Collaborative Project
Small-medium-scale focused research project (STREP)

Grant Agreement n. 224168 FP7-ICT-2007-2

WIDE

Decentralized Wireless Control of Large-Scale Systems

Starting date: 01 September 2008 Duration: 3 years

Deliverable number	**D2.1**
Title	**Loss and latency models for wide-area sensor networks**
Work package	WP2 – Wide-area WSN for distributed control
Due date	M24
Actual submission date	01/09/2010
Lead contractor for this deliverable	KTH
Author(s)	Mikael Johansson <mikaeljl@ee.kth.se> Riku Jäntti <riku.jantti@tkk.fi>
With the help of	Johannes Bleuel <Johannes.Bleuel@e-senza.de> Maurice Heemels <W.P.M.H.Heemels@tue.nl> Volkan Ungan ungan@kth.se
Revision	v1.0 (January 21, 2011)

Dissemination Level

PU	Public
PP	Restricted to other programme participants (including the Commission Services)
RE	Restricted to a group specified by the consortium (including the Commission Services)
→ **CO**	Confidential, only for members of the consortium (including the Commission Services)

Executive summary

This report discusses technologies and models for low power wireless industrial communication. The chapter provides a tutorial overview covering basic concepts and models for wireless propagation, medium access control, multi-hop networking, routing and transport protocols. Throughout, an effort is made to describe both key technologies and associated models of control-relevant characteristics such as latency and loss. Some existing and emerging specifications and standards, including Zigbee, WirelessHART and ISA100, are described in some detail, and links are made between the developed models and useful network abstractions for control design.

Contents

1 Introduction

Since the first application of wireless in industrial control almost a 100 years ago [41], the number of actual deployments has remained small. For a long time, the market has been limited to a specific target applications (e.g. wireless remote controls) engineered using customized technologies and sometimes even operating on licensed spectrum. There is a growing consensus that this trend is now about to change: the enormous success of short-range wireless in home and office applications has raised consumer confidence in wireless technologies; the emergence of standardized, low cost, low-power radios has made industrial wireless economically attractive compared with cabled sensors [28]. Intense efforts on wireless sensor networks [33] and networked control [4] indicate that a large class of industrial process could be reliably controlled despite deficiencies of the wireless medium. New standards for industrial wireless communication, such as WirelessHART and ISA100 have recently been approved, and stardards-compliant technology is hitting the market. All together, this raises expectations of a wide deployment of industrial wireless [38].

The development of controllers operating over wireless networks is inherently a problem of co-design, and a good system designs require insight and understanding of the traditionally separate disciplines of control and communications. This chapter reviews technologies and models for industrial wireless networking in an attempt to narrow the gap between models used in the theoretical control literature and the models that arise when emerging standards for industrial wireless networks are modelled using tools from communication theory. Covering a wide area of topics, from wireless propagation and medium access control to routing and transport protocols for multi-hop wireless networking, the chapter is admittedly broad rather than deep. Nevertheless, a range of useful models for control-relevant quantities such as latency and loss in industrial wireless networks are given, and several existing and emerging specifications and standards, including Zigbee, WirelessHART and ISA100, are described in some detail. Pointers to books and selected key publications are given throughout to guide the reader to good entry-points for more in-depth studies.

The chapter is organized as follows. We begin by a review of wireless propagation and models for the behavior of a single wireless link. We then review various medium access control techniques for sharing the wireless spectrum among several wireless links. Next, we discuss multi-hop wireless networking, routing and protocols for end-to-end communication. With this basic understanding of wireless communication at hand, we describe several specifications and standards for low-power wireless networking. Finally, we establish links between the developed models and useful network abstractions for control design.

2 Understanding the single link

Aim: describe the concepts of wireless propagation and outage, leading to reasonable models of the raw packet-error rates in the 2.4 GHz ISM-band

2.1 Wireless propagation and outage

Radio propagation refers to the way that radio waves behave when they are transmitted and propagate from one point to another. Radio waves are affected by the same phenomena as light waves, such as reflection, diffraction, absorption and scattering. The impact of the wireless channel on the transmitted signal is customarily divided into *large-scale effects* and *small-scale effects*. Large scale effects include signal attenuation due to the propagation distance and shadowing caused by large stationary obstacles in the radio-path. Small-scale effects are caused by signal reflections and movement of the receiver, transmitter or small objects in the radio path. Reflections cause multiple copies

of the signal to arrive at the receiver with different attenuation and propagation delay. The superposition of these signals at the receiver causes time dispersion of the original signal. In the frequency domain, this gives raise to frequency-selective fading where different frequency components of the signal experience different attenuation and phase shift. Movement of the transmitter, receiver or small objects in the environment cause dispersion of power in frequency. This is effect is called Doppler spread and gives rise to time-selective fading, *i.e.* it makes the channel response time-varying. In certain conditions, the signal may be absorbed or multi-path components may interfere each other destructively such that the signal vanishes completely at the receiver. This phenomenon is called *erasure fading*.

Small-scale effects

Coherence bandwidth is a statistical measurement of the range of frequencies over which the channel can be considered "flat", or in other words the approximate maximum bandwidth or frequency interval over which two frequencies of a signal are likely to experience comparable or correlated amplitude fading.

Because of multipath reflections, the impulse response of a wireless channel looks likes a series of pulses. It is customary to characterize the delay profile of the channel by observing the mean energy $p(\tau)$ for each delay τ. The maximum delay time spread is the total time interval during which reflections with significant energy arrive. A common parameter for quantifying the multipath behavior of the channel is the *root-mean-square (RMS) delay spread* σ_τ. The corresponding quantity in the frequency domain is the coherence bandwidth of the channel, the frequency range over which two transmitted sinusoidal signals are likely to see the phase shift and attenuation. The coherence bandwidth $B_{C,50}$ where the correlation of the channel exceeds 0.5 is approximately

$$B_{C,50} \approx \frac{1}{5\sigma_\tau}$$

If the bandwidth of the transmitted signal B is less than $B_{C,50}$, the fading is called frequency-non-selective or flat fading and can be modelled with impulse response having only one tap (one radio path); otherwise the channel fading is called frequency-selective. In frequency selective channel an equalizer (adaptive tapped delay line filter) is needed at the receivers in order to keep the bit error probability at tolerable levels. Such a filter is not needed in flat faded channel.

Measurements on the 1.3 GHz frequency band suggests that typical values for RMS delay spread in factory environment lie between 100 ns and 200 ns, but values as high as 300 ns are possible [48]. These values are one order of magnitude higher than what has been reported for office environments. Measurements on the 2.4 GHz band suggest that the mean RMS delay between 16 ns (mine in granite) to 85 ns (transformer station) translating to coherence bandwidth between 1250 kHz to 227 kHz [Kemp et. al. 2004]. The results also indicate that the RMS delay varies a lot between links in different positions within the same plant. The RMS delay spread observed in a site tends to follow a (truncated) normal distribution. In the two examples mentioned, the standard deviation was 8 ns and 89 ns, respectively. The mean values tend to be smaller than in office environment where measurement reports indicate RMS delay spreads between 45 ns and 420 ns [32]. Table 1 summarizes the measurement results reported in [26].

To understand how small-scale fading impacts radio performance, consider the IEEE 802.15.4 radio standard, in which receivers do not have equalizers. On the 2.4 GHz band 802.15.4-radios transmit 2 Mcps (chips per second) which translates to 2 Mhz half-power bandwidth B. To assure that the signal sees frequency-non-selective channel, it is customary to require that the coherence bandwidth is at least twice the signal bandwidth. By setting $B_{C,50} = 2 \times 2$ MHz, we get $\sigma_\tau = 50$ ns. Based on this analysis, a RMS delay spread exceeding 50 ns would cause concern. Simulation studies [16] suggest that IEEE 802.15.4 can achieve packet error rates below $5 \cdot 10^{-3}$ as long as the RMS delay does not

Table 1: Comparative table of sample mean and standard deviation of the RMS delay spread and the coherence bandwidth related to the mean RMS delay spread

Site	Mean RMS RMS delay spread (ns)	Standard deviation RMS delay spread (ns)	Coherence bandwidth $B_{C,50}$ (MHz)
Petrochemical plant	38	9	5.26
Transformer station	85	89	2.27
Manufacturing plant	44	24	4.55
Carpark amongst multi-story buildings	74	107	2.63
Mine in granite	16	8	12.50
Coal mine	23	9	8.70

exceed 400 ns. However, this comes with the cost of reduced range as up to 10dB higher signal-to-noise ratio is needed to compensate the effect of inter-symbol interference. The simulations also suggest that half-rate at 915 MHz would tolerate RMS delay spreads up to 800 ns. Comparing the tolerance of the IEEE 802.15.4 receiver to the values measured in industrial settings, we can conclude that in most cases the radio can operate reliably but in some environments the range could be limited and it is even possible to find environments where reliable operation is not possible. The low transmit power of IEEE 802.15.4 also helps in reducing the number of multipath reflected components.

In the wireless channel, many echoed copies of the transmitted signal appear almost with almost equal delay. The multipath model groups these delay paths into discrete set of resolvable paths, called taps, by summing all the signals received with almost the same delay. By applying the central limit theorem, we can model the signal as complex Gaussian random variable (it is customary to model the base band signal as complex random variables - modulated signal is then obtained by multiplying the signal with phasor $\exp(j2f_c t)$ rotating with the carrier frequency f_c and taking the real part of the resulting complex signal). The amplitude of the signal passing through single radio path follows either Rician or Rayleigh distribution depending whether there is significant line-of-sight (LOS) component in the signal or not. That is, whether the complex Gaussian variables have non-zero mean or not. In both cases, the phase of the received signal is uniformly distributed random variable. In case of Rayleigh fading, the signal power follows chi-square distribution with two degrees of freedom which happens to coincide with the exponential distribution. In the Rician case, the received signal power follows non-central chi-squared distribution with two degrees of freedom. Yet another distribution that is often utilized to characterize the lino-of-sight is the Nagagami-m distribution. For $m = 1$, the model coincides with Rayleigh distribution. The larger m is the less variations there are in the channel. Measurement results in industrial environment suggest that the firs tap in a line of sight channel, typically experience Nagagami-m distribution with m varying between 3 and 67 while the reflected paths follow Rayleigh distribution ($m = 1$) [62]. In case, m is integer, the cumulative probability density of the Nakagami-m faded signal strength $\tilde{P}_{rx}$ having mean $\bar{P}_{rx}$ follows Erlang-m distribution and can be written as

$$\Pr\left\{\tilde{P}_{rx} \leq P\right\} = 1 - \sum_{k=0}^{m-1} \left(\frac{k}{\bar{P}_{rx}}\right)^k \frac{1}{k!} \exp\left(-\frac{P}{\bar{P}_{rx}}\right) \stackrel{def}{=} F_P(P)$$

Doppler spread and *coherence time* are parameters that describe the time dispersive nature of the wireless channel in the small-scale region. When a node or reflectors in its environment is moving, the velocity of the moving node and/or environment causes a shift in the frequency of the signal transmitted along each signal path. This phenomenon is known as the Doppler shift. Although most wireless sensor networks tends to be immobile, the industrial environment often have moving reflectors such as overhead cranes, forklifts etc. The Doppler spread is the maximum Doppler shift

and is given by

$$B_D \approx \frac{v}{c} f_c$$

where v is the speed of the node, $c \approx 3 \cdot 10^8$ m/s denotes the speed of light and fc is the carrier frequency. Signals travelling along different paths can have different Doppler shifts, corresponding to different rates of change in phase. The difference in Doppler shifts between different signal components contributing to a single fading channel tap is known as the Doppler spread. The coherence time of the channel during which the channel stays essentially constant is approximately

$$T_c \approx \frac{0.423}{B_D}$$

In office environment, the Doppler spread varies from 2Hz to 20Hz which translates to coherence time from 20 to 200 ms [32]. These values are longer than the typical packet lengths in the wireless sensor networks, implying that the channel stays essentially constant during the transmission of single packet.

Large-scale effects

Large-scale effects can be understood by simply observing how the wireless channel affects the received power. *Frii's equation* relates the transmit power P_{tx} to the received power P_{rx} and is given in logarithmic form as

$$P_{\text{rx}}^{dBm} = P_{\text{tx}}^{dBm} + G_{\text{tx}}^{dBi} - L_p^{dB} + G_{\text{rx}}^{dBi}$$

where G_{tx}^{dBi} and G_{rx}^{dBi} are antenna gains at the transmitter and receiver, respectively, and L_p^{dB} is the *path loss*. The unit of the power is typically dBm, while the antenna gains are stated in dBi where the 'i' refers to isotropic antenna, i.e. an antenna that that radiates equally in all directions. One of the most simplest path loss models is the *single-slope empirical propagation model* given by

$$L_p^{dB} = L_0^{dBm} + n10\log_{10}(d)$$

The parameter $d > 1$ m denotes the distance between transmitter and receiver (in meters), n is the path loss exponent and L_0^{dB} is the path loss at one meter distance. The experienced path loss depends on the environment and typically the model parameters need to be tuned to match the environment: the parameter L_0^{dB} depends on the carrier frequency utilized; reflection from smooth surfaces can cause the direct path and reflected path to interfere such that the path loss exponent becomes equal to 4. The path loss exponent can be even less than 2 in some cases due to a wave guiding effect of the environment. Measurements conducted in factory environment suggest that the path loss varies from 1.49 measured for line of sight paths in light clutter (small number of echoes) to 2.81 in case of obstructed path and heavy clutter (large number of echoes) [48]. Slightly larger path loss exponents have been reported in [62] for metal processing facilities where reflections from metallic bodies are likely to happen. The reported path loss exponents range from 2.56 to 4.24.

The following two slope model

$$L_p^{dB} = \begin{cases} 40.2 + 20\log_{10}(d) & \text{if } d \leq 8 \text{ m} \\ 58.5 + 33\log_{10}(d/8) & \text{if } d \geq 8 \text{ m} \end{cases}$$

has been utilized by IEEE in the studies for the co-existence of various radios on the 2.4 GHz Industry, Science, Medical frequency band [21]. The model assumes line of sight path in the close proximity of the transmitter $(d < 8$ m $)$ and partially obstructed path for longer distances.

The impact of walls can be taken into account by increasing the pathloss. The attenuation cased by walls, floors and other objects which the radio wave can penetrate, depend on the dielectric properties of the materials of the object and frequency of the radio wave. The higher the frequency, the higher the pathloss. Majority of the empirical models available in literature consider building materials used in residential and office buildings. Models for residental and commerical buildings can be found from [22] Obstacles shadowing the radio path cause diffraction losses to the signal, the impact of which is typically modelled as lognormal random variable $\tilde{S}^{dB}$ called shadowing. Shadow fading has zero mean and its standard deviation varies between 3 to 20 dB based on the environment. Values 4 - 9 dB have been reported for factory environments [62]. In a given link, the shadow fading term stays essentially constant, S^{dB}, unless there are large changes in the node position with respect to the shadowing screens.

Outage

An important metric of the radio channel quality is the *Signal-to-Noise Ratio* (SNR). The instantaneous received SNR at time instant t can be written as

$$\tilde{\gamma}(t) = \frac{\tilde{g}(t)P_{\text{tx}}}{N_0 B}$$

Here N_0 is the noise power density given in terms of Watts per Herz (W/Hz) and B is the bandwidth of the signal. The random variable $\tilde{g}(t)$ is called the *link gain*. It is the ratio between received power and transmitted power. The link gain models both the short scale effects and large-scale effects. Assuming that the channel is *wide sense stationary* (i.e. the movement of the transmitters or receivers is slow compared to the time scale of interest), then the large scale effects are captured by the mean value of the channel gain

$$\overline{g}(t) = \mathbb{E}\{\gamma(t)\} = 10^{(G_{\text{tx}}^{dBi} - L_p^{dB} + S^{dB} + G_{\text{rx}}^{dBi})/10}$$

while the short scale effects explain the distribution of $\tilde{g}(t)$, $F_g(g) = F_P(gP_{tx})$ and its correlation properties. According to Clarke's model, the correlation can be expressed as

$$\rho(\tau) = \mathbb{E}\left\{(\tilde{g}(t) - \bar{g})\,(\tilde{g}(t+\tau) - \bar{g})\right\} = J_0\left(2\pi B_D\right)$$

where J_0 denotes the Bessel function of the first kind. The distribution of the received SNR is then $F_\gamma(\gamma) = F_g\left(\frac{N_0 B}{P_{tx}}\right)$.

Assuming that that the coherence time of the channel is large compared to the time it takes for the transmitter to transmit single information symbol (one symbol can consists of several bits depending on the utilized modulation method), the bit error probability can be expressed as monotonically decreasing function of the instantaneous SNR $p_{pe}(\gamma)$. If the coherence time of the channel is large compared to the packet length, then SNR stays essentially constant during the transmission of a single packet of b bytes. Hence, the packet error probability $p_{pe}(\gamma)$ is given by

$$p_{pe}(\gamma) = 1 - (1 - p_b e(\gamma))^{8b}$$

By fixing $p_{pe}(\gamma^{th}) = p_{pe}^{max}$, we can solve γ^{th} which guarantees that that as long as $\gamma \geq \gamma^{th}$, the packet error rate does not exceed p_{pe}^{max}. It is customary to express this threshold in terms of received power P_{rx} required to archive γ^{th}. This power value $P_s = \gamma^{th} N_0 B$ is called *receiver sensitivity*.

Example 1 (IEEE 802.15.4 O-QPSK PHY on 2.4 GHz band) *According to the IEEE 802.15.4 standard [20], the bit error probability conditioned on SNR (before despreading)* $\gamma(t) = \gamma$ *is*

$$p_{be}(\gamma) = \frac{1}{30}\sum_{j=2}^{16}\binom{16}{j}(-1)^j \exp\left(-20\left(1 - \frac{1}{j}\right)\gamma\right)$$

The resulting packet error probability PER is plotted in Figure 1 for 30 byte and 127 byte packets. It can be seen from the Figure, that the packet error rate is very high untill SNR exceeds certain threshold after which the PER drops quickly. Increasing physical layer packet size sifts the PER curve to the right. That is, for given SNR, the longer the packet, the higher the packet error probability - unless some measures are taken to protect the packet from bit errors. It can bee seen form Figure 1, that about 14 dB SNR is enough to obtain PER less than 0.001.

