KTH Electrical Engineering

Protocol Design for Control Applications using Wireless Sensor Networks

PANGUN PARK

Licentiate Thesis
Stockholm, Sweden 2009

Protocol Design for Control Applications
using Wireless Sensor Networks

TRITA-EE 2009:041
ISSN 1653-5146
ISBN 978-91-7415-441-5

Automatic Control Lab
School of Electrical Engineering
KTH (Royal Institute of Technology)
SE-100 44 Stockholm, Sweden

Printed by Universitetsservice US-AB.

Abstract

Given the potential benefits offered by wireless sensor networks (WSNs), they are becoming an appealing technology for process, manufacturing, and industrial control applications. In this thesis, we propose a novel approach to WSN protocol design for control applications. The protocols are designed to minimize the energy consumption of the network, while meeting reliability and packet delay requirements. The parameters of the protocol are selected by solving a constrained optimization problem, where the objective is to minimize the energy consumption and the constraints are the probability of successful packet reception and the communication delay. The proposed design methodology allows one to perform a systematic tradeoff between the control requirements of the application and the network energy consumption. An important step in the design process is the development of analytical expressions of the performance indicators. We apply the proposed approach to optimize the network for various communication protocols.

In Paper A, we present an adaptive IEEE 802.15.4 for energy efficient, reliable, and low latency packet transmission. The backoff mechanisms and retry limits of the standard are adapted to the estimated channel conditions. Numerical results show that the proposed protocol enhancement is efficient and ensures a longer lifetime of the network under different conditions. Furthermore, we investigate the robustness and sensitivity of the protocol to possible errors during the estimation process.

In Paper B, we investigate the design and optimization of duty-cycled WSNs with preamble sampling over IEEE 802.15.4. The analytical expressions of performance indicators are developed and used to optimize the duty-cycle of the nodes to minimize energy consumption while ensuring low latency and reliable packet transmissions. The optimization results in a significant reduction of the energy consumption compared to existing solutions.

The cross-layer protocol called Breath is proposed in Paper C. The protocol is suitable for control applications by using the constrained optimization framework proposed in the thesis. It is based on randomized routing, CSMA/CA MAC, and duty-cycling. The protocol is implemented and experimentally evaluated on a testbed, and it is compared with a standard IEEE 802.15.4 solution. Breath exhibits a good distribution of the work load among the network nodes, and ensures a long network lifetime.

Acknowledgments

First of all I would like to thank my supervisor Karl Henrik Johansson who accepted me as a PhD student at the Automatic Control Group. Specially, I appreciate his guidance and support throughout my research. I owe my gratitude to my co-supervisor Carlo Fischione, who inspired a lot of productive work. Without their valuable comments, this thesis would probably not exists today.

The main part of the thesis consists of three papers. I am indebted to the coauthors of these. The coauthors are Alvise Bonivento, Alberto Sangiovanni-Vincentelli, Carlo Fischione, Karl Henrik Johansson, Piergiuseppe Di Marco, and Sinem Coleri Ergen.

A special thanks to Mikael Skoglund and Adam Dunkels for being my reference group. I am very pleased with their productive comments for my research.

I would like to take the opportunity to thank Piergiuseppe Di Marco for all the interesting discussions we had, not only about research but also great meals. I am also particularly grateful to Phoebus Chen, who read and commented the thesis. Pablo Soldati, Chitrupa Ramesh, Jose Araujo, Haibo Zhang, all other people in the Automatic Control Lab, thanks for the supportive working environment.

Thanks to all the friends I met here in Sweden. I am grateful to Aram Anto for our great jogging in Lappis. I must thank Daeho, Hyunsil and all KOSAS members.

A great thank to my family in South Korea, for supporting me in all the time. Most of all I would like to thank my parents for their continuous presence, support and encouragement. I must thank SHJ, who gave me the valuable life. I must express my friends, Chanwoo, Seunwook, Jinho and Gibeum who gave me great pleasure in Korea.

The research described in this thesis is supported by the EU project FeedNetBack, Swedish Research Council, Swedish Strategic Research Foundation, and Swedish Governmental Agency for Innovation Systems.

Contents

Part I

1 Introduction

Wireless sensor networks (WSNs) are networks of tiny sensing devices for wireless communication, monitoring, control, and actuation. Given the potential benefits offered by these networks compared to wired networks, such as, simple deployment, low installation cost, lack of cabling, and high mobility, they are especially appealing for industrial control applications [1–3]. The use of advanced communication technologies, highly integrated control, and programming platforms can drastically increase the performance of industrial control systems. Emerson Process Management [4] estimates that WSNs enable cost savings of up to 90% compared to the deployment cost of wired field devices (see also [5–7]). Several market forecasts have recently predicted exponential growths in the sensor network market over the next few years, resulting in a multi-billion dollar market in the near future. [1]

Although WSNs provide a great advantage for process, manufacturing and industrial applications, there is not yet widespread use of WSNs in these domains. This is because the software for these applications is usually written by process engineers that are experts in process control technology, but know little of the network and sensing infrastructure that has to be deployed to support the control applications. On the other hand, the communication infrastructure is designed by communication engineers that know little about process control technology. Moreover, the adoption of wireless technology further complicates the design of these networks. Being able to satisfy the stringent requirements on communication performance over unreliable wireless communication channels is a difficult task.

The standard practice for control system design over communication networks is as follows: First, deploy the networked embedded system based on the experience and heuristic considerations. Then, tweak the software implementation of the control algorithm to meet the latency, bandwidth, and reliability offered by the network. This is far from ideal, because many control systems are highly cost sensitive, and using a non-optimized network is clearly expensive. Moreover, the complexity of large networked embedded systems continues to increase, making heuristic and experience-based design practices inadequate at best. To bridge this gap and derive a correct and efficient implementation, a system-level design approach has been proposed in [8,9]. By a system-level design for WSNs, the control algorithm designers impose a set of requirements on reliability, packet delay, and energy consumption that the communication infrastructure must then satisfy.

The main contribution of this thesis is a novel approach to the protocol design of WSNs for control applications. We especially focus on maximization of the network lifetime as objective function and reliability and delay aspects as constraints. In addition, the communication protocol must adapt its design parameters according to the traffic and channel conditions for control applications.

[1] ON World [5] predicts a total market for WSN industrial applications of 4.6B$ by 2011 and for the Smart Building scenario of 2.5B$ by 2011. IDTechEx [7] foresees a total market size for WSNs and active RFID of about 4B$ by 2012.

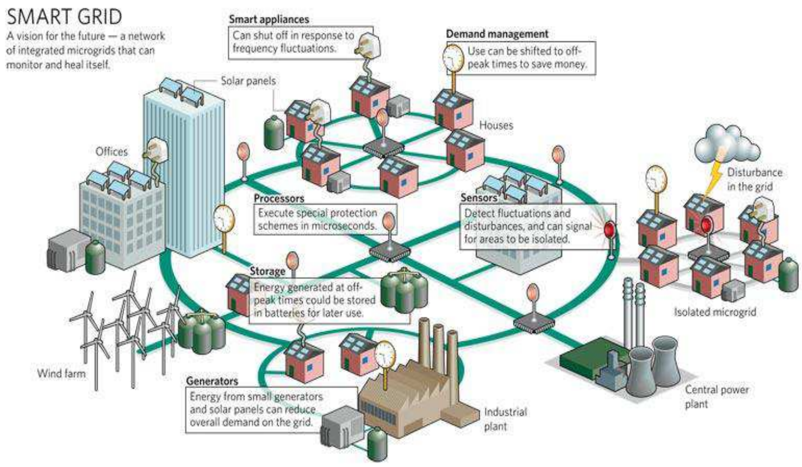


Figure 1: Structure of smart-grid network (U.S. department of energy, http://www.oe.energy.gov/.).

The remainder of this this chapter is organized as follows. In the next section, we motivate why WSNs are of interest. In Section 1.2 we present extra challenges WSNs impose on industrial control applications. Section 1.3 formulates the general mathematical problem used to design various protocols in this thesis. Finally, we present the contributions and an outline of the thesis.

1.1 Motivating Applications

We consider here two realistic scenarios where WSNs could be used.

Smart-grid technology

A recent interesting application of WSNs is to build the smart-grid infrastructure. WSNs are an integral part of the automated metering infrastructure and smart-grid plans. Smart-grid technology is designed to allow customers and utility companies to collaboratively manage power generation, delivery, storage, and energy consumption. One of the main messages from Bob Metcalfe, a co-inventor of Ethernet, is that the same type of innovation and entrepreneurship that built the Internet should be applied to building a smart-grid for a "squanderable abundance" of cheap and clean energy [10]. The *Enernet* is a new name for an internet-influenced vision of a smart-grid: The reason is that there will be lots of producers and consumers of energy, like the internet instead of a small number of dominant producers and a large number of consumers.

Figure 2: Wireless temperature transmitter of BP's Cherry Point plant [11], Emerson process management (http://www.emersonprocess.com).

Figure 1 describes the structure of a smart-grid network. Utilities are creating new technologies to make the power grid "intelligent" to automatically conserve energy. Solar panels and windmills will be mounted on houses to generate power. Smart appliances use WSNs to monitor how much electricity is being used and shut down when the power is too expensive. Consumers with remote control technologies can permit utilities to control their non-essential appliances, like pool pumps, turning them on and off to fine-tune the grid for maximum efficiency. Locally generated power avoids the power loss that occurs when you send electricity over long distance power lines. The efficient power lines route extra electricity from out-of-state utilities when demand spikes. Wireless devices let individual houses communicate with power utilities, swapping on-the-fly information about the current price and usage of electricity. Web and mobile phone interfaces allow consumers to see how much power their appliances are using when they are not at home and even allow consumers to turn them on or off remotely to reduce costs. When solar panels produce excess energy, the energy can be stored in batteries for use later at night.

Process control

Wireless technology can become a key technology in process control. In comparison to traditional wired sensors, wireless sensors provide advantages in the manufacturing environment, such as an increased flexibility for locating and reconfiguring sensors, elimination of wires in potentially hazardous locations, and easier of network maintenance.

