

Recent and Emerging Topics in Wireless Industrial Communications: A Selection

Andreas Willig, *Member, IEEE*

Invited Paper

Abstract—In this paper we discuss a selection of promising and interesting research areas in the design of protocols and systems for wireless industrial communications. We have selected topics that have either emerged as hot topics in the industrial communications community in the last few years (like wireless sensor networks), or which could be worthwhile research topics in the next few years (for example cooperative diversity techniques for error control, cognitive radio/opportunistic spectrum access for mitigation of external interferences).

Index Terms—Cognitive radio, IEEE 802154, industrial-QoS, spatial and cooperative diversity, ultra-wideband, wireless sensor networks, ZigBee.

I. INTRODUCTION

FOR most people the significance of wireless technologies comes from its ability to provide services like voice/video transmission or Internet access at places without cabled networking infrastructure or while being on the move. Wireless technologies have also been identified as a very attractive option for industrial and factory automation, distributed control systems, automotive systems and other kinds of networked embedded systems [1], [2], with mobility, reduced cabling and installation costs, reduced danger of breaking cables, and less hassle with connectors being important benefits. Some potentially interesting classes of industrial applications are closed-loop control involving mobile subsystems, coordination among mobile robots or autonomous vehicles, health monitoring of machines, tracking of parts and many more. An important characteristic in these application areas is that (wireless) data communications must satisfy tight real-time and reliability requirements *at the same time*, otherwise loss of time and money or even physical damage can result. To achieve this goal, on the one hand certain functionalities that are specific for wireless communications (like mobility management, quick handovers) must be considered, and on the other hand the unfriendly error properties of the wireless channel significantly challenge real-time and reliability. Consequently, significant research is needed to adapt existing wireless technologies and protocols to industrial settings, or, when this is not sufficient, to develop new ones. This research has been done with significant intensity for more than

one decade now, and in [1] a selective review and tutorial on research issues and approaches has been given.

This paper is a followup to [1]. Our main goal is to discuss a selection of promising and interesting research areas that received no or only limited coverage in [1] and in the industrial communications community. We start with a brief overview on the quality of service (QoS) features that have been added to the ubiquitous IEEE 802.11 wireless LAN standard (and resulting in IEEE 802.11e) and which are also interesting for use in industrial scenarios. Following this, the paper covers wireless sensor networks. Wireless sensor networks have recently received increased attention in the industrial communications community. They differ considerably from wireless LANs. Sensor networks support much lower data rates and much smaller transmit powers. More fundamental to the design of sensor networks, however, is that sensor nodes have a severely limited energy budget and consequently energy-efficiency is the single most important figure of merit. One consequence of this is that sensor nodes should spend most of their time in a sleep state in which they are not able to transmit or receive data. These properties do not favor the adoption of sensor networks in tight control loops. Instead, they are mostly considered for less time-critical monitoring tasks like for example monitoring machine health or leakage monitoring. We provide an introduction to important concepts of sensor networking and discuss a number (by far not all) of protocol design issues that are relevant for industrial applications.

In the second main part of this paper we discuss approaches that we believe can have a significant impact on the future design of wireless industrial communication protocols. In particular, we introduce recent techniques to mitigate channel fading and external interferences (two of the main reasons for the bad quality of the wireless channel!) that are currently hot topics in the wireless communications community and from which the industrial networking community can significantly benefit. For the sake of completeness we have also included a brief discussion of an existing commercial systems for wireless industrial communications: the WISA system from ABB.

We had to make choices on what to include in the paper. In terms of protocols we have mostly favored topics related to the lower layers of the industrial protocol stack (i.e., the MAC and the link-layer with its error control functionality) and their properties in terms of real-time and reliability. In terms of technologies we selected topics that have received only limited or no coverage at all in [1]. For example, we have included WSN technologies and IEEE 802.11e, but we have mostly left out Bluetooth or plain IEEE 802.11 technologies. We have furthermore

Manuscript received February 19, 2008; revised March 14, 2008. Paper no. TII-07-12-0198. R1.

The author is with the Telecommunication Networks Group (TKN), Technical University of Berlin, Berlin, Germany (e-mail: awillig@ieee.org).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TII.2008.923194

avored a tutorial-style exposition discussing fundamental issues and solution approaches over the detailed discussion and comparison of specific solutions from the literature. To help the reader to delve further into these solutions and approaches, we provide a fair number of references. We also aim to point out interesting research questions.

The paper is structured as follows: in Section II we provide a broad overview on the general research areas that need to be addressed for wireless industrial networking. Since the focus of the remaining paper is mostly on real-time and reliability properties, we also discuss appropriate performance measures. Following this, in Section III we describe the QoS enhancements to the IEEE 802.11 standard and point to an interesting research issue. In Section IV we begin our discussion of wireless sensor networks by explaining their fundamentals. In Section V we present the IEEE 802.15.4, ZigBee and ISA SP-100 standards for wireless sensor networking, since these can be expected to have significant impact in the industrial field. In Section VI we briefly look at the vast problem of providing real-time and reliability in multihop wireless sensor networks. The first part of the paper concludes with Section VII, in which the existing wireless industrial communication systems WISA is briefly reviewed.

In the second part we discuss more general research issues and selected topics from the field of wireless communications that are probably relevant for wireless industrial communication systems as well: spatial/cooperative diversity techniques in Section IX, the general issue of industrial QoS provisioning and analysis in Section X, the usage of cognitive radio techniques for mitigating external interferences in Section XI and the adoption of ultra-wideband technologies in Section XII. The paper ends with our conclusions in Section XIII.

II. OVERVIEW ON RESEARCH AREAS FOR WIRELESS INDUSTRIAL NETWORKING

To properly design networks and protocols for wireless industrial networking, several issues have to be addressed, including the following ones:

- Providing the required QoS in terms of reliability and real-time to applications: design of protocols and of wireless channel resource allocation schemes (frequencies, transmit power, rate [as determined by modulation and coding scheme], time budgets), as well as analysis schemes that evaluate achievable QoS over wireless fading channels. The relevant industrial-QoS measures are discussed in more depth in Section II-A.
- Engineering and network planning: a comprehensive set of methodologies and tools for network planning, dimensioning and configuration, as well as run-time fault and performance monitoring needs to be developed. This is tightly coupled to resource allocation. Some references for network planning in the industrial context are [3] and [4], a platform-based protocol framework allowing to adjust parameters according to pre-specified QoS levels is presented in [5].
- MAC protocol design: the MAC layer is a key functionality for (wireless) industrial communication systems, since it directly impacts the timeliness of packets. The goal is to find deterministic protocols that can support packet priorities or which allow fine-grained channel scheduling. We

discuss in Section III-B briefly the specific issue of priority enforcement on wireless channels, a technique that is needed to implement CAN-like protocols and to leverage existing work on schedulability analyses for CAN.

- Error-control schemes: error-control schemes directly impact the achievable reliability. Recent approaches to error control are addressed in more detail in Section IX.
- Routing and transport protocols: especially in multihop networks like wireless sensor networks (soft) real-time guarantees have to be provided over multiple hops. This will be discussed in more detail in Section VI.
- Application-layer protocols: application support protocols and the applications themselves must be designed with explicit consideration of the wireless channel properties in mind. One example is the research area of networked control systems (see [6] and [7]). An alternative view of application-layer protocols is taken for example in [8], where the authors argue that on top of commercial wireless hardware like IEEE 802.11 WLAN it is the application layer that must ensure appropriate real-time and reliability properties through specifically designed protocols.
- Hybrid wired/wireless systems: in many applications it is beneficial to adjoin wireless stations to existing wired networks and therefore to create *hybrid networks*. This has in general been discussed in [1] and [9], whereas specific networks and protocols have for example been considered in [10]–[13].
- Mobility support and handovers under real-time and reliability constraints. The design of suitable schemes, especially for hybrid systems, depends on the underlying MAC protocols. One example reference is [14], in which mobility and handovers are considered for hybrid wired/wireless PROFIBUS systems.
- Security and privacy: Security in general and wireless security in particular are vast research topics [15], security aspects for industrial networks are discussed in [16]. A brief account on some issues is provided in Section II-B.
- Energy consumption and energy-efficient design: To fully achieve the benefits of wireless communications, no cabling at all should be used, and this includes also the energy supply. This issue is discussed in more detail in Section IV.
- Scalability: in a factory plant both the number of wireless networks as well as the number of wireless nodes per network might become large. This is especially true for wireless sensor networks which must be significantly over-provisioned in order to give individual nodes enough opportunity to enter a low-energy sleep state while ensuring that the overall network has still enough awake nodes to achieve its task.

In the remainder of this section we discuss relevant quality of service measures for wireless industrial networks, followed by a brief account of security issues.

A. Quality of Service Measures for Wireless Industrial Networks

Industrial networks are in general designed to carry traffic that is dominated by exchanges of sensor readings and actuator commands between sensors/actuators on the one hand and (often centralized) controllers on the other hand [2], [17], [18]. Important characteristics of industrial traffic are the presence of

deadlines, high reliability requirements and the predominance of short packets [19]. For these exchanges two different interaction patterns have emerged: the master/slave model and the producer-(distributor)-consumer model, and for these patterns different QoS parameters are important. Since the wireless channel is random and time-varying, the classical deterministic performance measures like for example the worst-case transmission times should be replaced by probabilistic measures.

In the master/slave model (for example adopted in the PROFIBUS [20], [21]) information is exchanged between controller and sensors/actuators by using a unicast communication mechanism. In this setting the key aspect of service quality is to enable transmission of periodic or sporadic messages within pre-specified *deadlines* and in a *reliable* fashion. Reliability is especially important for critical alarm messages. There are different performance measures that *jointly* account for real-time and reliability. Such measures can be defined on different timescales: in terms of long-term averages or on much shorter timescales. One measure belonging to the first category is the *success probability*, defined as the long-term probability that a message can be successfully (i.e., acknowledged) transmitted within its deadline. Another one is the *capacity-vs-outage-probability* measure [22], denoting the transmission rates (and therefore the delay) so that the probability of not achieving this rate is below a pre-specified threshold. A third measure is the *delay-limited capacity*, it is defined as the capacity-vs-outage probability measure for a pre-specified outage probability of zero (see also the discussion in [23]).¹For short timescales formulations based on the notion of (m, k) -firm deadlines [24], [25] can be used, in which at most m out of any k consecutive packets are allowed to be lost, otherwise a deadline violation occurs.

In the producer-(distributor)-consumer model (for example adopted in WorldFIP [26], [27]) communication is based on unacknowledged broadcasts of data identifiers (by the distributor), to which the station possessing the identified data item (the *producer*) responds by broadcasting its current value. All nodes interested in this data (the *consumers*) copy the received value into an internal link-layer buffer for later delivery to the higher layers. This can be regarded as an instance of a publish/subscribe interaction pattern [28]. Here the most important performance measures are related to the degree by which all the consumers are able to simultaneously capture the data and to maintain consistent buffer states—this is referred to as *spatial consistency*.² Such a consistency requirement can be captured by the *agreement probability*, i.e., the probability that all consumers have reached agreement within a certain, pre-specified time window.

¹Technically, these measures are defined for stationary and frequency-flat block fading channels. In a block fading channel, the packet duration is smaller than the channel coherence time, i.e., the time during which the channel does not change its characteristics (appreciably). The term frequency-flat refers to a situation in which all the involved frequencies of a transmitted signal have the same attenuation level. For industrial networks with predominant small packet sizes a block fading channel is a reasonable assumption, and when the transmission rates are reasonably small then the channel is also frequency-flat.

²A similar consistency requirement is called relative temporal consistency: Many control applications require that the relevant sensors sample the environment nearly simultaneously in a pre-specified time window. In some industrial communication systems this is achieved by using explicit triggering signals, broadcast by the central controller. Again, the degree to which all the relevant sensors receive the trigger packet is important.

Those previously defined performance measures represent the prime performance metrics for wireless industrial communication systems—we refer to them as *industrial-QoS*. A range of secondary metrics can be devised that measure the efficiency of protocol mechanisms attempting to improve the primary performance measures. Examples of such secondary metrics are the protocol or memory overhead, computational overheads, additional interference created, and others.

B. Security

Today's automation networks tend to be more and more integrated with other networks, for example to allow cost-effective remote monitoring and maintenance of machine plants. There are many techniques to protect a network against attackers from outside, for example firewalls [15]. But when the network uses wireless transmission, an attacker that is close enough to the network can eavesdrop, it can insert malicious packets, or it can simply jam the wireless medium and distort any other transmission, this way challenging reliable and timely transmission.

Encryption can be used to prevent eavesdropping. To prevent insertion of malicious packets, mechanisms for ensuring authentication (“who sent this message?”) and message integrity (“is this the message originally sent?”) are needed [15], [29] to create mutual trust relationships between wireless stations. Such mechanisms are often implemented using shared secrets and public key cryptography, calling in turn for proper key distribution schemes. To avoid replay attacks proper sequence numbers/session keys have to be used. Some challenges for implementing security mechanisms in wireless industrial networks and wireless sensor networks are the following [30, Sec. 1.2]:

- Ensuring authentication and message integrity for each message requires *message integrity check* (MIC) fields in each message. To be effective, this field should have a reasonable minimal length, for example 16 bytes. However, since in most fieldbus systems the maximum allowable frame size is small (and many packets have only a few bytes anyway), the MIC fields account for significant fraction of overhead.
- Key distribution schemes introduce significant protocol and administrative overhead.
- In the case of hybrid systems, one has to take into account that often the legacy fieldbus protocols running in the wired stations do not have any security mechanisms.

III. IEEE 802.11E

The IEEE 802.11 family of wireless LAN (WLAN) standards is certainly predominant in the realm of WLAN technologies, and it has also been considered extensively in the context of wireless industrial communications, see for example [31]–[38]. In this paper we focus on aspects of the current standard that have gained increased importance since publication of [1] and which at the same time are especially interesting for industrial applications, namely the QoS support that is now part of the 2007 version of the standard [39] and which was formerly specified in a separate amendment [40].³ We therefore give a brief

³The 2007 version of the standard supersedes the 1999 version and its 2003 reaffirmation, and includes also previous amendments to the IEEE 802.11 standard like IEEE 802.11a (an OFDM physical layer for the 5.2-GHz ISM band), IEEE 802.11b (11 Mb/s physical layer for the 2.4-GHz band), IEEE 802.11g (a high-rate physical layer for the 2.4-GHz ISM band, including 54 Mb/s OFDM), IEEE 802.11i (security enhancements) and IEEE 802.11e (quality of service).

introduction into the new QoS mechanisms and point to some research issues.

A. IEEE 802.11E Quality-of-Service Support

The QoS functions are available in infrastructure IEEE 802.11 networks. They consist of a *hybrid coordination function* (HCF) that operates on top of a (modified) *distributed coordination function* (DCF) as it is known from the original IEEE 802.11 standard. For some parts of the hybrid coordination function a centralized control entity called *hybrid coordinator* (HC) is required which is co-located with an access point. Amongst other functionalities, the HC can perform *admission control*. More precisely, the stations in a QoS-enabled IEEE 802.11 network (henceforth simply referred to as network) can send a request to the HC asking him to schedule *transmit opportunities* (TXOP). A TXOP is a contiguous window of time which a station can use exclusively (i.e., without having to expect parallel transmissions from other network members) to transmit one or more frames including MAC layer acknowledgements to any station in the network. It is possible that a station requests the HC to grant these TXOPs periodically (for example for voice or video data streams) and upon receiving such a request the HC has to decide whether the new flow can be admitted without breaking guarantees for the already present flows. The precise admission control algorithm is not prescribed by the standard but left to the implementers.

The standard provides two different access schemes for providing QoS support: the *enhanced distributed coordination access* (EDCA) and the *hybrid coordination function controlled channel access* (HCCA). We describe both of them briefly but start with a reminder of the basic distributed coordination function (DCF) upon which both access schemes are built. Many details are left out, for example the RTS/CTS scheme.

1) *DCF and EDCF*: The (E)DCF belongs to the class of carrier-sense multiple access with collision-avoidance (CSMA-CA) protocols. It relies on the *physical carrier-sense* function of the underlying physical layer, which indicates the presence or absence of signals or ongoing transmissions on the wireless medium. In addition, a station performs a *virtual carrier-sense* operation. In this operation, the station maintains a special variable called *network allocation vector* (NAV). Most of the packets in IEEE 802.11 contain a duration field, which denotes the remaining time that is needed until the ongoing transaction (for example data plus acknowledge) is finished. Whenever a station receives a packet, it updates its NAV variable with the packets duration field in order to prevent own transmissions during the remaining transaction time. In summary, a carrier-sense operation indicates an idle channel only if the NAV is zero and the physical carrier-sense does not indicate any transmission activity.