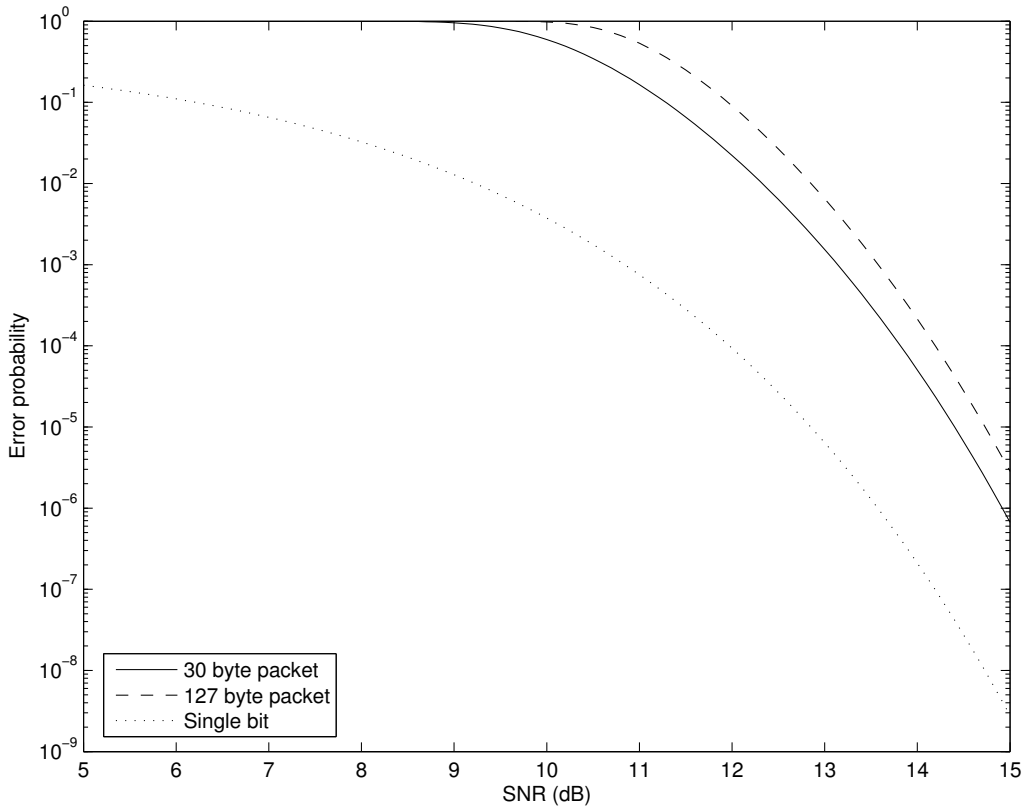


Figure 1: Packet error probability as a function of SNR after despreading for IEEE 802.15.4 O-QPSK PHY operating on 915 MHz band.

Due to the fact that the packet error probability has waterfall type of shape where after some threshold the packet drop probability falls down quickly, very simple model for the effect of channel induced packet drops can be utilized. The simple *outage* model assume that a packet is always successfully transmitted if the received power P_{rx} is above the receiver sensitivity; otherwise the packet is lost and we say that the link is in *outage*. Hence, in the average packet delivery ratio in the fading channel can be approximated as

$$p_{out} \approx F_P\left(P_s\right) \tag{1}$$

2.2 Markov models for the wireless channel

In wireless channels, the errors typically come in bursts. The simple outage model described above is not enough to characterize the time correlation of the errors. Markov models have been utilized to model the evolution of the channel SNR in time. A Finite State Markov Chain (FSMC) model of the fading channel partitions the received SNR into finite amount of intervals Each state corresponds to certain SNR interval. The transition probabilities between adjacent states are then selected to match the one sided level crossing rates of the fading process and the state probabilities simply refer to the probability of finding the received SNR in the corresponding interval [66]. There are, however, many ways of doing the partitioning and the choice will have an impact on the simulation results [23].

From application point of view, the instantaneous value of the channel SNR is not interesting. Rather we are interested in the time correlation between consecutive drops of packets. Hidden Markov Models (HMM) have been suggested for this end. In regular Markov chains, the state (e.g. SNR) is directly visible to the observer, and therefore the state transition probabilities are the only parameters. In a hidden Markov model, the state is not directly visible, but output dependent on the state (packet checksum ok or not) is visible. Each state has a probability distribution over the possible output values (packet error probability). The simplest HMM channel model is the Gilbert-Elliot (G-E) model [15, 12] . In the G-E model, the underlying discrete time Markov chain has only two states labeled as 0 or 'bad' and 1 or 'good'. In bad state, the packet is lost with probability $p_{pe,0}$ and in good state it is lost with probability $p_{pe,1}$. Let p_{01} and p_{10} denote the state transition probabilities between the states 0 and 1 and vice versa. Furthermore, let P_0 and P_1 denote the probabilities of finding the channel in bad and good state, respectively. In steady state, we have

$$\begin{aligned} P_0 &= \frac{1}{1+\frac{p_{01}}{p_{10}}} \\ P_1 &= \frac{p_{01}}{p_{10}}\frac{1}{1+\frac{p_{01}}{p_{10}}} \end{aligned} \tag{2}$$

The expected packet error rate of the channel is then $p_{pe} = P_0 p_{be,0} + P_1 p_{pe,1}$. Figure 2 illustrates the state transition diagram of the G-E model.

The error probabilities in the state and the transition probabilities can be fitted to measurement values. The state holding times of the Markov chain follow Geometric distribution with mean $1 - p_{ii}^{-1}$, $i = 0, 1$. The simplest, but somewhat artificial model is to have $p_{pe,0} = 0$ and $p_{pe_1} = 1$. In such case, the state holding time for state 1 is equivalent to the length of the error burst and the steady state probability P_1 gives directly the packet error probability. More realistic parameters can be obtained by fitting the model to the measurement data. The two-state G-E model can be viewed as first order approximation to the true channel in which only the mean length of the error burst is fitted to the data. If also higher moments need to be matched, then more states should be introduces. The G-E model is very popular among the researchers in the performance analysis of higher layer protocols (transport and application layer) due to its simplicity. See e.g. [23] and the references therein.

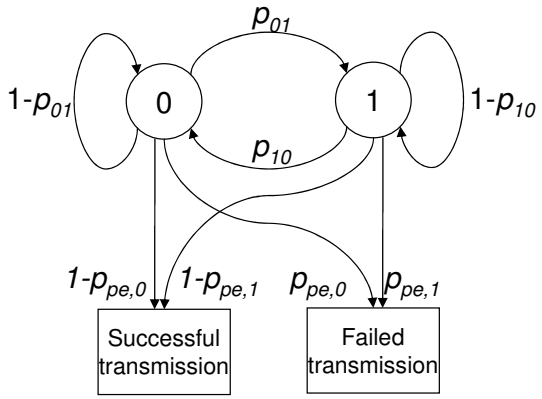


Figure 2: State transition diagram of the G-E HMM model.

2.3 The ISM band, co-existence and interference

Wireless devices are heavily regulated throughout the world. Most countries have allocated parts of the radio spectrum for open "license-free" use, while other parts of the spectrum can only be used with permission or "license". The license-free areas of the spectrum are also known as industrial, scientific, and medical (ISM) bands. The 2.4 GHz ISM band is available in most part of the world,

while ISM bands in lower frequencies vary between continents (e.g. Europe provides license-free access in the 433 and 869 MHz bands, while the Americas has opened up the 915MHz band). Each frequency band is split into channels: higher frequency bands are wider, which allows for wider channels and higher data rates, but lower frequencies give greater operating distance at a given transmit power, and a less crowded spectrum ensures more reliable operation. Radio communication systems that use the ISM bands are typically free for use by anyone, but typically operate under a ”license exempt” regime that sets limits on power, spectrum spreading techniques, or duty cycles to limit interference and make sure that no single device steals all bandwidth.

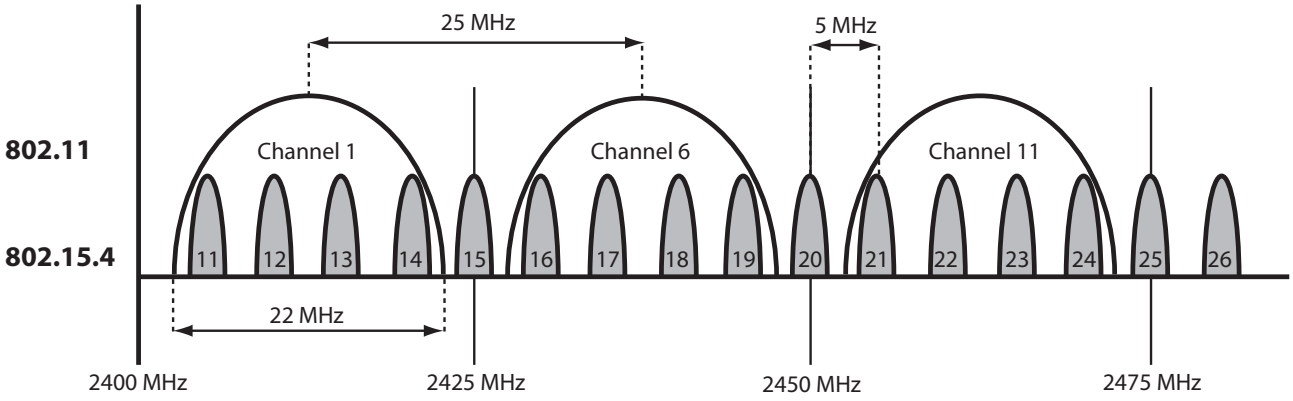


Figure 3: ISM band channels for 802.11 and 802.15.4. Only a handful of LR-WPAN channels do not overlap with WiFi.

In the 2.4 GHz ISM band, at least three wireless technologies are likely to coexist in current industrial deployments: wireless local area networks (WLAN, IEEE 802.11), Bluetooth (IEEE 802.15.1) and low-rate wireless personal area networks (LR-WPAN, IEEE 802.15.4). Figure 3 illustrates the associated wireless landscape. The 802.15.4-2006 standard defines sixteen channels, numbered 11 to 26, in the 2.4-GHz band. Each channel has a bandwidth of 2MHz and channels are separated by 5MHz. The IEEE 802.11b and 802.11g standards operate in fourteen channels available in the 2.4-GHz band, numbered 1 to 14, each with a bandwidth of 22 MHz and a channel separation of 5 MHz. The most common WLAN configuration is to enable the mutually orthogonal channels 1, 6, and 11 as in Figure 3. Finally, the IEEE 802.15.1 Bluetooth standard operates in 79 channels available worldwide in the 2.4-GHz band. Numbered 0 to 78, each channel has a bandwidth of one MHz and a channel separation of one MHz.

These three technologies operate using different output powers and offer communication over different ranges: WLAN output powers are typically around 20 dBm and operate within a 100m range; LR-WPAN decvices use a transmit power of 0dBm and operate within a 50 meter range; finally, Bluetooth output powers are generally less than 4 dBm and ranges below 10m for the more commonly used class 2 devices such as wireless headsets and keyboards. The less common class 1 devices can operate at up to 20 dBm and typically within a 100m range. The higher transmit powers used by WLAN and Bluetooth indicate that it could be difficult for 802.15.4-devices to coexist with these technologies. This has indeed been verified in practice e.g. in [56, 45]. both report that a saturated WLAN practically kills the 802.15.4 connection unless the operating frequencies are off-set with at least 7-10 MHz. However, the IEEE 802.15.4 radio can achieve over 90% packet delivery ratio in the presence of 802.11b/g WLAN interference if the packet delivery ratio of the WLAN is low (less than 100 packets/s) or the signal-to-interference ratio at the 802.15.4 receiver exceeds 15 dBs [31]. The upcoming IEEE 11n standard WLAN is expected to be even more severe source of interference due to multipoint transmission and wider bandwidth. The influence of Bluetooth on 802.15.4 is more limited (packet error rates around 10% were reported in [56]) mainly due to the fact that it uses an agile frequency hopping scheme and relatively narrow channels. In the converse direction, 802.15.4 has virtually no impact on 802.11b performance unless the channels have the same center frequency

and the 802.15.4 packets are very long [56].

If we are interested in modelling the impact of WLAN traffic on LR-WPAN communications, a reasonable first-order model (considering the power imbalance between the two technologies) is to assume that the WPAN packet is lost if there is a concurrent WLAN transmission on an overlapping channel. Simple and reasonably accurate semi-Markov models for WLAN interference, modelling the frequency and duration of transmit and idle periods, respectively can be found in [14, 60].

2.4 Means for increasing reliability

2.4.1 Error control

In most of the applications it is important to detect whether the received packet contain errors or not. Hence, it is customary to use error detection coding. The most common error detection coding scheme is the Cyclic Redundancy Check (CRC) CRCs are popular because they are simple to implement in binary hardware using shift registers, are easy to analyze mathematically, and are particularly good at detecting common errors caused by noise in transmission channels. The IEEE 802.15.4 standard specifies 16 bit CRC to be utilized. The probability of an undetected error is bounded above by 2^{-16}. In a plant having 50 nodes each transmitting one packet per minute, the average number of undetected errors is approximately one per day. If 32 bit CRC is used, errors occurs once per 200 years.

While CRC codes use redundant bits for detecting errors, Forward Error Correction (FEC) codes use the extra bits to correct some of the bit errors at the receiver hence increasing the robustness of the communications. The performance enhancement achievable by using FEC can be expressed in terms of *coding gain*. The coding gain expresses the difference in the required Signal-to-Noise with and without coding for given bit error rate. The coding gain can be utilized to increase range, decrease transmitter power consumption or to increase robustness against noise and interference. See e.g. [64]. In the outage model discussed earlier, the coding gain simply translates to lower SNR threshold value γ^{th}.

Automatic Repear reQuest (ARQ) algorithms are utilized to improve the transmission reliability at the cost of increased packet delay. The simplest and also most common ARQ method used in WSNs is Stop-and-Wait ARQ (SW-ARQ). In SW-ARQ, once a mote generates a packet is sets a timer. If the receiver receives the packet successfully, it will transmit an acknowledge (ACK) packet. In case it detects an error, it can issue negative acknowledgment (NACK). It is also possible, that the packet gets entirely lost in the wireless channel and the receiver is not able to detect the transmission at all. If the tranmitter receivers ACK before its timer expires, it resets the timer and proceeds to the next packet. If it receives NACK, it resets the timer and retransmit the packet. If neither ACK or NACK is received before the timer expires, the transmitter deduces that the packet must have got lost and will try to retransmit the packet. Typically, the number of retransmission attempts is limited; after which the packet is considered lost.

The packet transmission using SW-ARQ can be modeled by unsign the Markov chain model and adding a absorbing state 'packet received'. In case of two state G-E model, we the number of transmission attempts $\tilde{a}$ needed to successfully transmit a packet follows discrete phase type distribution with two phases:

$$\Pr\{\tilde{a} \leq a\} = 1 - \boldsymbol{\tau}\mathbf{T}^a\mathbf{1}$$

and $\boldsymbol{\tau} = (P_0, P_1)$ denotes the probability of starting the process from state 0 and 1, respectively. The matrix $\mathcal{T}$ is given by

$$\mathbf{T} = \begin{bmatrix} (1-p_{01})\,p_{pe,0} & p_{01}p_{pe,0} \\ p_{10}p_{pe,1} & (1-p_{10})\,p_{pe,1} \end{bmatrix}$$

The vector $\mathbf{1} = (1, 1)^T$ is a column vector of ones. Given that the maximum number of retransmission attempts in the system is a, we can can obtain the residual packet error probability p_{pe} after a transmission attempts assuming ideal error detection:

$$p_{pe}(a) = \Pr\{\tilde{a} > a\} = \boldsymbol{\tau}\mathbf{T}^a\mathbf{1}$$

Example 2 (Performance of ARQ) *The packet error characteristics for IEEE 802.15.4 [20] standard radio operating on the 2.4GHz band were measured in an industry assembly hall. Measurement results obtained from a 30m link between ground level and overhead crane suggested that the mean packet error probability is $p_{pe} = 0.133$ for 119 byte packets send by the rate of 10 packets/second. A G-E model was fitted to the packet delivery traces suggesting $p_{pe,0} = 0.02$ and $p_{pe,1} = 0.74$. The state holding times for good and bad state were 35.54 s and 6.58 s, respectively. Figure 1 show the packet error probability as a function of the number of allowed transmission attempts per packet for both the G-E model and a simple model that assumes that packet errors are uncorrelated and error occurs with probability p_{pe}. It can be seen form Figure 4 that the bursty nature of the channel makes reliable communication much more difficult than what the simple geometric model assuming independent errors would suggest.*