For example, Emerson Process Management [4] and BP [11] have a collaboration where they try to apply wireless technology to improve the performance of factory automation. They have expanded the Cherry Point Washington refinery deployment, installed a wireless sytem throughout its tank farm in the R&D facility in Naperville, Illinois, and are making installations in other refineries around the world. BP Cherry Point is a $225,000$ barrels per day refinery, and is the largest supplier of calcined coke to the aluminum industry. Emerson's wireless installation on the refinery's calciner unit monitors bearing and calciner coke temperatures to help prevent fan and conveyor failure (see Figure 2). Fans can cost up to US$100,000 to repair but, more importantly, can be down for up to ten days if fans fail with associated production losses.

1.2 Challenges of WSNs for Control Applications

The protocol design process for WSNs in industrial control applications encounters more challenges than the protocol design process for traditional communication networks, namely:

- **Reliability**: Sensor readings must be sent to the sink of the network with a given probability of success, because missing sensor readings could prevent the correct execution of control actions or decisions being the phenomena sensed. However, maximizing the reliability may increase the network energy consumption substantially [2]. Hence, the network designers need to consider the tradeoff between reliability and energy consumption.

- **Delay**: Sensor information must reach the sink within some deadline. Time delay is a very important QoS measurement since it influences on performance and stability of an industrial control system [3]. The delay jitter can be much difficult to compensate for, especially if the delay variability is large. Hence, a probabilistic delay requirement must be considered instead of using average packet delay. Furthermore, the packet delay requirement is important since the retransmission of data packet to maximize the reliability may increase the delay. The outdated packet is generally not useful for control applications [12].

- **Energy Efficiency**: The lack of battery replacement, which is essential for affordable WSN deployment, requires energy-efficient operations. Since high reliability and low delay may require significant energy consumption, the reliability and delay must be flexible design parameters that still meet the requirements. Note that controllers can usually tolerate a certain degree of packet losses and delay [3, 13–17]. Hence, the maximization of the reliability and minimization of the delay are not the optimal design strategies since these strategies will significantly decrease the network lifetime for the control applications.

- **Sensor Traffic Patterns**: The type and amount of data to be transmitted is also important when considering industrial control applications [3]. Control signals can be divided into two categories: real-time and event based. For

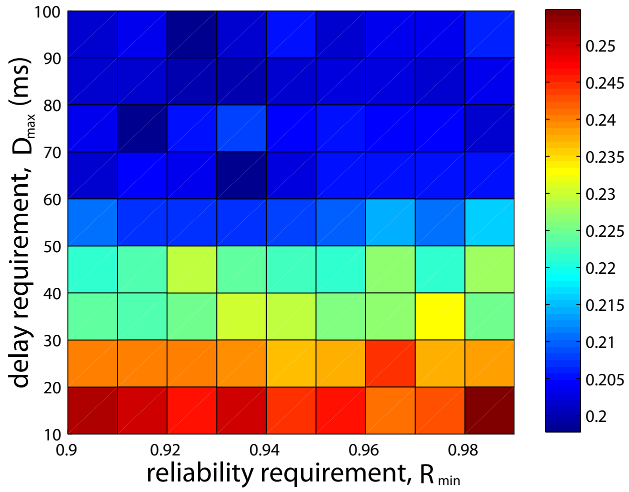


Figure 3: Power consumption of adaptive IEEE 802.15.4 with different reliability and average delay requirement.

real-time control, signals must be received within a specified time deadline for correct operation of the system. In order to support real-time control, networks must be able to guarantee the delay of a signal within a specified time deadline. Heavy traffic may be generated i.e., sensors send data very frequently. Event-based control signals are used by the controller to make decisions but do not have a time deadline. The decision is made if the system receives a signal or a timeout is reached. We remark here that some of the proposed protocol for environmental monitoring application, such as Dozer [18] and Fetch [19], operate in low traffic networks and can not handle the higher traffic loads of industrial control applications.

- **Adaptation**: The network operation should adapt to application requirement changes, time-varying wireless channels, and variations of the network topology. For instance, the set of application requirements may change dynamically and the communication protocol must adapt its parameters to satisfy the specific requests of the control actions. To support analysis-based design instead of experience-based design, it is essential to have analytical models describing the relation between the protocol parameters and performance indicators (reliability, delay, and energy consumption).

- **Scalability**: Since the processing resources on WSN nodes are limited, the calculations necessary to implement the protocol must be computationally light. These operations should be performed within the network, to avoid the burden of too much communication with a central coordinator. We note here that the tradeoff between tractability and accuracy of the analytical model is very important since sensor nodes [20, 21] have low computational capabilities. The protocol should also be able to adapt to variation in the network size, for example, size variations caused by the addition of new nodes.

As a consequence, the design of such networked systems has to take into account a large number of factors that ensure correct implementation. Starting from these requirements, it is important to design an efficient communication protocol that satisfies the constraints and optimizes the energy consumption.

Figure 3 reports a typical example of the power consumption of the network with different reliability and average delay requirements for adaptive IEEE 802.15.4 [22]. We clearly observe the tradeoff between the application requirements and power consumption of the network. Hence, the goal of the proposed design approach is to optimize the network behavior by considering the given constraints imposed by the application instead of just improving the reliability, delay, or energy efficiency without constraints. The objective function and requirements are used to solve a constrained optimization problem whose solution determines the policies and the parameters of the MAC and routing layer. In this thesis, we offer a complete design approach that embraces all the factors mentioned above.

1.3 Problem Formulation

In this section, we formulate a general constrained optimization problem for the protocol designs in this thesis.

Figure 4 shows the closed-loop tuning structure for multiple pairs of wireless communication, when several transmitters share the same wireless medium. Consider a network with N nodes. Node i considers the *application requirements* C_i chosen by the application and estimates the channel condition M_i to find the optimal parameters U_i of the protocol for the optimization of the WSN. Examples of C_i are the reliability and delay imposed by the control applications, and battery life time. The channel condition indicator M_i may denote the collision, busy channel probability, and channel accessing probability estimated by the transmitter. According to the estimated channel condition M_i, transmitter i computes and updates its optimal protocol parameter U_i. In addition, we denote the performance indicators as F, e.g., the reliability, delay and power consumption offered by the network.

We want to minimize the total energy consumption for transmitting and receiving packets, denoted by $E_{\rm tot}(\mathbf{u})$ where $\mathbf{u}$ is a vector of decision variables. The application requirements impose constraints on the probability of successful packet

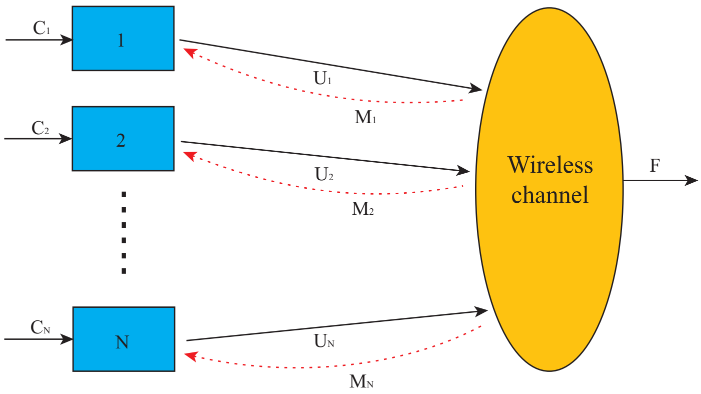


Figure 4: The network structure for the optimization problem described in (1).

delivery (reliability) and packet delay. Hence, the optimization problem is

$$\begin{aligned} \min_{\mathbf{u}} \quad & E_{\text{tot}}(\mathbf{u}) && \text{(1a)} \\ \text{s.t.} \quad & \mathbf{u} \in \mathcal{R}\,, && \text{(1b)} \\ & \mathbf{u} \in \mathcal{D}\,, && \text{(1c)} \\ & \mathbf{u} \in \mathcal{H}\,. && \text{(1d)} \end{aligned}$$

The decision variables $\mathbf{u}$ are the protocol parameters of the PHY, MAC, and routing layer. In (1b) and (1c), $\mathcal{R}$ and $\mathcal{D}$ are the feasible sets for the protocol parameters that meet the reliability and delay constraints, respectively. In addition, the feasible set $\mathcal{H}$ is due to physical layer properties of the hardware platform or limitations of the protocol standards. In the protocols proposed in this thesis, Problem (1) is a mixed integer-real optimization problem, because it is common that $\mathbf{u}$ may take on both real and integer numbers. We need to model Problem (1), along with a strategy to achieve the optimal solution, $\mathbf{u}^*$. As we will see later, the system complexity prevents us from deriving an exact expression such as reliability, delay, and energy consumption. Approximations will be used to get tractable analytical models for in-network processing. Note that this constrained optimization problem may be a local or global optimization problem.

Next, we present an example of a local optimization problem and an example of a global optimization problem used to design our protocols.

Example 1

Paper A presents a local optimization problem for IEEE 802.15.4 for reliable and timely communication. This protocol considers a star network topology with a personal area network coordinator, and N nodes with beacon enabled slotted

CSMA/CA and acknowledgements. It minimizes the power consumption while meeting the reliability and delay constraints without any significant modifications of the IEEE 802.15.4 standard. Each node solves the optimization problem by estimating the channel condition, i.e., busy channel probability and channel accessing probability. The local constrained optimization problem at node i is

$$\min_{\mathbf{u}_i} \quad E_{\mathrm{tot},i}(\mathbf{u}_i, \mathbf{u}_{-i}) \tag{2a}$$

$$\text{s.t.} \quad \mathcal{R}_i = \{\mathbf{u}_i \,|\, R_i(\mathbf{u}_i, \mathbf{u}_{-i}) \geq \Omega\}\,, \tag{2b}$$

$$\mathcal{D}_i = \{\mathbf{u}_i \,|\, \Pr[D_i(\mathbf{u}_i, \mathbf{u}_{-i}) \leq \tau] \geq \Delta\}\,, \tag{2c}$$

where $\mathbf{u}_{-i} = [\mathbf{u}_1, \ldots, \mathbf{u}_{i-1}, \mathbf{u}_{i+1}, \ldots, \mathbf{u}_N]$, $E_{\mathrm{tot},i}$ is the energy consumption and $\mathcal{R}_i$, and $\mathcal{D}_i$ are the feasible sets for the protocol parameters that meet the reliability and delay constraints of node i, respectively. Note that the objective function and constraints are also functions of the decision variables of the other nodes in the network. The decision variables are the MAC parameters related to the backoff mechanism and the maximum number of retransmissions. Each node updates its optimal protocol parameters by solving the local optimization problem. R_i is the reliability from the source node to the sink, and Ω is the minimum desired probability. D_i is a random variable describing the delay when transmitting a packet from the source node to the sink. τ is the desired maximum delay, and Δ is the minimum probability with which such a maximum delay should be achieved. We remark that τ, Δ, and Ω are the *application requirements*, and **u** are the *protocol parameters* that must be adapted to the traffic regime, wireless channel conditions, and application requirements for an efficient network.