For the following description please refer also to Fig. 1. When a new packet shall be transmitted, the station performs a carrier-sense operation. When the medium is idle for a certain amount of time called *inter-frame space*, the station starts to transmit. The packet at hand is associated with one of four pre-defined *access classes* and the inter-frame space that the station depends on this access class. For the i -class the inter-frame space

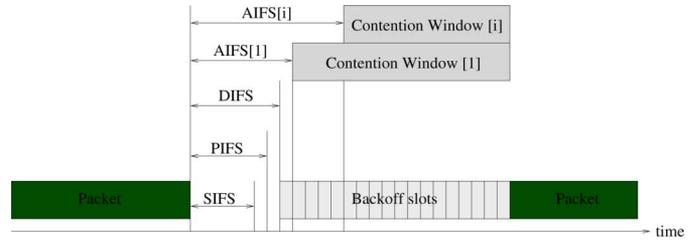


Fig. 1. Timing of the IEEE 802.11 DCF.

is called $AIFS[i]$. These $AIFS[i]$ values can be configured, but are required to be distinct.

When the medium is busy, the station enters the *backoff mode*. It station first waits until the medium is idle for a time of at least distributed inter-frame space (DIFS). At this time the backoff slots start. If the station was not in backoff mode before, it draws a random number out of the current *contention window* and sets a counter with this number. In the following, during each idle slot the counter is decremented by one and if it reaches zero, the station starts to transmit. If the carrier-sense mechanism indicates a busy medium, the process of decrementing the backoff counter is suspended, and it is resumed later on once the medium becomes idle again.

The contention window size is dynamic. It is initialized with a pre-configured value CW_{min} . Whenever a packet transmission is not successful, the contention window size is doubled, until a maximum value CW_{max} is reached. This is useful when transmission failures are interpreted as resulting from channel collisions, since an increase in contention window size leads to increased average backoff times and therefore to a reduction of the pressure on the channel.

We are now in the position to briefly explain the difference between the classical DCF and the EDCF. In the DCF there are no access classes and all stations use the DIFS instead of the $AIFS[i]$ inter-frame spaces. Furthermore, all stations use the same values for CW_{min} and CW_{max} . This means that all stations have the same chance to access the channel. In the EDCF for each access class i separate values of $AIFS[i]$, CW_{min} and CW_{max} can be configured at the access points, which then distributes these values to all stations in its beacons. By choosing proper values for these parameters it is possible to ensure with high probability that a packet of a better access class wins over a contender having a packet of lower access classes. In other words: it is possible to achieve *stochastic prioritization* or *service differentiation*. A further difference between the DCF and the EDCF is the following: in the DCF winning contention acquires the right to transmit *one packet*. In the EDCF the winner receives a TXOP which, as explained above, may encompass several packets. The maximum duration of a TXOP is a pre-configured value. However, only packets belonging to the access class for which the TXOP was won are allowed to be transmitted during a TXOP.

A number of performance analyses (e.g., regarding throughput) of the DCF and EDCF are available in the literature [34], [41]–[45]. These studies confirm also that the EDCA mechanism is indeed capable of achieving stochastic prioritization.

2) *HCCA*: In the HCCA a central entity, the HC, is responsible for coordinating access to the medium. The HC receives reservation requests for TXOPs, grants or rejects these requests (admission control) and is responsible for actually scheduling the TXOPs for all attached stations. The channel is, however, not all the time under full control of the HC. Instead, scheduled TXOPs alternate with phases in which stations contend for the medium using the EDCF. The reservation requests can either be sent as separate packets (using the EDCF) or they can be piggy-backed onto data packets that are sent during a scheduled TXOP.

To gain control over the channel the HC uses also the basic DCF mechanism, but the medium has only to be idle for PIFS time (which is smaller than DIFS and all the AIFS[i] values) and after that time the HC starts with its transmissions. Specifically, to start a TXOP the HC sends a poll packet to the owner of the TXOP and the owner then uses the TXOP to transmit one or multiple packets according to the results of a local scheduling policy (which, however, typically prefers better access classes over lower ones). The admission control policy and the scheduling policies applied in the HC and the stations are not specified in the standard.

The HCCA is rather complex as there are rich interactions with other features of the IEEE 802.11 protocol like for example the rate adaptation feature of some of the physical layers. It is similarly complex as the point coordination function of the original IEEE 802.11 standard, which is still present in the 2007 version and which can be operated jointly with the HCCA. The author is not aware of any existing implementation of the HCCA.

One weakness of the HCCA that prevents it from achieving perfectly periodic services is that at the scheduled time of TXOPs the medium might still be busy from a previous TXOP obtained by the EDCF mechanism and the HC has to wait a random time until the packet end before he gains access to the medium. A key research issue in the HCCA is the design of appropriate admission control scheduling policies, which has for example been done in [46]–[48]. Its usage in industrial scenarios has been considered in [33] and [49].

B. Research Issues

Assuming that HCCA implementations will not become widely available during the next few years (PCF implementations actually never did), it makes sense to concentrate on improvements of the DCF and the EDCF.

One very interesting and relevant problem is the deterministic priority enforcement on the wireless channel. The problem is defined as follows: given two devices A and B which have packets ready for transmission at the same time and whose transmission ranges overlap, it should be deterministically ensured that A 's packet can be transmitted before B 's when A 's packet has a higher priority and A 's receiver is in the overlap area of A 's and B 's ranges. In other words: B 's less important packet should not block A 's transmission. In a probabilistic variant of priority enforcement the protocol must make it more probable that A wins than that B wins. This probabilistic variant is indeed implemented by the EDCF, supporting four different priorities.

Some industrial and automotive communication systems like for example CAN [50] employ a MAC layer technique in which to each packet a priority value is assigned and these priorities are then used to resolve contention among different stations. When the priority enforcement is deterministic, a proper assignment

of priorities to packets allows to perform deterministic schedulability analyses [51], [52].

In the CAN protocol, a bitwise priority-arbitration technique is used for collision resolution: a contending station awaits the end of an ongoing transmission (if any) and then enters contention phase. The contention phase is driven by the priority of the packets ready in the contenders and proceeds bit-by-bit. A contender transmits the value of the current priority bit and simultaneously receives feedback from the channel (which is guaranteed to have a well-defined level and can be converted back into a bit). When the own transmitted bit and the bit read back from the channel differ, the station has lost contention and defers, otherwise the station proceeds with the next bit.

This approach cannot be directly implemented with commercial wireless transceivers, since it requires full-duplex operation of the transceiver. Wireless transceivers are in general half-duplex: they are not able to transmit and receive simultaneously on the same channel because their own signals would drown all signals from other stations. Because of this fact, most wireless transceivers share some circuitry between transmit and receive path, which naturally prevents that they work in parallel.

Therefore, alternative mechanisms are needed to enforce packet priorities on the channel. There are several options for (almost) deterministic priority enforcement in fully meshed networks. By the rules of (persistent) CSMA protocols, a contention cycle starts at the end of a previous packet. One possibility is to let contending nodes send jamming signals of length according to the priority of their current packet. Afterwards, a node switches to receive mode and performs a carrier-sense operation to check whether any other contender emits a longer jamming signal. If so, the listening station defers and the other station has won the contention. This approach, while having been used in the HIPERLAN-I standard [53], is unfortunately not implementable with commercial IEEE 802.11 transceivers as these do not offer the generation of jamming signals. A complementary and more practical approach is to let nodes listen on the channel for a time proportional to their packets priority (the more important the packet, the shorter a station listens for other stations and the earlier the station starts to transmit its packet). In both cases, the maximum length of bursts/listening periods is linear in the number of priorities that can be supported. In the wireless dominance protocol (WiDom) approach presented in [54] a bitwise priority arbitration scheme is mimicked by providing one time slot for each priority bit, i.e., n slots for n priority bits. During such a timeslot a station having a dominant bit transmits, while stations with recessive bits receive. When a recessive station receives a signal, it has lost contention and gives up. With this approach, the number of priorities that can be distinguished is 2^n , while the duration of the contention resolution period is linear in n .

However, all these approaches share some problems and therefore need additional research:

- The above discussed deterministic schemes all rely on carrier-sensing and are thus vulnerable to external interferences—any unwanted signal that is detected by the carrier-sense algorithm leads to deferral of transmissions and higher probabilities of deadline misses.
- They do not work in hidden-terminal situations, i.e., in settings where two stations having packets of different

priorities want to transmit to the same receiver but cannot sense each others signals. In this situation the receiver needs to help with resolving priorities. An important goal is to design suitable schemes with small overhead to achieve this, and to assess their scalability in terms of the number of contending stations and the number of supported priorities.

IV. WIRELESS SENSOR NETWORK TECHNOLOGY

In this section we review the fundamentals of wireless sensor network technology and their potential for industrial applications.

Wireless sensor networks (WSNs) [30], [55]–[60] consist of a large number of small, energy- and resource-constrained *sensor nodes*.⁴ An individual sensor node is composed of sensor circuitry (for example temperature or humidity), a microcontroller, some small amount of RAM and program memory, a wireless transceiver and an energy supply, most often a battery. Some nodes might also be attached to actuators, in this case sensor networks are sometimes referred to as wireless sensor- and actuator networks (WSAN). However, in the following we use the term WSN to include actuator nodes as well. While each individual sensor node has limited computational capabilities, the nodes can communicate wirelessly and perform collaborative signal processing tasks. A driving vision for sensor networks can be identified in the concept of “smart dust” [62]: sensor nodes shall become so small and so cheap that they can be embedded almost everywhere and could make our environment intelligent. With wireless sensor networks it is possible to collect much more real-time data than was possible before, from places which are hazardous or otherwise inaccessible for wired technologies.

Wireless sensor networks can be used in many ways in industrial and factory automation [63], automotive applications have also been considered [64]. An important class of applications is monitoring of equipment and machinery health, using for example vibration, heat or thermal sensors. This can help to detect upcoming machine failures and to trigger a preventive maintenance cycle before an often more costly repair is needed. Sensor networks can also be useful for leakage or radiation monitoring in chemical plants. In this general class of monitoring applications the sensors are typically not part of any control loop and therefore the timeliness requirements are not extremely hard, but reliability is an important issue. There are also activities to use wireless sensor networks in distributed and process control applications [65]–[67].

A. Architecture

Sensor network architectures have a lot in common with the architecture of ad hoc networks, but there are also some important differences. In general ad hoc networks there are no distinguished stations. All stations potentially run different applications and communicate with each other in a peer-to-peer fashion. In sensor networks all nodes cooperate to fulfill a common task. A sensor network is designed *as a whole* to run a single or very few related applications, which involves sensing the physical environment and collaborative processing of sampled data. Sensor networks have more structure than ad hoc networks. There are typically a few *sink nodes* present, to

which the sensor nodes report their data. The sink nodes can configure and control the operation of the sensor nodes, they provide the interface to human users and they can serve as gateways to other networks. In control applications they are the natural point for the controller. Sink nodes are often more powerful than ordinary sensor nodes, they could for example be PDAs, laptops or desktop computers.

The application-induced interaction patterns between sensor nodes reflect the presence of sinks: communication typically happens either between closeby sensor nodes (for purposes of collaborative signal processing) or between sensors and sink nodes. It occurs only rarely that a sensor node communicates with a sensor node more than a few hops away (a notable example are distributed or geographic hash tables in sensor networks [68]).

Sensor readings are often only meaningful when location and time of observation are known, too. This requires that the geographical position of sensors (either absolute or with respect to some coordinate system) or their logical position (“room FT 131”) is known and that in addition the sensors are time-synchronized with each other. The design of both localization protocols [69], [70] and time synchronization protocols [71]–[74] is challenged by the lack of GPS receivers for reasons of costs, form factor and energy consumption of such devices. Often, only a small fraction of the nodes will be equipped with GPS receivers and the remaining nodes have to infer their position from additional range or angle measurements and subsequent iterative trilateration.

Two of the core challenges in any sensor network design are energy-efficiency and scalability, discussed next.

B. Energy Efficiency

In most cases sensor nodes use batteries for energy supply. Batteries have a finite lifetime, although it is sometimes possible to prolong this lifetime by combining energy-harvesting techniques [75] (in which secondary batteries are recharged by extracting energy from the environment, for example from vibrations or heat gradients) with proper node-level power-management strategies [76], [77], or by using battery management schemes in which the time pattern of drawing energy is carefully chosen to exploit battery self-recharge effects [78], [79]. A further option are mechanisms for wireless power transmission, like for example the magnetic-field-induced power supply available for the WISA wireless fieldbus system [80].

As a result of finite node lifetime, energy-efficiency can be considered as the single most important design goal for sensor network hardware, algorithms, protocols and applications [81]. The fact that an individual sensor nodes’ energy can be depleted, together with the low reliability of nodes (resulting from cost constraints), calls in general for substantial overprovisioning of the network in terms of redundant nodes. When the nodes coordinate their activities properly, the useful lifetime of the overall network (i.e., the time where it can still fulfill the given task subject to pre-defined quality requirements) can be substantially larger than the individual node lifetimes.

Energy- and power-efficiency has a significant influence on protocol design. For many node designs the wireless transceiver requires the largest share of the overall power budget. Depending on the actual combination of microcontroller and transceiver, the microcontroller can execute several hundreds

⁴The presentation in this section is in parts based on [61].

or even thousands of instructions with the same energy as needed for transmitting one bit on the wireless channel. Hence: *computation is cheaper than communication*. This observation is a key motivation for *in-network processing*: intermediate nodes could for example aggregate the readings from multiple sensor packets into one single packet and transmit only the resulting packet to the remote sink. This reduces the amount of bits to be transmitted and thus saves transceiver energy [82], [83], on the other hand the aggregated data is more important and needs better (more energy-consuming) protection through error-control mechanisms [84]. Aggregation is in turn a key example for *data-centric design*: not only the application but also the protocols know the data that is transported, and this *data can influence protocol decisions*.

Due to the small transmit power (often in the order of 1 mW as compared to typical WLAN output powers of 100 mW), the energy costs for transmitting and receiving packets have the same order of magnitude. For several transceiver designs, receiving requires approximately the same energy as transmitting. As an example, the datasheet for the IEEE 802.15.4-compliant ChipCon CC2420 transceiver [85] lists current consumptions of 18.8 mA (which for supply voltages of 2.1–3.6 Volts corresponds to a power consumption of 39.5–67.7 mW) for receiving, and 17.4 mA for transmitting. There might also be the situation that the transceiver is listening on the channel for incoming packets but without actually receiving something. This is referred to as the *idle state*. The energy consumed in the idle state is often a significant fraction (50% and more) of the energy consumed in receive mode. To reduce the average power consumption, sensor nodes should switch selected hardware components like the transceiver into a *sleep state* with ultra-low power consumption for most of the time, or even switch them off completely. The fraction of time in which the transceiver is not in a sleep state but in idle, receive or transmit state is called its *duty cycle*. Clearly, the smaller the duty cycle, the longer the lifetime of the sensor node. However, a sleeping sensor node can neither participate in packet forwarding nor can it sense the environment if it has also switched off the sensor circuitry. To let individual sensor nodes sleep, redundant nodes must be present that can take over their duties, for example to ensure that the network is still connected and environmental stimuli are properly observed. Accordingly, neighbored nodes have to coordinate their sleeping schedules. A number of tradeoffs are involved here: the coordination costs energy, routing protocols have to examine more possible routes and have to deal with time-variable topologies, there are potentially more contenders for channel bandwidth and so forth. A further complication comes from the fact that switching forth and back from the sleep mode incurs energy costs as well: depending on the precise nature of the wakeup process (it might for example be necessary to fully re-initialize the sleeping device) there can also be significant energy costs associated with it, since the wakeup costs time and power [86], [87]. Therefore, the rate of alternation between sleep and awake states should be properly controlled as well.

Since the radiated power of an individual sensor node is often small, the communication range is restricted to a few meters and the network must operate in *multihop* mode when information is to be transmitted over larger distances. Therefore, energy-efficient routing is a key design issue in sensor networks [55, Chap. 11]. Also, the data rates offered by low-power transceivers are

constrained to the range of tens to hundreds of kbits per second. One example is again the ChipCon CC2420 transceiver [85], offering 250 kBit/s as the highest transmission rate.

As a general design rule, the protocols and applications in a sensor network should be *jointly* designed towards high energy efficiency. It is often possible to reach more energy-efficient protocol designs with *cross-layer approaches* [88] instead of relying on traditional protocol layering principles, but cross-layer design is also more challenging because of the larger potential for unwanted interactions between protocol components [89].

C. Scalability

The need for redundancy increases the number of nodes in a sensor networks. When large areas have to be observed, the number of required nodes can be in the range of thousands or even tens of thousands.⁵ Scalability is therefore a second critical concern, and one of the immediate results is that individual sensor nodes and the overall sensor network should be self-organizing—manually configuring hundreds or thousands of nodes is not an option.