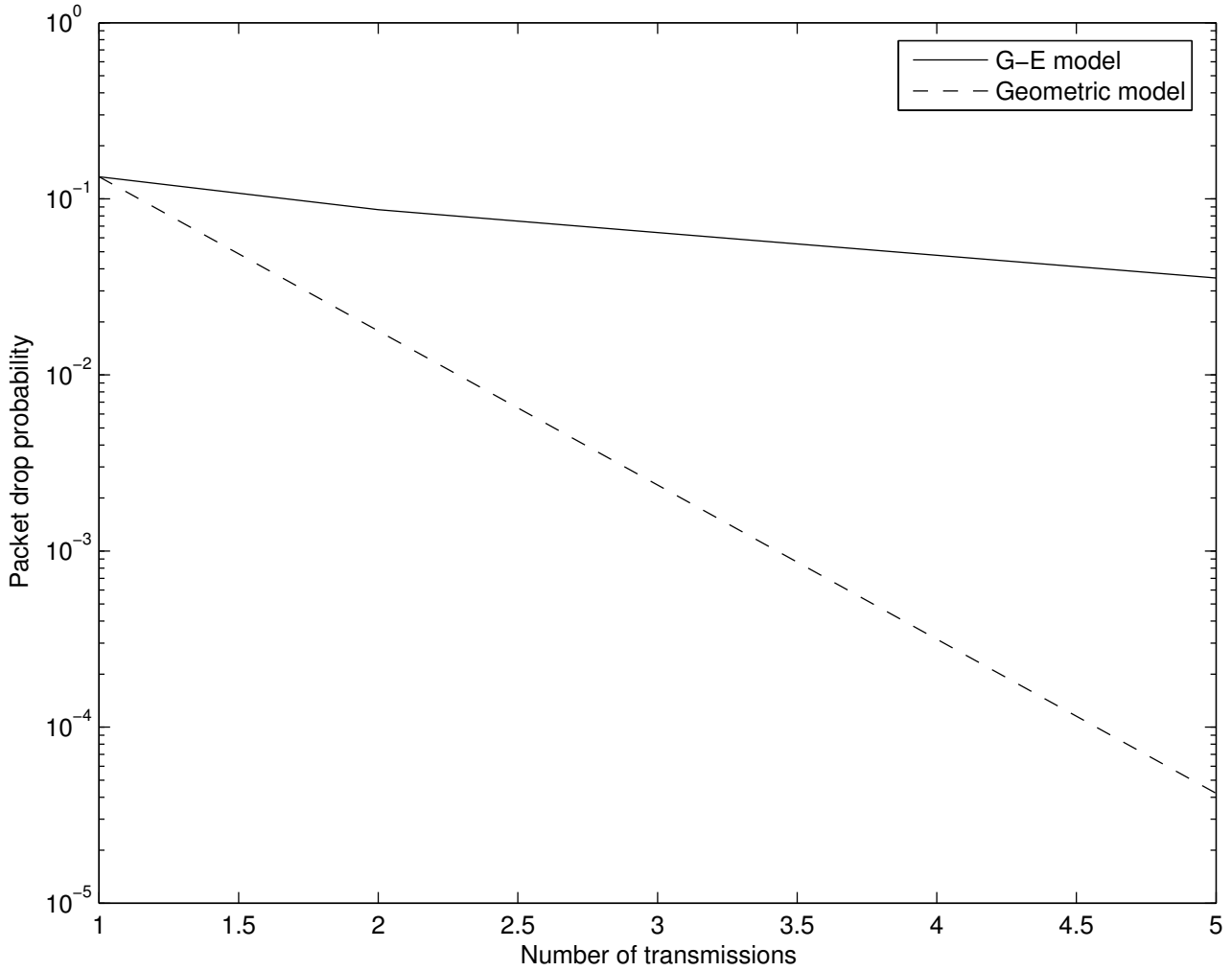


Figure 4: Packet error probability as a function of allowed number of transmission attempts for channel with error burst (G-E model) and independent errors (Geometric).

FEC can be combined with ARQ to obtain Hybrid ARQ (HARQ). HARQ schemes aim to exploit the advantages of both FEC and ARQ by incrementally increasing the error resiliency of the packet

through retransmissions. The HARQ schemes are typically divided into two types. Type I-HARQ adds FEC on top of ARQ. The transmission starts with low number of code bits or possibly with CRC only. Once the packet fails, the retransmitted packet is encoded with stronger code. In type II, first uncoded packet with CRC is sent. Once the packet fails, the transmitter only sends code bits (incremental redundancy). The code bits are typically selected such that the packet is self-decodable and together with the original packet forms 1/2 rate code. The receiver can first try to check if the retransmitted packet was successful - if not, it then decodes the codeword composed by combining the original packet and the retransmitted one. Measurement resutls with MicaZ nodes employing IEEE 802.15.4 radio, indicate that HARQ-II is the best to reduce latency and energy consumption [Akyildiz 2009].

2.4.2 Diversity techniques

Due to frequency dependent multi-path fading certain frequency channels can be severely faded. In wireless sensor networks, the channel is typically slowly changing which indicates that once the channel is deeply faded, the situation is likely to remain the same for long time causing many consecutive packets to get drop. The fading seen by two frequency channels that are spaced more than the coherence band from each other see independent fading. This fact can be exploited to decorrelate the transmission of the consecutive packets by utilizing different channel for each transmission attempt. *Frequency hopping* FH is a diversity technique, in which the synchronized transceivers changes channels according to some predefined *hopping pattern*. It is still possible that some channels are deeply faded or contain a lot of interference. In such situations, the system should avoid bad channels. To avoid these problems, adaptive FH techniques (AFH) have been developed. The basic idea implemented by these schemes is to remove from the hopping sequence frequencies experiencing bad channel conditions (for instance high packet error rate), consequently reducing the number of utilized frequency bands. To this purpose the resulting algorithms make use of a channel classification procedure that is basically required to identify channels unsuitable for packet transmissions. This procedure however introduces delays. Moreover, channels removed from the hopping sequence need to be periodically checked to verify if they still dont meet the specified performance requirements. An alternative to channel classification is to use probabilistic (pseudo random) frequency hopping where the probability of utilizing certain channel is constantly updated based on the estimated packet error rate for the channel. The use of FH in wireless sensor networks is discussed e.g. in [59].

In some applications, the number of available channels for sensor network operation can be very limited. For instance, in case of heavy WLAN traffic, there is only one channel available for IEEE 802.15.4 which is completely free from the WLAN. Spatial diversity can be exploited by utilizing multiple antennas on the sensor motes. If the antennas are at least half a wavelength ($\frac{\lambda}{2}$) apart from each other, and there is enough scattering in the environment, then each antenna would see independent fading. Just like in case of FH, we cab decorrelate the transmission of the consecutive packets by utilizing different antenna pairs for each transmission attempt. In low mobility scenarios, simple electro-mechanical RF switch can be utilized to select between multiple antennas. *Antenna switching* in WSN setting has been studied in [53] where it was demonstrated by laboratory scale tests.

If the receiver has multiple RF units, it is possible to utilize *receiver selection diversity*. In the simple outage model the packet error rate of single transmission was $p_{out} = F_P(P_s)$ whit M-fold selection diversity (receiver has M radios), the outage becomes $p_{out}(M) = F_P^M(P_s))$ that is, the outage probability decreases exponentially as the number of receiver increase.

Example 3 (Spatial diversity) *Example 2 illustrated G-E parameters for one particular 30 m link in an industry assembly hall. The measurements were done using two transmit antennas and four receiving antennas. The antennas were placed half an wavelength from each other. The resulting*

8 single input - single output spatial links were tested together. The resulting packet error rates (PER) for the links (labeled L1,L2,....L8) are plotted in Figure 5. The figure also shows result for random antenna switching (RS) that selects the transmit and receive antenna randomly for each transmission. The resulting PER is the average of individual link PERs. Ideal antenna switching (IS) selects the transmit - receiver antenna pair that minimizes PER. If there is separate receiver for each antenna at the sink node, we can utilize receiver diversity (D1 and D2 corresponding to the choice of transmitter antenna). In case receiver selection diversity is utilized, packet is only lost if none of the four receivers were able to capture the packet. The receiver diversity scheme can be combined with random transmit antenna switching (RSD) and ideal transmit antenna selection (ISD).

The results indicate that in certain radio environments, the channel quality is very sensitive to the placing of transmit and receiver antennas. If the node size is not a limiting factor, then antenna diversity is a viable option to obtain diversity gains and increase the communication reliability.

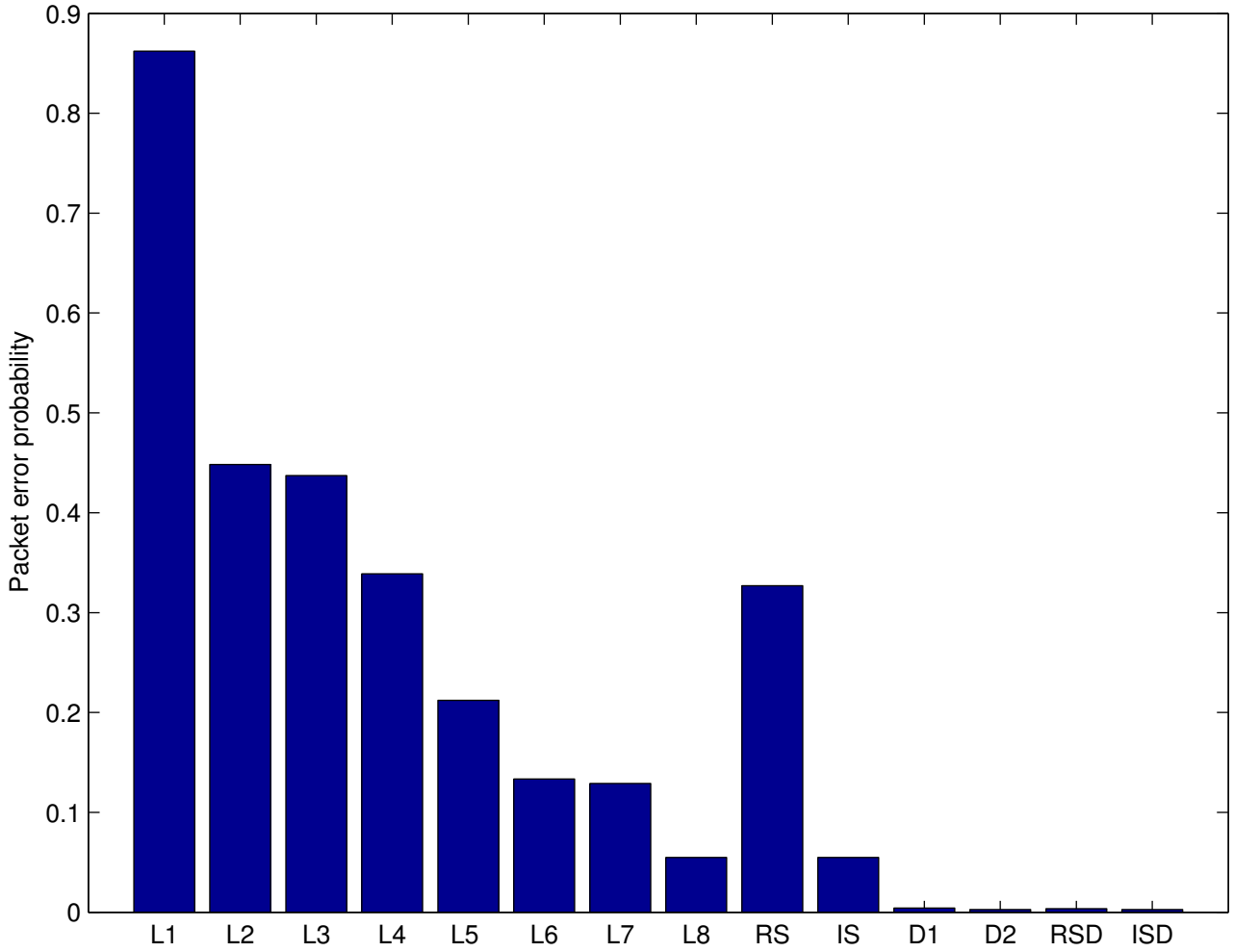


Figure 5: Packet error probability for various transmit antenna and receiver antenna configurations.

3 Multiple links: medium access control

Aim: to understand various MAC protocols, their advantages and disadvantages, and how channel utilization and delay can be analyzed.

The wireless medium is inherently broadcast, in the sense that a transmitter can be heard by multiple

receivers and a receiver can hear many transmitters. If two transmitters transmit at the same time, their signals may interfere and become unrecoverable by the receivers, leading the associated packets to be lost. The role of the *medium access control* (MAC) protocol is to define a set of rules for how to share the medium between transmitters to avoid interference and communicate efficiently.

Although many MAC protocols exist, they can be broadly categorized as scheduled and contention-based. In scheduled approaches, the medium is divided into fixed non-overlapping portions (in frequency, time, etc) and assigned to individual transmitter pairs once and for all. In contention-based approaches, on the other hand, transmitters try to access the medium when they have traffic and release the channel when all data is sent. Roughly speaking, scheduled methods tend to work best when the traffic (the number of users and their traffic patterns) is predictable, while contention-based approaches are more efficient when the traffic is unpredictable.

To quantify this intuition, we need to specify the user traffic on the network. It is customary to assume that sensors generate packets according to a stationary ergodic process (e.g. that packet arrival times follow a Poisson distribution with a given fixed intensity). We say that the traffic is *light* when the probability that a node has multiple packets in its buffer is negligible and *saturated* when this probability approaches one. Classical analysis of MAC schemes assumes that the the arrival process of packets to nodes are mutually *independent*. However, in most sensor applications, sensor reading events tend to be *correlated*. An extreme situation, common in control scenarios, is when all sensors generate packets simultaneously. We will review some of the underlying principles of common MAC schemes next.

3.1 Scheduled medium access: TDMA and FDMA

The most common scheduled medium access techniques are time-division multiple access (TDMA) and frequency-division multiple access (FDMA). In TDMA, the time axis is divided into time slots of fixed length, and transmitter-receiver pairs are allocated a fixed number of time slots in which they can communicate. FDMA, on the other hand, divides the available communication bandwidth into disjoint frequency bands ("channels"), and allocates a fixed number of channels to each user. The two techniques can be combined into multi-channel TDMA protocols, where the communication channel is divided in frequency and space. Such an approach is natural in 802.15.4 networks, since the ISM band is channelized.

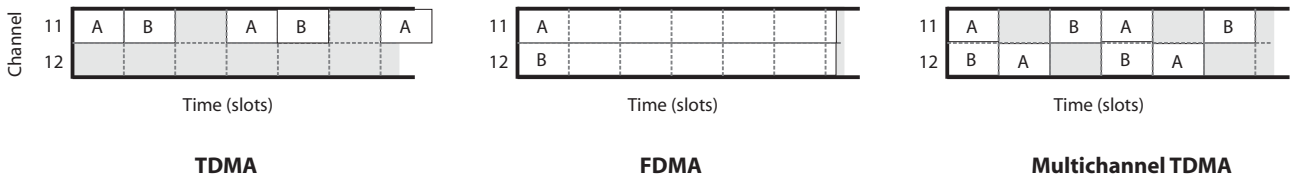


Figure 6: TDMA, FDMA and multi-channel TDMA for two transmitters (A and B) on 802.15.4-channels 11 and 12 in the ISM band.

Neither TDMA nor FDMA is able to use the full channel resource for communication. TDMA requires that all nodes are synchronized to a common clock. Typically, some bandwidth is lost to synchronization messages between nodes and time slots are made slightly longer than what is required to perform a single data transaction to allow for slight synchronization errors. Similarly, FDMA requires guard bands between channels to avoid co-channel interference. For example, the 802.15.4 channels only occupy 32 out of 100 MHz in the 2.4 GHz ISM band. Similarly, out of the 10ms slot time used by the WirelessHART multi-channel TDMA protocol (see § 4.4), only 5.088 ms are used for data and acknowledgement transmissions.

Another performance loss occurs when there is a mismatch between the traffic characteristic and the deterministic transmission opportunities of scheduled access:

Example 4 (TDMA) *To make this discussion more specific, consider a system with N sensors, each generating traffic according to a Poisson process with intensity g/N. The medium access is scheduled using TDMA with a frame of N slots of equal length T. Since a packet needs to wait until the next time slot allocated to that specific sensor appears in the schedule, the average delay is given by*

$$D = T\left(1 + \frac{N}{2(1-gT)}\right) \tag{3}$$

if the throughput (or busy probability) gT does not exceed one [51]. As we will see shortly, contention-based MAC schemes, are sometimes a better match with random traffic.

3.2 Contention-based medium access: Aloha, CSMA and beyond.

While scheduled medium access such as TDMA provides deterministic latency (under perfect channel conditions), it does come with some implementation overhead: clocks need to be synchronized, topology information is needed for efficient transmission scheduling, and nodes joining and/or leaving the network dynamically typically means that the complete schedule needs to be recomputed and redistributed. Many of these issues are eliminated in contention-based medium access schemes.

The basic idea of contention-based medium access is the one of conflict resolution: users simply transmit whenever they have a packet to send and follow a specific set of rules to resolve conflicts when collisions occur. One of the simplest contention-resolution protocols is the Aloha-protocol [1]. In its slotted variant ("Slotted Aloha") [50], nodes are assumed to be synchronized and time is subdivided into slots, each of sufficient length to transmit a single packet. Whenever users have data to transmit, they will attempt to do so at the start of the next slot. If a collision occurs, then the node becomes backlogged. Backlogged nodes transmit in each slot with probability q until successful. The protocol performance can be optimized by tuning the access probability q. If the system is saturated, *i.e.* all N nodes have enough packets to attempt to access the medium in every slot, then one should make sure that $q < 1/N$ to avoid excessive collisions. In light traffic conditions, higher access probabilities might be advantageous, see e.g. [47].

In many multi-access channels, it is possible for a node to detect when other nodes are transmitting (after a slight propagation and detection delay). When this *carrier sensing* functionality is available, it is natural to check if the medium is free before transmitting. This leads to the so-called *p-persistent CSMA protocol*, which works as follows: whenever a node has data to send, it first performs carrier sense; if the channel is busy, the node does not transmit in the current slot; if the medium is idle, then the node transmits with probability p (and refrains from transmitting with probability 1-p). An alternative CSMA variant is the *non-persistent CSMA*. Also here, nodes perform carrier sense when they have data to send and only attempt to transmit their data if the channel is considered idle. If the channel is busy, on the other hand, the node waits a random time before trying to access the channel again. The random wait mechanism is often implemented using a *back-off counter*: nodes initialize a counter by drawing a random number in the interval [0,CW] (CW for *contention window*) and decrement the counter in each slot. When the counter reaches zero, the node performs carrier sense again in hope of gaining channel access. To understand the relationship between the contention window size in non-persistent CSMA and the access probability in p-persistent CSMA, it is useful to consider the average waiting time after a detected collision. This waiting time is CW/2 for non-persistent CSMA and 1/p for p-persistent CSMA. Setting p=2/CW typically results in comparable performance of the two approaches [61]. Note, however, that the access probability distributions differ (uniform over time in non-persistent CSMA and geometrically distributed in p-persistent) and that this can influence the performance of certain applications [73].