Example 2

In Paper C, a global optimization problem is introduced to optimize the wake-up rate and the number of hops in the network. The cross-layer solution, called Breath, is designed for industrial control applications where source nodes attached to the plant must transmit information via multi-hop routing to a sink. The protocol is based on randomized routing, medium access control, and duty-cycling to minimize the energy consumption, while meeting reliability and packet delay constraints. The optimization problem is

$$\min_{\mathbf{u}} \quad E_{\mathrm{tot}}(\mathbf{u}) \tag{3a}$$

$$\text{s.t.} \quad \mathcal{R} = \{\mathbf{u} \,|\, R(\mathbf{u}) \geq \Omega\}\,, \tag{3b}$$

$$\mathcal{D} = \{\mathbf{u} \,|\, \Pr[D(\mathbf{u}) \leq \tau] \geq \Delta\}\,, \tag{3c}$$

where E_{tot} is the energy consumption and $\mathcal{R}, \mathcal{D}$ are the feasible sets for the protocol parameters that meet the reliability and delay constraints of the entire network, respectively. The decision variables of the wake-up rate and the number of hops are achieved by collaboration between the nodes in the network.

1.4 Contributions of the Thesis

This thesis presents a novel protocol design based on a systematic modelling and optimization to guarantee explicitly reliability and delay requirements in WSNs. We especially focus on maximization of the network lifetime by taking into account the tradeoff between energy consumption and application requirements with the dynamic and continuous adaptation of the network operations to the traffic and channel conditions for control applications. In Paper A, an adaptive IEEE 802.15.4 medium access control (MAC) protocol for minimizing the power consumption while guaranteeing reliability and delay constraints is presented. In Paper B, we investigate the unslotted IEEE 802.15.4 MAC with duty-cycled wireless sensor networks. The analytical expressions of delay and packet reception probabilities, and energy consumption are used for the efficient design and optimization of these resource-constrained networks. In Paper C, the cross-layer protocol called Breath is presented and experimentally evaluated on a test-bed with off-the-shelf wireless sensor nodes. It is based on randomized routing, medium access control, and duty-cycling jointly optimized for energy efficiency. The included papers are summarized in the following.

Paper A

P. Park, P. Di Marco, C. Fischione, and K. H. Johansson, "Adaptive IEEE 802.15.4 Protocol for Reliable and Timely Communications," *IEEE/ACM Transactions on Networking*, 2009, submitted.

The IEEE 802.15.4 standard for wireless sensor networks can support energy efficient, reliable, and timely packet transmission by tuning the medium access control parameters *macMinBE*, *macMaxCSMABackoffs*, and *macMaxFrameRetries*. Such a tuning is difficult, because simple and accurate models of the relations of these parameters on the probability of successful packet transmission, packet delay and energy consumption are not available. Moreover, it is not clear how to adapt the parameters to the changes of the network and traffic regimes by simple algorithms that can run on resource-constrained nodes. In this paper, a generalized Markov chain is proposed to model these relations by simple expressions without giving up the accuracy. In contrast to previous work, the presence of limited number of retransmissions, acknowledgments, unsaturated traffic and packet size is accounted for. The analysis is then used to propose an adaptive algorithm for minimizing the power consumption while guaranteeing reliability and delay constraints in the packet transmission. The algorithm does not require any modification of the IEEE 802.15.4 standard and can be easily implemented one existing network nodes. Numerical results show that the analysis is accurate, that the proposed algorithm satisfies reliability and delay constraints, and ensures a longer lifetime of the network under both stationary and transient network conditions.

Paper B

C. Fischione, P. Park, S. C. Ergen, K. H. Johansson, and A. Sangiovanni-Vincentelli, "Analytical Modeling and Optimization of Duty-cycles in Preamble-based IEEE 802.15.4 Wireless Sensor Networks," *IEEE/ACM Transactions on Networking*, 2009, submitted.

The efficient design and optimization of duty-cycled wireless sensor networks with preamble sampling random medium access control (MAC) greatly is based on accurate modeling of delay, packet reception probabilities, and energy consumption. The challenges for modeling are the random MAC and the sleep policy of the receivers that makes it impossible to determine the exact time of data packet transmission. A novel approach to the modeling of the delay, reliability, and energy consumption is proposed for a clustered network topology with unslotted IEEE 802.15.4 and preamble sampling MAC. The analysis developed in this paper gives expressions of the delay, reliability and energy consumption as a function of sleep time, listening time, traffic rate and MAC parameters. These expressions can then be effectively used to optimize the duty-cycle of the nodes. The optimization ensures a significant reduction of the energy consumption compared to existing solutions in the literature. Monte Carlo simulations using the ns-2 simulation tool demonstrate the validity of the analysis.

Paper C

P. Park, C. Fischione, A. Bonivento, K. H. Johansson, and A. Sangiovanni-Vincentelli, "Breath: an Adaptive Protocol for Industrial Control Applications using Wireless Sensor Networks," *IEEE Transactions on Mobile Computing*, 2009, submitted.

Energy-efficient, reliable and timely data transmission is essential for wireless sensor networks (WSNs) employed in scenarios where plant information must be available for control applications. To reach a maximum efficiency, cross layer interaction is a major design paradigm to exploit the complex interaction among the layers of the protocol stack. This is challenging because latency, reliability, and energy are at odds, and resource constrained nodes support only simple algorithms. In this paper, the novel protocol Breath is proposed for control applications. Breath is designed for WSNs where source nodes attached to a plant must transmit information via multi-hop routing to a sink. Breath ensures a desired packet delivery and delay probabilities while minimizing the energy consumption of the network. The protocol is based on randomized routing, medium access control, and duty-cycling jointly optimized for energy efficiency. The design approach relies on a constrained optimization problem, whereby the objective function is the energy consumption and the constraints are the packet reliability and delay. The challenging part is the modelling of the interactions among the layers by simple expressions of adequate accuracy, which are then used for the optimization by in-network processing. The optimal working point of the protocol is achieved by a simple algorithm, which adapts to traffic variations and channel conditions with negligible overhead. The protocol has been implemented and experimentally evaluated on a test-bed with off-the-shelf wireless sensor nodes, and it has been compared with a standard IEEE 802.15.4 solution. Analytical and experimental results show that Breath is tunable and meets reliability and delay requirements. Breath exhibits a good distribution of the working load, thus ensuring a long lifetime of the network. Therefore, Breath is a good candidate for efficient, reliable, and timely data gathering for control applications.

Other Related Papers

The following publications, although not included in this thesis, contain material that is similar or related to the aforementioned contributions:

– Investigations on IEEE 802.15.4 standard:

- P. Park, C. Fischione and K. H. Johansson, "Performance analysis of GTS allocation in Beacon enabled IEEE 802.15.4", in *IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, Rome, Italy, July, 2009.

- C. Fischione, S. Coleri Ergen, P. Park, K. H. Johansson and A. Sangiovanni-Vincentelli, "Medium Access Control Analytical Modeling and Optimization in Unslotted IEEE 802.15.4 Wireless Sensor Networks", in *IEEE SECON*, Rome, Italy, July, 2009.

- P. Park, P. Di Marco, P. Soldati, C. Fischione and K. H. Johansson, "A Generalized Markov Chain Model For Effective Analysis of Slotted IEEE 802.15.4", in *IEEE International Conference on Mobile Ad-hoc and Sensor Systems (MASS)*, Macau, P. R. C., 2009, (**Best Paper Award**).

- P. Park, P. Di Marco, C. Fischione and K. H. Johansson, "Delay Analysis of Slotted IEEE 802.15.4 with a Finite Retry Limit and Unsaturated Traffic", in *IEEE International Conference on Communications (ICC)*, 2010, submitted.

– Cross-layer solution and application:

- P. Park, C. Fischione, A. Bonivento, K. H. Johansson and A. Sangiovanni-Vincentelli, "Breath: a Self-Adapting Protocol for Wireless Sensor Networks in Control and Automation", in *IEEE SECON*, San Francisco, USA, June, 2008.

- P. Di Marco, P. Park, C. Fischione and K. H. Johansson, "TREnD: a Timely, Reliable, Energy-efficient and Dynamic WSN Protocol for Control Applications", in *IEEE ICC*, 2010, submitted.
- E. Witrant, P. Park, M. Johansson, C. Fischione and K. H. Johansson, "Predictive Control Over Wireless Multi-Hop Networks", in *IEEE Multiconference on System and Control*, Singapore, 2007.

– Transmit power control of WSN:

- B. Z. Ares, P. Park, C. Fischione, A. Speranzon and K. H. Johansson, "On Power Control for Wireless Sensor Networks: System Model, Middleware Component and Experimental Evaluation", in *European Control Conference (ECC)*, Greece, 2007.
- P. Park, C. Fischione and K. H. Johansson, "A simple power control algorithm for wireless ad-hoc networks", in *IFAC World Congress*, Seoul, Korea, July, 2008.

Contributions by the author

The thesis is partially based on papers written with co-authors. In the joint papers the author has actively contributed both to the development of the theory as well as the paper writing.

1.5 Thesis Outline

The remaining of this thesis is organized in two parts. The first one, comprising Sections 2.1 through 2.3, highlights and briefly summarizes the background studies: Section 2.1 presents the characteristics and challenges of networked control system (NCS); Section 2.2 describes the system-level design approach to interconnect the control applications and communication network; Section 2.3 summarizes the characteristics of the protocols that are relevant for the category of applications we are concerned in this thesis. Concluding remarks and open issues are outlined in Section 3. The second part contains verbatim copies of all the papers included in the thesis.

2 Background

In this section, we briefly summarize the related work: Section 2.1 presents the features and challenges of networked control systems; Section 2.2 introduces the system-level design approach for WSNs. In addition, Section 2.3 presents the current standardization work of WSNs and compares the main features of the protocols that are relevant for the category of applications we are concerned in this thesis. Details follow in the sequel.