The large number of sensor nodes mandates that in general *distributed* algorithms and protocols should be preferred over *centralized* ones whenever possible. Otherwise, traffic hot spots would emerge around centralized servers and all the traffic heading to/coming from those servers would have to go over the sensor nodes in their vicinity. As a result, these would quickly run out of energy and eventually it becomes impossible to reach the servers. It depends on several factors how pronounced this problem is for the special case of sink nodes. When the sink is mobile, sensor nodes are part of a hot spot only for a short time. In networks where the sensors are not transmitting continuously but only upon occurrence of an external event, the problem is not as pronounced as it would be for the periodic transmissions from the sensors.

Scalability issues also motivate the need for *topology control* in sensor networks [55, Chap. 10], [90]. The underlying problem is that in densely deployed networks with several thousands of nodes where each node might have many one-hop neighbors, any routing algorithm would have to consider a very large number of routes. The goal for topology control algorithms is to explicitly reduce the number of available routes. One approach is to reduce the numbers of neighbors of individual nodes, for example by controlling their transmit power. In another class of approaches, additional structure is imposed on a sensor network. For example, the sensor field is partitioned into clusters such that simple cluster members communicate only with their clusterhead and routing then takes place among clusterheads. The construction of backbones is another class of approaches.

V. IEEE 802.15.4, ZIGBEE, AND ISA SP-100

The IEEE 802.15.4 standard describes physical layers and MAC layers for low-energy and low-rate wireless sensor networks and wireless personal area networks (WPAN). The complementary ZigBee standard covers the networking and application support layer. These standards are expected to play a key

⁵The authors of [80] assume that in automotive assembly plants up to 100.000 input/output points are present. To replace wires for a non-vanishing fraction of these points would require thousands of sensor nodes.

role for the deployment of WSN technologies in various application areas, including industrial applications.

A. IEEE 802.15.4 Standard

The first version of the IEEE 802.15.4 standard was finalized in October 2003, an updated version appeared at the end of 2006 [91], adding new modulation schemes at the physical layer (the IEEE 802.15.4a standard [92] adds further physical layers) and correcting a number of MAC-layer issues. There are now different implementations of the standard available, both commercial⁶ and open-source implementations [93]. It covers the physical layer and the MAC layer of a low-rate wireless sensor network. The target applications of IEEE 802.15.4 are sensor networks, building automation, connecting devices to computers and others.

On the physical layer the 2006 version of the standard operates in three different frequency bands, offering different rates: in the 868.0–/868.6-MHz band at rates of 20/100/250 kBit/s, in the 902– to 928-MHz band at rates of 40 and 250 kBit/s, and in the 2.4– to 2.4835-GHz band at a rate of 250 kBit/s. In the 868-MHz band only a single channel is available, while ten channels are available in the 915-MHz band and 16 channels in the 2.4-GHz band. An IEEE 802.15.4 network selects one of these channels and performs all operations on the selected channel, frequency hopping is not used. For the ultra-wideband physical layer specified in IEEE 802.15.4a the channelization is different, see Section XII.

The IEEE 802.15.4 MAC layer is built on the assumption that in the envisioned target application areas it is often possible to attach at least *some* of the sensor nodes to permanent power sources. This is exploited by designing the MAC protocol in an asymmetric fashion: there are different types of nodes and these have different responsibilities. More complex protocol operations are pushed to the energy-rich nodes. In fact, IEEE 802.15.4 distinguishes two types of nodes: *full function devices* (FFD) can operate in three different roles: as a *PAN coordinator*, a *coordinator* or as a *device*. In contrast, a *reduced function device* (RFD) can only act in the role of a device. There is only a single PAN coordinator in a network, but there could be several coordinators. The PAN coordinator starts the network and controls its main operational parameters (e.g., center frequency, duty cycle). The PAN coordinators can communicate in a peer-to-peer fashion or they can construct tree or mesh networks. In contrast, devices have to associate to a PAN coordinator and only exchange packets with this, thus forming a star network with the coordinator. The standard leaves the decision on how mesh or tree networks are constructed to upper layers. The hardware/software protocol implementation of the RFDs is significantly less complex than for the FFDs.

The protocol offers two different modes: the *unbeaconed mode* and the *beaconed mode*. For both modes we focus on the case of a star network, where a number of devices are associated with a coordinator. In the unbeaconed mode the coordinator has to be awake all the time while the associated devices can sleep at their discretion. In the uplink direction (from devices to the coordinator) the devices use an unslotted nonpersistent

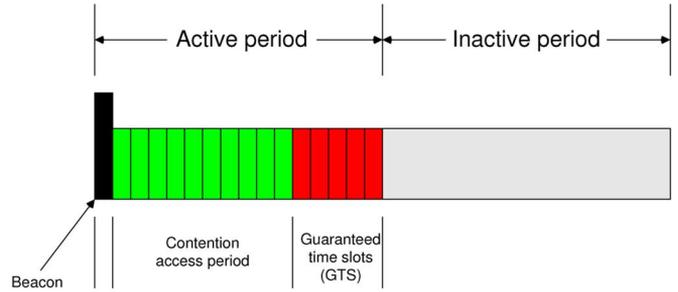


Fig. 2. Superframe structure of IEEE 802.15.4.

CSMA-variant (or an ALOHA scheme when using the UWB physical layer). The protocol inserts a random waiting time (a *backoff time*) before the actual carrier-sense operation is carried out. This is a collision-avoidance mechanism, it aims to reduce the probability that two or more nodes wishing to transmit at the same time (for example, because they are triggered by the same external event) collide. A coordinator buffers downlink packets and the protocol is arranged so that the devices have to explicitly retrieve those packets from the coordinator. A device sends a poll request using the CSMA-protocol, to which the coordinator responds with a MAC-layer acknowledgement. If a downlink packet is present, the coordinator transmits it in response to this request—the standard allows to perform this data transfer either with or without invoking the CSMA mechanism. By this approach, a device can follow its own sleep schedule without needing to coordinate its activities with anyone else. It depends on timing constraints, on the amount of buffer space available to the coordinator and the buffer management strategy for how long a device can sleep between successive poll operations.

In the beaconed mode the time is subdivided into consecutive *superframes*, the structure of a superframe is shown in Fig. 2. The superframe is subdivided into an active period and an inactive period. The length of the superframe and the relative length of the active period within a superframe (i.e., the duty cycle) are configurable—the parameter for the length of a superframe is called *beacon order* and the one for the duration of the active period is called *superframe order*.

In the inactive period all nodes, including the coordinator, can sleep. The active period is subdivided into 16 slots. In the first slot the coordinator broadcasts a *beacon packet* without performing any prior carrier-sense operation. At the end of the active period a maximum of seven *guaranteed time slots* (GTS) can be exclusively allocated to associated devices. The allocation happens upon explicit request and during a GTS slot the owning station simply transmits or receives packets without performing any carrier-sensing operation. In the remaining slots of the active period (called contention access period, CAP) the nodes compete for the medium using a slotted CSMA-scheme. To reduce the probability of collisions at the slot boundaries, a random waiting time and a prolonged carrier-sense operation are used. The CAP slots are entirely used for uplink transmissions, whereas the guaranteed time slots can be used for both downlink and uplink packets. Please note that there are no separate time slots for downlink transmissions to nodes having no

⁶For example the Z-Stack of Texas Instruments, see <http://focus.ti.com/docs/toolsw/folders/print/z-stack.html>.

guaranteed time slot. Instead, a device is notified in a beacon packet that there is pending data, and in turn it invokes the same poll mechanism as in the unbeaconed mode. Please note also that both the unslotted and the slotted CSMA schemes have no mechanisms for eliminating hidden-terminal problems. The coordinator's transceiver has to be switched on during the entire active period, but devices can spend the largest part of the active period in sleep state, they only need to receive the beacon packets.

B. ZigBee Standard

The ZigBee standard was prepared by an industry consortium, the ZigBee alliance.⁷ The first version of the standard appeared in 2005, the second version appeared in December 2006 [94]. As of this writing, the ZigBee alliance works on a new version, ZigBee 2007. ZigBee covers the networking layer and application layer of sensor network applications and is defined to work on top of a modified version of the 2003 IEEE 802.15.4 standard. In addition, the ZigBee standard describes a number of security services.

ZigBee allows to create different kinds of networks: in *star networks* the *ZigBee coordinator* starts the network and all the other network members (the *end devices*) are directly associated with the ZigBee coordinator. The ZigBee coordinator is co-located with the PAN coordinator of the underlying IEEE 802.15.4 network. In the two other types of networks there is a third type of ZigBee nodes present, the *ZigBee routers*, to which end devices can associate, but which can also associate themselves to other ZigBee routers or the ZigBee coordinator. The end devices can associate to any ZigBee router or coordinator in their vicinity. In *tree networks* the ZigBee routers form a tree that is rooted at the ZigBee coordinator, whereas in *mesh networks* the network topology might be a general mesh involving ZigBee routers and the ZigBee coordinator.

The ZigBee standard provides two different routing schemes, but there is no transport protocol that ensures end-to-end reliability (specifically there are no end-to-end acknowledgements), only hop-by-hop reliability is available from the MAC retransmission of the underlying IEEE 802.15.4 protocol. The first routing scheme is *tree routing*, in which all packets travel along the edges of the tree network. More specifically, an end device generating a packet sends it to the ZigBee router/coordinator to which it is associated. If the destination address refers to a child of the router/coordinator, then the packet is routed downwards the tree. Otherwise, the packet is first routed upwards to the ZigBee coordinator, which then starts routing it downwards using the correct branch. The correct branch can be inferred from the destination address—the address allocation to ZigBee routers and end devices is tightly coupled to the structure of the tree (“Cskip” addressing algorithm). The tree routing approach is a natural choice for networks in which all of the traffic is directed towards a sink node. A second advantage is that the ZigBee routers do not need any routing tables, they only need to know their parent router and the child routers as well as their associated address ranges.

For networks with peer-to-peer communication another routing scheme is used. In fact, the tree routing scheme can become very inefficient when two ZigBee routers in different branches of the tree but in mutual radio range want to communicate with each other—in the tree routing scheme all packets would have to travel through the ZigBee coordinator. Therefore, an additional routing scheme can be used in parallel to tree routing. This additional routing protocol leverages mesh networks and is based on the *ad hoc* on-demand distance vector (AODV) routing protocol [95], [96]. AODV is a reactive routing protocol, i.e., routes are only computed when they are needed the first time (“on demand”). This is in contrast to proactive or table-driven routing protocols, in which all routers constantly maintain a routing table even when no packets need to be transmitted. Reactive routing protocols are favorable when communication happens very rarely. In this case it is more energy-efficient to avoid *a priori* setup and maintenance of a routing table and instead to find routes and create routing tables on demand. The drawback of reactive protocols, however, is that the route discovery adds additional latency to packet transmissions. This additional latency does not occur in table-driven protocols. In ZigBee's AODV variant the intermediate nodes involved in an active route keep their routing table entries persistently. To deal with changes in the topology (for example failure of intermediate nodes) route repair mechanisms are included in the protocol. Therefore, the ZigBee AODV variant is more suited for static networks.

On the application layer ZigBee introduces a number of concepts. ZigBee provides an application support sub-layer (APS) on top of the network layer. The APS provides a number of end-points (similar to TCP/IP ports), to which up to 240 application objects can be attached. A dedicated end-point is reserved for the ZigBee device object (ZDO), which is responsible for service and device discovery and the management of so-called *binding tables*. The space of possible application objects is structured into so-called *profiles* which typically cover a class of applications, for example home automation. The set of possible devices within a profile is structured into *clusters*, which may for example represent inputs or outputs of real devices. A *binding* connects different clusters and can be considered as an abstraction for data transport. An example is the connection of a light switch (to be more precise: its digital output) to a light bulb (its digital input). Bindings can be made in different ways. With a *direct binding* the device owning the light switch (the source device) is configured with the address of the device containing the light bulb (the target device), and tree routing or mesh routing can be used for packet transmissions. While mesh routing can be beneficial in terms of performance, direct binding has the disadvantage that a change in the network address of a target device potentially needs modifications in many source devices talking to this target. Furthermore, the source device needs storage for potentially many target devices, for example when one light switch is responsible for many light bulbs. When using *indirect bindings* the ZigBee coordinator can store bindings in a table. A source device simply forwards its packet the tree upwards until a router/coordinator with an appropriate binding is found. After resolving this binding (looking up the destination addresses), the packet is forwarded to the right target

⁷<http://www.zigbee.org>.

device(s). Using a centralized binding table eases the maintenance of bindings.

Some general considerations about ZigBee in industrial applications can be found for example in [97]–[99]. The upcoming version of ZigBee⁸ adopts a frequency-hopping scheme: a ZigBee network can switch to another channel if too much interference is perceived on the present one.

C. ISA SP-100

The Instrumentation, Systems, and Automation Society (ISA) is currently working on a series of standards addressing the adoption of wireless technologies in different industries. In a first step the process industry is addressed with the upcoming ISA-SP100.11a standard, which is expected to be published in the second half of 2008.⁹ Later revisions of the standard will also address factory automation and building automation applications.¹⁰

ISA-SP100.11a addresses noncritical process applications that can tolerate delays up to 100 ms. Since it leverages the IEEE 802.15.4 standard in the 2006 version [91], it inherits some of its properties: low rates (up to 250 kBit/s) and low implementation complexity for simple end devices. Similar to the upcoming version of ZigBee, ISA-SP100.11a will adopt a frequency-hopping scheme in the 16 frequency channels offered by IEEE 802.15.4 in the 2.4-GHz ISM band. The goal is to improve robustness against static or frequency-hopping interferers, and also again to improve the reliability by exploiting frequency diversity. In addition, further frequency management schemes like blacklisting (i.e., configuring frequency bands that shall never be used in frequency hopping) or adaptive frequency hopping (learning which frequency bands are crowded) will be available. On the MAC layer ISA-SP100.11a again leverages from the IEEE 802.15.4 standard in the 2006 version [91], but in addition a data-link layer and an adaptation layer between MAC and data link layer are introduced. The data link layer controls the frequency hopping and adds a TDMA scheme.

ISA-SP100.11a networks can have a rich structure. On the one hand, wireless networks can be multihop networks having a mesh topology built among devices with routing capabilities, to which simple nonrouting devices are associated. Such a network is also called a *mesh*. On the other hand, different meshes can be interconnected via a backbone, and routing support for inter-mesh communications is available. Furthermore, a mesh can have one or more gateways to other types of plant networks. ISA-SP100.11a offers also some transport layer functionality, including unacknowledged (best effort) as well as end-to-end acknowledged data transfer, flow control and services like fragmentation and reassembly.

It will be very interesting to see the future development of ISA-SP100.11a, especially as compared to ZigBee. Both standards target overlapping application areas and are based on the same underlying wireless technology.

⁸See http://www.zigbee.org/imwp/idms/popups/pop_download.asp?contentID=11925.

⁹<http://www.isa.org/isasp100/>.

¹⁰The information contained in this subsection is based on the little available published information: http://www.isa.org/source/ISA100_Big_Picture.pdf and http://www.isa.org/source/ISA100.11a_Release1_Status.ppt.

D. Research Issues for IEEE 802.15.4 in Industrial Environments

A first key research issue pertains to the performance of IEEE 802.15.4. Although a number performance assessments have already been conducted (some examples are [100]–[102]), to the best of our knowledge there are not many publications available (one example is [103]) that study the achievable industrial-QoS performance as defined in Section II-A of these protocols in industrial environments.

The IEEE 802.15.4 MAC scheme has some deficiencies that specifically harm its real-time and reliability properties. Consequently, research activities with the goal to overcome these deficiencies have started in the past few years. A first deficiency is the inability of the IEEE 802.15.4 MAC protocol to deterministically enforce priorities on the channel, see the discussion in Section III-B. The authors of [104] have suggested a scheme for stochastic priority enforcement that modifies certain parameters of the IEEE 802.15.4 CSMA-MAC protocol. In particular, the length of the prolonged carrier-sensing period and the distribution of the backoff times are influenced by their approach. In the same vein it is conceivable to vary the maximum number of MAC-layer retransmissions on a per-packet basis to provide differentiated reliability levels. There is, however, a problem with this kind of approaches: the data transmission service primitives of the IEEE 802.15.4 MAC protocol do not provide any means to express priorities or to set the relevant MAC operational parameters *on a per-packet* basis. The only way to change these parameters is through management service primitives, but for these no timing guarantees are given in the standard. It is not specified whether such a management service primitive is handled instantaneously or with any delay. This leads to the unfortunate situation that the kind of per-packet parameter control required for approaches like presented in [104] might work for some implementations and not for others.

A second fundamental problem of the IEEE 802.15.4 standard is the possibility for *beacon collisions* in the beamed mode when two or more co-located IEEE 802.15.4 WPANs work in the same frequency band. As explained above, a coordinator does neither perform any carrier-sensing operations when transmitting its beacon, nor provides the IEEE 802.15.4 standard any other means to detect and resolve such a situation.¹¹ A similar problem are GTS slot collisions of co-located WPANs, i.e., scenarios in which the GTS slots overlap even though the beacons do not.