Since the optimal access probability depends on the traffic load, it is natural to try to automatically adapt the contention window size based on the perceived congestion. The most common solution is the *binary exponential backoff* (BEB) algorithm. Here, the contention window is doubled each

time a collision occurs (until it reaches a maximum allowed value) and reset to its minimum value at the beginning of each new transmission. Many successful wireless standards, such as 802.11 and 802.15.4, use medium access control mechanisms that are based on non-persistent CSMA with BEB. However, the actual implementation varies in many (sometimes subtle) ways from the basic description above. For example, in 802.11, the backoff timer only elapses when the medium is idle. Since listening consumes significant power, 802.15.4 decrements the backoff counter irrespectively of the state of the medium, but then performs carrier sensing in two consecutive slots to judge if the medium is idle. It is important to re-iterate that these mechanisms were introduced as a means to adapt to an unknown traffic situation. If the traffic is known, then the binary backoff offers few, if any, advantages over an appropriately tuned fixed contention window [34]. Naturally, one can apply similar ideas to p-persistent CSMA, e.g. halving the access probability at collision events [7]. Such approaches are sometimes called *predictive p-persistent CSMA*, and standardized in *e.g.* the LON Talk protocol of the ANSI/EIA 709.1 standard [37].

The following example illustrates the performance of contention-based MAC protocol under saturated random traffic.

Example 5 (Slotted ALOHA and CSMA delay under Poisson traffic) *We assume that packets arrive at the rate g packet per time unit. The transmission time of the packet is T. Let $G = gT$ denote the traffic volume in terms of packets. In slotted ALOHA, transmissions can start only in the beginning of a slot. The slot length is equal to the transmission time of the packet T. Under the Poisson arrival assumption and large backoff range, the consecutive transmission attempts become independent of each other and follows geometric distribution with mean $1/p_s$. The packet delay stays bounded as long as the probability of successful transmission fulfills $p_s = exp(-G) > 0.5$ which implies upper bound for the offered traffic $G = gT < ln2$. Let ω denote the initial CW window size. The expected delay of the protocol have been derived in [72] and it is given by*

$$D = \frac{T}{2}\left(\frac{3}{p_s} + \frac{\omega p_s}{1-2(1-p_s)} - \omega\right) \tag{4}$$

In case of slotted CSMA, we assume that the time is divided into slots of length $\tau = aT$, $a \leq 1$. The slot length is selected to conside with the carrier sensing time. As in the slotted ALOHA case, transmissions can only start in the beginning of a new slot. A sensor wishing to transmit must sense the channel to be free for the period of one slot. If multiple sensors generated packets during the same slot, a collision will occur. The probability that the transmission is successful in case CSMA is utilized was derived in [27]

$$p_s = \frac{ae^{-aG}}{1+a-e^{-aG}}$$

For CSMA with BEB, the expected access delay stays bounded if $p_s > 0.5$ and has been derived in [72]:

$$D = \frac{T}{2}\left(\frac{a\omega p_s}{1-2\left(1-p_s\right)} + \frac{2+5a}{p_s} - \frac{2+4a}{p_s}p_b - a\left(4+\omega\right)\right) \tag{5}$$

The throughput for slotted ALOHA and CSMA is given by $S = p_s G < 1$.

The delay - throughput characteristics of the three protocols are shown in Figure 7. It can be seen that the maximum throughput of the ALOHA and CSMA is much less than what can be achieved with TDMA, but for low load, both contention based protocols can offer significantly shorter delay. It should be noted, however, that this conclusion is only valid for Poisson traffic. If all sensors generate packets at the same time, CSMA will result in multiple collisions and leading to significantly longer transmission times than scheduled transmission.

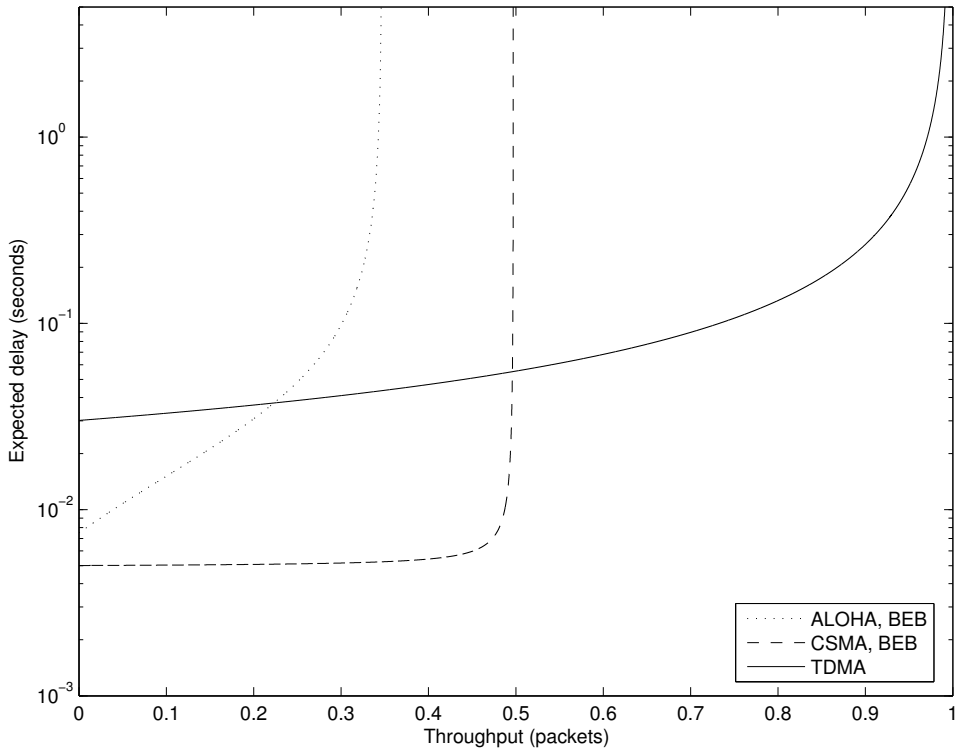


Figure 7: Throughput-Delay characteristics of TDMA (solid line), CSMA with BEB (dashed line) and ALOHA with BEP (dotted line).

Example 6 (Delay under correlated traffic) *In control systems, traffic is seldom completely random, but is typically correlated. For example, control loops often request synchronous samples from multiple sensors. To study this scenario, consider an idealized case where N sensors each generate a single packet at the same time and attempt to transmit these over a shared channel. For slotted Aloha, assuming that all sensor uses the same transmit probability p, the delay distribution as been derived in [61] and is given by*

$$\mathbf{Pr}(D = d) = \sum_{i=1}^{N} a_i \sum_{j=1}^{N} \frac{(a_j - 1)^N}{\prod_{k \neq j}(a_j - a_k)} (1 - a_j)^d$$

with $a_k = k(1-p)p^{k-1}$. While similar expressions can be derived also for CSMA with fixed contention-window, it appears hard to derive closed-form expressions for CSMA with BEB. Figure 8 demonstrates the latency distributions when $N = 10$ sensors simultaneously content for access using slotted Aloha (with $p = 0.1$), CSMA (with $CW = 20$) and CSMA with Binary Exponential Backoff, respectively.

Our discussion of CSMA protocols has assumed that carrier sensing is able to reliably detect conflicting transmissions. In reality, two phenomena known as the *hidden terminal* and *exposed terminal* problem, respectively, makes carrier sensing error-prone and degrades the performance of CSMA protocols. The hidden terminal problem occurs when two nodes which are outside each other's transmission range attempt to communicate with the same destination node. Since the two transmitter nodes cannot sense the carrier, they consider the medium to be idle and might transmit simultaneously, leading to collisions. The exposed terminal problem, on the other hand, occurs when two transmitters which are within transmission range of each other try to communicate with two different, well separated, receivers. Carrier sensing then blocks both transmissions although they would not have collided. The exposed and hidden terminal problems can be alleviated to some extent by implementing a two-way "request-to-send/clear-to-send" handshake between transmitter and receiver prior to the actual data transmission (see, *e.g.*, [70] for a discussion).

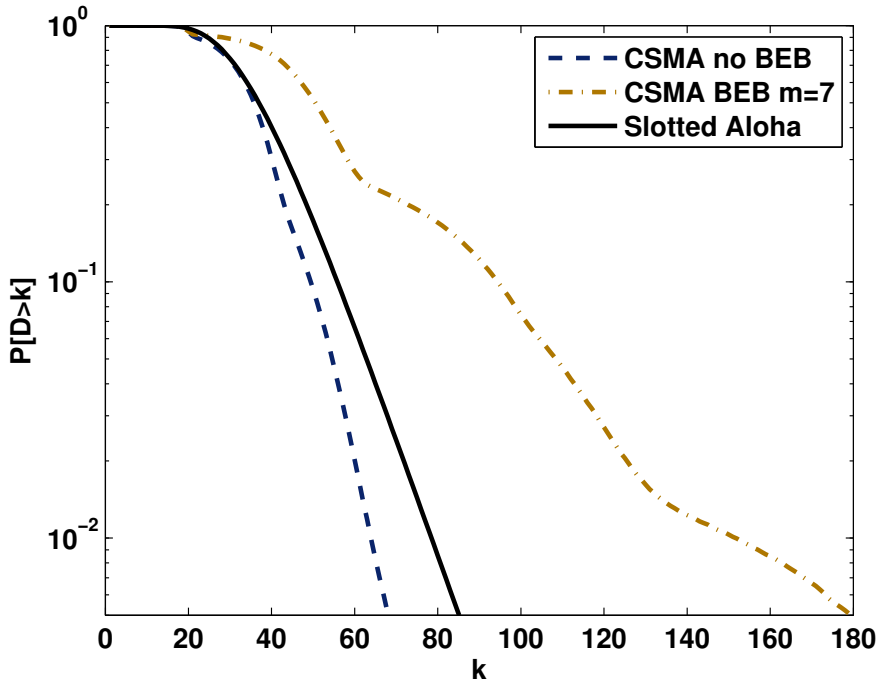


Figure 8: Delay distribution for N sources that start contending for access at the same time (from [61]).

3.3 Dynamic access scheduling via polling and reservation.

While Aloha and CSMA allow nodes to arbitrate channel access under varying traffic conditions, the presence of collisions limits their performance at high traffic load. Better channel utilization can be achieved by scheduling nodes so that no collisions occur. In dynamically scheduled medium access schemes, nodes reserve the channel before transmitting. Channel access is coordinated by a central node and can be based on either polling (the coordinator asks nodes if they have data, as in the 802.11 point coordination function PCF) or reservation requests (nodes demand channel access from the coordinator, as in the 802.15.4 Guaranteed Time Slot service GTS). Dynamically scheduled medium access can be efficient if the reservation overhead is low, but is typically restricted to very simple node topologies. The following example provides some more details about the dynamically

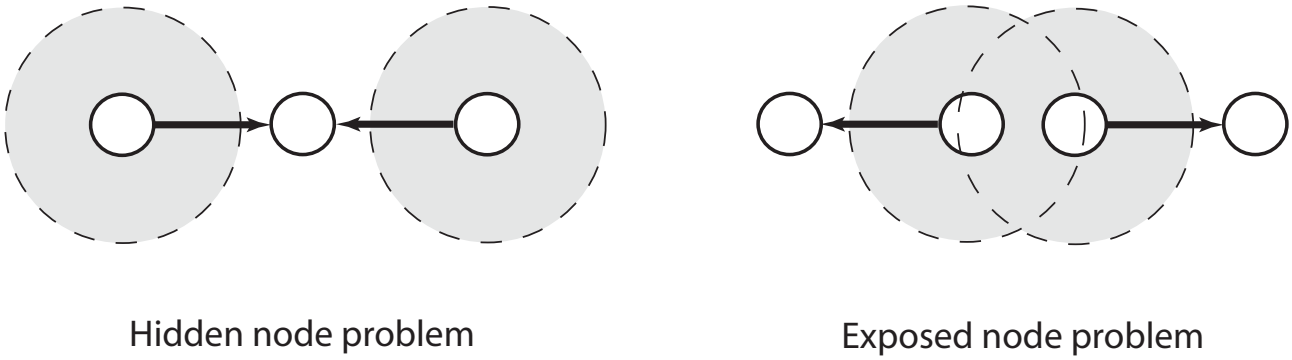


Figure 9: Pictorial illustration of hidden and exposed node problem. Carrier sensing blocks nodes from transmitting within the carrier sensing range (marked in light grey) of an ongoing transmission. Hidden nodes may receive simultaneous transmissions despite collisions, and exposed nodes are block although simultaneous transmissions would be possible without collisions.

scheduled access in 802.15.4.

Example 7 (Dynamic scheduling in 802.15.4) *The beacon-enabled mode of the 802.15.4 MAC supports a combination of contention-based and dynamically scheduled access. As illustrated in Figure 10, the MAC operates in a cyclic fashion where an initial beacon triggers a sequence of contention-based access, scheduled access, and sleep periods. To reserve time slots in the contention-free period, nodes send reservation requests to the coordinator during the contention-access period. The requests can be for multiple time slots and extend over multiple superframes, until deallocated by the coordinator or the node itself.*

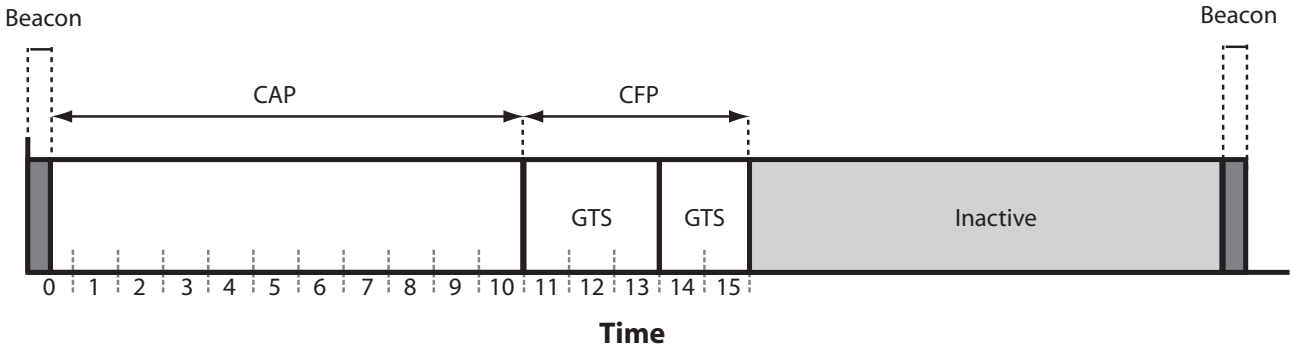


Figure 10: Superframe structure of 802.15.4 in beacon-enabled mode. After the beacon, the active part of the superframe is divided into a contention access period and a contention free period.

The performance of the IEEE 802.15.4 beacon enabled mode has been analyzed e.g. in [36, 40]. The results indicate that the default parameter settings given in the standard lead to low performance and abrupt change from non-saturation to saturation regime, and that queue stability is limited by the amount of downstream traffic from coordinator to nodes. One observation, elaborated in [6], is that delay-sensitive control applications would have been better served if the CFP and CAP would have switched place in the superframe.

An noticeable alternative to the IEEE standards is the Z-MAC protocol [49], which combines TDMA with a mechanisms that allows backlogged nodes to detect and utilize vacant time slots. The basic idea is that the node assigned to a specific time slot has priority and starts its transmission at the slot boundary, while other backlogged nodes wait a short time before performing their carrier sense. In this way, the non-assigned nodes can sense the carrier of the prioritized node and refrain from transmission if the slot is occupied. Similar ideas for improving the efficiency of TDMA have been included in the ISA100 standard described later in this chapter.

3.4 Energy-efficient medium access control

Although the power consumption of wireless nodes can be reduced by avoiding collisions and unnecessary protocol overhead, the largest energy savings can be obtained by turning off the transceivers when they are not needed for communication. For example, in TDMA it is clear that nodes can go to sleep whenever they are not scheduled to transmit or receive. If further energy savings are needed and the traffic rate is not that high, one can introduce a duty-cycle where nodes are awake for communication D% of the duty cycle period and asleep the rest of the time. Many MAC protocols for sensor networks require nodes to synchronize their duty cycles, but rely on contention-based medium access in the access period [74]. Duty-cycling becomes more challenging when nodes are not synchronized, and is then typically based on *preamble sampling*: whenever a node has data to send, it starts transmitting a train of short preambles. When a receiver wakes up, it samples the medium. If no preamble is detected, the receiver goes back to sleep. If a preamble is detected, the

receiver compares its ID with the one embedded in the preamble to see if it is the intended receiver of the pending data transmission. If it is the intended receiver, it sends an acknowledge message to the transmitter which aborts the preamble strobe and starts transmitting the actual data packet. Nodes that sample the preamble strobe but detect that they are not the intended receiver go back to sleep. Representative examples of such MAC protocols are B-MAC [46] and X-MAC [8].