2.1 Networked Control Systems

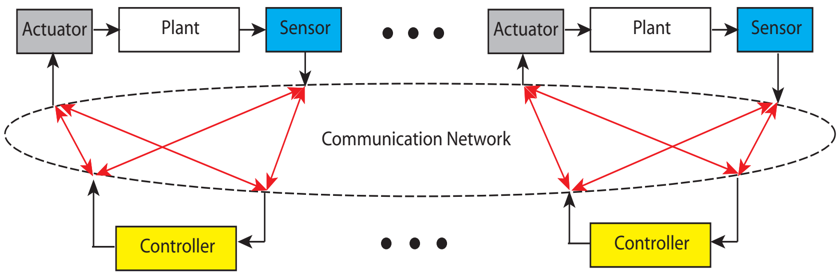


Figure 5: General networked control systems structure.

Networked control systems (NCSs) are spatially distributed systems in which the sensors, actuators, and controllers connect through a communication network instead by traditional point-to-point connections, as shown in Figure 5. The significant advantages over traditional control architectures include reduced wiring and cost, increased modularity, easier maintenance, and high flexibility and reconfigurability [1, 2]. Networked control has become an enabling technology for many military, commercial and industrial applications such as mobile sensor networks [23], remote surgery [24], industrial automation [25]. Wireless communication is playing an increasingly important role in NCSs. Transmitting sensor measurements and control commands over wireless links allows rapid deployment, flexible installation, fully mobile operation and prevents cable problems in industrial control applications.

Figure 5 depicts the general structure of NCSs where a plant is remotely commissioned over a network. Outputs of the plant are sampled at periodic intervals by the sensor and forwarded to the controller through a network. When the controller receives the measurements, a new control command is computed. The control is forwarded to the actuator attached to the plant. Research on NCSs sometimes considers structures simpler than the general one depicted in Figure 5. For example, controllers may be collocated with the corresponding actuators. It is also common

to consider single feedback loops closed over a network [26, 27].
There has been much research effort to design controllers that are robust to communication faults. Limited communication resources, e.g., bandwidth limit, random packet delay, and data dropouts may cause undesirable behavior of the system. Accurate modelling of communication networks requires heavy computation load and can still be hard. Regarding protocol design for communication networks, there has been much research on deterministic performance of control networks using token passing bus and controller area network (CAN) bus architectures. Comparatively, much less work on wireless NCSs has considered protocols for the recently developed standard such as IEEE 802.11 [28], Bluetooth and 802.15.4 [29]. Many papers have been written about networked control: extensive research on the impact on system performance and stability of the network and protocols can be found in [13, 30–33]. Furthermore, the papers [3, 12, 14] focus on the effects of data sampling, network delay, and packet dropouts on the stability of the resulting closed-loop NCSs. We summarize here the important network quality measures for NCSs.

- **Bandwidth**: Bandwidth is the information-carrying capacity of a communication channel. There is recent research on the problem of determining the minimum bit rate that is needed to stabilize a linear system [34, 35]. The data rate of a network must be considered together with the packet size and overhead since data are encapsulated into packets. Notice that the size of the headers depends on the protocol design of the communication network.

- **Sampling and Delay**: The time delay of data on the network is the total time between the data being available at the source node (e.g., sampled from sensors) and being available at the sink node (e.g., received at the controller). This process is significantly different from the usual periodic sampling in digital control system. The overall delay between sampling and receiving can be highly variable because both the network access delays (i.e., the time it takes for a shared network to accept data) and the transmission delays (i.e., the time during which data are in transit inside the network) depend on highly variable network conditions such as congestion and channel quality. In some NCSs, the data transmitted are time stamped, which means that the receiver may have an estimate of the delay duration and can take an appropriate corrective action. Many research results have attempted to characterize a maximum upper bound on the sampling interval for which stability can be guaranteed. These results implicitly attempt to minimize the packet rate that is needed to stabilize a system through feedback. Furthermore, the delay jitter needs to be considered since it can be much more difficult to compensate for, especially if the variability is large.

- **Packet Dropout**: Another significant factor is the reliability of the network in NCSs compared to standard digital control system, e.g., packet loss of the wireless channel. Packet dropouts result from communication errors in the physical layer of the wireless link or from buffer overflows due to congestion of the network. Longer forward delays result in packet reordering, which essentially amounts to a packet dropout if the receiver discards "outdated" arrivals which is also critical factor for time-bounded traffic such as audio and video [36]. Reliable transmission protocols, such as RMST [37] and MMSPEED [38] of WSNs, guarantee the eventual delivery of packets by using an acknowledgement mechanism. However, these protocols may not be appropriate for NCSs since repeated retransmissions of old data is generally not useful for control applications. Maximizing the reliability may increase substantially the network energy consumption [2]. Furthermore, in general, there is tradeoff between the reliability of a network and the delay of packet delivery [22]. We remark here that controllers can usually tolerate a certain degree of packet losses and delay [3, 13–17].

When using networks for control applications, the selection of networks for a particular application is important to asses determinism and balancing as QoS parameters [3]. In the next section, we summarize the system-level design which provides us a design paradigm for control application and communication network.

2.2 System-Level Design

System-level design [9, 39] is used in systems that consist of several components like semiconductors, cars, airplanes, buildings, telecommunication systems, and biological systems. Hence, specifications are given in terms of functionality. In system-level design, one has to consider

- an objective function that expresses the desirable features of the design,
- constraints on the design and on the individual components available for implementation.

Platform-based design (PBD) [40–44] is a common approach in system-level design to reduce development time and manufacturing costs by using system-on-chips. It reuses the whole platform instead of individual cores and thus significantly improves the reusability and design efficiency. The central idea of the PBD methodology is to achieve performance constraints through mapping adjustment and architecture model refinement. Platforms allow several users to customize the same basic platform into different products. Platforms are particularly useful in standards-based markets where some basic features must be supported but other features must be customized to differentiate products. This property makes PBD an important actor to bridge the gap between control and communication designers in WSNs. Many high-volume markets of wireless network are standards driven. This trend encourages PBD since a communication designer chooses implementation of standards functions and a control designer add features for different applications such as process control, factory automation using WSNs. This partition

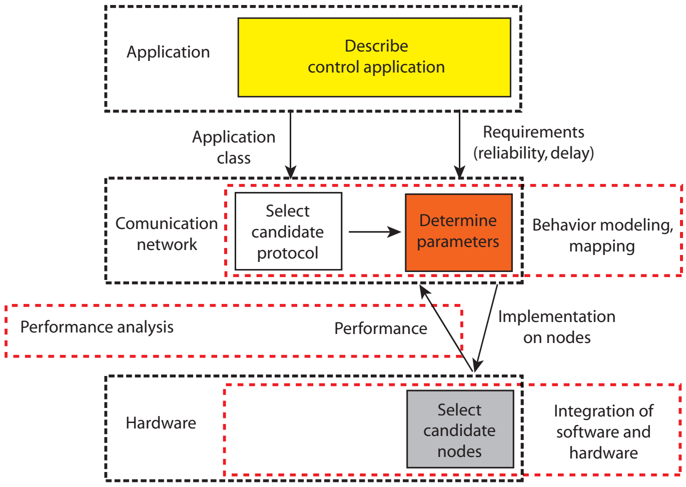


Figure 6: Design flow of PBD for wireless sensor networks. The red box describes the related step in PBD design flow.

matches well two phases of PBD: the design of the platform and the use of the platform [41]. Hence, PBD methodology can contribute to the solution of these problems of control application using WSNs focusing the effort on the definition of a clear set of abstraction layers.

Figure 6 shows the design flow of PBD for control applications using WSNs. The output of PBD produces an implementation of the control network that is optimized such that the network satisfies the given constraints while minimizing the overall cost of the network implementation. The process of mapping the application description to a communication protocol instance and eventually to a hardware platform instance goes through a set of steps.

The application designer describes the control application independently from the communication protocols or hardware platforms. The control designer generates a set of constrains that the communication network and hardware must satisfy to ensure correct functionality. Note that it is often important to assess determinism as a QoS parameter, specifically evaluating whether packet delay can be predicted and bounded. Balancing QoS parameters of NCS is investigated in [3].

In the next step, PBD defines a possible set of modules of communication library and the interfaces that these communication protocols offer to the control applications. Behavior modelling is a key step to select the communication protocol since behavior model details the function constraints of the specifications and it will be used to develop object code running on hardware. This is one of reasons that we develop the analytical modelling in the design process. The model is an efficient tool for PBD in terms of computation load and development time compared to experience-based or simulation-based approach. Once the communication protocol is selected, it must be implemented on a set of physical nodes such as Tmote [20] and MICA [21]. This step integrates the software refinement of the communication protocols and the hardware refinement of the sensor nodes.

Designers use performance analysis to evaluate the architecture model, mapping, and interfaces and adjust the design decisions of the communication protocols by looking at the analysis results. The accuracy of performance directly affects the real performance of the whole system. Inaccurate information affects the result of the performance analysis. Any unrealistic requirements must be gotten rid of by either mapping adjustment or architecture revising. Based upon the requirements, designers can either purchase design customized modules such as communication protocols or hardware platforms. Note that application designers when using this approach can adapt to a new implementation platform exploiting the advantages of the technology without having to pay the price of redesigning their applications.

2.3 Protocol Design for WSNs

During last years, many protocols for WSNs have been proposed for a variety application, such as area, environmental monitoring and industrial network, both in academia (e.g., [2,45]) and industry (e.g., [20,46,47]). In addition, new protocols may continue to emerge to address niches where a unique QoS balance is needed. WSN promises to dramatically expand the number of devices in a plant that can be connected wirelessly. We will first introduce the most practical and promising standards and an existing commercial systems for the industrial communication community. In the second part we discuss interesting protocols that are relevant for the category of applications we are concerned in this thesis.

There have been many contributions to the standardized low-power protocols such as IEEE 802.11 [28], 802.15.4 [29], Zigbee [48], ISA SP-100 [49], WirelessHART [50], 6loWPAN [51] and routing over low power and lossy networks (ROLL) [52]. The IEEE 802.11 family of wireless LAN (WLAN) standards is composed of a number of specifications that primarily define the physical and MAC layers of the realm of WLAN technologies, and it has also been considered extensively in the context of wireless industrial communications, see [53,54]. Similar to other standards from the IEEE 802.x series, the IEEE 802.11 MAC suggest the IEEE 802.2 logical link control (LLC) [55] as a standard interface to higher layers. Since IEEE 802.11 is a WLAN standard, its key intentions are to provide high throughput and a continuous network connection.

