-time period. Thanks to the short cycle time, a delayed ACK strategy has been adopted without affecting performance; an ACK bitmap has been added in the BEACON with a negligible overhead. Also synchronizing strategies have been improved; low-power oscillators: provided by adopted transceiver has been used to wake up nodes in the proximity of beacon reception and data transmission, with a worst case jitter of about 5.5 μs.

The basic idea is to realize a WSNs coordinator wired network that synchronizes time slots collocation. This is not a crucial point, since in the considered scenario coordinators are already interconnected with a wired backbone (infrastructure WSNs) and performances of Real-Time Ethernet allows for synchronization on the order of 1 μs, well below application requirements. In particular, a beacon jitter has been measured on the order of 3.5 μs.

The architecture of the proposed system is shown in Fig. 15. The arbiter that schedules communications works at the MAC level, since strategies at the PHY level cannot be considered without hardware modifications. For this reason the proposed arbiter has been called “MetaMAC”. In the simplest implementation, it receives a [time, frequency, area] description of networks from coordinators and simply orthogonalizes their needs delaying beacon instants. A prototype has been realized and tested, allowing the coexistence of two wireless thermocouple networks occupying the same channel and the same area. This result was obtained simply delaying BEACON of a quantity equal to 3 ms.

8. Conclusions

In this paper an overview of technologies available for sensor networking in industrial applications has been presented. The state of the art has been summarized describing opportunities and limits as applied to “real world” case studies.

Fieldbuses gained more and more success in the last few years thanks to their advantages in terms of scalability and ease of installation and thanks to their ability to furnish tailored answers to industrial field requirements. The actual frontier is the convergence towards a unified standard solution, which seems to be the adoption of the so-called Real Time Ethernet.

On the other side, advancements in communications integrated circuits are making wireless an attractive alternative, at least in some monitoring applications. Besides lower cost, due to the absence of cabling, the main advantages are flexibility and redundancy, just to cite a few of them.

References


Please cite this article as: A. Flammini, et al., Wired and wireless sensor networks for industrial applications, Microelectron. J (2008), doi:10.1016/j.mejo.2008.08.012


[89] IEEE Standard for Information Technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements, Part 11: Medium Access Control (MAC) and Physical Layer (PHY) specifications. Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements—IEEE802.11e.


[94] [http://www.802.org/19](http://www.802.org/19).


[96] IEEE Recommended Practice for Information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements, Part 15.2: Coexistence of Wireless Personal Area Networks with other wireless devices operating in unlicensed frequency bands.