4 From single links to network: the upper networking layers

4.1 Topologies and multi-hop communications

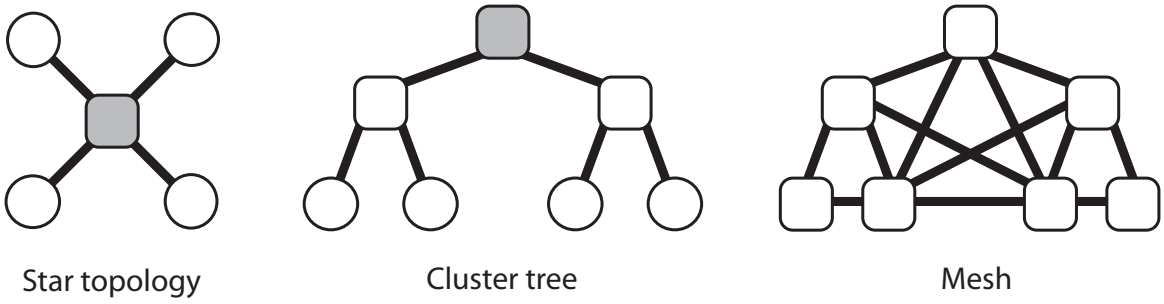


Figure 11: Star, cluster tree and mesh topologies. Squares denote devices with routing functionality.

Two or more nodes that communicate form a network. The simplest network is the single receiver-transmitter pair (the single link). Although interesting phenomena occur already when multiple wireless links co-exist within transmission range, the first non-trivial network topology is the *star topology*; see Figure 11. Here, a single node maintains connection with all other nodes in the network. In many cases, the central node coordinates transmissions and might also work as the gateway to external networks. The star topology has at least two disadvantages: it has limited coverage (essentially limited by the maximum distance over which a transmitter-receiver pair can sustain reliable communication) and limited reliability (limited by the reliability of a single link). The range can be extended by forming a *cluster tree*. Here, the children of the central node in the initial star topology also have the option to act as cluster heads, maintaining a star topology with their own children, etc; see Figure 11. In this way, data might be routed over multiple hops, and coverage is increased (at the expense of increased latency). However, there is still only one path from each source to the destination and the failure of a single intermediate node destroys the communication for all its children. The *mesh topology* allows addressing both coverage and reliability issues. In a mesh topology, all nodes participate in the forwarding of data packets from sources to destinations. By letting nodes maintain communication with more than one neighbor, it is possible to set up multiple paths from each source to their destination, and quickly divert data to a new path if an intermediate router fails. The following set of examples illustrate key aspects of multi-hop networking: the energy consumption for delivering a packet is likely to increase; the latency and loss rate both increase as the number of hops between source and destination increases; retransmissions allow to improve reliability at the price of increased delay, but are less effective when packet erasures occur in burst; multi-path diversity allows to increase reliability; and the interaction between medium access control and multi-path routing can severely degrade the system performance.

Example 8 (Energy and latency cost of multi-hop networks [13]) *Consider the three node topology in Figure 12 (left). Transmitter A has the possibility to send data directly to the sink C, or forward data via the intermediate node B. The intermediate route requires two transmissions ($A \to B$, $B \to C$) which doubles the end-to-end delay compared to the direct transmission. However, since the path gain decays as $d^{-\alpha}$, the shorter links can reduce their transmit power by a factor $2^{-\alpha}$ while*

maintaining the same received power as the direct link, multi-hop transmissions could potentially result in energy savings. Unfortunately, the energy gains often disappear when we account for the fact that radios consume significant energy also in reception mode. In [13], measurements on a 802.15.4 radio revealed a receive energy of 0.68 *times the maximal transmit energy* $P_{\max}$*. Even with a path loss of* $\alpha = 4$ *the difference between the energy cost of the direct path* $P_{Tx} + P_{Rx} = 1.68P_{\max}$ *and that of the two-hop path* $2 \cdot 2^{-4}P_{\max} + 2 \cdot 0.68P_{\max} = 1.485P_{\max}$ *is very limited, and no energy gain is possible when* $\alpha \leq 2.64$*. As Figure 12(right) reveals the energy gains disappear even earlier when the intermediate* B *node is moved from its ideal position in the middle of the nodes.*

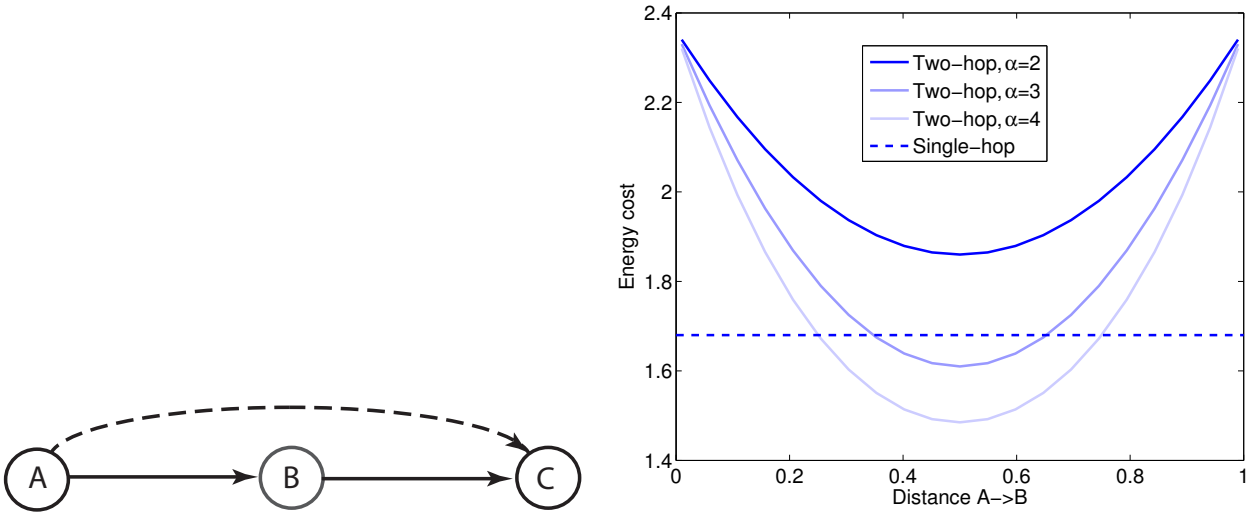


Figure 12: Node A can transmit directly to C or send data via the intermediate node B (left). The right figure demonstrates the total energy cost for communication under various path loss coefficients α and for various placement of the intermediate node.

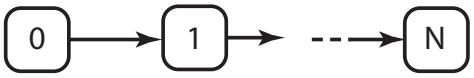
Example 9 (Latency and reliability of multi-hop communications) *To understand some of the latency and reliability issues of multi-hop networking better, consider the transmission of a single packet across the* $(N+1)$*-node topology shown in Figure 13. Assume a slotted system where a single time slot admits the transmission of one packet, and that link transmissions fail independently with probability* p*. The minimum latency is* N *time slots, and the probability of successful end-to-end communication in* N *time slots is* $(1-p)^N$*, which decays rapidly with* N*.*

The reliability can be increased if we allow a longer latency, so that nodes can retransmit a message if it does not receive a packet acknowledgement from the intended receiver. If we allocate two consecutive time slots to each link (corresponding to primary transmission and retransmission, respectively), then the probability of reaching the sink in $2N$ *time slots is* $(1-p^2)^N$*,* i.e. *the reliability has been increased a factor* $(1+p)^N$*. If we can allocate retransmission attempts dynamically (*e.g. *to let a node retransmit until successful) the probability of successful end-to-end communication with latency* $D > N$ *is given by [58]*

$$(1-p)^N \sum_{r=0}^{D-N} \binom{N+r-1}{N-1} p^r$$

When losses are bursty, the effect of immediate retransmissions diminishes, since the probability that the retransmission will also fail is high. The latency-loss curves for dynamic retransmissions on a line where link losses are given by the Gilbert-Elliot model are shown in Figure 13. Note that a significant latency is needed to ensure high reliability.

Example 10 (The mesh forwarding advantage) *As discussed earlier in this chapter, reliability can be improved by exploiting diversity. Rather than retrying a transmission on a bursty link, with high*



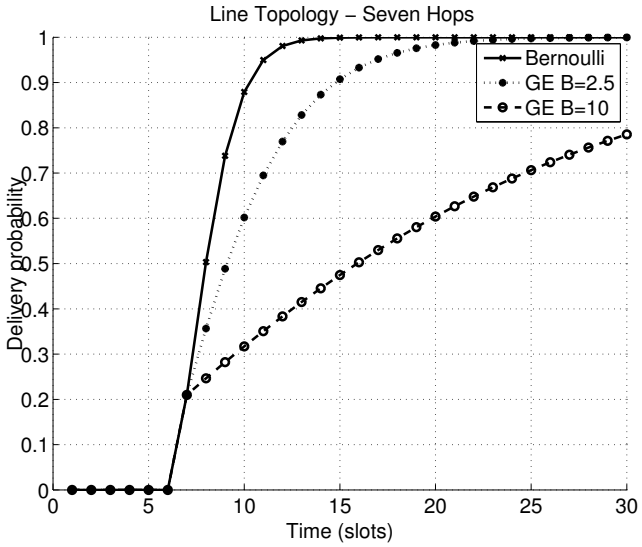
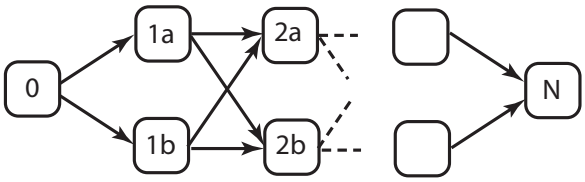


Figure 13: Single packet transmissions over N hops (left). The latency distributions for a link loss probability of $p = 0.2$ for line of length 4 and 8, respectively.

probability of continued loss, we could schedule the retransmission on another link (possibly also on another frequency). To illustrate the ideas, consider the "tube" topology shown in Figure 14(left) and assume that erasure events on links are dictated by independent (two-state) Markov chains, parameterized according to Gilbert-Elliot. The achievable loss-latency curve, shown in Figure 14(right) demonstrate an increased reliability compared to the line, with a significant reliability boost for low latencies [75].



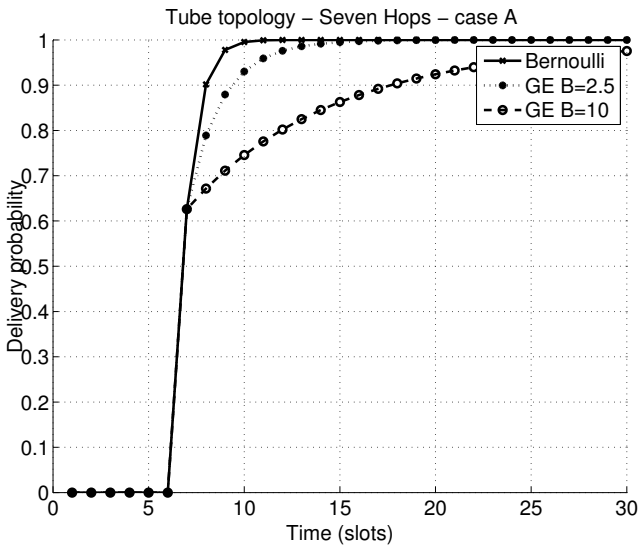


Figure 14: Single packet transmissions over an N-hop tube (left). The latency distributions for a link loss probability of $p = 0.2$ for line of length 4 and 8, respectively.

Example 11 (Multiple data streams and interaction with MAC) *To consider the interaction of medium access control and multi-hop networking, consider the problem of synchronized data collection from multiple sensors using the so-called* convergecast *operation. In convergecast, all nodes have a single packet that should be transmitted to the sink. Thus, if we consider a simple line with N sensors*

then convergecast on a single channel needs a total of

$$T_{\mathrm{TDMA}} = \sum_{n=1}^{N} n = \frac{N(N+1)}{2}$$

time slots. The latency can be decreased to

$$T_{\mathrm{MCH}} = 2(N-1)$$

time slots when we use multi-channel TDMA to allow for parallel transmissions [57]. Figure 15(right) demonstrates how the performance degrades significantly when nodes use slotted ALOHA to contend for channel access. Even using the optimal transmit probability, the average latency is 94.5 *time slots compared with the* 15 *slots required for single-channel TDMA and* 9 *slots for multi-channel TDMA.*

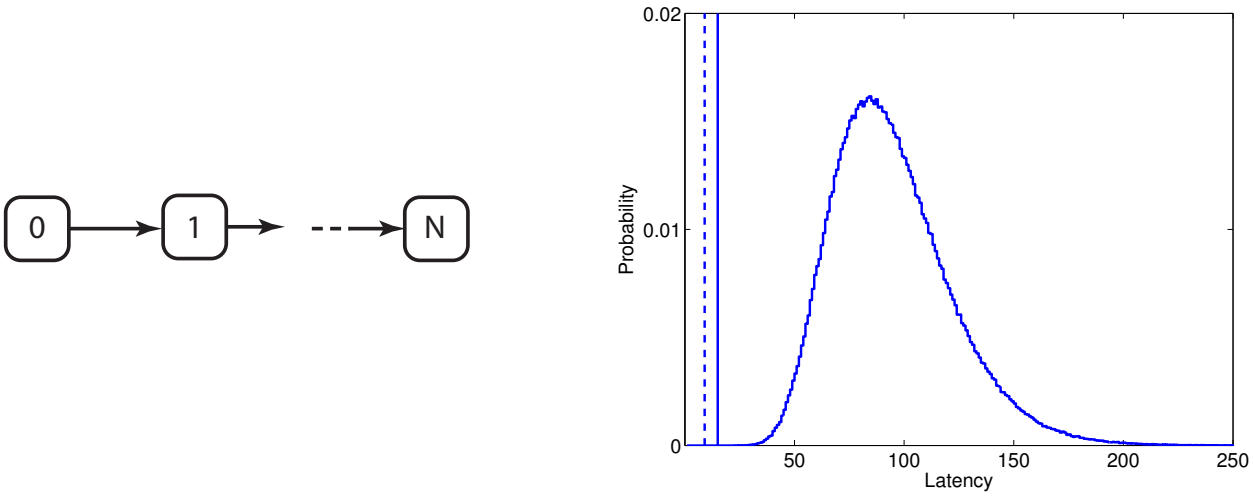


Figure 15: Convergecast operation using N sensors in line topology (left). For a 5 node case, the right plot reveals the latency distribution for TDMA and multi-channel TDMA (deterministic at 9 and 15 time slots, respectively) and the latency distribution for completion of convergecast under the ALOHA protocol.

4.2 Routing.

Routing is the process of selecting paths along which to send data traffic. Data packets are then forwarded from sources to their final destinations via the intermediate nodes on the selected routes. Most routing algorithms use a single network path between sources and destinations, while multipath routing protocols maintain several alternative paths to improve reliability. Routing protocols are further classified as either reactive (on-demand) or pro-active (table-driven) [52]. Table-driven routing protocols attempt to maintain consistent up-do-date routing information from each node to every other node in the network, while on-demand routing protocols gather such information only when needed.

Routing on resource-constrained low-power radio nodes is challenging: constrained memory and processing power limits the size of routing tables that can be stored, low-power radios experience higher loss rates than high-end technologies, and limited power supply requires energy-optimized implementations and low overhead traffic. The wireless sensing community has developed many routing protocols that respect these design constraints and exploit the traffic and data characteristics of sensor networks to optimize performance (see, *e.g.*, [2] for a survey). We will not go into the details of the specific protocols, but rather focus on the resulting routing topology and its performance.

Link-metrics and shortest path routing

Paths are typically selected based on some quality metric, such as latency or reliability. These path metrics can often be written as the sum of the costs for using each link in the path. For example, the end-to-end latency of a path is the sum of the transmission times for its individual links. When link costs are additive, the best paths can be found by solving a shortest path problem using, for example, the algorithms due to Dijkstra or Bellman-Ford (see, *e.g.*, [5]). When path costs are multiplicative, a simple transformation allows for casting the optimal route calculation as a classical shortest path problem. The following example illustrates this idea.

Example 12 (Maximum reliable path as a shortest path problem) *Consider the problem of finding the path with the maximum packet delivery probability. If links fail independently with probability p_l, then the reliability of path P, r_P, can be written as $r_P = \prod_{l \in P}(1 - p_l)$. Since the logarithm is monotone, the path that maximizes r_P will also maximize $log(r_P) = \sum_l \log(1 - p_l)$. Now, maximizing $\sum_l \log(1 - p_l)$ is the same as minimizing $\sum_l -\log(1 - p_l)$, so the most reliable path can be found by solving a shortest path problem with link weights equal to $-\log(1 - p_l)$.*

Related to this model of link reliability is the expected transmission count, or ETX for short [11]. The aim of ETX is to increase overall network throughput by finding paths with the lowest expected end-to-end latency. Observing that to avoid link level retransmissions, both the data packet and its link-level acknowledgement must be successfully received, the expected number of transmissions for a given link is approximated by $\mathrm{ETx} = 1/(d_f \cdot d_r)$. Here, d_f and d_r represent the delivery ratios in the forward (data) and reverse (acknowledgement) direction of the link. Note that under Bernoulli losses and symmetric links $d_f = d_r = (1 - p_l)$, in which case ETX and $-\log(1 - p_l)$ have a similar qualitative behavior.