The ZigBee standard is prepared by an industry consortium, the ZigBee alliance. ZigBee covers the networking layer and application layer of WSN applications and is defined to work on top of a modified version of the IEEE 802.15.4 standard. ZigBee allows to create different kinds of networks: in star networks the ZigBee coordinator starts the network and all the other network members (the end devices) are directly associated with the ZigBee coordinator. The ZigBee coordinator is co-located with the personal area network (PAN) coordinator of the underlying IEEE 802.15.4 network. In tree networks the ZigBee routers form a tree that is rooted at the ZigBee coordinator, whereas in mesh networks the network topology might be a general mesh involving ZigBee routers and the ZigBee coordinator.

The Instrumentation, Systems, and Automation Society (ISA) [49] is currently working on a series of standards addressing the adoption of wireless technologies in different industries. ISA-SP100.11a addresses noncritical process applications that can tolerate delays up to 100 ms. Since it leverages the IEEE 802.15.4 standard, it inherits some of its properties: low rates (up to 250 kBit/s) and low implementation complexity for simple end devices. In addition a data-link layer and an adaptation layer between MAC and data link layer are introduced. The data link layer controls the frequency hopping and adds a TDMA scheme. We remark that both standards [48, 49] target overlapping application areas and are based on the same underlying wireless technology.

WirelessHART is a promising solution for the replacement of the wired HART protocol in industrial contexts. Power consumption is not a main concern in WirelessHART, whereas the data link layer is based on Time Division Multiple Access (TDMA), which requires time synchronization and pre-scheduled fixed length time-slots by a centralized network manager. Such a manager should update the schedule frequently to consider reliability and delay requirements and dynamic changes of the network, which demands complex hardware equipments, and this is in contrast with the necessity of simple protocols able to work with limited energy and computing resources.

ROLL is focused on routing issues for Low power and Lossy networks (LLNs). LLNs are made up of many embedded devices with limited power, memory, and processing resources. They are interconnected by a variety of links, such as IEEE 802.15.4, Bluetooth, Low Power WiFi, wired or other low power PLC (Powerline Communication) links. LLNs are transitioning to an end-to-end IP-based solution to avoid the problem of non-interoperable networks interconnected by protocol translation gateways and proxies. The working group focuses on routing solutions for a subset of these: industrial, connected home, building and urban sensor networks for which routing requirements have been specified. These application specific routing requirements will be used for protocol design. The framework will take into consideration various aspects including high reliability in the presence of time varying loss characteristics and connectivity while permitting low power operation with very modest memory and CPU pressure in networks potentially comprising a very large number of nodes.

As we discussed, many promising standards and an existing commercial system are based on the IEEE 802.15.4 standard. According to a recent survey, this standard already represents more than 50% of the market [5]. IEEE 802.15.4 radio standard and ZigBee emerge as the prevalent choice for industrial and smart building applications. IEEE 802.15.4 is standard for a low data rate solution with efficient energy consumption and very low complexity. There are also many discussions of IEEE 802.15.4 for the routing over low power and lossy networks in the internet engineering task force (IETF) working group [52]. Recently, many task groups launch IEEE 802.15.4 family for specific applications in WSN. Summarizing, the task groups of IEEE 802.15.4:

- IEEE 802.15.4 a
  It specifies two additional PHYs using Ultra-wideband (UWB) and Chirp Spread Spectrum (CSS) which is an amendment to IEEE 802.15.4 to provide communications and high precision ranging/location capability, high aggregate throughput, and ultra low power, scalability to data rates, longer range, and lower cost.

- IEEE 802.15.4 e
  The intent of this amendment is to enhance and add functionality to the IEEE 802.15.4 MAC that are required to enable those application spaces: factory automation, process automation, asset tracking, general sensor control (industrial/commercial, including building automation), home medical health/monitor, telecom application, neighborhood area networks, audio.

- IEEE 802.15.4 f
  It define new wireless Physical (PHY) layer(s) and enhancements to the IEEE 802.15.4 standard MAC layer which are required to support new PHY for active RFID system bi-directional and location determination applications.

- IEEE 802.15.4 g
  It creates a PHY amendment to IEEE 802.15.4 to provide a global standard that facilitates very large scale process control applications such as the utility smart-grid [56] networks.

Now, we discuss the interesting protocols that have been developed in the recent years relevant for the category of applications we are concerned in this thesis. In Tab. 1, we summarize the characteristics of the relevant protocols. In the table, we have evidenced performance indications as energy, reliability, and delay have been included in the protocol design and validation, and whether a cross-layer approach has been adopted. We discuss these protocols in the following.
GAF, SPAN and X-MAC [57–59] consider the energy efficiency as a performance indicator, which is attained by algorithms under the routing layer and above the MAC layer (bridge layer) or in the MAC layer. Simulation results of reliability and delay are reported in [57, 58], but these protocols have not been designed out of an analytical modelling of reliability and delay, so that there is not control of them. One of the first protocol for WSNs designed to offer a high reliability is RMST [37], but no energy consumption of the network and delay have been accounted for. The same lack of energy efficiency and delay requirements can be found in the reliable solutions presented in [19, 60, 63]. Dozer [18] comprises the MAC and routing layer to minimize the energy consumption while maximizing the reliability of the network, but an analytical approach has not been followed. Specially, Fetch [19] and Dozer [18] are designed for monitoring application, which

Table 1: Protocol comparison. The circle denotes that a protocol is designed by considering the indication of the column, but it has not been validated experimentally. The circle with plus denotes that the protocol is designed including the indication of the column plus experimental validation. The dot denotes that the protocol does not include in the design the indication of the column, but simulation or experiment results include them. The term “bridge” means that the protocol is designed between MAC and routing layer.

Protocol	Energy	Reliability	Delay	Layer
GAF [57]	○	•	•	bridge
SPAN [58]	○	•	•	bridge
XMAC [59]	⊕	•	•	MAC
Flush [60]		⊕		MAC
Fetch [19]	•	⊕	•	phy, MAC, routing
GERAF [61]	○		○	MAC, routing
Dozer [18]	⊕	⊕		MAC, routing
MMSPEED [38]		○	○	routing
Breath [62]	⊕	⊕	⊕	phy, MAC, routing

mainly deals with low traffic regime than industrial control application. The latency of Fetch [19] is significantly dependent on the depth of the nodes in the routing tree and is around some hundred seconds. In addition, experimental results of Dozer [18] show good energy efficiency and reliability under very low traffic intensity with the data sampling interval 120 s, but the delay in the packet delivery is not considered, which is essential for industrial control applications [12], [13]. Energy efficiency with a delay requirement for a MAC and randomized routing is considered in GERAF [61], without simulation or experimental validation.

The purposes of the protocols mentioned above [57–61, 63] are the maximization of the energy efficiency or reliability, or just minimization of the delay, without considering simultaneously application requirement in terms of reliability and delay in the packet delivery. In other words, none of these protocol supports explicitly an adaptation to the changes of these typical industrial control application requirements. However, industrial control applications are able to cope with a certain degree of packet losses and delay [3, 13–17], which implies that the approaches followed in the protocols mentioned above are not the ideal solution for these applications. The maximization of the energy efficiency and reliability may give long delay, which are dangerous for the stability of closed loop control systems. Analogously, the maximization of the reliability may be energy demanding and may give long delay, all of which are not tolerable for the control applications we are concerned in this thesis.

MMSPEED [38] and SERAN [64] are appealing for industrial control applications. However, MMSPEED is not energy efficient because it considers a routing technique with an optimization of reliability and delay without energy constraints. The protocol satisfies a high reliability requirement by using duplicated packets over a multi-path routing. However, duplicated packets increase the traffic load with negative effect on the stability and energy efficiency of the network. In SERAN, a system level design methodology has been presented for industrial applications, but even though SERAN allows the network to operate with low energy consumption subject to delay requirements, it does not consider tunable reliability requirements nor duty-cycling policies, which are essential to reduce energy consumption. Furthermore, SERAN focuses on low traffic networks. These characteristics limit the performance of SERAN both in term of energy and reliability in our application setup.

Given the availability of numerous techniques to reduce energy consumption, ensure reliability and low delays, a cross-layer optimization is a natural approach to integrate the protocol layers. Some cross-layer design challenges of the physical, MAC and network layers to minimize the energy consumption of WSNs have been surveyed in [65–67]. Many of the cross-layer solutions proposed in the literature are hardly useful for the application domain we are targeting, because they require sophisticated processing resources, or instantaneous global network knowledge, which are out of reach of the node's capabilities. Network design can be formulated as an optimization problem. However, as it was noted in [68], the complex interdependence of the decision variables (sleep disciplines, clustering, MAC, routing, power control, etc.) lead to difficult problems even in simple network topologies, where the analytical relations describing packet reception rate, delay and energy consumption may be expressions highly nonlinear. Such a difficulty is further exacerbated when considering non-TDMA scheme [69]. Hence, a design approach that offers a computationally attractive solution by simplifications of adequate accuracy is also important.

3 Conclusions

In this thesis, we propose a novel protocol design that network satisfies the specifications of the application designer while minimizing an objective function. Specifically, the object function is energy consumption and the constraints are reliability, delay aspect and hardware infrastructure. The main idea of the proposed protocol design is to apply the tradeoff between the application requirements and energy consumption of the network instead of just improving the reliability, delay or energy efficiency. In the design process, we consider an analytical expression of the total energy consumption of the network, as well as reliability and delay for the packet delivery. This seems suitable for many control applications as they require guarantees for stability and performance of system. We briefly summarize the key contributions presented in the three papers.

In Paper A, an adaptive IEEE 802.15.4 is introduced for energy efficient, reliable, and low latency packet transmission. The backoff mechanisms and retry limits of the standard are adapted to the estimated channel conditions. Numerical results show that the proposed protocol enhancement ensures a longer lifetime of the network while guaranteeing application requirements under both stationary and transient conditions. Furthermore, we investigate the robustness and sensitivity of the protocol to possible errors during the estimation process.

In Paper B, we investigate the efficient design and optimization of duty-cycled WSNs with preamble sampling. The analysis gives expressions of the delay, reliability and energy consumption as functions of sleep time, listening time, traffic rate and MAC parameters. These expressions can then be used to optimize the duty-cycle of the nodes to minimize energy consumption while ensuring low latency and reliable packet transmissions. The optimization results in a significant reduction of the energy consumption compared to existing solutions.