These problems have also been identified in the ZigBee standard [94], where beacon scheduling is introduced for tree topologies (only in these the usage of the IEEE 802.15.4 beamed mode is allowed). More specifically, the ZigBee standard foresees that the ZigBee coordinator assigns parameter values for the beacon order and the superframe order, as well as relative positions of the respective beacons in time (their phase) to each network member.¹² While the standard does not prescribe

¹¹Letting coordinators detect beacon collisions is not helpful anyway, for example when the coordinators are hidden to each other but their transmission ranges overlap at some network member.

¹²The association of a phase shift to an underlying IEEE 802.15.4 MAC requires an extension to the MAC standard. Furthermore, all ZigBee network members must be properly time-synchronized.

any specific allocation scheme, it should be ensured that the active periods of ZigBee coordinators/routers in mutual range do not overlap at any time. Specific allocation schemes have been discussed in [105] and [106], while the issue of GTS slot collisions is investigated in [107].

Some further problems with the IEEE 802.15.4 standard are as follows:

- The data exchange is not very efficient. First, there is no possibility to transmit data in acknowledgement packets as it can be done in PROFIBUS—when the coordinator wants to send output data to a device and wants to receive input data from the same device, four packets are needed in total, two data packets and two acknowledgements. Even worse, such a mutual data exchange with a node that has no pre-allocated GTS slot (there are only seven of them) would require six packets in total.
- The error control mechanisms are not very sophisticated—coding is not available in the 2006 version¹³ of the standard, nor are there any physical layer adaptation mechanisms that allow to vary physical layer parameters (like transmit power) from retransmission to retransmission in response to channel feedback. To facilitate the design of error control mechanisms, a better understanding of the error characteristics of IEEE 802.15.4-compliant transceivers in industrial environments will be useful [108], [109].
- The management of GTS slots is limited: a GTS slot is allocated to one device in every superframe, there is no means to allocate a slot to two or more devices in an alternating fashion, nor is there any means for devices to express corresponding allocation requests.

VI. REAL-TIME AND RELIABILITY IN MULTIHOP WSNs

In this section we discuss end-to-end real-time and reliability¹⁴ in multihop wireless sensor networks, which are key properties when sensor networks shall be used in, for example, distributed control applications [60], [67], [111], [112]. In multihop sensor networks timeliness and reliability are not only harmed by channel errors, but many of the problems that are well-known from the Internet are present as well. Two notable examples are problems related to congestion in sensor networks [113] and subsequent packet losses, or problems related to queueing delays.

A key theme in wireless sensor networks is to jointly design protocols and application to just fulfill the mission of the sensor network and not to attempt to provide very general and flexible platforms that can be used for many applications. Therefore, there has not emerged a single transport and routing solution that satisfies all needs, but a number of specialized solutions for different settings have appeared: one class of protocols

¹³The ultra-wideband physical layer specified in IEEE 802.15.4a includes error control coding.

¹⁴By the term reliability we refer to reliability of data packet delivery. With wireless sensor networks, however, reliability has some further important aspects: Is there any chance to actually detect the phenomena the network is supposed to detect? Are there enough sensors of the right modalities present? Is the area of interest sufficiently covered by sensor nodes? This is the subject of the *coverage and deployment problem* [110]. The sensors are cheap and their readings can thus be noisy, giving rise to the *information accuracy problem*.

addresses the reliable delivery of large blocks of program code from a sink node to all nodes in the sensor network to facilitate network re-programming [114]–[116], whereas others are concerned with the reverse direction, i.e., reliable delivery of environmental data from the sensors to the sinks [117]–[119]. The available protocols can be classified according to different criteria (compare [55, Chap. 13]):

- size of the items to be delivered: individual packets, larger but finite blocks of data, or infinite (periodic) streams of data;
- desired level of guarantee: guaranteed or stochastic delivery (within a prescribed deadline);¹⁵
- involved communication partners and direction: sensors to sink, sink to sensors, or sensor to sensor.

Before continuing, we want to make two general observations about real-time and reliability in wireless sensor networks:

- There is a clear tradeoff between reliability and good real-time properties on the one hand and energy-consumption on the other hand. The higher the reliability requirements, the higher the efforts for retransmissions, coding, redundant packets or other error control schemes. Although the average per-node energy consumption can be decreased by providing many redundant nodes that share the work, there is still the overhead required to coordinate the redundant nodes.
- Traditionally, multihop real-time networks rely on advance resource-reservation in intermediate nodes. This approach is not easily applicable in wireless sensor networks because intermediate sensor nodes sleep most of the time and are in general very unreliable—state in sensor networks should be considered as ephemeral. Therefore, either resource reservation is dropped, the reservation methods are adapted to the high network variability (for example by introducing methods for path repair), or the data packets themselves carry additional control information that lets a forwarder node decide on the proper resource allocation for the packet at hand.

To limit the discussion, we concentrate in the remainder of this section on the delivery of individual packets or finite blocks of data from sensors to a sink. For these two problems we discuss some important points in the protocol design space.

A. Single-Packet Delivery

For the case of single-packet delivery most of the available literature focuses on stochastic guarantees and accordingly the most important performance measure in terms of reliability is the probability of packet delivery at the sink.

1) *Usage of Acknowledgements:* Retransmissions are a very important building block to achieve reliable end-to-end transmission. In the case of single packet delivery only positive acknowledgements can be used, i.e., the receiving node must confirm successful reception instead of indicating a failure in case of unsuccessful reception—in many situations the receiving node simply cannot be aware of the existence

¹⁵The word “guarantee” is of course not to be interpreted in a strict sense; it simply means that in the deterministic case higher effort is spent to ensure delivery than in the stochastic case.

of a packet that it did not receive. Retransmissions can be made either hop-by-hop (at the link-layer), end-to-end (transport-layer) or in a combined fashion. Some studies have shown that link-layer acknowledgements and link-layer retransmissions should definitely be used [55, Chap. 13],[120], since relying only on end-to-end acknowledgements has significant drawbacks: 1) without per-hop link-layer retransmissions the probability that the packet reaches the sink is much lower, and therefore many end-to-end retransmissions are needed on average; 2) setting the retransmission timer at the source sensor requires global network information (e.g., hop distance to destination); and 3) the distribution of work among forwarders is uneven: forwarders that are very close to the source sensor are involved in almost every end-to-end trial, while forwarders close to the destination are used much less often. It should be noted that link-layer acknowledgements can either be explicit acknowledgement packets or *implicit acknowledgements*. In the latter, after A has forwarded a packet to B , node B does not send an acknowledgement but forwards the data packet further. When A overhears the data packet, it accepts this as an acknowledgement. This approach saves the acknowledgement packet but is only useful if the data packets are themselves small, since it is less probable that A can successfully overhear a long data packet than to overhear a short acknowledgement packet. If no end-to-end retransmissions are used at all, link-layer retransmissions might be combined with additional redundancy based on error-control coding [120], [121].

2) *Networks Without Routing Infrastructure*: When no routing structure is available at all, the only reasonable strategy available to the sensor node is to flood the whole network with the data packet in order to reach the sink.¹⁶ Flooding achieves a very high delivery probability, but unfortunately at very high energy costs [122]. It is therefore only an alternative for very infrequent messages that are extremely urgent—like for example messages indicating a leakage.

3) *Networks With Routing Infrastructure—General Considerations*: The majority of schemes assume the presence of routing structures. With routing in place, several options exist to improve the reliability.

When only one path is available (like for example in tree routing), then the source and the forwarders might apply MAC layer retransmissions or coding. They could also send multiple copies of the packet to the next forwarder, and the forwarders can behave in the same way (as in [121]). This corresponds to a simple repetition code. However, the single-path approach has the problem of being vulnerable to prolonged error bursts on one of the involved channels.

When multiple paths (either disjoint paths or braided paths) are available, these can be used in different ways. One option is to let the source sensor use a default path, for example a particularly energy-efficient path between source sensor and sink.¹⁷ If

¹⁶There are other strategies available as well, for example based on gossiping or random walks. However, we do not consider those as reasonable in the context of real-time systems.

¹⁷It is in general in sensor networks not a good idea to insist on using the same path forever, even if it is at setup time the lowest-energy path. By repeatedly using this path, the forwarder nodes quickly run out of energy. It is therefore advisable to either switch the path from time to time or to employ randomization in the path selection, leading to paths that are longer on average [123].

the default path fails, the source selects another path as default path. However, this approach requires feedback from the sink and therefore end-to-end acknowledgements. Another option is to let intermediate nodes that face problems on a link perform a local route repair operation.

The second option is to let the source send identical copies of the packet on multiple paths—again a simple repetition code. When the forwarders do not themselves create multiple offspring packets, then the initial number of paths must be carefully chosen by the source sensor to balance out the probability that at least one copy reaches the sink versus the network resources used. When forwarders can create offspring packets, their generation rate must be properly controlled to avoid explosion or extinction. For example, in the ReInForM scheme [122] each forwarder node locally decides on the number of offspring packets based on the following information: 1) its hop distance to the sink; 2) a local estimate of the packet error probability; and 3) control information contained in the packet. The local error probability and the number of hops to the sink is used to obtain an estimate of the probability of losing a packet on a single path. Based on this, the number of offspring packets and paths is chosen so that a pre-specified delivery probability at the sink is maintained. One of the effects of including the remaining number of hops into the calculation is that forwarders closer to the source sensor create more offspring paths than forwarders close to the destination.

A third option is to let the source node split the packet into k smaller fragments. Out of these k fragments a number of $m > k$ fragments is created by applying a certain coding scheme known as erasure codes or fountain codes [124], [125]. These codes have the property that it suffices for the receiver to successfully receive any k out of the m coded fragments to successfully decode the packet. This has for example been explored in [120].

4) *Tree Networks*: One often-found routing structure in sensor networks are trees rooted in sink nodes—it can be expected that this structure will be common in industrial settings. During a tree setup phase each node learns shortest or best routes to one or several sinks—more specifically, a node does not need to learn entire routes, but it suffices to learn about those immediate neighbors through which the root can be easily reached. In this sense the nodes do not need to maintain global information. In the collection tree protocol (CTP) that is delivered with the TinyOS 2.0 sensor node operating system [126] a sensor node maintains a path to a sink that has the lowest accumulated costs, where the per-link costs reflect the average number of trials needed to successfully transmit a packet over this link (called expected number of transmissions—ETX—in [126]; see also [127]). At the same time a sensor node maintains a neighbor table in which selected neighbors and their ETX values are stored. When for a transmitting node its preferred link breaks down, it can consult the table and choose another neighbor. A similar approach has been evaluated in [120].

5) *Combining Real-Time and Reliability—MMSPEED*: The MMSPEED scheme [119] is an extension of the SPEED scheme [128]. A key feature of MMSPEED is that it addresses both real-time and reliability in a separated manner. The MMSPEED protocol rests on a few important assumptions: 1) all nodes

know their geographical location, 2) the location of the packet destination is known, and 3) the underlying MAC protocol allows to prioritize between different classes at least stochastically (see Section III-B). The basic concept of *speed* can be explained as follows: suppose a forwarder X transmits a packet to a forwarder Y being d meters closer to the destination than X is. Furthermore, X possesses an estimate of the time t it needs to transmit the packet to Y , for example including queuing delays, MAC delays and propagation times. Then the achievable packet speed of forwarding to Y is given by d/t . The network offers a number of pre-defined speed levels to the application. A source node selects the required speed according to deadline and distance to destination. The data packet is transmitted labeled with the relative deadline, elapsed time and destination position. When receiving a packet, a forwarder makes a new choice on the required speed according to the elapsed time (as compared to deadline) and his own distance to the destination. If the required speed is higher than highest available speed, the packet is dropped. Otherwise, the packet is placed into a packet queue associated to the selected speed level, and an appropriate forwarder supporting this speed is selected. One such queue corresponds to one MAC priority level and higher speeds are assigned to higher priorities. A forwarder serves the queues in priority order, within a queue FIFO service is applied. Immediately before the packet is transferred to the chosen forwarder, the packets elapsed time field is increased by the time that the packet spent in the current forwarder. This overall approach allows for some compensation: when time is saved on the first few hops, the last hops have some more time at their disposal.

MMSPEED controls the reliability by sending copies of the packet over multiple paths in parallel. More specifically, each forwarder node decides locally about the number of offspring paths based on the currently observed error conditions for all its immediate neighbors.

B. Block Delivery

In the case of block delivery a large data block (for example a picture taken by a camera) is split into a finite number of *fragments*, each fragment corresponding to one packet. The available protocol designs aim to ensure that blocks of packets are *completely* received by the sink. One important performance measure is consequently the probability that this goal is reached within a pre-specified time.

1) *Usage of Acknowledgements*: As compared to single-packet delivery there are further options. Most notably, the usage of negative acknowledgements (NACKs) can be incorporated into the protocols. This potentially reduces the number of acknowledgement packets: over links with very low packet loss rates it is more economic to send a NACK packet in the unlikely case that a packet is not received than to send a positive acknowledgement (ACK) packet upon every correctly received packet (i.e., very often). How does a receiving forwarder know that packets are missing and that it should send a NACK packet? One approach is to label all packets of the block with sequence numbers. If a receiving forwarder gets packets 1, 2, 3, 4, 6, 7, ... from its successor, then it knows

that packet 5 is missing in the moment in which packet 6 is received. Correspondingly, it sends a NACK packet, requesting retransmission of packet 5. This criterion for loss detection fails in two different cases: when the last fragment is outstanding or when no fragments have been received at all. The first problem can be circumvented by adding a field indicating the total number of fragments to each fragment, the latter problem must be solved by other means. For example, in GARUDA [115] high-powered energy signals are used to indicate the presence of a block.

A NACK can be regarded as a retransmission request in a selective-repeat-type ARQ protocol. When any successor node on the forwarding path has cached the requested fragment, it can intercept the NACK and send the missing fragment to the requesting node. Otherwise, the successor node forwards the NACK further towards the source sensor. When caching is used in the network, NACKs do not need to travel the whole path back to the source sensor. A caching node is also called a *recovery server* [115], [129]. In an extreme case, all nodes in the network spend some buffer for caching.

2) *One Example—RMST*: The RMST scheme [117] adds reliable data transfer to directed diffusion [130]. Directed diffusion can be regarded as a publish/subscribe mechanism [28] in which the sink subscribes to data which it specifies not by giving the addresses of the sensors producing the data, but by directly describing the data it wants in terms of attributes. After disseminating the interest message via flooding into the network (“interest propagation phase”), the directed diffusion protocols create in a distributed fashion a single path from sensor nodes publishing the data to the sink (“enforcement phase”).

The RMST scheme combines several mechanisms to enforce reliability:

- MAC-layer retransmissions.
- In RMST’s *cached mode* the sink node and all intermediate nodes on an enforced path cache fragments and check their cache periodically for missing fragments. When a node A misses some fragments, it generates a NACK message specifying the missing fragments, and forwards it along the path down to the source sensor. When a node on this path, say B , has one or more of the missing fragments, it transmits them back to A . If B has not transmitted all missing fragments, it B propagates the NACK further towards the source (those fragments which B could serve are removed from the NACK), otherwise the NACK is suppressed. In the *noncached mode* of RMST only the sink node has such a cache; therefore, NACK’s have to travel back to the source node.
- The source sensor sends out the whole data block periodically until the sink explicitly unsubscribes. This gives intermediate nodes an opportunity to complete their cache contents.

By investigating different combinations of the above mechanisms for the total number of bytes (data plus overhead) needed to transmit 50 fragments of 100 bytes size, it showed up that MAC-layer retransmissions are helpful in case of higher packet loss rates, but interestingly using the cached mode *without*

MAC-layer retransmissions (and thus without MAC layer overhead like acknowledgements) is the cheapest approach (given that *all* intermediate nodes cache fragments).

C. Research Issues

Multihop protocols for wireless sensor networks are a vast research area and many groups are active here. Therefore, we narrow down the discussion to a topic which we believe is especially interesting for the industrial community.

As we already mentioned, in our opinion tree-like networks, where a number of sensor nodes form a tree topology and transmit their data to the sink node at the root of the tree, will be important in industrial settings. For this kind of networks some interesting research questions could be:

- Design of scalable tree protocols that do not rely on individual sensor nodes to fill positions in the tree but that allows several nodes to share this burden while ensuring that enough nodes are awake and the required service can be performed (for example [5]). This is especially important for tree positions close to the sink, which are traversed by many packets.
- Design of mobility management schemes of trees which account for the movement of whole subtrees with the corresponding sudden changes in the load situation of old and new subtree attachment points.
- A precise understanding of the performance of tree protocols in terms of real-time, reliability and energy consumption, and the dependency of these performance measures on certain deployment parameters like the sensor node density, the average number of potential parents in the tree, etc. It can be expected that these relationships depend crucially on interactions of the routing and forwarding protocol with lower-layer protocols like MAC protocols and link-layer error control.
- Methods and tools for proper planning of protocols for tree-like networks. One example of such an approach is given in [5].