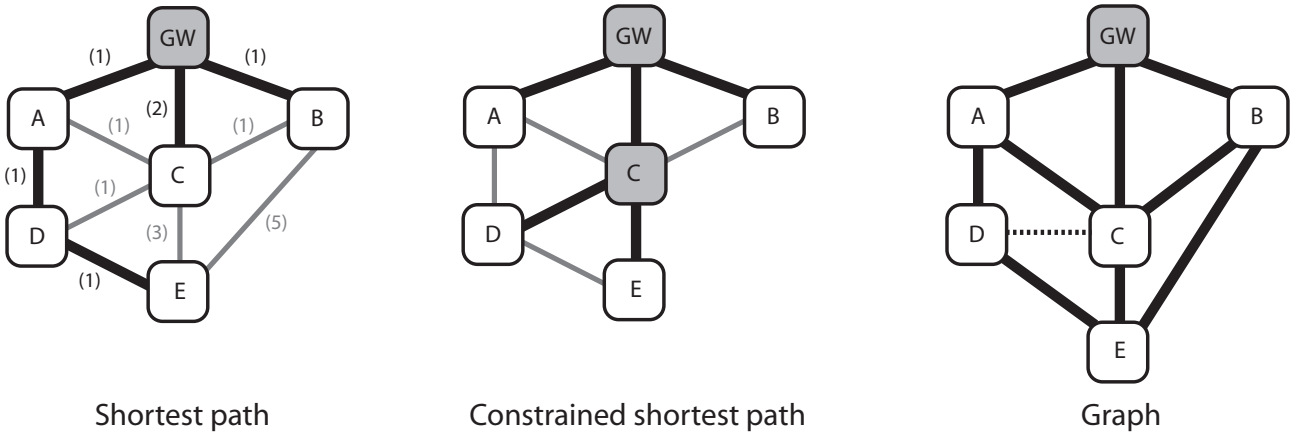


Figure 16: Routing strategies. Numbers in parenthesis indicate ETX for the links. The leftmost figure demonstrates shortest-path routing in the ETX metric, where thicker lines indicate links selected for forwarding data towards the gateway (GW). The middle figure demonstrates constrained shortest path routing. Here, device C is the only battery operated field device allowed to forward data, which shifts the permissible shortest paths (again, indicated with thick lines). Finally, the rightmost figure demonstrates graph-based routing. Here, nodes are organized into “levels” according to their ETX distance to the sink, and additional links are introduced to neighbors on higher levels. The dashed line indicates a sibling link between devices on the same level, which may by used if care is taken to avoid routing loops.

Constraint-based routing

In some cases, it is useful to exclude certain links or nodes from the routing process. For example, we might want to make sure that battery operated nodes do not need to forward other nodes' packets. Such considerations are included in constrained shortest-path routing protocols. Here, nodes or links that do not satisfy certain constraints are removed from the routing topology, and the routes are found by computing the shortest paths in this reduced graph. For example, the most reliable path that does not use battery-operated devices can be computed by first removing the battery-driven devices from the graph representing the network topology, and then solving a shortest path problem with link weights equal to $-\log(1-p_l)$. To balance reliability and latency, we could consider removing all unreliable links from the connectivity graph and then find the routing paths with minimum hop distance to the destination (for a comparison of this approach and the use of ETX, see [68]). A more in-depth discussion about uses of constrained shortest paths in wireless sensor networks can be found in ¡ietf-roll-routing-metrics¿

Graph-based routing

The routes computed by a constrained shortest path routing protocol form trees rooted in the destination. A drawback with this structure is if one node breaks, the full branch from this node to its leaves breaks. One way to come around this problem is to set up multiple (partially) disjoint trees rooted in the destination, and switch from one tree to another when a failure is detected. An alternative is to set-up a *directed acyclic graph* (DAG), in which nodes might have multiple parents, and forward messages along this graph. Since the graph is directed and does not contain any cycles, each successful message transmission guarantees forward progress towards the destination. Note that the forwarding policy now becomes a critical part of the routing algorithm: not only does the preferred parent matter, but also the decision about whether to retransmit on the same link or switch to a secondary route, etc. [75]. Setting up the graph is also non-trivial, but in many cases a reasonable graph can be built from the shortest path tree (in some metric) by connecting nodes in the DAG with neighbors that are closer to the shortest path tree root (in e.g. number of hops) than the node itself.

Example 13 (Markov-models for unicast forwarding) *When link losses are independent, it is easy to construct Markov models for the latency of a single packet traversing a routing graph from source node s to destination node d under a given schedule. The basic idea is to construct a Markov chain whose n^{th} state corresponds to the packet being buffered at node n. The state transition depends on the schedule (and hence on time): if the packet is located at node n and one of its outgoing links l is scheduled for transmission at time t, then the packet is forwarded to the end-node of link $l$with probability $(1-p_l)$ and stay in node n with probability p_l.*

To define the Markov chain in a compact way, consider network with N nodes and L links whose underlying topology is represented by a node-arc incidence matrix $A=[a_{nl}] \in \mathbb{R}^{N\times L}$ where, $a_{nl}=-1$ if link l is outgoing from node n, $a_{nl}=1$ if link l is incoming to node n and $a_{nl}=0$ otherwise. For convenient notation, we also define $A_m = max(-A,0)$. Let $p \in \mathbb{R}^L$ be the vector of loss probabilities for links and let $P=diag(p)$. Introduce the diagonal matrix $S(t)=[s_{lm}(t)] \in 0,1^{L\times L}$ to encode the scheduling decision at time t: $s_{ll}(t)=1$ if link l is scheduled for transmission at time t while all other entries are zero. Let $\pi(t)$ be the probability distribution for the Markov chain at time t. Then,

$$\pi(t+1) = \left(A(I_{L\times L}-P)S(t)A_m^T + I_{N\times N}\right)\pi(t)$$

Since the packet is known to be at state s at time zero, we have that $\pi_n(0)=0$ for $n \neq s$ and $\pi_s(0)=1$. Similar models can also be developed for correlated losses and more complex communication patterns [9, 43, 44].

On-demand routing

The most well-known on-demand routing protocol is the Ad hoc on-demand distance vector routing protocol, AODV [42]. When a source node needs to send a packet to a destination to which it does not already have a route, it initiates a route discovery process: it broadcasts a route request packet to its neighbors, who forward the request to their neighbors, etc. The process continues until either the destination or an intermediate node with ”fresh enough” route to the destination receives the request. The destination or the intermediate node then transmits a route reply message to the neighbor from which it received the first route request. The route reply message is then transmitted back along the reverse path, and nodes on this path set up forward route entries in their routing tables pointing to the node from which they received the route reply. Associated with each route is a timer which will cause a deletion of the entry if it is not used within the specified lifetime. By letting intermediate nodes wait in a controlled fashion before forwarding route requests and/or letting the destination wait for multiple route requests (that have arrived over different paths) and picking the best path, AODV can come close to generating (constrained or unconstrained) shortest paths [13].

4.3 Transport layer protocols and traffic patterns

The physical layer, data-link and network layers described so far provide all the essential functionalities required for moving information between remote hosts in the network. The role of the transport layer is now to “package” this functionality into an transparent end-to-end communication service for applications. While the wireless sensor networking community has developed a wide variety of transport protocols (see, *e.g.*, [65]), we will focus on the solutions that appear in the industrial wireless standards.

The most basic transport-layer protocol in the Internet is the *user datagram protocol (UDP)*. The UPD protocol simply adds information about the source and destination application process (“ports”) along with a checksum bit for error detection, and passes the resulting packet to the network layer for forwarding to the end host. The protocol neither provides support for informing the source about lost or corrupted packets detected at the end-host, nor functionality for handling retransmissions. If such functionality is needed, it must be implemented by the application.

The *transmission control protocol (TCP)* adds reliability and flow control functionalities to the end-to-end communication primitive. In contrast to UDP, TCP is a connection-oriented protocol: before any communication can take place, and end-to-end connection has to be established. In addition to a checksum, TCP packets also include sequence numbers to allow the end host to detect lost messages and to reorganize packets that are delivered out of order. The end-host informs the source about which packets it has received by returning acknowledgement messages. If the source does not receive an acknowledgement within a given time-out period, it assumes that the associated packet is lost and retransmits the data. Another important feature of TCP is flow control: the protocol automatically adjusts the rate at which packets are sent, so as to avoid congestion in the network. However, TCP’s flow control mechanism was originally designed for fixed networks and has performance problems in case of wireless transmissions [18]. TCP treats lost packets as an indication of congestion and decreases the flow rate when no acknowledgements are returned. While optical networks are essentially free from transmission errors, most of the packet drops in wireless networks are due to fluctuating channel quality and not from congestion-induced buffer overflows. The flow control of TCP thus unnecessarily decreases the flow rate every time a packet gets lost in a wireless link, leading to poor utilization of the radio resources. Another known performance problem is related to the interaction between TCP protocol and BEB in multi-hop networks. Single-hop flows will capture a majority of the transmission opportunities leaving the multi-hop flows to starve. The problems of the TCP in multi-hop environments are discussed in, e.g., [71]. The rich feature set of TCP makes it rather heavy-weight to implement and most real-time communication protocols rely on UDP for end-to-end communication.

Finally, note that contrary to fixed networks where point-to-point (unicast) communication is the most common traffic pattern, wireless sensor and actuator networks often rely on many-to-one (*e.g.* sensors to gateway) and one-to-many (*e.g.* controller to actuators) communication. A large body of work in the wireless sensor networks literature considers theory and tailored system solutions for convergecast (many-to-one) and multicast (one-to-many) solutions.

4.4 Standards and specifications for industrial wireless networking

Aim: To list some existing standards, and to describe which techniques/functionalities from above that they explore.

Zigbee PRO

Zigbee exists since 2004 and was one of the first attempts to provide a low-power radio standard for home automation and industrial control. Its focus was on providing a self-organizing scalable and secure short-range wireless solution with a battery life up to two years. However, Zigbee has some serious reliability flaws, which has stopped it from making its way into a true industrial standard. Zigbee remains, to date, in specification form.

Zigbee builds on the physical and MAC-layer of IEEE 802.15.4, and specifies the behavior of the higher protocol layers. There are two types of nodes in a Zigbee network: full-function devices and reduced-function devices. Full-function devices can route messages and act as network coordinator, while reduced-function devices can only communicate directly with a full-function device. The standard supports mesh networking among full-function devices and routing is performed using the AODV protocol. Hence, a node can send packets to any other node in the network, but need to execute the algorithms for route discovery and route maintenance as discussed earlier in this chapter. In a Zigbee network, all nodes share the same channel, and there is no frequency hopping. The only way to attempt to provide reliable operation is to scan the spectrum for a channel with low interference before deploying the network. Since AODV is a single-path routing protocol (unless we allow us the time to re-discover a route when we detect a breakage) neither frequency nor path diversity is supported in Zigbee and the overall reliability is typically low. Zigbee can operate in both beaconed and non-beaconed mode. Beaconing allows some degree of synchronization among nodes, but the medium access is generally contention-based (CSMA/CA). There is an option to use guaranteed time slots, but the support for this is not mandatory and might break interoperability of nodes. As we have seen in Example 11, the lack of synchronized communication can seriously increase latency under correlated traffic. Moreover, the uncoordinated operation of nodes also means that devices need to stay awake practically all the time and that the energy consumption is higher than needed. The latency and reliability issues, together with the security concerns detailed in [29], are some of the key reasons for the limited industrial success of Zigbee.

WirelessHART

WirelessHART is an extension of wired HART, a transaction-oriented communication protocol for monitoring and control applications, including equipment and process monitoring, advanced diagnostics and closed-loop control. As illustrated in Figure 17, the basic elements of a WirelessHART network include: *field devices* connected to the process equipment and able to source, sink and forward packets on behalf of other devices in the network; *gateways*, possibly equipped with multiple access points, enable communication between host applications and field devices; a *network manager* responsible for configuring the network, health monitoring, managing routing tables and scheduling communication between devices. WirelessHART networks may also include *adapters*

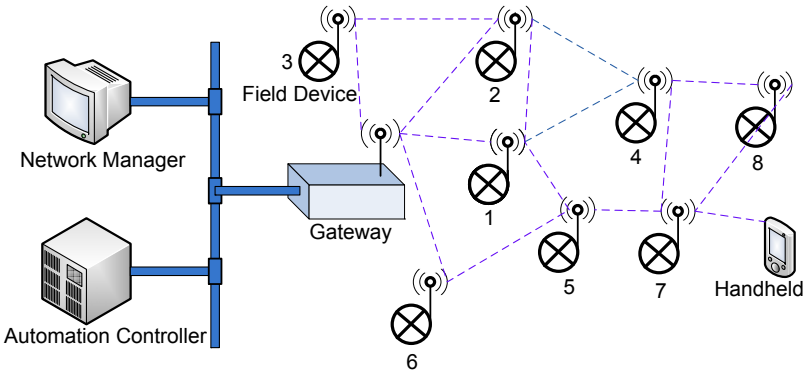


Figure 17: Example of wirelessHART network infrastructure.

for connecting to existing HART-compatible devices and handhelds to configure, maintain or control plant assets. An important restriction of WirelessHART is that device-to-device communication is not supported: all data must pass through the gateway.

Like Zigbee, WirelessHART is based on the 802.15.4-2006 physical layer, but it is restricted to operate on 15 channels (number 11-26) of the 2.4GHz ISM band. The medium access is controlled using a multi-channel TDMA MAC. A time slot is 10 ms, which allows for a complete data packet and associated acknowledgement transaction. Transmissions are scheduled on one of the 15 logical channels, which are mapped onto the physical channel by a channel hopping sequence

$$ch := \ldots$$

The transmission schedule is organized into multiple *superframes*. A superframe is simply a collection of links, each assigned to a specific time slot and a specific channel, see Figure 18 for an illustration. Superframes repeat in time, and should have periods that form a harmonic chain (e.g.

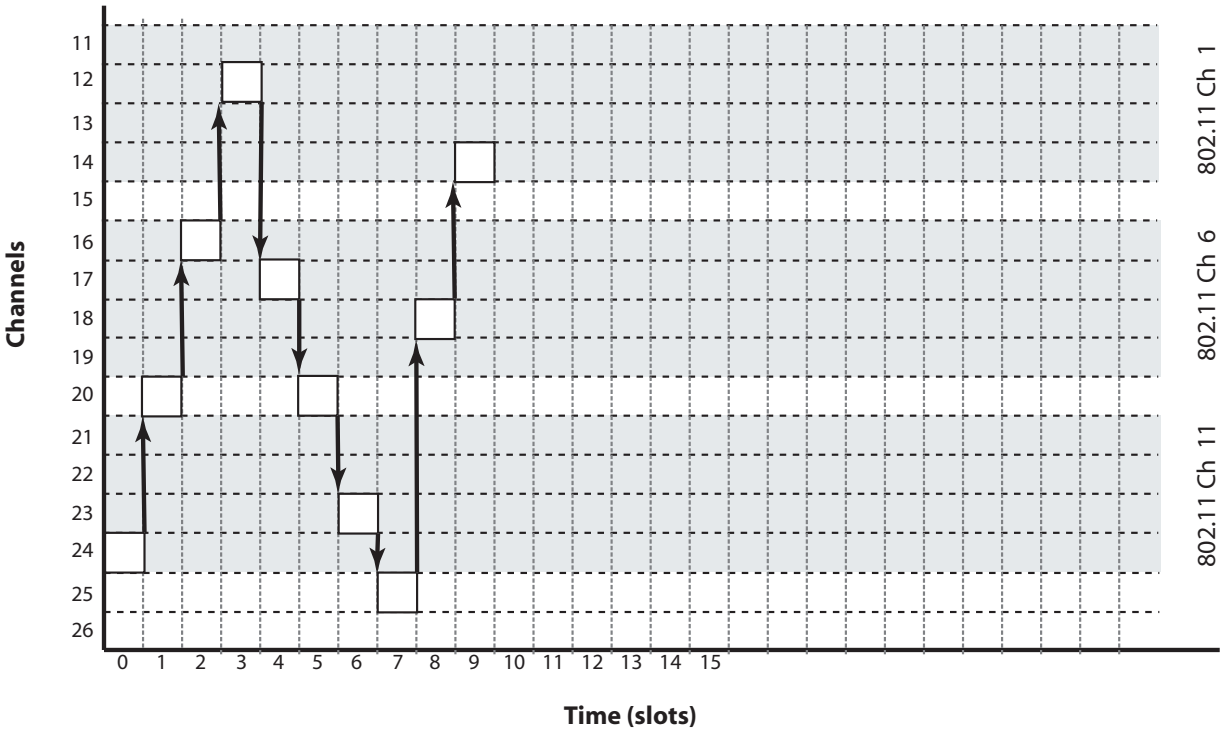


Figure 18: WirelessHART schedules transmissions in time and frequency.

$1, 2, 4, 8, 16, \ldots$ or any other period that conforms to an expression on the form ab^n), so that all superframes are aligned at the start of the longest superframe. Each superframe is typically constructed

to accommodate for communication among a specific group of devices and allows to run the whole network at different duty cycles. Additional superframes might be defined to support traffic at different scan rates, to handle alarm traffic, etc. At least one superframe is always enabled while additional superframes can be added and removed. The standard also specifies procedures for link arbitration to resolve conflicts that may appear if a network device participates in multiple superframes.

The schedule supports three different types of slots: dedicated slots, with a single transmitter-receiver pair; shared slots, where many transmitters but only a single receiver are scheduled; and broadcast slots, where one transmitter and multiple receivers are scheduled. Dedicated and shared slot transmissions are acknowledged, *i.e.*, the receiver transmits a short acknowledgement packet to the transmitter in the same time slot, while broadcast packets are unacknowledged. WirelessHART uses non-persistent CSMA to resolve conflicts in shared slots. If no acknowledgement is received, the backoff exponent of the link is incremented and the backoff counter is initialized with a random number in the associated backoff window. The backoff counter is decremented in every time the same link is scheduled in a shared slot, and the data is retransmitted when the backoff counter reaches zero. Note that this is a completely different time-scale than the back-off counter in the 802.15.4 CSMA and that back-off times in shared WirelessHART slots can be significant.

WirelessHART supports two routing paradigms: graph and source routing. In graph routing, network nodes store multiple DAG:s. Packets are tagged with the ID of the graph along which they should be forwarded. In source routing, on the other hand, the complete path is encoded into the packet. Source routing specifies a single path between source and destination, and is thus fragile to link failures, while graph routing supports multiple forwarding paths to improve reliability.

The transport layer provides both unacknowledged and acknowledged end-to-end communication. The unacknowledged service is, much like UDP, a basic protocol without acknowledgement packets and without guarantees of packet ordering at the destination node. The acknowledged service, on the other hand, provides end-to-end acknowledgements and guarantees packet ordering (but not advanced TCP features such as flow control).

The standard supports methods to maintain network-wide time synchronization, network formation and topology maintenance, security features on both link level and network level, and much more. Since these features do not have a direct influence on the end-to-end performance of the network, we do not describe these in detail here, but refer to the standard documents.