In Paper C, the cross-layer protocol called Breath is proposed for control applications by using the proposed constrained optimization. The protocol optimizes the wake-up rate of duty-cycling mechanism and the number of hops in the network. The optimal working point of the protocol is achieved by a simple algorithm, which adapts to traffic variations and channel conditions with negligible overhead. The protocol is implemented and experimentally evaluated on a testbed with off-the-shelf wireless sensor nodes, and it is compared with a standard IEEE 802.15.4 solution. Breath exhibits a good fairness of the work load, thus ensuring a long network lifetime.

4 Future Work

The thesis presents original contributions to WSN protocol design for industrial control applications, but it also suggests many research issues for further studies. We here outline possible directions for both short-term and long-term future work.

We are currently investigating the extension of the design methodology to consider general topology of mesh networks (ad-hoc and wireless sensor networks) and to interact with other control applications. It is important to implement and evaluate our protocols for industrial control applications, see [15]. The tradeoff between tractability and accuracy is important since a long computation time does not allow practical usage of the protocol.

As mentioned in the Paper A, the stability analysis of adaptive tuning algorithms is an interesting issue. Such a study can be generalized for the stability analysis of a variety protocols of WSNs. It would be interesting to extend the proposed constrained optimization problem by a game theoretic approach.

The IEEE 802.15.4 has superframe structure to support the random access scheme (CSMA-CA algorithm) and TDMA allocation mechanisms. Although there are many investigations for various mechanisms of random access and TDMA allocation, hybrid MAC modelling is not satisfactorily considered for IEEE 802.15.4 standard. Since two mechanisms influence each other in the superframe structure, proper hybrid MAC modelling is instrumental for the efficient design for IEEE 802.15.4. One of the critical questions related to IEEE 802.15.4 is the feasibility and limitation of standard for industrial, smart building applications and smart-grid technology. If it is not a good solution, then we need to propose and analyze the modifications of the current standard. The integration with routing layer like 6loWPAN [51] and ROLL [52] is also important to study.

The interconnection of the control algorithm and the communication network is a critical part. Performance analysis with various methodologies like component-based, cross-layer design should be studied for wireless sensor and actuator network.

References

[1] A. Willig, K. Matheus, and A. Wolisz, "Wireless technology in industrial networks," *Proceedings of the IEEE*, 2005.

[2] A. Willig, "Recent and emerging topics in wireless industrial communication," *IEEE Transactions on Industrial Informatics*, vol. 4, no. 2, pp. 102–124, 2008.

[3] J. R. Moyne and D. M. Tilbury, "The emergence of industrial control networks for manufacturing control, diagnostics, and safety data," *Proceedings of the IEEE*, 2007.

[4] *Emerson Process Management*. [Online]. Available: http://www.emersonprocess.com

[5] *WSN for Smart Industries: A Market Dynamics Report*, OnWorld, 2007. [Online]. Available: http://www.onworld.com/smartindustries/index.html

[6] *Network Operations Platforms: a New Segment in an Expanded Wireless Sensor & Control Network Market Category*, Network Operations Platform White paper.

[7] P. Harrop, "Wireless sensor networks 2009-2019," IDTechEx, Tech. Rep., 2008.

[8] A. Bonivento, L. Carloni, and A. Sangiovanni-Vincentelli, "Platform-based design for wireless sensor networks," *Mobile Netw. Appl.*, 2006.

[9] A. Sangiovanni-Vincentelli, "Quo Vadis, SLD? reasoning about the trends and challenges of system level design," *Proceedings of the IEEE*, vol. 95, no. 3, 2007.

[10] B. Metcalfe, "Internet history applied to solving energy," in *Green and Network (Green:Net)*, 2009. [Online]. Available: http://www.youtube.com/watch?v=FmMd8dxxOQQ

[11] *BP*. [Online]. Available: http://www.bp.com

[12] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, 2007.

[13] W. Zhang and M. S. Braniky and S. M. Phillips, "Stability of Networked Control Systems," *IEEE Control Systems Magazine*, 2001.

[14] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, "Foundations of control and estimation over lossy networks," *Proceedings of the IEEE*, 2007.

[15] E. Witrant, P. Park, M. Johansson, C. Fischione, and K. H. Johansson, "Predictive control over wireless multi-hop networks," in *IEEE MSC*, 2007.

[16] A. Speranzon, C. Fischione, and K. H. Johansson, "Distributed and collaborative estimation over wireless sensor networks," in *IEEE Conference on Decision and Control*, 2006.

[17] A. Speranzon, C. Fischione, K. H. Johansson, and A. Sangiovanni-Vincentelli, "A distributed minimum variance estimator for sensor networks," *IEEE Journal on Selected Areas in Communications, Special Issue on Control and Communication*, vol. 26, pp. 609–621, May 2008.

[18] N. Burri, P. von Rickenbach, and R. Wattenhofer, "Dozer: ultra-low power data gathering in sensor networks," in *ACM/IEEE IPSN*, 2007.

[19] G. Werner-Allen, K. Lorincz, J. Johnson, J. Lees, and M. Welsh., "Fidelity and yield in a volcano monitoring sensor network," in *USENIX OSDI*, 2006.

[20] *Tmote Sky Data Sheet*, Moteiv, San Francisco, CA, 2006. [Online]. Available: http://www.moteiv.com/products/docs/tmote-sky-datasheet.pdf

[21] *MICA motes*. [Online]. Available: http://www.xbow.com

[22] P. Park, P. D. Marco, C. Fischione, and K. H. Johansson, "Adaptive IEEE 802.15.4 protocol for reliable and timely communication," *ACM/IEEE Transactions on Networking*, 2009, submitted.

[23] P. Ogren, E. Fiorelli, and N. E. Leonard, "Cooperative control of mobile sensor networks: Adaptive gradient climbing in a distributed environment," *IEEE Transactions on Automatic Control*, vol. 49, no. 8, pp. 1292–1302, 2004.

[24] C. Meng, T. Wang, W. Chou, S. Luan, Y. Zhang, and Z. Tian, "Remote surgery case: robot-assisted teleneurosurgery," in *IEEE ICRA*, vol. 1, 2004, pp. 819–823.

[25] J. Garcia, F. Palomo, A. Luque, C. Aracil, J. Quero, D. Carrion, F. Gamiz, P. Revilla, J. Perez-Tinao, M. Moreno, P. Robles, and L. Franquelo, "Reconfigurable distributed network control system for industrial plant automation," *IEEE Transactions on Industrial Electronics*, vol. 51, no. 6, pp. 1168–1180, 2004.

[26] L. A. Montestruque and P. J. Antsaklis, "On the model-based control of networked systems," *Automatica*, p. 18371843, 2003.

[27] P. Seiler and R. Sengupta, "Analysis of communication losses in vehicle control problems," in *Proceedings of the American Control Conference*, 2001, pp. 1491–1496.

[28] *IEEE Std 802.11 Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE, 1999. [Online]. Available: http://www.ieee802.org/11

[29] *IEEE Std 802.15.4-2996, September, Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)*, IEEE, 2006. [Online]. Available: http://www.ieee802.org/15

[30] G. Walsh, H. Ye, and L. Bushnell, "Stability analysis of networked control systems," *IEEE Transactions on Control Systems Technology*, pp. 2876–2880, 1999.

[31] F. L. Lian, J. Moyne, and D. Tilbury, “Analysis and modeling of networked control systems: MIMO case with multiple time delays,” in *Proceedings of the ACC*, vol. 6, 2001, pp. 4306–4312.

[32] R. Luck and A. Ray, “An observer-based compensator for distributed delays,” *Automatica*, vol. 26, no. 5, pp. 903–908, 1990.

[33] O. Beldiman, G. Walsh, and L. Bushnell, “Predictors for networked control systems,” in *Proceedings of the American Control Conference*, vol. 4, 2000, pp. 2347–2351.

[34] N. Elia and S. K. Mitter, “Stabilization of linear systems with limited information,” *IEEE Transactions on Automatic Control*, vol. 46, no. 9, p. 13841400, 2001.

[35] W. S. Wong and R. W. Brockett, “Systems with finite communication bandwidth constraints-II: Stabilization with limited information feedback,” *IEEE Transactions on Automatic Control*, vol. 44, no. 5, 1999.

[36] A. Khaladj, M. Rahgozar, and N. Yazdani, “Reducing the variations in delay in IEEE 802.11 DCF,” in *Proc. of IEEE APCC*, 2005, pp. 755–758.

[37] F. Stann and J. Heidemann, “RMST: Reliable Data Transport in Sensor Networks,” in *IEEE SNPA*, 2003.

[38] E. Felemban, C. G. Lee, and E. Eylem, “MMSPEED: Multipath multi-speed protocol for QoS guarantee of reliability and timeliness in wireless sensor networks,” *IEEE Transactions on Mobile Computing*, vol. 5, no. 6, pp. 738–754, 2006.

[39] A. Sangiovanni-Vincentelli, “Is a unified methodology for system-level design possible?” *IEEE Design and Test of Computers*, vol. 25, no. 4, 2008.

[40] A. Sangiovanni-Vincentelli and G. Martin, “Defining platform-based design,” *EEDesign of EETimes*, 2002.

[41] K. Keutzer, S. Malik, A. R. Newton, J. M. Rabaey, and A. Sangiovanni-Vincentelli, “System-level design: orthogonalization of concerns and platform-based design,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, pp. 1523–1543, 2000.

[42] A. Sangiovanni-Vincentelli and G. Martin, “Platformbased design and software design methodology for embedded systems,” *IEEE Design and Test of Computers*, pp. 23–33, 2001.

[43] J. A. Leijten, J. L. V. Meerbergen, A. H. Timmer, and J. A. G. Jess, “Prophid: a platformbased design method,” *Design Automation for Embedded Systems*, pp. 5–37, 2000.

[44] M. Kawarabayashi, J.-Q. Lu, K. Goto, and P. W. Fung, “System level design methodology for system on a chip,” *NEC Research and Development*, pp. 248–252, 2000.

[45] C. Intanagonwiwat and R. Govindan and D. Estrin, "Directed diffusion: A scalable and robust communication paradigm for sensor networks," in *ACM MobiCOM*, 2000.

[46] P. Buonadonna, D. Gay, J. Hellerstein, W. Hong, and S. Madden, "TASK: Sensor network in a box," Intel Research Lab Report, Tech. Rep., 2007.