VII. WISA WIRELESS INDUSTRIAL COMMUNICATION SYSTEMS

The wireless interface for sensors and actuators (WISA) system, described in [80], was developed by ABB and is now commercially available. The system provides solutions for wireless communications and for wireless power transmission (using a magnetic-field-induced power supply).

A WISA network is a star network, where a number of wireless devices are associated to a base station (BS). The envisaged wireless devices provide input/output points towards the underlying manufacturing process, and the BS is expected to be connected to a central controller. Up to 120 I/O devices can be connected to a BS.

The physical layer is based on Bluetooth/IEEE 802.15.1 [131], it provides raw data rates of 1 MBit/s at 1 mW output power (this gives a transmission range in the order of ten meters), operates in the 2.4-GHz ISM band and employs frequency hopping.

Uplink and downlink traffic are in WISA separated by frequency-division duplex (FDD), i.e., uplink and downlink traffic can be transmitted truly in parallel over different frequency bands. In the downlink direction the BS is the only transmitter and it transmits continuously. In the uplink direction a time-division multiple access (TDMA) scheme is combined with multichannel transmission. More specifically, the time is divided into superframes of 2048 μ s length. A superframe is again subdivided into 30 slots, each slot can accommodate a packet of 64 bit length (including physical layer preamble, checksum and control fields). To support up to 120 stations, four uplink channels are used in parallel. A single wireless device is exclusive allocated a transmit opportunity in one of the four uplink channels and in one fixed packet slot. Together with a single channel for the downlink direction, five channels are active at any time in the WISA system. The wireless devices can be equipped with standard Bluetooth transceivers, while the BS needs a modified transceiver that can receive on four channels in parallel.

Frequency hopping is performed for every superframe, and all five active channels hop synchronously according to a well-known hopping pattern. The hopping pattern aims to maintain a frequency separation between subsequent jumps that exceeds the coherence bandwidth (found to be a few tens of MHz) and the bandwidth of static IEEE 802.11 networks of 22 MHz. By this, the channels in subsequent superframes experience channel fades independently of each other. This frequency diversity improves the transmission reliability, which is further increased by acknowledgements and retransmissions in the next superframe. Furthermore, when the hopping sequences are properly chosen, multiple WISA networks can operate at the same place without creating excessive interference to each other.

The performance measurements for WISA presented in [80] indicated that for a 60 nodes network with 70 cm distance between wireless devices and BS the fraction of packets that need four or more retransmissions (corresponding to delays of 8 ms and more) is below 10^{-4} for a range of different interference situations (e.g., interference from other WISA networks, from static IEEE 802.11g interferers, from Bluetooth).

VIII. SOME FUTURE OPPORTUNITIES FOR SYSTEM AND PROTOCOL DESIGN

In Sections IX–XII we turn our attention to some theoretical and technological developments that have so far not been so massively considered in the industrial community, but which in our opinion really should. They all address, in the one form or the other, the core problem of wireless industrial networking: providing the required levels of timeliness and reliability despite the unfriendly error properties of the wireless channel.

IX. COMBATING CHANNEL FADING WITH SPATIAL/COOPERATIVE DIVERSITY TECHNIQUES

In realistic environments the wireless channel suffers from phenomena like path loss and shadowing, multipath propagation and thermal noise [132]–[134]. On the digital level, these phenomena can lead to bit errors or even to the total loss of packets due to a receivers failure to acquire carrier or symbol synchronization [135]. The precise error characteristics depend on the

specific wireless technology, the carrier frequency, the distance and propagation environment between transmitter and receiver, and other factors. Measurements of the error characteristics in industrial environments have been presented in a number of studies and for different types of wireless technologies: Bluetooth [136], IEEE 802.11b [135], IEEE 802.11a [31] or IEEE 802.15.4 [109].

The wireless channel is often *time-varying*, with variations caused by mobility of wireless stations or in the propagation environment. The signal strength at the receiver can vary over time, thus creating a *fading channel* [137]. The channel variations can happen on different timescales. *Fast fading* is caused by a combination of mobility and multipath propagation, since movements give rise to quick changes in the mutual interference created at the receiver by differently delayed and phase-shifted copies of the same transmitted waveform. Fast fading typically happens on timescales of milliseconds to tens of milliseconds. On the other hand, when mobility leads to situations where obstacles are placed or removed from wave propagation paths between transmitter and receiver, additional attenuation results, also referred to as *shadowing*. The timescales involved are typically on the order of seconds to tens of seconds. Furthermore, when the distance between transmitter and receiver changes significantly, the attenuation coming from path loss can also change significantly. Given pedestrian-speed mobility, this often happens on the timescale of tens of seconds to minutes. This type of time-variation is also referred to as *slow fading*.

Another very important property of practical wireless channels is their *location-dependency*. With multipath propagation even a slight movement of the receiver suffices to change the absolute value and the statistics of the received signal strength (signal envelope). The degree of correlation in the signal envelope when varying the spatial position is captured in the spatial autocovariance function [138, Sec. 5]. Depending on the precise propagation characteristics, a distance of roughly half a wavelength between two positions can already provide a decorrelation of signal envelopes. For systems operating in the frequency range of several hundred MHz to a few GHz, movements of a few meters to centimeters suffice to achieve this (for the 2.4-GHz ISM bands the wavelength is ≈ 12.5 cm), thereby not changing the path loss/average channel attenuation. In summary, the wireless channel can be considered location-dependent, time-varying and random, and often significant error rates can be observed.

The classical error control methods (error control coding, retransmissions [139]–[141]) operate on a single wireless channel and add redundancy to the data packets to improve reliability. A common problem of all the classical methods is that their usefulness is seriously degraded over fading channels: when packet transmission starts in a deep fade and the deep fade lasts beyond the packet deadline, then these methods are usually not sufficient to successfully transmit the packet. A fundamental approach to circumvent this problem is the exploitation of *spatial diversity* [142]. In this class of mechanisms the single-channel restriction between a wireless transmitter and receiver is removed and information is transmitted over multiple spatial channels. The hope behind this is that due to sufficient spatial separation between antennas these channels

become stochastically independent and are with only small probability all in a deep fade at the same time.

In the remainder of this section we give a brief introduction to spatial diversity in general and to relaying, an instantiation of cooperative diversity, in particular. We also point out some research issues.

A. Fundamentals of Spatial Diversity and Cooperative Diversity

In *spatial diversity schemes* independent realizations of a transmitted signal are obtained from multiple antennas placed at geographically sufficiently separated locations. In the *single-user case* only a single transmitter node and receiver node are considered, and at least one of them has multiple antennas which can be perfectly controlled, thus achieving perfect cooperation among them. Recent multiple-input, multiple-output (MIMO) techniques [143], [144] like the upcoming IEEE 802.11n belong to this class.

There are two different types of gains that can be achieved from spatial diversity [145]. The first one, called *multiplexing gain*, characterizes the increase in achievable transmission rates over a fixed bandwidth. When transmitter and receiver use the same number N of antennas (with sufficient spacing), then the achievable channel capacity over fading channels increases at least linearly with N as $N \rightarrow \infty$ [142], [146]. The second one, called the *diversity gain*, is a measure of an increase in reception reliability. On Rayleigh fading channels the bit/symbol error probability depends on the signal-to-noise ratio (SNR) roughly like $1/\text{SNR}$ for large SNR values. A diversity gain of $n \geq 1$ is achieved when this dependency changes to $1/\text{SNR}^n$. In a scenario with a single transmit antenna and n receive antennas the full diversity gain of n can be achieved when fading is independent across the n antennas [145]. It is a fundamental result [145] that there are tradeoffs between these two gains—for an increased diversity gain there has to be a sacrifice in multiplexing gain. One approach to give these tradeoffs into the hand of a designer is to use space-time codes [147]–[149] which can be configured for both types of gains.

In the *multiuser case* further nodes are involved in a transmission between a transmitter and receiver—this is also often referred to as *cooperative diversity* [150], [151]. These external nodes can help with transmission or reception of data. However, as opposed to the MIMO case where all antennas are attached to a single node and can be perfectly controlled, multiple nodes now need to *coordinate* their antennas. Depending on the specific coordination scheme, losses in multiplexing or diversity gain might result from (lack of) coordination. One example are relaying techniques (discussed below), another one are cooperative MIMO approaches (see for example [152] and [153]). In cooperative MIMO, two groups of nodes form a virtual transmit and receive antenna array, respectively. When a transmitter node wants to transmit a packet to a receiver node, the transmitting node sets up a virtual transmit array of neighbored nodes, disseminates the packet to the array members, and each array member transmits a copy of the packet. On the other side, the receiver sets up a receive array. The receive array members receive (parts of) the incoming packets and forward their observations to the ultimate receiver, which then can try to decode the packet. The receive array members can try to arrange their forwarding activities so that their waveforms

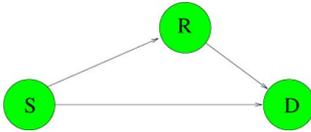


Fig. 3. Basic relaying operation.

combine coherently at the receiver (i.e., they arrive all with the same phase) and an improved signal-to-noise ratio is achieved. This, however, requires precise knowledge of the propagation environment as well as extremely precise time synchronization, which is practically impossible to achieve and maintain in face of mobility. A much less demanding option is to let the receive array members use orthogonal channels, for example they could transmit their observation packets separated in time but on the same frequency. To achieve this, a medium access control protocol is required.

In practical terms, for industrial applications multiuser techniques like the relaying approach discussed below are attractive. The individual nodes need only a single antenna, which reduces system complexity as compared to MIMO systems. The latter require significant signal processing complexity at the receiver side. Secondly, with multiuser techniques the spacing between the antennas can be larger than for true multiantenna nodes, where the typically small form factor of the node puts practical constraints on the number and spacing of mountable antennas. This larger spacing can be beneficial when obstacles block the direct line-of-sight between transmitter and receiver, since a third-party node can be used as a “detour” for transmitting information.

B. Concept of Relaying

The concept of *relaying* is not new, the first theoretical works date back to the seventies (see for example [154]; see also [155, Chap. 15]). In relaying schemes, there are a number of *relay nodes* that help in the transmission between a neighbored sender-receiver pair. All involved nodes can be single-antenna nodes. The relay nodes possibly receive the sender's packet and can for example unconditionally forward their observations to the receiver node, or they could assist the sender with performing retransmissions when the receiver has not received the packet.¹⁸ Relaying can hence be tightly coupled to ARQ protocols. A lot of information-theoretic research has been carried out to investigate multiplexing and diversity gains achievable with relaying (e.g., [156], [157]). In the last years, there have also been significant activities towards practical integration of relaying into wireless protocols, see for example [158]. In [159]–[161] different proposals for industrial settings have been discussed.

In its simplest form, the relaying ARQ channel consists of three nodes S , D and R ; see also Fig. 3. The source S wants to send a packet towards the destination D . A third node, the relay R picks up S 's signals and forwards its observations to D , who can combine R 's observations with his own ones to decode the packet.

¹⁸It should be noted that relaying, as described in this paper, is considered a link-layer technique, in which third-party nodes (the relayers) are used to stabilize an existing link between a transmitter and receiver. This usage of the term “relaying” should not be confused with the usage in multihop networks, which often refers to mere forwarding to connect non-neighbored nodes.

C. Research Issues for Relaying

The author is convinced that the incorporation of spatial diversity approaches into error-control for wireless industrial networks is a very promising approach to significantly improve the reliability and real-time properties, and that it is a very interesting area of future research. It is shown in [160] and [161] that relaying schemes can give appreciable benefits for industrial communication systems by showing that the success probability (as defined in Section II-A) can be significantly increased on fading channels as compared to classical error control schemes.

In more realistic cases, the broadcast property of the wireless medium (or sleeping activities of sensor nodes) leads to situations where a random number of relayers can be present. The key research issue here is how to coordinate and control the activities of these relayers in a deadline-aware fashion [160], [162]. When relaying is integrated into an ARQ protocol (stated differently: when the activity of the relayers depends on feedback from the destination), then all relayers need a consistent view on the destination feedback. Unfortunately, some relayers can be mutual hidden terminals (i.e., they are not able to receive packets from each other), or a relay can be hidden from the destination.

A second important research issue are rules and heuristics for network planning and deployment that give hints for good positions of relayers. It is intuitively clear that a relay is most helpful when placed right between source and destination, and that it is harmful when it is even farther away from the destination than the source node. In industrial networks it might be helpful to deploy extra nodes that do nothing else but to help with relaying, and for these nodes good positions need to be identified.

X. QUALITY OF SERVICE PROVISIONING AND ANALYSIS

In a narrow sense, the ultimate goal of research in the design of protocols for wireless industrial networks is to provide satisfying levels of industrial-QoS. A necessary precondition for this is to collect detailed measurements about wireless industrial channels (for example [31], [109], [135], and [136]). To support this, community initiatives for collecting wireless traces, similar to the CRAWDAD initiative at the University of Dartmouth might be useful.¹⁹

More generally, a satisfying theory and a set of accompanying tools for wireless resource management is needed. Both the theory and the tools should specifically address deadlines and high reliability requirements, and they should be aware of the different types of disturbances found on a wireless channel (noise, fading and attenuation, external interference). Consideration of deadlines and reliability makes resource management for wireless industrial communication systems different from classical resource management, where mostly the network throughput coupled with fairness constraints is in the focus. Wireless resource management is in general concerned with the assignment of channel resources like the transmit power, frequency, rates (as determined by modulation and coding scheme) or time budgets to stations [163]–[165]. A resource

¹⁹<http://crawdada.cs.dartmouth.edu/>.

management framework can be used in different ways: 1) in pre-deployment planning of the network to come up with an initial deployment plan; and 2) it can help guiding the dynamic network adaptation/reconfiguration that is needed to maintain the desired industrial-QoS in face of inevitable changes in the external interference situation or average signal attenuation levels.

In the network planning phase, a resource management generates an initial allocation that specifies for example 1) the association of wireless nodes to different wireless networks; 2) the assignment of wireless networks to one of a set of available orthogonal frequency bands; and 3) the transmit powers and rates that should be used in the different wireless networks (carefully taking into consideration the resulting mutual interference) For a candidate initial allocation each wireless network must be individually evaluated whether the real-time flows allocated to this network can receive the desired industrial-QoS. If not, a new candidate allocation must be generated. To assess an individual network for its achievable industrial-QoS, wireless channel effects like fading or time-varying external interference should be taken into account. Different resource allocations could be tested and the overall network design could be optimized to minimize cost or frequency usage subject to fulfillment of industrial-QoS properties (see for example [3], who optimize the number of base stations and their placement using a genetic algorithm; another reference is [166]).

Classical real-time scheduling theory [167] and the various schedulability analyses for classical fieldbus systems [168], [169] provide schedulability criteria that can be used to assess a network's capability to carry a given set of real-time streams. A set of streams is said to be schedulable when all deadlines can be deterministically met. In some works the analysis has been extended to include transient channel errors. For example, in [52] scheduling analysis and reliability analysis are combined to assess the effects of transient errors on CAN. In the author's view, this kind of investigations could be a useful starting point to extend scheduling theory towards better consideration of wireless channels. As already mentioned in Section II-A, the deterministic view on schedulability of a set of real-time streams should be replaced by probabilistic industrial-QoS measures that allow for losses—for example by specifying that a set of real-time streams is schedulable (for fixed assumptions on the wireless channel error process) when all of its streams achieve a long-term success probability of 99%.

To make this kind of statements, the dynamics of the wireless channel should be explicitly considered in the scheduling analysis. Appropriate classes of stochastic models for channel dynamics in industrial scenarios need to be identified, and theories need to be developed that allow to calculate the industrial-QoS of a set of real-time flows relative to a chosen model. The model type will assumedly depend on the considered timescales and the nature of the channel variations (slow fading versus shadowing versus fast fading; see Section IX). One class of stochastic models that has been successfully applied to short-term fading are finite-state Markov channels (FSMC) [170]–[175], in which the signal strength at a receiver is varied according to a

Markov chain with a finite number of states. FSMC models are on the one hand complex enough to model a wide range of wireless channels and on the other hand they are still simple enough to be treated analytically. In the authors view, a systematic development of probabilistic scheduling criteria for FSMC channels (a “scheduling theory for fading channels”) could represent a significant step towards the creation of network planning tools, since they allow to assess achievable industrial-QoS over a wide range of channel behaviors.

It should also be mentioned that in the industrial communications community the formalism of network calculus [176]–[178] has also been used for *a priori* analyses of worst-case timings, for example in industrial switched-Ethernet networks [179]. A recent line of research in the networking and information theory community works on the development of formalisms for *stochastic network calculus*, in which deterministic end-to-end time bounds are replaced by stochastic ones (for example [180]). It remains to be seen how successful such formalisms can be adapted to the domain of wireless industrial communications.

It is clear that the channel models used during the network planning process (and also for later admission control of dynamically arriving real-time streams) should represent the actual channels found in an industrial plant. It is therefore desirable to develop methodologies that enable quick online model identification and parameterization from actual measurements.