ISA 100

While WirelessHART formally only standardizes wireless communication with HART devices, the ISA standard was created with the more ambitious aim of providing a single wireless standard that supports multiple legacy application layer protocols. However, as the standardization work has proceeded, ISA100 and WirelessHART have evolved into rather similar specifications and there is now an ongoing effort in the standard committees to see if and how the two standards could “converge” into one. In this section, we will review ISA100 and highlight some of the features that differ from WirelessHART.

Like WirelessHART, ISA100 is based on the IEEE 802.15.4-2006 physical layer and the data link layer uses a multi-channel TDMA medium access control. However, in ISA100 different superframes can use different time slot lengths, ranging from the 10 ms of WirelessHART up to 12 ms. Longer time slot lengths give some extra margin against poor device synchronization, but perhaps more importantly this feature enables prioritized access in shared time slots and advanced transmission primitives beyond unicast and broadcast. Prioritized access in shared slots is implemented by transmitting high-priority data at the very beginning of time slots, while waiting a short time before performing CCA and attempting to transmit lower-priority data. The *duocast* primitive allows a field device to broadcast a data packet to multiple access points (backbone routers in the ISA100 terminology) and receive serial acknowledgements within the same time slot. The duocast transaction is considered successful if at

least one link-level acknowledgement is received, and allows low-latency communication with high probability of first success. There is no restriction of individual superframe lengths, but the system realigns time slots of all superframes every 250 ms by the insertion of short idle periods. Frequency hopping can be performed on a per time-slot basis ("slotted hopping", like in WirelessHART) or with longer intervals ("slow hopping") where the same channel offset is used for multiple consecutive transmissions. The slower hopping can be useful, for example at network discovery, since it allows for less accurate synchronization among nodes. Similarly to WirelessHART, ISA100 supports both source and graph routing. The network layer of ISA100 is based on the IETF 6LoWPAN standard, and the transport layer provides end-to-end encrypted communication through UDP datagrams as outlined in IETF RFCs RFC 768 and 2460.

Alternative and emerging standards

While WirelessHART and ISA100 appear to be able to provide the industrial-grade reliability and security that Zigbee was lacking, they still have a serious drawback: they both rely on centralized network configuration and resource management. The centralized management comes with large overhead for collecting network health information and disseminating routing graphs and transmission schedules to nodes, and the times required for a node to join a WirelessHART network can be substantial. It is thus natural to ask whether it is possible to create distributed routing and MAC protocols that could exploit diversity and offer the reliability required by control applications. No such standard exists at the writing of this book chapter, but (at least) two noticeable efforts are currently taking place towards that aim. The first one is the work on the Internet Engineering Task Force (IETF) on Routing over Low-power and Lossy Networks (ROLL). This effort attempts to develop and standardize a routing protocol that includes distributed DAG construction and maintenance. Complementing this work, the IEEE 802.15.4e attempts to amend the 802.15.4-2006 standard to better cater for industrial needs, and the current draft standard includes a time-slotted channel hopping medium access control.

5 Control relevant models of latency and loss

Aim: show how the existing models used in the control literature apply to various scenarios above.

The purpose of this section is not to provide an extensive overview of design methodologies for networked control (such expositions can be found elsewhere in this book and in several survey papers, *e.g.* [17, 63]). Rather, our aim is to provide the link between the communication network models derived earlier and abstractions that have been used in the theoretical control community.

The network as a delay

The initial work on networked control, *e.g.* [30, 39], which focused on wired control loops, advocated to model the network as a time-varying delay. The general set-up, shown in Figure 19, includes delays from sensors to controller $\tau_{\rm sc}(t)$ and from controller to actuator $\tau_{ca}(t)$.

If we assume that

$$\tau_{sc}(t) + \tau_{ca}(t) \leq h$$

Then, a discrete-time controller with sampling period h can be developed that does not need to consider the possibility of packet loss. The time-varying delays can be dealt with using varying degrees of sophistication.

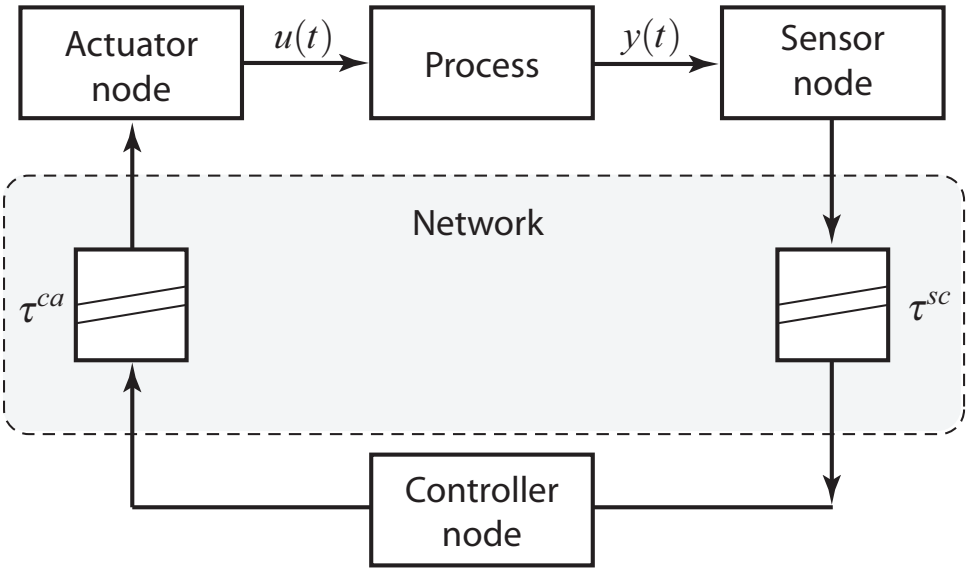


Figure 19: The network abstracted as a time-varying delay.

The simplest is to let the sensor and controller work synchronously with sampling period h, and insert a buffer in the actuator node that waits until the end of the sampling interval before applying the control [30]. In this way, all stochasticity is removed and replaced by a fixed one-sample information delay in the control loop. Hence, a wealth of classical control design techniques apply.

Since the control performance generally degrades with increasing information delay, it is natural to let the controller be event-driven, in the sense that it computes the control signal as soon as senor data arrives and immediately transmits the command to the actuator [39]. Since the sensor-controller delay is known at the controller, a time-varying Kalman filter allows to optimally estimate the process state. In this case, we do not need to know the latency distribution

$$\mathbb{P}r(\tau_{\rm sc} = \kappa)$$

to design the algorithm, only to evaluate its performance. Suboptimal schemes that use a fixed gain Kalman filter, however, will need access to latency distributions in their design [10]. The latency distributions are also needed for computing the appropriate controller gains, since the delays between controller and actuator are not known to the controller (see, *e.g.* [24] for alternative solutions).

When it is not possible to guarantee a communication delay less than the sampling interval, it is natural to consider designs that allow

$$\tau_{\rm sc}(t) + \tau_{ca}(t) \leq Kh$$

for some fixed $K > 1$. A new problem that appears is that packets might now arrive out of order. Similarly as above, one can use buffers to remove the stochasticity in the control loop. Although the advantage of acting as quickly as possible when new data arrives increases when K increases, dealing with out-of-order delivery increases implementation complexity. The optimal filter, for example, now needs buffers to store the last K sensor packets received and the covariance matrix of the time-varying Kalman filter at time $t - Kh$, and it needs to run K iterations of the Riccati equations when new data arrives [54].

When the latencies are upper-bounded, then one can also consider the networked control problem from a robust control perspective. One useful result here is the jitter-margin [25] which relates the maximum tolerable network delay to the "bandwidth" of the complementary sensitivity function (see also [35] for less conservative results).

As we have seen, however, natural models for wireless control systems are stochastic and do not

allow any deterministic guarantees on latencies. The possibility of packet loss should thus be considered.

The network as an erasure link

One other class of models disregard communication delay and only consider whether or not packets arrive at their destination; see Figure 20.

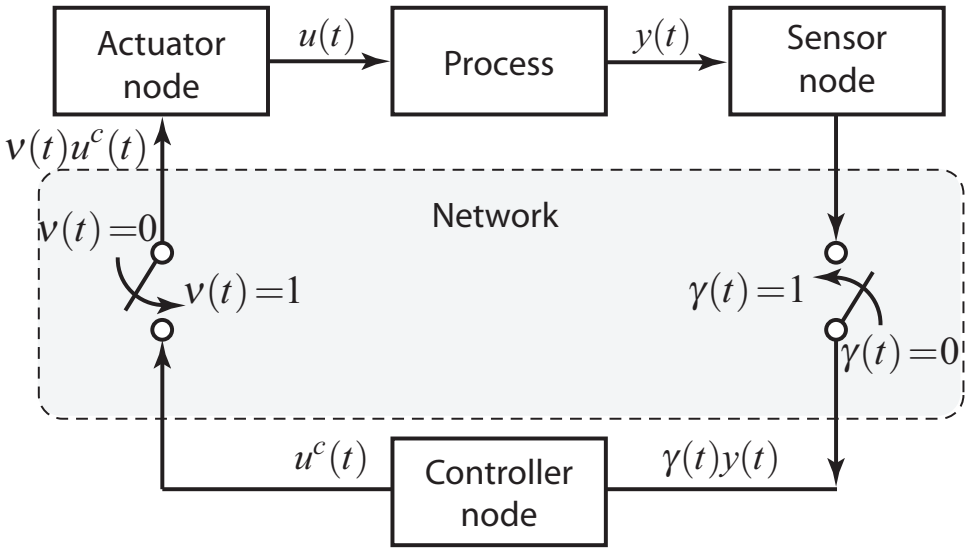


Figure 20: The network abstracted as an erasure link.

The most common model in the theoretical control literature is the Bernoulli loss process, which assumes that losses occur independently with a fixed probability p_{loss}. The loss probability can be computed from the latency distributions derived earlier, as

$$p_{\text{loss}} = \int_{\kappa=h}^{\infty} \mathbb{P}r(\tau = \kappa)\, d\kappa \tag{6}$$

For this type of model, bounds on the tolerable loss probability can be computed and some optimal controllers and estimators can be derived [55] and co-design of network and controllers are possible [44]. The independent loss model is reasonable if the sampling times are longer than the coherence time of the wireless channels, but not appropriate for modelling single links at the short time-scale.

In the shorter time-scale of a few consecutive transmissions, the loss process on a given link tends to be correlated [67] and more complex loss models, such as higher-order Markov models should be used. In this case, the natural modeling framework is the one of jump-linear systems [10, 69]. Some specific results for estimation over Gilbert-Elliot models also exist [19, 3].

Models which assume a bounded delay of hK also applies if one can guarantee that the network produces no more than K consecutive losses. However, once again, no such deterministic guarantees can be derived from the natural stochastic models of loss described earlier in this document.

Combined latency-loss models

In many cases, the most natural models for control is to pay careful attention to the shorter delays, in order to act quickly when the network incurs a short delay, but then put an upper bound on the tolerable latency after which the packet is considered as lost. Such a mechanism is easily implemented by retransmission and buffering policies in the single link and using time-to-live counters for

in the multi-hop scenario. Also in this case, the appropriate modelling framework is one of jump-linear systems [69]. Once again, the latency distributions using various technologies can be derived using the techniques described earlier in this document and the loss probability calculation (6).

It is natural to relate the time-to-live not only to the network but also to the sampling interval and characteristics of the process at hand. Networking-controller co-design procedures following this philosophy are presented in, *e.g.* [47, 44].

6 Application to the Barcelona water distribution network

So far, we have presented a wide range of theoretical models for capturing the effects of the wireless channel, accounting for propagation, interference, and co-existence with other devices (both within the low-power 802.15.4 family as well as other ISM-band technologies). We believe that these models are important, since they help to understand and quantify the most important issues for designing and dimensioning low-power wireless communications solutions for industrial control.

Now, for a given scenario, there is typically some effect that is dominating and needs detailed modeling attention while the others can be captured using simple abstract models. For example, in the sensor placement and network dimensioning phase, propagation effects are critical. For small footprint protocols that do not implement global synchronization and resource allocation, the co-existence and medium access effects might be the most critical factor (especially in dense deployments) while such issues are completely removed in a well-dimensioned *Wireless*HART or ISA100 deployment.

6.1 Loss models for the small-scale demo

In this section, we will discuss the appropriate models for the small-scale water distribution demo, described in detail in Devlierable 5.3. This demo considers the replacement of a wired valve controller in a pumping station with a wireless solution. In its basic configuration, a down-stream flow sensor reports measurements to the controller node that computes the appropriate control action and sends the computed command wirelessly to the valve. The sensor and actuator positions are fixed due to physical constraints of the pumping station, and well within range for wireless communication with a controller node. We hence decided not to deploy any relay and with the fixed positions of tranceivers it makes more sense to measure actual packet loss rates rather than trying to carefully model the propagation. The pumping station is relatively free from external interference and there is no other co-existing wireless technology in close range. The set-up with a single time-triggered sensor, and event-triggered communication from the controller to the actuator before the next sampling instant (the controller-actuator communication occurs as soon as the controller has computed the new actuator command or at a relatively tight time-out in case no sensor data was received) implies that there is no contention latencies or losses that need careful modelling. Altogether, this led us to base our modeling work on real loss traces on the relevant time-scales.

The current wired PID controller is running at a sampling rate of 400 ms, which should accomodate sensor-to-controller and controller-to-actuator communication plus the additional time needed for overhead and calculations in sensor, actuator and controller node. Hence, we need to consider communication rates of at least 10 packets per second. This is also compatible with the recommendation in the *Wireless*HART standard that half the superframe (200 ms in this case) is left for overhead communication. On the other hand, the time slot lengths in *Wireless*HART are 10ms per default. Hence, we focus our measurements of interpacket arrival times in the range 10-100 ms and collect traces from the sensor and actuator positions to the persumptive controller node position.

We run multiple experiments during our site visit, focusing on interpacket times of 10, 15 and 20 ms, and running experiments on multiple channels (12, 19 and 26 in this case). Figure 21 shows some representative loss traces from the pumping station, taken with 10 and 15 ms interpacket times,

respectively. The traces were collected using the TelosB sensor network platform using a simple loss trace program running on the Contiki operating system. While interpacket times of 15 ms or more were reliable with above 95% packet reception rates, the reliability droped significantly when we tried to communicate at 10ms with recption rates below 65%. We can also see an increased burstiness at the 10 ms time scale.

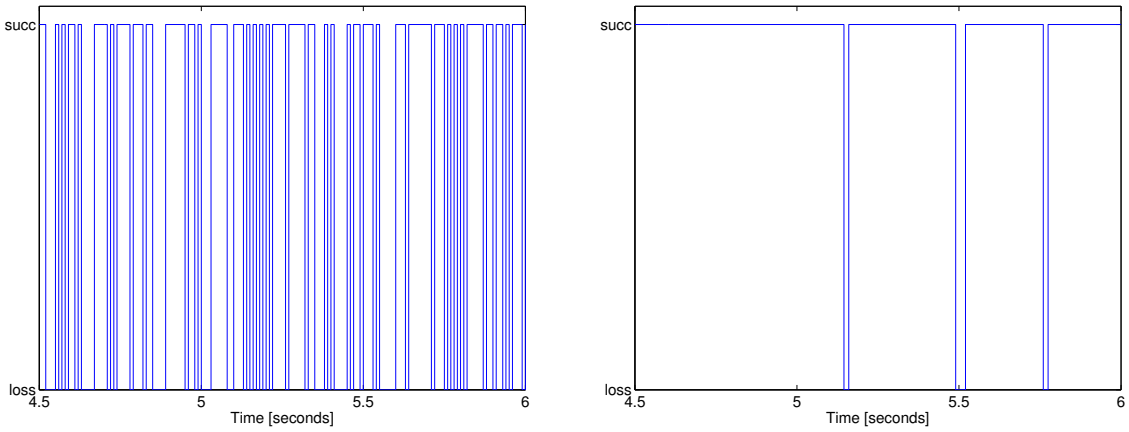


Figure 21: Part of the loss traces collected for 10 ms (left) and 15 ms (right) interpacket times.

We fit a simplified two-state Gilbert-Elliot model to the loss traces. The model is simplified compared to the Gilbert-Elliot model presented earlier in this report in the sense that the observation process is deterministic: packets are always succesfully received when the model is in the good state, and always lost when the model is in the bad state. We use the parameterization shown in Figure 22.

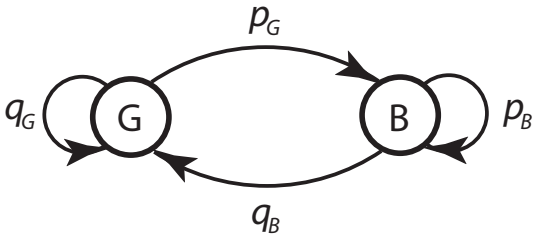


Figure 22: State transition diagram for the simplified Gilbert-Elliot model.

In this case, the model parameters can be estimated from counts of the transitions (from good to bad state and bad to good state, respectively) that occur in the loss traces. Let t_{GG}, t_{GB}, t_{BB} and t_{BG} be the number of transisions from good to good, good to bad, bad to bad, and bad to good state, respectively. Then the model parameters are

$$p_G = \frac{t_{GB}}{t_{GB} + t_{GG}}, \quad q_G = 1 - p_G$$
$$q_B = \frac{t_{BG}}{t_{BG} + t_{BB}}, \quad p_B = 1 - q_B$$

Now, analyzing the trace collected at 10 ms interpacket times, we find that $p_G = 0.3341$, $q_B = 0.5907$ and the packet reception ratio (the number of received packets over the total number of packets) is 63.87%. The average burst length is 1.69 packets, and the largest burst-length recorded was 12 consecutive losses.

Considering now the trace collected at 15 ms interpacket times, we found $p_G = 0.00195$, $q_B = 0.8205$, and the packet reception ratio 97.67%. The average burst length was 1.21 packet and the largest burst length in the trace was 5 consecutive losses.