[47] R. Steigman, and J. Endresen, "Introduction to WISA and WPS, WISA-wireless interface for sensors and actuators and WPS-wireless proximity switches," *White paper*, 2004. [Online]. Available: http://www.eit.uni-kl.de/litz/WISA.pdf

[48] *The ZigBee Alliance* . [Online]. Available: http://www.zigbee.org

[49] *ISA SP-100*. [Online]. Available: http://www.isa.org/

[50] *Wirelesshart data sheet*, IEEE, 2007. [Online]. Available: http://www.hartcomm2.org/hartprotocol/wirelesshart/wirelesshartmain.html

[51] *IPv6 over Low power WPAN (6lowpan)*. [Online]. Available: http://www.ietf.org/dyn/wg/charter/6lowpan-charter.html

[52] *Routing Over Low power and Lossy networks*, Internet Engineering Task Force (IETF). [Online]. Available: http://www.ietf.org/dyn/wg/charter/roll-charter.html

[53] S. Lee, K. C. Lee, M. H. Lee, and F. Harashima, "Integration of mobile vehicles for automated material handling using profibus and IEEE 802.11 networks," *IEEE Trans. Ind. Electron.*, vol. 49, no. 3, 2002.

[54] R. Moraes, F. Vasques, P. Portugal, and J. A. Fonseca, "VTP-CSMA: A virtual token passing approach for real-time communication in IEEE 802.11 wireless networks," *IEEE Trans. Ind. Informat.*, vol. 3, no. 3, 2007.

[55] *International Standard ISO/IEC 8802-2:1998: Information TechnologyTelecommunications and Information Exchange Between SystemsLocal and Metropolitan Area NetworksSpecific RequirementsPart 2: Logical Link Control*. [Online]. Available: http://grouper.ieee.org/groups/802/2/

[56] S. M. Amin and B. F. Wollenberg, "Toward a smart grid," *IEEE Power and Energy Magazine*, vol. 3, no. 5, pp. 34–38, 2005.

[57] Y. Xu, J. Heidemann, and D. Estrin, "Geography-informed energy conservation for ad hoc routing," in *ACM MobiCom*, vol. pp. 70-84, 2001.

[58] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris, "Span: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks," in *ACM MobiCom*, 2001.

[59] E. A. M. Buettner, G. Yee and R. Han, "X-MAC: A short preamble mac protocol for duty-cycled wireless sensor networks," in *ACM SenSys*, 2006.

[60] S. Kim, R. Fonseca, P. Dutta, A. Tavakoli, D. Culler, P. Levis, S. Shenker, and I. Stoica, "Flush: a reliable bulk transport protocol for multihop wireless networks," in *ACM SenSys*, 2007.

[61] M. Zorzi and R. R. Rao, "Energy and latency performance of geographic random forwarding for ad hoc and sensor networks," in *IEEE WCNC*, 2003.

[62] P. Park, C. Fischione, A. Bonivento, K. H. Johansson, and A. Sangiovanni-Vincentelli, "Breath: an adaptive protocol for industrial control applicationa using wireless sensor networks," *IEEE Transactions on Mobile Computing*, 2009, submitted.

[63] O. B. Akan and F. Akyildiz, "Event-to-sink reliable transport in wireless sensor networks," *IEEE Transactions on Networking*, vol. 13, no. 5, pp. 1003–1016, 2005.

[64] A. Bonivento, C. Fischione, L. Necchi, F. Pianegiani, A. Sangiovanni-Vincentelli, "System level design for clustered wireless sensor networks," *IEEE Transactions on Industrial Informatics*, 2007.

[65] J. Misic, S. Shafi, and V. Misic, "Cross-layer activity management in an 802.15.4 sensor network," *IEEE Communications Magazine*, vol. 44, no. 1, pp. 131–136, 2006.

[66] L. van Hoesel, T. Nieberg, J. Wu, and P. Havinga, "Prolonging the lifetime of wireless sensor networks by cross-layer interaction," *IEEE Wireless Communications*, vol. 11, no. 6, pp. 78–86, 2004.

[67] J. Van Greuen, D. Petrovic, A. Bonivento, J. Rabaey, K. Ramchandran, and A. Sangiovanni-Vincentelli, "Adaptive sleep discipline for energy conservation and robustness in dense sensor networks," in *IEEE ICC*, 2004.

[68] R. Cristescu, B. Beferull-Lozano, M. Vetterli, and R. Wattenhofer, "Network correlated data gathering with explicit communication: NP-completeness and algorithms," *ACM/IEEE Transactions on Networking*, vol. 14, no. 1, 2006.

[69] P. Chen and S. Sastry, "Latency and connectivity analysis tools for wireless mesh networks," in *ACM/IEEE ROBOCOMM*, 2007.

Part II

Paper A

Adaptive IEEE 802.15.4 Protocol for Reliable and Timely Communications

P. Park, P. Di Marco, C. Fischione and K. H. Johansson

Submitted to
IEEE/ACM Transactions on Networking

Adaptive IEEE 802.15.4 Protocol for Reliable and Timely Communications

P. Park, P. Di Marco, C. Fischione and K. H. Johansson

Abstract

The IEEE 802.15.4 standard for wireless sensor networks can support energy efficient, reliable, and timely packet transmission by tuning the medium access control parameters macMinBE, macMaxCSMABackoffs*, and* macMaxFrameRetries*. Such a tuning is difficult, because simple and accurate models of the relations of these parameters on the probability of successful packet transmission, packet delay and energy consumption are not available. Moreover, it is not clear how to adapt the parameters to the changes of the network and traffic regimes by simple algorithms that can run on resource-constrained nodes. In this paper, a generalized Markov chain is proposed to model these relations by simple expressions without giving up the accuracy. In contrast to previous work, the presence of limited number of retransmissions, acknowledgments, unsaturated traffic and packet size is accounted for. The analysis is then used to propose an adaptive algorithm for minimizing the power consumption while guaranteeing reliability and delay constraints in the packet transmission. The algorithm does not require any modification of the IEEE 802.15.4 standard and can be easily implemented one existing network nodes. Numerical results show that the analysis is accurate, that the proposed algorithm satisfies reliability and delay constraints, and ensures a longer lifetime of the network under both stationary and transient network conditions.*

Index Terms–IEEE 802.15.4 standard, Markov chain model, Optimization.

1 Introduction

The IEEE 802.15.4 standard has received considerable attention as a major low data rate and low power protocol for wireless sensor network (WSN) applications in industry, control, home automation, health care, and smart grids [1, 2]. Many of these applications require that packets are received with a given probability of success. In addition to such a reliability constraint, other applications ask for timely packet delivery [3]. It is known that IEEE 802.15.4 may have poor performance in terms of power consumption, reliability and delay [4], unless the medium access control (MAC) parameters are properly selected. It follows that 1) it is essential to characterize the performance to understand the protocol limitations, and 2) it is instrumental to tune the IEEE 802.15.4 parameters to enhance the network lifetime and improve the quality of the service experienced by the applications running on top of the network.

This paper focuses on the modelling and optimization of the performance metrics (reliability, delay, power consumption) for IEEE 802.15.4 WSNs. We show that existing analytical studies of IEEE 802.15.4 are not adequate to capture the real-world protocol behavior, where there are retry limits to send packets, acknowledgement (ACK), and unsaturated traffic. We use this modelling to pose a novel optimization problem where the objective function is the power consumption of the network, subject to reliability and delay constraints of the packet delivery. Our aim is the design of distributed and adaptive algorithms that are simple to implement on sensor nodes, flexible, scalable, and able to provide high quality of service for WSNs applications.

The remainder of this paper is as follows. In Section 2, we summarize existing work for the analytical model and adaptive tuning of IEEE 802.15.4. Section 3 presents the main contributions of the paper. In Section 4, we propose a generalized Markov chain model of CSMA/CA with retry limits and unsaturated traffic regime. The optimization problem to adapt the MAC parameters is investigated in Setion 5. Practical issues on how to implement the algorithm on sensors are discussed in Section 6. Numerical results achieved during stationary and transitionary conditions are reported in Section 7. Finally, Section 8 concludes the paper.

2 Related Work

The modelling of IEEE 802.15.4 is related to IEEE 802.11 [5]. We first discuss the literature concerning the analysis of IEEE 802.11 and 802.15.4, then we review previous work about adaptive MAC mechanisms for these protocols.

2.1 Analytical Model of MAC

Both IEEE 802.11 and 802.15.4 are based on a MAC that uses a binary exponential backoff scheme. Bianchi's model describes the basic functionalities of the IEEE 802.11 through a Markov chain under saturated traffic and ideal channel conditions [6]. Extensions of this model have been used to analyze the packet reception rate [7], the delay [8], the MAC layer service time [9] and throughput [10] of IEEE 802.11.

The analysis of the packet delay, throughput, and power consumption of IEEE 802.15.4 WSNs has been the focus of several simulations-based studies

e.g., [11], [12], and some more recent analytical works, e.g. [4], [13]– [16]. Inspired by Bianchi's work, a Markov model for IEEE 802.15.4 and an extension with ACK mechanism have been proposed in [4] and [13]. A modified Markov model including retransmissions with finite retry limits has been studied in [15] as an attempt to model the slotted carrier sense multiple access with collision avoidance (CSMA/CA) mechanism. However, the analysis gives inaccurate results because the power consumption and throughput expressions under unsaturated traffic with finite retry limits show a weak matching with simulation results.
In [16], a throughput analysis has been performed by an extension of the Markov chain model proposed [14]. The superframe structure, ACK, and retransmissions are considered. However, the proposed Markov chain does not model the length of data and ACK packets, which is crucial to analyze the performance metrics for IEEE 802.15.4 networks with low data rate. Furthermore, in [14] the power consumption, reliability, and delay performance are not investigated. We remark here that all analytical models available from the literature use numerical methods to solve nonlinear equations [4], [13]– [16], which is a major drawback for in-network processing [17].