A second important line of research does not apply in the network planning phase, but during network runtime. In many industrial settings wireless channels, even when their characteristics have been precisely modeled from measurements, will change over time, and the network needs to adapt its resource usage to those changes. This requires on the one hand criteria that allow a node to detect (significant) changes in the channel, for example from continuously monitoring the actual industrial-QoS and applying (to-be-designed) detectors for significant changes. On the other hand, the network must have the ability to re-configure, for example by exploiting software-defined radio concepts to change to another center frequency (see Section XI).

XI. INTERFERENCE MITIGATION USING COGNITIVE RADIO TECHNIQUES

Multipath fading is only one source of channel errors, external interferences are another one. External interferences are created by other transmitters operating in the same or in neighbored frequency bands. Other transmitters can be stations of the same or different wireless technologies working in the same band, or in industrial environments it could be machinery like arc welders, motors or power electronics.²⁰

This problem is especially pronounced in the license-free ISM (industrial, scientific and medical) bands. A range of different technologies share the 2.4-GHz ISM band: IEEE 802.11 WLANs, Bluetooth/IEEE 802.15.1 WPANs, or IEEE 802.15.4

²⁰In [80] it was stated that many high-power industrial devices like arc welding machines, frequency converters etc. do not create appreciable interference beyond 1.8 GHz.

WPANs, and each person is free to buy wireless equipment working in this band and to transmit data (only subject to limitations in transmit power or power spectral density masks). In this situation it is very hard to maintain a given level of industrial-QoS over time when the 2.4-GHz ISM band is used—the interference level can change already if an employee brings a Bluetooth phone. The coexistence of different technologies on the same frequency band and the resulting performance impacts have been investigated extensively, see for example [80], [181]–[183] and in addition [1]. For critical industrial applications it is, however, more attractive to circumvent the interference problem than to live with it.

One possible solution, the allocation of an exclusive frequency band involves interactions with regulatory bodies, like for example the Federal Communications Commission (FCC), and is in general a lengthy and costly process. Another, more immediate solution would be to use other unlicensed frequency bands, like the 5-GHz ISM bands, which are currently less crowded. However, in the higher bands the frequency regulation is not fully harmonized in all countries, and furthermore this only postpones the problem for a while. The concept of *cognitive radio* [184]–[189] provides an approach to circumvent this problem.

The concept of cognitive radios is in turn built on the more fundamental concept of *software-defined radios* (SDR) [190], which, roughly speaking, follows the idea to perform (almost) all physical layer signal processing functions in software instead of hardware, for example on a digital signal processor. This gives substantial flexibility, in that it is much easier for software-implemented signal processing algorithms to expose control knobs to higher layers, or to be quickly replaced by other signal processing algorithms with a simple software update. These features allow quick reconfiguration of a SDR-based transceiver. This can for example concern modulation schemes, coding schemes, or center frequencies.

With *cognitive radios*, the idea is to exploit the flexibility of SDR by reconfiguring the radio according to the current state of the wireless terminal and the perceived state of its external environment. In full generality, the environment state can be anything for which the node possesses sensors, more specifically the radio could provide sensing mechanisms to check for the presence of signals in certain frequency bands.

In this context, one of the main uses of cognitive radio is *opportunistic spectrum access*. The motivation behind this idea comes from two observations: 1) electromagnetic spectrum is a scarce resource and license-free spectrum is crowded; and 2) if a spectrum analyzer is placed at a certain location, one will notice that many exclusively allocated bands are used only intermittently—there are *spectrum holes* which position depends on time and location [186]. A cognitive radio node can exploit these holes: if at his current operating frequency band and position the interference situation degrades, it can sense other frequency bands and seek for new spectrum holes. There is, however, an important constraint: the activity of the cognitive radio nodes must not create any interference to the primary/licensed user when it comes back. To quickly detect the return of the primary user, all nodes in a cognitive radio network must continuously sense the currently used frequency band for the spectral signa-

tures of the primary user and quickly agree on another common center frequency to be used in the future. To minimize disturbance to a primary user, all this should happen within very short time. Instead of specifically seeking for the primary users spectral characteristics, a cognitive radio node could also try to detect the presence of *any* other system in the same band, whether primary or secondary user.

In the author's view the concept of opportunistic spectrum access can bring significant benefits to wireless industrial networks, since it provides a very promising way to deal with external interferences. Much of the research in opportunistic spectrum access deals with physical layer issues like the development of quick and reliable methods for sensing large portions of the spectrum, methods for detecting the return of primary users, etc. With respect to protocols there is a need for signaling protocols within a cognitive radio network to signal presence of interferers to all nodes and to agree on another frequency band. In an industrial setting it is conceivable that different networks on the same industrial site (and belonging to the same administrative domain!) *cooperate* with each other by exchanging measurements or negotiating spectrum usage. Besides the protocols, also suitable *policies* for spectrum sensing (which piece to scan next, which network member performs the scan), for choosing the next frequency and for the cooperation among different networks in the same industrial site are worthy research topics.

It should be noted that the concept of opportunistic spectrum access is more general than frequency-agile approaches like the one that will for example be adopted in the upcoming version of ZigBee or in ISA-SP100.11a. The key difference is that ZigBee is still confined to the set of channels provided by the underlying IEEE 802.15.4 (all in the ISM bands), whereas cognitive radio approaches are not.

The IEEE 802.22 working group on wireless regional area networks is currently working on a standard for a wireless medium access control and physical layer that allows opportunistic spectrum access in spectrum that is exclusively allocated to the TV Broadcast Service.²¹

XII. ULTRA-WIDEBAND (UWB) TECHNOLOGIES

Ultra-wideband (UWB) technologies [191]–[195] provide another answer to the problem of interference, more specifically: of narrowband interference. According to the FCC, UWB communications takes place when the bandwidth W used for the communications is more than 20% of the carrier frequency used, but at least W should be 500 MHz. In 2002, the FCC allowed the usage of UWB devices for communications and measurement systems in the range between 3.1 and 10 GHz.²² As compared to the cognitive radio approach, UWB stations are not required to sense the medium for the return of the primary (licensed) user, but are allowed to fully use the spectrum as long as it is ensured that at the location of a UWB transmitter the power spectral density of its signal is below a spectral mask defined by the FCC. According to this specification, the UWB signal must not exceed -41 dBm radiated power. By this limitation, UWB signals are kept below the level of

²¹<http://www.ieee802.org/22/>.

²²http://www.fcc.gov/Bureaus/Engineering_Technology/News_Releases/2002/nret0203.html.

unintentionally generated interference that any licensed user must accept from other devices, and furthermore UWB is confined to short-range transmissions. Due to the large bandwidth, UWB is robust against narrowband interferences from licensed bands, and theoretically very high data rates are achievable over short distances. UWB technology is conceived for high-rate personal-area networking, and for lower-rate cable replacement applications.

There are different approaches to UWB radio. In the *impulse radio approach* no carrier frequency is used, but impulses that are very narrow in time and correspondingly occupy a very large frequency range.²³ Controlled variations in the relative positions of those pulses in time can be used to transmit information (pulse-position modulation). To overlay multiple UWB networks at the same physical location, this can be coupled with a time-hopping spread spectrum technique [196]. One example for an impulse-based UWB standard is the UWB physical layer specified in the IEEE 802.15.4a standard [92]. It is specified to work on four channels in the range from 3.1 to 4.8 GHz, eleven channels from 6.0 to 10.8 GHz and one channel in the range of 249.6 to 749.6 MHz. The availability of these channels depends on local regulations. The standard specifies that the ALOHA medium access control protocol²⁴ is used together with the UWB PHY. The UWB PHY uses error-correction coding and supports different data rates (110, 850, 1.700, 6.810 and 27.240 kBit/s).

In the *multiband approach* the available spectrum is subdivided into subbands, for example of 500-MHz width. Within such a subband different modulations can be used, including pulse modulation, the more traditional carrier-based modulations, or orthogonal frequency division multiplexing (OFDM, the technique used in IEEE 802.11a [197]). To facilitate coexistence of several UWB networks at the same location, a frequency-hopping scheme can be adopted.

The perspectives of UWB in industrial applications have been discussed in [194]. Physical-layer models for UWB channels in industrial environments have been presented in [195]. A UWB physical layer and protocol stack which can be used (with the help of protocol adaptation layers) as a basis for different short-range networking technologies (Bluetooth, IEEE 1394, IP) is currently pushed by the WiMedia alliance (<http://www.wimedia.org>) and in parts already finalized as an ECMA standard. It remains to be assessed in the future to which degree the anticipated advantages of UWB (robustness against narrow-band interference and multipath fading) can be materialized in wireless industrial communication systems built on them, and what the achievable industrial-QoS over this technologies really is.

XIII. CONCLUSIONS

In the view of the author there is a broad potential for future research in wireless industrial communication systems. One source of this research is the adoption of new communication

²³As a side note, the usage of pulses also allows very precise distance measurements (ranging) between communicating nodes. This can be very attractive for industrial applications.

²⁴In ALOHA, a newly arriving packet is transmitted instantly without performing any carrier-sense operations. When the transmitter does not receive an immediate acknowledgement, it waits for a random backoff time and retransmits.

technologies like wireless sensor networks or UWB technologies, another source is to identify promising approaches from the wireless communications and networking community like cooperative diversity schemes or, not mentioned in the paper, network coding [198], [199]. The challenges ahead are not only in the design of protocols leveraging these approaches, but also the systematic development of all the tools and methodologies necessary for proper dimensioning, planning and configuration of wireless industrial networks, and good methods to adapt these to varying environmental conditions.

ACKNOWLEDGMENT

The author would like to thank the reviewers and the associate editor for their constructive and valuable comments.

REFERENCES

- [1] A. Willig, K. Matheus, and A. Wolisz, "Wireless technology in industrial networks," *Proc. IEEE*, vol. 93, no. 6, pp. 1130–1151, Jun. 2005.
- [2] J. R. Moyne and D. M. Tilbury, "The emergence of industrial control networks for manufacturing control, diagnostics, and safety data," *Proc. IEEE*, vol. 95, no. 1, pp. 29–47, Jan. 2007.
- [3] T. M. Chan, K. F. Man, K. S. Tang, and S. Kwong, "A jumping-genes paradigm for optimizing factory WLAN network," *IEEE Trans. Ind. Informat.*, vol. 3, no. 1, pp. 33–43, Feb. 2007.
- [4] S. Ivanov, E. Nett, and S. Schemmer, "Planning available WLAN in dynamic production environments," in *Proc. 7th IFAC Int. Conf. Fieldbuses and Networks in Industrial and Embedded Systems (FET'2007)*, Toulouse, France, Nov. 2007.
- [5] A. Bonivento, C. Fischione, L. Necchi, F. Pianegiani, and A. Sangiovanni-Vincentelli, "System level design for clustered wireless sensor networks," *IEEE Trans. Ind. Informat.*, vol. 3, no. 3, pp. 202–214, Aug. 2007.
- [6] V. authors, "Special issue on networked control systems," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1421–1603, Sep. 2004.
- [7] V. authors, "Special issue on the technology of networked control systems," *Proc. IEEE*, vol. 95, no. 1, pp. 9–312, Jan. 2007.
- [8] F. D. Pellegrini, D. Miorandi, S. Vitturi, and A. Zanella, "On the use of wireless networks at low level of factory automation systems," *IEEE Trans. Ind. Informat.*, vol. 2, no. 2, pp. 129–143, May 2006.
- [9] J.-D. Decotignie, "Interconnection of wireline and wireless fieldbuses," in *The Industrial Information Technology Handbook*, R. Zurawski, Ed. Boca Raton, FL: CRC, 2005.
- [10] M. F. Alves, "Real-time communications over hybrid wired/wireless Profibus-based networks," Ph.D., Univ. Porto, Fac. Eng., Porto, Portugal, 2003.
- [11] D. Miorandi and S. Vitturi, "Hybrid wired/wireless implementations of PROFIBUS DP: A Feasibility study based on Ethernet and Bluetooth," *Comput. Commun.*, vol. 27, pp. 946–960, Jun. 2004.
- [12] P. B. Sousa and L. L. Ferreira, "Hybrid wired/wireless PROFIBUS architectures: Comparative performance analysis in an error-prone environment," in *Proc. 7th IFAC Int. Conf. Fieldbuses and Networks in Industrial and Embedded Systems (FET'2007)*, Toulouse, France, Nov. 2007.
- [13] M. Alves and E. Tovar, "Engineering PROFIBUS networks with heterogeneous transmission media," *Comput. Commun.*, vol. 30, no. 1, pp. 17–32, Dec. 2006.
- [14] E. Tovar and M. Alves, "Real-time communications over wired/wireless PROFIBUS networks supporting inter-cell mobility," *Comp. Netw.*, vol. 51, no. 11, pp. 2994–3012, Aug. 2007.
- [15] G. Schäfer, *Security in Fixed and Wireless Networks—An Introduction to Securing Data Communications*. Chichester, U.K.: Wiley, 2003.
- [16] D. Dzung, M. Naedele, T. P. v. Hoff, and M. Crevatin, "Security for industrial communication systems," *Proc. IEEE*, vol. 93, no. 6, pp. 1152–1177, Jun. 2005.
- [17] J.-P. Thomesse, "Fieldbus technology in industrial automation," *Proc. IEEE*, vol. 93, no. 6, pp. 1073–1101, Jun. 2005.
- [18] R. Zurawski, Ed., *The Industrial Communication Technology Handbook*. Boca Raton, Florida: CRC, 2005.
- [19] J. Jasperneite and P. Neumann, "Measurement, analysis and modeling of real-time source data traffic in factory communication systems," in *Proc. 2000 IEEE Int. Workshop Factory Communication Systems (WFCS)*, Sep. 2000.
- [20] Union Technique de l'Electricité, *General Purpose Field Communication System, EN 50170, Volume 2: PROFIBUS*, 1996.