Comparing the results of this trace with other traces collected at 15 ms interpacket arrival times, we found very close matches also when the traces were collected on other communication channels (when repeating the experiment with 15 ms interpacket arrivals on another channels, we found $p_G = 0.0199$, $q_B = 0.8179$, and packet reception ratio of 97.62%, for example). We hence feel that the models and parameters are sufficiently accurate for control design.

6.2 Loss models for the wide-area system

For the wide-area system, we will have to rely on GRPS communication to provide city-wide coverage. In the near future, GPRS could be replaced by 3G, WiMAX or 4G data communications, but these technologies were not available to us for experimentation and evaluation at the writing of this report.

Since the MPC system operates with a sampling interval of one hour (e.g. every hour, data is collected from the measurement points across the city, a new control action is computed, and the new set-points are disseminated to the actuators), we have judged that simple GPRS connections are feasible. It is also likely that operators will start to prioritize machine-to-machine communication better in their network and that current latencies will be brought down in the near future (compare, *e.g.*, with the 20bn devices 2020 vision of Ericsson).

Our experience with GPRS connections indicate that delays up to 10 seconds are common, while the reliability is relatively high. We hence propose to use a simple model of a fixed 10 second delay, combined with a small loss probability (to model that connections could be dropped and no readmission to the cellular infrastructure would be possible before the time-out where the MPC controller decides to proceed with the current data). Although this model is crude, it appears futile to try to model the internal workings and policies of the public cellular infrastructure in Bareclona.

7 Conclusions

This deliverable has given an overview of technologies and models for wireless networking for control and monitoring applications. Starting with physical signal transmission of radio signals, we have described theoretical models for propagation, fading, and outage. These models describe how packets on a single isolated link can get corrupted and impossible to decode at the receiver. We shortly discussed various diversity techniques for minimizing the probability of outage, and coding techniques for correcting occasional bit errors. We have also discussed how the co-existence with other technologies on the license-free ISM band creates interference for low-power radios. Next, we described various methods for sharing the spectrum among several transmitters. These medium access control methods could broadly be characterized as scheduled or contention-based: scheduled medium access gives predictable channel access for predictable traffic, but tends to be inefficient when the traffic itself is random and uncorrelated. Contention-based medium access, on the other hand, is more efficient in dealing with random traffic but creates random and sometimes unnecessarily long latencies when traffic is correlated and predictable. From the single link, we moved to multi-hop networks to illustrate several critical issues such as increased end-to-end latency and decreased reliability. We demonstrated how multi-path diversity improved reliability, especially in the case of correlated losses. We also described various routing paradigms, including single path and multi-path (graph-based) routing and relevant link metrics. With this basic understanding, we described several standards and specifications for low-power industrial wireless communication. A short section describing how control-relevant models of latency and loss could be developed based on the individual models that we have described has also been included. Finally, we applied the models to the small-scale demonstrator in the Barcelona water distribution network and developed models for control design based on real loss traces.

Acknowledgement Luca Stabellini kindly let us use Figure 8 from his paper [61]. Zhenhua Zou generated the Figures 13 and 14 based on our recent work. Joonas Pesonen, Pablo Soldati, Haibo Zhang and Olaf Landsiedel gave several constructive comments that helped to improve the quality of the text.

References

[1] N. Abramson. The ALOHA system – another alternative for computer communications. In *Proc. of the Fall Joint Computer Conference*, pages 281–285, 1970.

[2] K. Akkaya and M. Younis. A survey on routing protocols for wireless sensor networks. *Ad Hoc Networks*, 3:325–349, 2005.

[3] Peter Almstrom, Maben Rabi, and Mikael Johansson. Networked state estimation over a gilbert-elliot type channel. In *Proc. IEEE Conference on Decision and Control*, Shanghai, China, December 2009.

[4] P. Antsaklis and Eds. J. Baillieul. Special issue on technology of networked control systems. *Proceedings of the IEEE*, 95(1), January 2007.

[5] D. Bertsekas and R. Gallager. *Data Networks*. Prentice Hall, 1987.

[6] G. Bhatti, A. Mehta, Z. Sahinoglu, J. Zhang, and R. Viswanathan. Modified beacon-enabled ieee 802.15.4 mac for lower latency. In *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*, pages 1 –5, 30 2008-dec. 4 2008.

[7] P. Buchholz and J. Plonnigs. Analytical analysis of access-schemes of CSMA type. In *Proc. IEEE International Workshop on Factory Communication Systems*, Vienna, Austria, 2004.

[8] Michael Buettner, Gary V. Yee, Eric Anderson, and Richard Han. X-mac: a short preamble mac protocol for duty-cycled wireless sensor networks. In *SenSys '06: Proceedings of the 4th international conference on Embedded networked sensor systems*, pages 307–320, 2006.

[9] Phoebus Chen and Shankar Sastry. Latency and connectivity analysis tools for wireless mesh networks. In *ROBOCOMM*, page 33, 2007.

[10] O.L.V. Costa, M.D. Fragoso, and R.P. Marques. *Discrete-Time Markov Jump Linear Systems*. Springer, 2005.

[11] D. S. J. De Couto. *High-Throughput Routing for Multi-Hop Wireless Networks*. PhD thesis, MIT, 2004.

[12] E. O. Elliot. Estimates of error rates for codes on burst-noise channels. *Bell SystemTechnology Journal*, 39:1253 – 1264, 1963.

[13] C. Gao. *Thesis*. PhD thesis, University of Vaasa, November 2009.

[14] S. Geirhofer, L. Tong, and B. M. Sadler. Dynamic spectrum access in WLAN channels: Empirical model and its stochastic analysis. In *Proceedings of the First International Workshop on Technology and Policy for Accessing Spectrum*, Boston, MA, 2006.

[15] E. N. Gilbert. Capacity of a burst-noise channel. *Bell Systems Technology Journal*, 39:1253 – 1264, 1960.

[16] P. Gorday. 802.15.4 multipath. *IEEE 803.15-4-0337-00-004b*, 2004.

[17] J.P. Hespanha, P. Naghshtabrizi, and Yonggang Xu. A survey of recent results in networked control systems. *Proceedings of the IEEE*, 95(1):138–162, Jan. 2007.

[18] G. Holland and N. Vaidya. Analysis of TCP performance over mobile ad hoc networks. *Wireless Networks*, 8(2):275–288, 2002.

[19] M. Huang and S. Dey. Stability of Kalman filtering with Markovian packet losses. *Automatica*, 43:598–607, 2007.

[20] 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). *IEEE Std 802.15.4-2003*, 2003.

[21] Ieee recommended practice for information technology - telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements part 15.2: Coexistence of wireless personal area networks with other wireless devices operating in unlicensed frequency bands. *IEEE Std 802.15.2-2003*, 2003.

[22] ITU-R. Propagation data and prediction methods for the planning of indoor radiocommunication systems and radio local area networks in the frequency range 900 mhz to 100 ghz. *Recommendation ITU-R P.1238-6*, 2009.

[23] P. Krishnamurty J. Arauz and M. A. Labrador. Discrete rayleigh fading channel modelling. *Wireless Communications and Mobile Computing*, 4:413 – 423, 2004.

[24] M. Johansson and L. Xiao. On optimal control over packet-oriented communication links. In *38th Annual Allerton Conference on Communication, Control, and Computing*, Monicello, Il, October 2000.

[25] C.-Y. Kao and B. Lincoln. Simple stability criteria for systems with time-varying delays. *Automatica*, 40(8):1429–1434, 2004.

[26] A. Kemp and E. Bryant. Channel sounding of industrial sites in the 2.4 ghz ism band. *Wireless Personal Communications*, 31:235 – 248, 2004.

[27] L. Kleinrock and F. Tobagi. Packet switching in radio channels: Part i–carrier sense multiple-access modes and their throughput-delay characteristics. *Communications, IEEE Transactions on*, 23(12):1400 – 1416, dec 1975.

[28] K. Koumpis, L. Hanna, M. Andersson, and M. Johansson. Wireless industrial control and monitoring beyond cable replacement. In *Proc. 2nd PROFIBUS International Conference*, Coombe Abbey, UK, June 2005.

[29] T. Lennvall, S. Svensson, and F. Hekland. A comparison of wirelesshart and zigbee for industrial applications. In *Factory Communication Systems, 2008. WFCS 2008. IEEE International Workshop on*, pages 85 –88, may 2008.

[30] R. Luck and A. Ray. An observer-based compensator for distributed delays. *Automatica*, 25(6):903–908, 1990.

[31] A. Mahmood M. M. A. Hossain and R. Jantti. Channel ranking algorithms for cognitive coexistence of ieee 802.15.4. In *Proc. IEEE PIMRC 2009, Tokyo, Japan*, 2009.

[32] H. MacLeod, C. Loadman, and Z. Chen. Experimental studies of the 2.4-ghz ism wireless indoor channel. pages 63 – 68, may 2005.

[33] P.J. Marrón, D. Minder, and the Embedded WiSeNts Consortium. Embedded WiSeNts research roadmap. Logos Verlag, Berlin, 2006.

[34] K. Medepalli and F. A. Tobagi. Towards performance modeling of IEEE 802.11 based wireless networks: a unified framework and its applications. In *Proc. IEEE Infocom*, Bercelona, Spain, April 2006.

[35] L. Mirkin. Some remarks on the use of time-varying delay to model sample-and-hold circuits. *IEEE Trans. Automat. Control*, 52(6):1109–1112, 2007.

[36] Jelena Misic, Shairmina Shafi, and Vojislav B. Misic. Performance of a beacon enabled ieee 802.15.4 cluster with downlink and uplink traffic. *IEEE Trans. Parallel Distrib. Syst.*, 17(4):361–376, 2006.

[37] M. Miskowicz, M. Sapor, M. Zych, and W. Latawiec. Performance analysis of predictive p-persistent csma protocol for control networks. In *Factory Communication Systems, 2002. 4th IEEE International Workshop on*, pages 249 – 256, 2002.

[38] J. Morse. Market pulse: wireless in industrial systems: cauthious enthusiasm. *Industrial Embedded Systems*, 2(7):10–11, 2006.

[39] J. Nilsson. *Real-time control systems with delays*. PhD thesis, Lund Institute of Technology, Lund, Sweden, 1998.

[40] Pangun Park, C. Fischione, and K.H. Johansson. Performance analysis of gts allocation in beacon enabled ieee 802.15.4. In *Sensor, Mesh and Ad Hoc Communications and Networks, 2009. SECON '09. 6th Annual IEEE Communications Society Conference on*, pages 1 –9, june 2009.

[41] A. Pérez-Yuste. Early developments of wireless remote control: The Telekino of Torres-Quevedo. *Proceedings of the IEEE*, 96(1):186–190, January 2008.

[42] C.E. Perkins and M.E. Royer. Ad-hoc on demand distance vector routing. In *Proc. 2nd IEEE Workshop on Mobile Computing Systems and Applications*, number 90–100, New Orleans, LA, 1999.

[43] Joonas Pesonen. Stochastic estimation and control over wirelessHART networks: Theory and implementation. Master's thesis, Royal Institute of Technology, February 2010.

[44] Joonas Pesonen, Haibo Zhang, Pablo Soldati, and Mikael Johansson. Methodology and tools for controller-networking codesign in wirelesshart. In *Proc. 14th IEEE International Conference on Emerging Techonologies and Factory Automation*, Palma di Mallorca, Spain, 2009.

[45] M. Petrova, J. Riihijarvi, P. Mahonen, and S. Labella. Performance study of ieee 802.15.4 using measurements and simulations. volume 1, pages 487 –492, april 2006.

[46] J. Polastre, J. Hill, and D. Culler. Versatile low power media access for wireless sensor networks. In *SenSys '04: Proceedings of the 2nd international conference on Embedded networked sensor systems*, pages 95–107, 2004.

[47] M. Rabi, L. Stabellini, A. Proutiere, and M. Johansson. Networked estimation under contention-based medium access. *International Journal of Robust and Nonlinear Control*, 2010.

[48] T.S. Rappaport. Characterization of uhf multipath radio channels in factory buildings. *Antennas and Propagation, IEEE Transactions on*, 37(8):1058 –1069, aug 1989.

[49] Injong Rhee, Ajit Warrier, Mahesh Aia, Jeongki Min, and Mihail L. Sichitiu. Z-mac: a hybrid mac for wireless sensor networks. *IEEE/ACM Trans. Netw.*, 16(3):511–524, 2008.

[50] L. G. Roberts. *Computer communication networks*, chapter Dynamic allocation of satellite capacity through packet reservation. Noordhoff Internat Publishing, Groningen, Netherlands, 1972.

[51] R. Rom and M. Sidi. *Multiple access protocols – performance and analysis*. Springer-Verlag, 1989. Available online: http://webee.technion.ac.il/people/rom/PDF/MAP.pdf.

[52] E. M. Royer and C.-K. Toh. A review of current routing protocols for ad hoc mobile wireless networks. *IEEE Personal Communications*, pages 46–55, April 1999.

[53] R. Jantti S. Nethi and V. Nassi. Time and antenna diversity in wireless sensor and actuator networks. In *Proc. IEEE MICC 2009, Kuala Lumpur, Malaysia*, 2009.

[54] L. Schenato. Optimal estimation in networked control systems subject to random delay and packet drop. *Automatic Control, IEEE Transactions on*, 53(5):1311–1317, June 2008.

[55] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S.S. Sastry. Foundations of control and estimation over lossy networks. *Proceedings of the IEEE*, 95(1):163–187, Jan. 2007.

[56] A. Sikora and V.F. Groza. Coexistence of ieee802.15.4 with other systems in the 2.4 ghz-ism-band. volume 3, pages 1786 –1791, may 2005.

[57] Pablo Soldati, Haibo Zhang, and Mikael Johansson. Deadline-constrained transmission scheduling and data evacuation in wirelesshart networks. In *Proc. European Control Conference*, Budapest, Hungary, September 2009.

[58] Pablo Soldati, Haibo Zhang, Zhenhua Zou, and Mikael Johansson. Optimal routing and scheduling of deadline-constrained traffic over lossy networks. In *IEEE Globecom*, Miami, Florida, December 2010.

[59] L. Stabellini. Design of reliable communication solutions for wireless sensor networks, licentate thesis. Technical report, Royal Institute of Technology (KTH), 2009.

[60] L. Stabellini. Quantifying and modeling spectrum opportunities in a real wireless environment. In *Proc. IEEE WCNC*, Sydney, Australia, April 2010.

[61] L. Stabellini and A. Proutiere. Evaluating delay and energy in sensor networks with sporadic and correlated traffic. In *The 7th Scandinavian Workshop on Wireless Ad-hoc and Sensor Networks*, Stockholm, Sweden, 2009.

[62] E. Tanghe, W. Joseph, L. Verloock, L. Martens, H. Capoen, K. Van Herwegen, and W. Vantomme. The industrial indoor channel: large-scale and temporal fading at 900, 2400, and 5200 mhz. *Wireless Communications, IEEE Transactions on*, 7(7):2740 –2751, july 2008.

[63] Yodyium Tipsuwan and Mo-Yuen Chow. Control methodologies in networked control systems. *Control Engineering Practice*, 11(10):1099 – 1111, 2003. Special Section on Control Methods for Telecommunication.

[64] M.C. Vuran and I.F. Akyildiz. Error control in wireless sensor networks: A cross layer analysis. *Networking, IEEE/ACM Transactions on*, 17(4):1186 –1199, aug. 2009.

[65] Chonggang Wang, K. Sohraby, Bo Li, M. Daneshmand, and Yueming Hu. A survey of transport protocols for wireless sensor networks. *Network, IEEE*, 20(3):34 – 40, may-june 2006.

[66] Hong Shen Wang and N. Moayeri. Finite-state markov channel-a useful model for radio communication channels. *Vehicular Technology, IEEE Transactions on*, 44(1):163 –171, feb 1995.

[67] A. Willig. A new class of packet- and bit-level models for wireless channels. In *Personal, Indoor and Mobile Radio Communications, 2002. The 13th IEEE International Symposium on*, volume 5, pages 2434–2440 vol.5, Sept. 2002.

[68] A. Woo, T. Tong, and D. Culler. Taming the underlying challenges in reliable multihop routing in sensor networks. In *ACM Sensys*, Los Angeles, CA, 2003.

[69] Lin Xiao, A. Hassibi, and J.P. How. Control with random communication delays via a discrete-time jump system approach. In *American Control Conference, 2000. Proceedings of the 2000*, volume 3, pages 2199–2204 vol.3, 2000.

[70] Kaixin Xu, M. Gerla, and Sang Bae. How effective is the ieee 802.11 rts/cts handshake in ad hoc networks. In *Global Telecommunications Conference, 2002. GLOBECOM '02. IEEE*, volume 1, pages 72 – 76 vol.1, nov. 2002.

[71] S. Xu and T. Saadawi. Does the IEEE 802.11 MAC protocol work well in multihop wireless ad hoc networks. *IEEE Communications Magazine*, June 2001.

[72] Yang Yang and T.-S.P. Yum. Delay distributions of slotted ALOHA and CSMA. *Communications, IEEE Transactions on*, 51(11):1846 – 1857, nov. 2003.

[73] K. Jamieson Y.C. Tay and H. Balakrishnan. Collision-minimizing csma and its applications to wireless sensor networks. *IEEE Journal on Selected Areas in Communications*, 22(6):1048–1057, August 2004.

[74] Wei Ye, J. Heidemann, and D. Estrin. Medium access control with coordinated adaptive sleeping for wireless sensor networks. *Networking, IEEE/ACM Transactions on*, 12(3):493 – 506, june 2004.

[75] Zhenhua Zou, Pablo Soldati, Haibo Zhang, and Mikael Johansson. Delay-constrained maximum-reliability routing over lossy links. In *IEEE Conference on Decision and Control*, Atlanta, GA, December 2010.