- [21] U. Jecht, W. Stripf, and P. Wenzel, "Profibus—Open solutions for the world of automation," in *The Industrial Information Technology Handbook*, R. Zurawski, Ed. Boca Raton, FL: CRC, 2005.
- [22] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 359–378, May 1994.
- [23] R. A. Berry and R. G. Gallager, "Communication over fading channels with delay constraints," *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1135–1149, May 2002.
- [24] M. Hamdaoui and P. Ramanathan, "A dynamic priority assignment technique for streams with (m, k) -firm deadlines," *IEEE Trans. Comput.*, vol. 44, no. 12, pp. 1443–1451, Dec. 1995.
- [25] M. Hamdaoui and P. Ramanathan, "Evaluating dynamic failure probability for streams with (m, k) -firm deadlines," *IEEE Trans. Comput.*, vol. 46, no. 12, pp. 1325–1337, Dec. 1997.
- [26] Union Technique de l'Electricité, *General Purpose Field Communication System, EN 50170, Volume 3: WorldFIP*, 1996.
- [27] J.-P. Thomesse, "The worldfip fieldbus," in *The Industrial Information Technology Handbook*, R. Zurawski, Ed. Boca Raton, FL: CRC, 2005.
- [28] P. T. Eugster, P. A. Felber, R. Guerraoui, and A.-M. Kermarrec, "The many faces of publish/subscribe," *ACM Comput. Surv.*, vol. 35, no. 2, pp. 114–131, Jun. 2003.
- [29] A. Boukerche, "Security and fraud detection in mobile and wireless networks," in *Handbook of Wireless Networks and Mobile Computing*, I. Stojmenovic, Ed. New York: Wiley, 2002, pp. 309–323.
- [30] E. H. Callaway, *Wireless Sensor Networks—Architectures and Protocols*. Boca Raton, FL: Auerbach, 2003.
- [31] D. Brevi, D. Mazzocchi, R. Scopigno, A. Bonivento, R. Calcagno, and F. Rusina, "A methodology for the analysis of 802.11a links in industrial environments," in *Proc. 2006 IEEE Int. Workshop Factory Communication Systems (WFCS)*, Jul. 2006.
- [32] S. Cavalleri and D. Panno, "On the integration of fieldbus traffic within IEEE 802.11 wireless LAN," in *Proc. 1997 IEEE Int. Workshop Factory Communication Systems (WFCS'97)*, Barcelona, Spain, 1997.
- [33] S. P. Karanam, H. Trsek, and J. Jasperneite, "Potential of the HCCA scheme defined in IEEE802.11E for QoS enabled industrial wireless networks," in *Proc. 2006 IEEE Int. Workshop Factory Communication Systems (WFCS)*, Jul. 2006.
- [34] D. Koscielnik, "Simulation study of IEEE 802.11e wireless LAN—Enhancements for real time applications," in *Proc. IEEE Int. Symp. Industrial Electronics (ISIE)*, Montreal, QC, Canada, Jul. 2006.
- [35] N. Krommenacker and V. Lecuire, "Building industrial communication systems based on IEEE 802.11g technology," in *Proc. 10th IEEE Int. Conf. Emerging Technologies and Factory Automation (ETFA 2005)*, Catania, Italy, Sep. 2005.
- [36] S. Lee, K. C. Lee, M. H. Lee, and F. Harashima, "Integration of mobile vehicles for automated material handling using Profibus and IEEE 802.11 networks," *IEEE Trans. Ind. Electron.*, vol. 49, no. 3, pp. 693–701, Jun. 2002.
- [37] D. Miorandi and S. Vitturi, "Performance analysis of producer/consumer protocols over IEEE 802.11 wireless links," in *Proc. IEEE Workshop Factory Communication Systems (WFCS)*, Vienna, Austria, Sep. 2004.
- [38] R. Moraes, F. Vasques, P. Portugal, and J. A. Fonseca, "VTP-CSMA: A virtual token passing approach for real-time communication in IEEE 802.11 wireless networks," *IEEE Trans. Ind. Informat.*, vol. 3, no. 3, pp. 215–224, Aug. 2007.
- [39] IEEE Standard for Information Technology—Telecommunications and Information Exchange Between Systems—Local and Metropolitan Area Networks—Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE Computer Society, sponsored by the LAN/MAN Standards Committee, 2007.
- [40] IEEE Standard for Information Technology—Telecommunications and information exchange between systems—Local and Metropolitan networks—Specific requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements, LAN/MAN Standards Committee of the IEEE Computer Society, 2005.
- [41] P. Ferre, A. Doufexi, A. Nix, and D. Bull, "Throughput analysis of IEEE 802.11 and IEEE 802.11e MAC," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, 2004, Mar. 2004, pp. 783–788.
- [42] Z. n. Kong, D. H. K. Tsang, B. Bensaou, and D. Gao, "Performance analysis of IEEE 802.11e contention-based channel access," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 10, pp. 2095–2106, Dec. 2004.
- [43] J. del Pavon and S. Shankar, "Impact of frame size, number of stations and mobility on the throughput performance of IEEE 802.11e," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, 2004, Mar. 2004, pp. 789–795.
- [44] J. W. Robinson and T. S. Randhawa, "Saturation throughput analysis of IEEE 802.11e enhanced distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 5, pp. 917–928, Jun. 2004.
- [45] B. Xiang, M. Yu-Ming, and X. Jun, "Performance investigation of IEEE802.11e EDCA under non-saturation condition based on the M/G/1/K model," in *Proc. 2nd IEEE Conf. Industrial Electronics and Applications (ICIEA 2007)*, Harbin, China, May 2007.
- [46] G. Boggia, P. Camarda, L. A. Grieco, and S. Mascolo, "Feedback-based control for providing real-time services with the 802.11e MAC," *IEEE/ACM Trans. Netw.*, vol. 15, no. 2, pp. 323–333, Apr. 2007.
- [47] M. M. Rashid, E. Hossain, and V. K. Bhargava, "Queueing analysis of 802.11e HCCA with variable bit rate traffic," in *Proc. IEEE Int. Conf. Communications (ICC)*, Istanbul, Turkey, Jun. 2006.
- [48] M. M. Rashid, E. Hossain, and V. K. Bhargava, "HCCA scheduler design for guaranteed QoS in IEEE 802.11e based WLANs," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, Hong Kong, Mar. 2007.
- [49] H. Trsek, J. Jasperneite, and S. P. Karanam, "A simulation case study of the new IEEE 802.11e HCCA mechanism in industrial wireless networks," in *Proc. 11th IEEE Int. Conf. Emerging Technologies and Factory Automation (ETFA 2006)*, Prague, Czech Republic, Sep. 2006.
- [50] ISO Standard 11898—Road Vehicle—Interchange of Digital Information—Controller Area Network (CAN) for High-Speed Communication, ISO—International Organization for Standardization, International Organization for Standardization, 1993.
- [51] K. W. Tindell, H. Hansson, and A. J. Wellings, "Analysing real-time communications: Controller area network (CAN)," in *Proc. 15th Real-Time Systems Symp.*, 1994.
- [52] H. Hansson, T. Nolte, C. Norström, and S. Punnekat, "Integrating reliability and timing analysis of can-based systems," *IEEE Trans. Ind. Electron.*, vol. 49, no. 6, pp. 1240–1250, Dec. 2002.
- [53] High Performance Radio Local Area Network (HIPERLAN)—Draft Standard, ETSI, 1996.
- [54] N. Pereira, B. Andersson, and E. Tovar, "WiDom: A dominance protocol for wireless medium access," *IEEE Trans. Ind. Informat.*, vol. 3, no. 2, pp. 120–130, May 2007.
- [55] H. Karl and A. Willig, *Protocols and Architectures for Wireless Sensor Networks*. Chichester, U.K.: Wiley, 2005.
- [56] G. Pottie and W. Kaiser, *Principles of Embedded Networked Systems Design*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [57] S. S. Iyengar and R. R. Brooks, Eds., *Distributed Sensor Networks*. Boca Raton, FL: Chapman & Hall/CRC, 2005.
- [58] F. Zhao and L. Guibas, *Wireless Sensor Networks—An Information Processing Approach*. Amsterdam, The Netherlands: Elsevier/Morgan-Kaufman, 2004.
- [59] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Comp. Netw.*, vol. 38, pp. 393–422, 2002.
- [60] I. F. Akyildiz and I. H. Kasimoglu, "Wireless sensor and actor networks: Research challenges," *Ad Hoc Netw.*, vol. 2, pp. 351–367, 2004.
- [61] A. Willig, "Wireless sensor networks: Concept, challenges, approaches and applications," *E&I, Verbandszeitschrift desÖsterreichischen Verbandes für Elektrotechnik und Informationstechnik*, 2006.
- [62] J. M. Kahn, R. H. Katz, and K. S. J. Pister, "Next century challenges: Mobile networking for smart dust," in *Proc. ACM/IEEE Int. Conf. Mobile Computing and Networking (MobiCom 99)*, Seattle, WA, Aug. 1999.
- [63] K. S. Low, W. N. N. Win, and M. J. Er, "Wireless sensor networks for industrial environments," in *Proc. Int. Conf. Computational Intelligence for Modelling, Control and Automation, 2005 and Int. Conf. Intelligent Agents, Web Technologies and Internet Commerce*, Nov. 2005.
- [64] H. Ramamurthy, D. Lal, B. S. Prabhu, and R. Gadh, "Rewins: A distributed multi-RF sensor control network for industrial automation," in *Proc. Wireless Telecommunications Symp.*, Apr. 2005.
- [65] M. D. Lemmon, Q. Ling, and Y. Sun, "Overload management in sensor-actuator networks used for spatially-distributed control systems," in *Proc. 1st Int. Conf. Embedded Networked Sensor Systems (SenSys)*, Los Angeles, CA, Nov. 2003, pp. 162–170.
- [66] V. Singhvi, A. Krause, C. Guestrin, J. H. Garret, and H. S. Matthews, "Intelligent light control using sensor networks," in *Proc. 3rd Int. Conf. Embedded Networked Sensor Systems (SenSys)*, San Diego, CA, Nov. 2005.
- [67] B. Sinopoli, C. Sharp, L. Schenato, S. Schaffert, and S. S. Sastry, "Distributed control applications within sensor networks," *Proc. IEEE*, vol. 91, no. 8, pp. 1235–1246, Aug. 2003.

- [68] S. Ratnasamy, B. Karp, L. Yin, F. Yu, D. Estrin, R. Govindan, and S. Shenker, "GHT: A geographic hash table for data-centric storage," in *Proc. First ACM Int. Workshop Wireless Sensor Networks and Applications (WSNA'02)*, Atlanta, GA, 2002.
- [69] K. Langendoen and N. Reijers, "Distributed localization in wireless sensor networks: A quantitative comparison," *Comp. Netw.*, vol. 42, pp. 499–518, Aug. 2003.
- [70] A. Savvides, W. L. Garber, R. L. Moses, and M. B. Srivastava, "An analysis of error inducing parameters in multihop sensor node localization," *IEEE Trans. Mobile Comput.*, vol. 4, no. 6, pp. 567–577, Nov. 2005.
- [71] H. Dai and R. Han, "TSync: A lightweight bidirectional time synchronization service for wireless sensor networks," *ACM SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 8, no. 1, pp. 125–139, Jan. 2004.
- [72] J. Elson, L. Girod, and D. Estrin, in *Proc. 5th Symp. Operating Systems Design and Implementation (OSDI 2002)*, Dec. 2002.
- [73] W. Su and I. F. Akyildiz, "Time-diffusion synchronization protocol for wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 13, no. 2, pp. 384–397, Apr. 2005.
- [74] J. v. Greunen and J. Rabaey, "Lightweight time synchronization for sensor networks," in *Proc. 2nd ACM Int. Workshop Wireless Sensor Networks and Applications (WSNA)*, San Diego, CA, Sep. 2003.
- [75] S. Roundy, D. Steingart, L. Frechette, P. Wright, and J. Rabaey, "Power sources for wireless sensor networks," in *Proc. Wireless Sensor Networks, 1st Eur. Workshop (EWSN 2004)*, Berlin, Germany, Jan. 2004.
- [76] R. M. Passos, C. J. N. Coelho, A. A. F. Loureiro, and R. A. F. Mini, "Dynamic power management in wireless sensor networks: An application-driven approach," in *Proc. 2nd Annu. Conf. Wireless On-demand Network Systems and Services (WONS)*, 2005.
- [77] A. Bogliolo, L. Benini, E. Lattanzi, and G. D. Micheli, "Specification and analysis of power-managed systems," *Proc. IEEE*, vol. 92, no. 8, pp. 1308–1346, Aug. 2004.
- [78] C.-F. Chiasserini and R. R. Rao, "Energy efficient battery management," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 5, pp. 1235–1245, Jul. 2001.
- [79] S. Sarkar and M. Adamu, "A framework for optimal battery management for wireless nodes," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 2, pp. 179–188, Feb. 2003.
- [80] G. Scheible, D. Dzung, J. Endresen, and J.-E. Frey, "Unplugged but connected—Design and implementation of a truly wireless real-time sensor/actuator interface," *IEEE Ind. Electron. Mag.*, vol. 1, no. 2, pp. 25–34, 2007.
- [81] A. J. Goldsmith and S. B. Wicker, "Design challenges for energy-constrained ad hoc wireless networks," *IEEE Wireless Commun.*, vol. 9, no. 4, pp. 8–27, Aug. 2002.
- [82] B. Krishnamachari, D. Estrin, and S. Wicker, "The impact of data aggregation in wireless sensor networks," in *Proc. Workshops 22nd IEEE Int. Conf. Distributed Computing Systems*, Vienna, Austria, Jul. 2002.
- [83] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "Tag: A tiny aggregation service for ad-hoc sensor networks," in *Proc. 5th Symp. Operating System Design and Implementation (OSDI 2002)*, Dec. 2002.
- [84] A. Köpke, H. Karl, and M. Löbbers, "Using energy where it counts: Protecting important messages in the link layer," in *Proc. Eur. Workshop Wireless Sensor Networks (EWSN)*, Istanbul, Turkey, Jan. 2005.
- [85] Chipcon, 2.4 GHz IEEE 802.15.4/ZigBee-Ready RF Transceiver, Chipcon Products from Texas Instruments, 2004, CC2420 Data Sheet.
- [86] R. Min, M. Bhardwaj, S.-H. Cho, N. Ickes, E. Shih, A. Sinha, A. Wang, and A. Chandrakasan, "Energy-centric enabling technologies for wireless sensor networks," *IEEE Wireless Commun.*, vol. 9, no. 4, pp. 28–39, Aug. 2002.
- [87] R. Min and A. Chandrakasan, "A framework for energy-scalable communication in high-density wireless networks," in *Proc. ISLPED '02*, 2002, pp. 36–41.
- [88] V. Srivastava and M. Motani, "Cross-layer design: A survey and the road ahead," *IEEE Commun. Mag.*, vol. 43, no. 12, pp. 112–119, Dec. 2005.
- [89] V. Kawadia and P. R. Kumar, "A cautionary perspective on cross-layer design," *IEEE Wireless Commun.*, vol. 12, no. 1, pp. 3–11, Feb. 2005.
- [90] P. Santi, "Topology control in wireless ad hoc and sensor networks," *ACM Comput. Surv.*, vol. 37, no. 2, pp. 164–194, Jun. 2005.
- [91] IEEE Standard for Information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements—Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low Rate Wireless Personal Area Networks (LR-WPANs), LAN/MAN Standards Committee of the IEEE Computer Society, 2006.
- [92] IEEE Standard for Information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements—Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low Rate Wireless Personal Area Networks (WPANs)—Amendment 1: Add Alternate PHYs, LAN/MAN Standards Committee of the IEEE Computer Society, 2006.
- [93] J. Flora and P. Bonnet, *Never Mind the Standard: Here is the Tinyos 802.15.4 Stack*, Dept. Comput. Sci., Univ. Copenhagen, 2006, Tech. Rep.
- [94] ZigBee Standards Organization, *ZigBee Specification*, 2006.
- [95] C. Perkins, E. Belding-Royer, and S. Das, Ad Hoc On-Demand Distance Vector (AODV) Routing, IETF Network Working Group, 2003, RFC 3561 (Experimental RFC).
- [96] C. E. Perkins and E. M. Royer, "Ad hoc on-demand distance vector routing," in *Proc. 2nd IEEE Workshop Mobile Computing Systems and Applications (WMCSA'99)*, New Orleans, LA, Feb. 1999.
- [97] D. Egan, "The emergence of ZigBee in building automation and industrial control," *Comput. Control Eng. J.*, vol. 16, no. 2, pp. 14–19, Apr. 2005.
- [98] N. Baker, "ZigBee and Bluetooth strengths and weaknesses for industrial applications," *Comput. Control Eng. J.*, vol. 16, no. 2, pp. 20–25, Apr. 2005.
- [99] T. J. Park, Y. J. Chon, D. K. Park, and S. H. Hong, "Bacnet over ZigBee, a new approach to wireless datalink channel for bacnet," in *Proc. 5th IEEE Int. Conf. Industrial Informatics (INDIN)*, Jun. 2007.
- [100] A. Koubaa, M. Alves, and E. Tovar, "A comprehensive simulation study of slotted CSMA/CA for IEEE 802.15.4 wireless sensor networks," in *Proc. 2006 IEEE Int. Workshop Factory Communication Systems (WFCS)*, Jul. 2006.
- [101] G. Lu, B. Krishnamachari, and C. S. Raghavendra, "Performance evaluation of the IEEE 802.15.4 MAC for low-rate low-power wireless networks," in *Proc. 2004 IEEE Int. Conf. Performance, Computing, and Communications*, Apr. 2004, pp. 701–706.
- [102] M. Petrova, J. Riihijarvi, P. Mähönen, and S. Labella, "Performance study of IEEE 802.15.4 using measurements and simulations," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, 2006.
- [103] M. Bertocco, G. Gamba, A. Sona, and S. Vitturi, "Performance measurements of CSMA/CA-based wireless sensor networks for industrial applications," in *Proc. IEEE Instrumentation and Measurement Technology Conf. (IMTC)*, Warsaw, Poland, May 2007.
- [104] A. Koubaa, M. Alves, B. Nefzi, and Y.-Q. Song, "Improving the IEEE 802.15.4 slotted CSMA/CA MAC for time-critical events in wireless sensor networks," in *Proc. Workshop Real-Time Networks (RTN 2006), Satellite Workshop to ECRTS 2006*, Jul. 2006.
- [105] A. Koubaa, M. Alves, M. Attia, and A. V. Nieuwenhuysse, "Collision-free beacon scheduling mechanisms for IEEE 802.15.4/ZigBee cluster-tree wireless sensor networks," in *Proc. 7th Int. Workshop Applications and Services in Wireless Networks (ASWN2007)*, May 2007.
- [106] A. Koubaa, A. Cunha, and M. Alves, "A time division beacon scheduling mechanism for IEEE 802.15.4/ZigBee cluster-tree wireless sensor networks," in *Proc. 19th Euromicro Conf. Real-Time Systems (ECRTS 2007)*, Jul. 2007.
- [107] J. Franconme, G. Mercier, and T. Val, "Beacon synchronization for GTS collision avoidance in an IEEE 802.15.4 meshed network," in *Proc. 7th IFAC Int. Conf. Fieldbuses and Networks in Industrial and Embedded Systems (FET'2007)*, Toulouse, France, Nov. 2007.
- [108] D. Sexton, M. Mahony, M. Lapinski, and J. Werb, "Radio channel quality in industrial wireless sensor networks," in *Proc. Sensors for Industry Conf. (SIcon/05)*, Feb. 2005, pp. 88–94.
- [109] L. Tang, K.-C. Wang, Y. Huang, and F. Gu, "Channel characterization and link quality assessment of IEEE 802.15.4-Compliant radio for factory environments," *IEEE Trans. Ind. Informat.*, vol. 3, no. 2, pp. 99–110, May 2007.
- [110] B. Liu and D. Towsley, "A study on the coverage of large-scale sensor networks," in *Proc. 1st IEEE Int. Conf. Mobile Ad-hoc and Sensor Systems (MASS'04)*, 2004.
- [111] J. A. Stankovic, T. F. Abdelzaher, C. Lu, L. Sha, and J. C. Hou, "Real-time communication and coordination in embedded sensor networks," *Proc. IEEE*, vol. 91, no. 7, pp. 1002–1022, Jul. 2003.
- [112] J. Baillieul and P. J. Antsaklis, "Control and communication challenges in networked real-time systems," *Proc. IEEE*, vol. 95, no. 1, pp. 9–28, Jan. 2007.
- [113] S. Tilak, N. B. Abu-Ghazaleh, and W. Heinzelman, "Infrastructure tradeoffs for sensor networks," in *Proc. 1st ACM Int. Workshop Sensor Networks and Applications (WSNA)*, Atlanta, GA, Sep. 2002.
- [114] C.-Y. Wan, A. T. Campbell, and L. Krishnamurthy, "Pump-slowly, fetch-quickly (PSFQ): A reliable transport protocol for sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 862–872, Apr. 2005.

- [115] S.-J. Park, R. Vedantham, R. Sivakumar, and I. F. Akyildiz, "A scalable approach for reliable downstream data delivery in wireless sensor networks," in *Proc. Symp. Mobile Ad Hoc Networking & Computing (MobiHoc'04)*, Tokyo, Japan, May 2004.
- [116] C.-J. Liang, R. Musaloiu-E, and A. Terzis, "Typhoon: A reliable data dissemination protocol for wireless sensor networks," in *Proc. Wireless Sensor Networks, 5th Eur. Workshop (EWSN 2008)*, Bologna, Italy, Jan. 2008.
- [117] F. Stamm and J. Heidemann, "RMST: Reliable data transport in sensor networks," in *Proc. 1st IEEE Int. Workshop Sensor Network Protocols and Applications (SNPA)*, Anchorage, AK, May 2003.
- [118] B. Akan and I. F. Akyildiz, "Event-to-Sink reliable transport in wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 13, no. 5, pp. 1003–1016, Oct. 2005.
- [119] E. Felemban, C.-G. Lee, and E. Ekici, "MMSPEED: Multipath multi-speed protocol for QoS guarantee of reliability and timeliness in wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 5, no. 6, pp. 738–754, Jun. 2006.
- [120] S. Kim, R. Fonseca, and D. Culler, "Reliable transfer on wireless sensor networks," in *Proc. First IEEE Int. Conf. Sensor and Ad hoc Communications and Networks (SECON 2004)*, Santa Clara, CA, Oct. 2004.
- [121] B. Deb, S. Bhatnagar, and B. Nath, "Information assurance in sensor networks," in *Proc. 2nd ACM Int. Workshop Wireless Sensor Networks and Applications (WSNA)*, San Diego, CA, Sep. 2003.
- [122] B. Deb, S. Bhatnagar, and B. Nath, "Reinform: Reliable information forwarding using multiple paths in sensor networks," in *Proc. 28th Annual IEEE Conf. Local Computer Networks (LCN 2003)*, Bonn, Germany, Oct. 2003.
- [123] R. C. Shah and J. M. Rabaey, "Energy aware routing for low energy ad hoc sensor networks," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, Orlando, FL, Mar. 2002.
- [124] L. Rizzo, "Effective erasure codes for reliable computer communication protocols," *ACM Comp. Commun. Rev.*, vol. 27, no. 2, pp. 24–36, Apr. 1997.
- [125] J. W. Byers, M. Luby, and M. Mitzenmacher, "A digital fountain approach to asynchronous reliable multicast," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 8, pp. 1528–1540, Oct. 2002.
- [126] R. Fonseca, O. Gnawali, K. Jamieson, S. Kim, P. Levis, and A. Woo, "The Collection Tree Protocol (CTP)," TEP 123, TinyOS Network Working Group, 2006.
- [127] D. S. J. D. Couto, D. Aguayo, J. Bicket, and R. Morris, "A high-throughput path metric for multi-hop wireless routing," in *Proc. Ninth Annu. Int. Conf. Mobile Computing and Networking (MobiCom)*, San Diego, CA, Sep. 2003.
- [128] T. He, J. A. Stankovic, C. Lu, and T. Abdelzaher, "Speed: A stateless protocol for real-time communication in sensor networks," in *Proc. 23rd Int. Conf. Distributed Computing Systems (ICDCS'03)*, Providence, RI, May 2003.
- [129] S.-J. Park and R. Sivakumar, "Poster: Sink-to-sensors reliability in sensor networks," in *Proc. 4th ACM Int. Symp. Mobile Ad Hoc Networking and Computing (MOBIHOC)*, Annapolis, MD, Jun. 2003.
- [130] C. Intanagonwivat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, "Directed diffusion for wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 11, no. 1, pp. 2–16, Feb. 2003.
- [131] IEEE Standard for Information Technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements—Part 15.1: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for wireless personal area networks (WPANs), IEEE Computer Society—Sponsored by the LAN/MAN Standards Committee, 2005.
- [132] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [133] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [134] H. Bai and M. Atiquzzaman, "Error modeling schemes for fading channels in wireless communications: A survey," *IEEE Commun. Surv. Tutor.*, vol. 5, no. 2, pp. 2–9, 2003.
- [135] A. Willig, M. Kubisch, C. Hoene, and A. Wolisz, "Measurements of a wireless link in an industrial environment using an IEEE 802.11-compliant physical layer," *IEEE Trans. Ind. Electron.*, vol. 49, no. 6, pp. 1265–1282, Dec. 2002.
- [136] U. Bilstrup and P.-A. Wiberg, "Bluetooth in industrial environment," in *Proc. 2000 IEEE Workshop Factory Communication Systems, WFCS'2000*, Porto, Portugal, 2000, pp. 239–246.
- [137] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.
- [138] T. S. Rappaport, *Wireless Communications—Principles and Practice*. Upper Saddle River, NJ: Prentice-Hall, 2002.
- [139] H. Liu, H. Ma, M. E. Zarki, and S. Gupta, "Error control schemes for networks: An overview," *MONET—Mobile Netw. Appl.*, vol. 2, no. 2, pp. 167–182, 1997.
- [140] D. Haccoun and S. Pierre, "Automatic repeat request," in *The Communications Handbook*, J. D. Gibson, Ed. Boca Raton, FL: CRC/IEEE Press, 1996, pp. 181–198.
- [141] E. Biglieri, *Coding for Wireless Channels*. New York: Springer, 2005.
- [142] S. N. Diggavi, N. Al-Dhahir, A. Stamoulis, and A. R. Calderbank, "Great expectations: The value of spatial diversity in wireless networks," *Proc. IEEE*, vol. 92, no. 2, pp. 219–270, Feb. 2004.
- [143] H. Bölcskei, "MIMO-OFDM wireless systems: Basics, perspectives and challenges," *IEEE Wireless Commun.*, vol. 13, no. 4, pp. 31–37, Aug. 2006.
- [144] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bölcskei, "An overview of MIMO communications—A key to gigabit wireless," *Proc. IEEE*, vol. 92, no. 2, pp. 198–218, Feb. 2004.
- [145] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.
- [146] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.
- [147] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.
- [148] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.
- [149] T. H. Liew and L. Hanzo, "Space-time codes and concatenated channel codes for wireless communications," *Proc. IEEE*, vol. 90, no. 2, pp. 187–219, Feb. 2002.
- [150] A. Scaglione, D. L. Goeckel, and J. N. Laneman, "Cooperative communications in mobile ad hoc networks," *IEEE Signal Process. Mag.*, vol. 23, no. 5, pp. 18–29, Sep. 2006.
- [151] "Special issue on models, theory, and codes for relaying and cooperation in communication networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3297–3842, Oct. 2007.
- [152] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1089–1098, Aug. 2004.
- [153] A. d. Coso, U. Spagnolini, and C. Ibars, "Cooperative distributed MIMO channels in wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 402–414, Feb. 2007.
- [154] T. M. Cover and A. A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.
- [155] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.
- [156] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.
- [157] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behaviour," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
- [158] H. Zhu and G. Cao, "RDCCF: A relay-enabled medium access control protocol for wireless ad hoc networks," *IEEE Trans. Mobile Comput.*, vol. 5, no. 9, pp. 1201–1214, Sep. 2006.
- [159] A. Willig, "Polling-based MAC protocols for improving realtime performance in a wireless PROFIBUS," *IEEE Trans. Ind. Electron.*, vol. 50, no. 4, pp. 806–817, Aug. 2003.
- [160] A. Willig, "How to exploit spatial diversity in wireless industrial networks," in *Proc. 7th IFAC Int. Conf. Fieldbuses and Networks in Industrial and Embedded Systems (FET'2007)*, Toulouse, France, Nov. 2007.
- [161] E. Uhlemann and A. Willig, "Joint design of relay and packet combining schemes for wireless industrial networks," in *Proc. IEEE Vehicular Technology Conf. (VTC), Spring 08*, Marina Bay, Singapore, May 2008.
- [162] A. Nosratinia and T. E. Hunter, "Grouping and partner selection in cooperative wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 369–378, Feb. 2007.
- [163] A. Hills and B. Friday, "Radio resource management in wireless LANs," *IEEE Commun. Mag.*, vol. 42, no. 12, pp. S9–S14, Dec. 2004.
- [164] L. Georgiadis, M. Neely, and L. Tassiulas, *Resource Allocation and Cross Layer Control in Wireless Networks*. Hanover, MA: Now, 2006.
- [165] X. Wang, G. B. Giannakis, and A. G. Marques, "A unified approach to QoS-guaranteed scheduling for channel-adaptive wireless networks," *Proc. IEEE*, vol. 95, no. 12, pp. 2410–2431, Dec. 2007.

- [166] K.-S. Tang, K.-F. Man, and S. Kwong, "Wireless communication network design in IC factory," *IEEE Trans. Ind. Electron.*, vol. 48, no. 2, pp. 452–459, Apr. 2001.
- [167] J. W. S. Liu, *Real-Time Systems*. Upper Saddle River, NJ: Prentice-Hall, 2000.
- [168] E. Tovar, "Supporting real-time communications with standard factory-floor networks," Ph.D. dissertation, Dept. Elect. Eng., Univ. Porto, Porto, Portugal, 1999.
- [169] E. Tovar and F. Vasques, "Real-time fieldbus communications using Profibus networks," *IEEE Trans. Ind. Electron.*, vol. 46, no. 6, pp. 1241–1251, Dec. 1999.
- [170] J. Arauz and P. Krishnamurthy, "Markov modeling of 802.11 channels," in *Proc. 58th IEEE Vehicular Technology Conf. (VTC) 2003-Fall*, Oct. 2003, pp. 771–775.
- [171] F. Babich, O. E. Kelly, and G. Lombardi, "Generalized Markov modeling for flat fading," *IEEE Trans. Commun.*, vol. 48, no. 4, pp. 547–551, Apr. 2000.
- [172] A. J. Goldsmith and P. P. Varaiya, "Capacity, mutual information, and coding for finite-state Markov channels," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 868–886, May 1996.
- [173] M. Hassan, M. M. Krunz, and I. Matta, "Markov-based channel characterization for tractable performance analysis in wireless packet networks," *IEEE Trans. Wireless Commun.*, vol. 3, no. 3, pp. 821–831, May 2004.
- [174] H. Wang and N. Moayeri, "Finite state Markov channel—A useful model for radio communication channels," *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 163–171, Feb. 1995.
- [175] M. Zorzi, R. R. Rao, and L. B. Milstein, "On the accuracy of a first-order Markov model for data block transmission on fading channels," in *Proc. IEEE ICUPC'95*, Nov. 1995, pp. 211–215.
- [176] C.-S. Chang, *Performance Guarantees in Communication Networks*, ser. Telecommunication Networks and Computer Systems. London, U.K.: Springer Verlag, 2000.
- [177] R. L. Cruz, "A calculus for network delay, part I: Network elements in isolation," *IEEE Trans. Inf. Theory*, vol. 37, no. 1, pp. 114–131, Jan. 1991.
- [178] R. L. Cruz, "A calculus for network delay, part II: Network analysis," *IEEE Trans. Inf. Theory*, vol. 37, no. 1, pp. 132–141, Jan. 1991.
- [179] J.-P. Georges, E. Rondeau, and T. Divoux, "Evaluation of switched ethernet in an industrial context by using the network calculus," in *Proc. 4th IEEE Workshop Factory Communication Systems 2002 (WFCS 2002)*, Vasteras, Sweden, 2002.
- [180] Y. Jiang, "A basic stochastic network calculus," in *Proc. ACM SIGCOMM Conf. Network Architectures and Protocols*, Pisa, Italy, Sep. 2006.
- [181] A. K. Arumugam, A. Doufexi, A. R. Nix, and P. N. Fletcher, "An investigation of the coexistence of 802.11g WLAN and high data rate Bluetooth enabled consumer electronic devices in indoor home and office environments," *IEEE Trans. Consum. Electron.*, vol. 49, no. 3, pp. 587–596, Aug. 2003.
- [182] C.-F. Chiasserini and R. R. Rao, "Coexistence mechanisms for interference mitigation in the 2.4-GHz ISM band," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 964–975, Sep. 2003.
- [183] I. Howitt and J. A. Gutierrez, "IEEE 802.15.4 low rate—Wireless personal area network coexistence issues," in *Proc. Wireless Communications and Networking Conf. 2003 (WCNC 2003)*, New Orleans, LA, Mar. 2003, pp. 1481–1486.
- [184] J. Mitola, III and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Personal Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.
- [185] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.
- [186] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
- [187] F. H. P. Fitzek and M. D. Katz, Eds., *Cognitive Wireless Networks—Concepts, Methodologies and Visions Inspiring the Age of Enlightenment of Wireless Communications*. Dordrecht, The Netherlands: Springer, 2007.
- [188] "Special issue on cognitive wireless networks," *IEEE Wireless Commun.*, vol. 14, no. 4, pp. 4–52, Aug. 2007.
- [189] Q. Zhao and A. Swami, "A decision-theoretic framework for opportunistic spectrum access," *IEEE Wireless Commun.*, vol. 4, no. 4, pp. 14–20, Aug. 2007.
- [190] J. Mitola, "The software radio architecture," *IEEE Commun. Mag.*, vol. 33, no. 5, pp. 26–38, May 1995.
- [191] D. Porcino and W. Hirt, "Ultra-wideband radio technology: Potential and challenges ahead," *IEEE Commun. Mag.*, vol. 41, no. 7, pp. 66–74, Jul. 2003.
- [192] L. Yang and G. B. Giannakis, "Ultra-wideband communications—An idea whose time has come," *IEEE Signal Process. Mag.*, vol. 21, no. 6, pp. 26–54, Nov. 2004.
- [193] S. Roy, J. R. Foerster, V. S. Somayazulu, and D. G. Leeper, "Ultrawideband radio design: The promise of high-speed, short-range wireless connectivity," *Proc. IEEE*, vol. 92, no. 2, pp. 295–311, Feb. 2004.
- [194] G. P. Hancke and B. Allen, "Ultrawideband as an industrial wireless solution," *IEEE Pervasive Comput. Mag.*, vol. 5, no. 4, pp. 78–85, Oct. 2006.
- [195] J. Karedral, S. Wyne, P. Almers, F. Tufvesson, and A. F. Molisch, "A measurement-based statistical model for industrial ultra-wideband channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 3028–3037, Aug. 2007.
- [196] M. Z. Win and R. A. Scholtz, "Ultra-wide bandwidth time-hopping spread-spectrum impulse radio for wireless multiple-access communications," *IEEE Trans. Commun.*, vol. 48, no. 4, pp. 679–691, Apr. 2000.
- [197] IEEE Standard for Telecommunications and Information Exchange Between Systems—LAN/MAN Specific Requirements—Part 11: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High Speed Physical Layer in the 5 GHz Band, LAN/MAN Standards Committee of the IEEE Computer Society, 1999.
- [198] R. Ahlswede, N. Cai, R. Li S.-Y., and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.
- [199] R. W. Yeung, S.-Y. R. Li, N. Cai, and Z. Zhang, *Network Coding Theory*. Hanover, MA: Now, 2006.



Andreas Willig (M'97) received the Dr.-Ing. degree in electrical engineering from the Technical University Berlin, Berlin, Germany, in 2002 and the diploma degree in computer science from the University of Bremen, Bremen, Germany, in 1994.

He has been a Senior Researcher with the Telecommunication Networks Group (TKN) at the Technical University of Berlin since April 2005. From 2002 to 2005, he was an Assistant Professor with the Hasso-Plattner-Institut at the University of Potsdam, Potsdam, Germany. His research interests

include wireless networks, fieldbus and real-time systems, and *ad hoc* and sensor networks, all with specific focus on protocol design and performance aspects.