2.2 Adaptive Tuning of MAC

Several algorithms to tune the MAC of IEEE 802.11 and IEEE 802.15.4 protocols have been proposed. The algorithms can be grouped in those based on the use of physical layer measurements, and those based on the use of link-layer information.
An adaptive tuning based on physical layer measurements has been investigated in [18]– [20], where a p-persistent IEEE 802.11 protocol has been considered to optimize the average backoff window size. The channel access probability p value that maximizes the throughput or minimize the power consumption is derived. This algorithm and its scalability to the network size have been applied also to IEEE 802.15.4 [19]. However, such an approach is not suitable in IEEE 802.15.4, because the channel sensing mechanism, the optional acknowledgement (ACK), and retransmission mechanisms are hard to be approximated by a p-persistent MAC. Furthermore, in [19] and [20] a saturated traffic regime is assumed, which is a scenario of reduced interest for typical WSNs applications.
Link-based optimizations for IEEE 802.11 and 802.15.4 have been investigated in [21]– [25], where simple window adjustment mechanisms that are based on ACK transmissions have been considered. In these papers, the algorithms adapt the contention window size depending on the successful packet transmission, packet collision and channel sensing state, but the algorithms are not grounded on an analytical study. In [21], different backoff algorithms are presented to improve the channel throughput and the fairness of channel usage for IEEE 802.11. A fair backoff algorithm is also studied in [22, 23]. A link-based algorithm of the IEEE 802.15.4 random backoff mechanism to maximize the throughput has been presented in [24]. In [25], a dynamic tuning algorithm of the contention window size is evaluated on goodput, reliability, and average delay.

An IEEE 802.15.4 enhancement based on the use of link-layer information has some drawback. First, it requires a modification of the standard. Then, although link-based mechanisms are simple to implement, the ACK mechanism use may be a costly solution since it introduces large overhead for small size packets. For instance, alarm messages in industrial control application have just 1 byte size whereas the ACK has a size of 11 bytes and the ACK mechanism requires the extra waiting time to receive it.
. Moreover, link-based algorithms adapt the MAC parameters for each received ACK, which mean a slow (inefficient) adaptation to network, traffic, and channel changes.

3 Original Contribution

We consider a star network with a personal area network (PAN) coordinator, and N nodes with beacon enabled slotted CSMA/CA and ACK. The important parameters of CSMA/CA algorithm are the minimum value of the backoff exponent *macMinBE*, the maximum value of the backoff exponent *macMaxBE*, the maximum number of backoffs *macMaxCSMABackoffs* and the maximum number of retries *macMaxFrameRetries*, see details of IEEE 802.15.4 in [1, 26].
In this paper, we propose a novel modelling and adaptive tuning of IEEE 802.15.4 for reliable and timely communication while minimizing the energy consumption. The protocol is optimized dynamically by a constrained optimization problem. The objective function, denoted by $E_{\text{tot}}(\mathbf{V})$, is the total energy consumption for transmitting and receiving packets of the node. The constraints are given by the probability of successful packet delivery (reliability) and average delay. The constrained optimization problem is

$$\min_{\mathbf{V}} E_{\text{tot}}(\mathbf{V}) \tag{1a}$$

$$\text{s.t. } R(\mathbf{V}) \geq R_{\min}\,, \tag{1b}$$

$$D(\mathbf{V}) \leq D_{\max}\,, \tag{1c}$$

$$\mathbf{V}_0 \leq \mathbf{V} \leq \mathbf{V}_m\,. \tag{1d}$$

The decision variables $\mathbf{V} = (m_0, m, n)$ are $m_0 \triangleq$ *macMinBE*, $m \triangleq$ *macMaxCSMABackoffs*, $n \triangleq$ *macMaxFrameRetries* of standard. $R(\mathbf{V})$ is the reliability, and $R_{\min}$ is the minimum desired probability. $D(\mathbf{V})$ is the average delay for a successfully received packet, and $D_{\max}$ is the desired maximum average delay. The constraint $\mathbf{V}_0 \leq \mathbf{V} \leq \mathbf{V}_m$ is due to the limited range of the MAC parameters.
Main contributions are the following: 1) the modelling of the relation between the MAC parameters of IEEE 802.15.4 and the selected performance metrics, 2) the modelling of simple approximate relations to characterize the operations of the MAC by computationally affordable algorithms, 3) formulation and solution of a novel optimization problem for the MAC parameters, 4) discussion on a practical

implementation of the optimization by an adaptive algorithm and 5) performance evaluations of the algorithm by simulation of both stationary and transient conditions of the network.

Unlike previous work, we propose a generalized Markov model of the exponential backoff process including retry limits, acknowledgements and unsaturated traffic regime. However, the numerical evaluation of these performance metrics asks in general for heavy computations. This is a drawback when using them to optimize the IEEE 802.15.4 MAC parameters by in-network processing [17], because a complex computation is out of reach for limited sensing devices. Therefore, we devise a simplified and effective method that reduces drastically the computational complexity while ensuring a satisfactory accuracy.

Based on our novel modelling, we propose an adaptive tuning of MAC parameters that uses the physical layer measurement of channel sensing. The adaptive IEEE 802.15.4 is furnished with two distinctive features: it does not require any modification of the existing standard, and it makes a global optimization of the MAC parameters. Specifically, in contrast to link-based adaptation [21]– [25], our algorithm does not require ACK mechanism and RTS/CTS handshakes (and related standard modification). In contrast to [18]– [20], we do not use the inaccurate p-persistent approximation and the modification of the standard therein proposed, and we do not require any hardware modification to make an estimate of the signal to noise ratio. Our adaptive tuning optimizes the available MAC parameters all at once, and not some of them, as proposed in [18]– [25], and the literatures therein cited.

The proposed adaptive IEEE 802.15.4 improves the power efficiency substantially while guaranteeing reliability and delay constraints. The adaptation is achieved by distributed asynchronous iterations that require simple information of channel condition, the number of nodes of the network and the traffic load. The convergence is fast and robust to errors in the estimation process of the channel condition, number of nodes and traffic load. A good fairness is also achieved.

4 Analytical Modelling of IEEE 802.15.4

In a star network, all nodes N contend to send data to the PAN coordinator, which is the data sink. Throughout this paper we consider applications where nodes asynchronously generate packets with probability $1-q$. We model an unsaturated traffic by assuming that a node stays for $L_0 S_b$ s without generating packets, where L_0 is an integer number and S_b is the time unit *aUnitBackoffPeriod* (corresponding to 20 symbols). The data packet transmission is successful if an ACK packet is received.

In such a scenario, we propose an effective analytical model of the slotted CSMA/CA by a Markov chain. The chain gives us the objective function, energy (1a), and constraints on reliability (1b) and delay (1c) of the optimization problem (1). Monte Carlo simulations validate the proposed modelling.

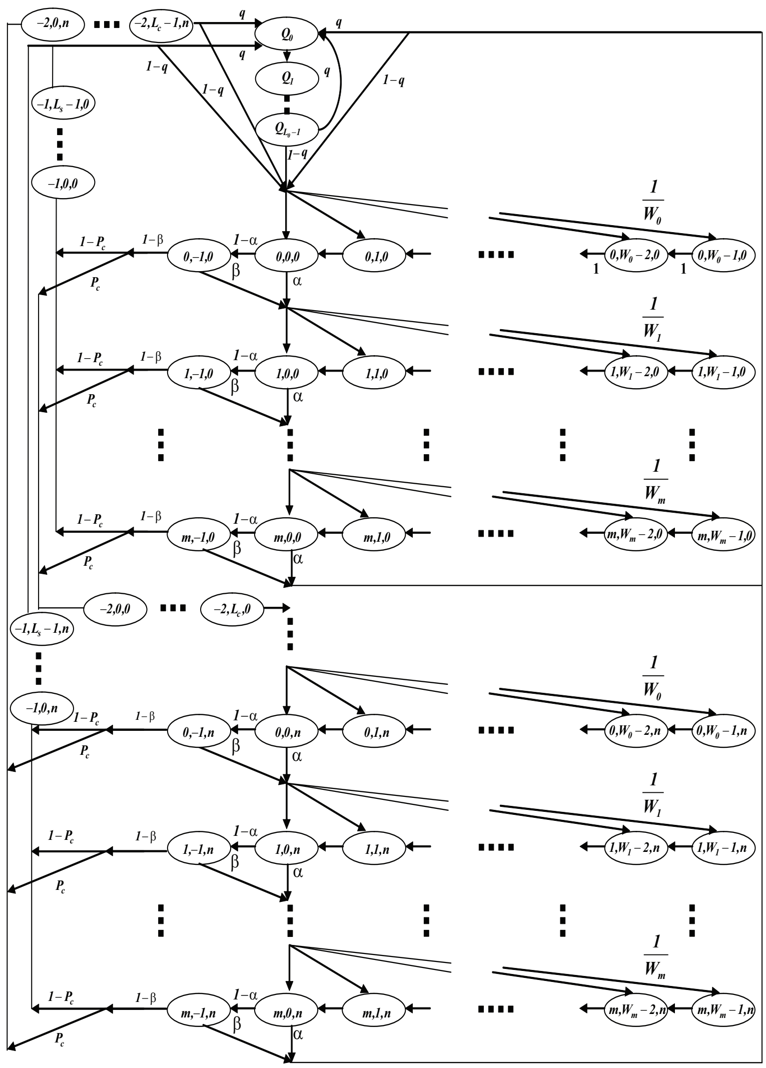


Figure 1: Markov chain model for CSMA/CA algorithm for IEEE 802.15.4

4.1 Markov Chain Model

In this section, we develop a generalized Markov chain model of the slotted CSMA/CA mechanism of beacon enabled IEEE 802.15.4. Compared to previous results, e.g., [4], [13]– [16], the novelty of this chain consists in the modelling of the retry limits for each packet transmission, the inclusion of unsaturated traffic regimes, and packet size.

Let $s(t), c(t)$ and $r(t)$ be the stochastic processes representing the backoff stage, the state of the backoff counter and the state of retransmission counter at time t experienced by a node to transmit a packet, as summarized in Fig. 1. By assuming independent probability that nodes start sensing, the stationary probability τ that a node attempts a first carrier sensing in a randomly chosen slot time is constant and independent of other nodes, and the tuple $(s(t), c(t), r(t)$ is a three dimensional Markov chain. Recall that we denote the MAC parameters by $W_0 \triangleq 2^{\mathit{macMinBE}}, m_0 \triangleq \mathit{macMinBE}, m_b \triangleq \mathit{macMaxBE}, m \triangleq \mathit{macMaxCSMABackoffs}, n \triangleq \mathit{macMaxFrameRetries}$. The states from $(i, W_m - 1, j)$ to $(i, W_0 - 1, j)$ represent the backoff states. States $(Q_0, \dots, Q_{L_0-1})$ consider the idle state when the packet queue is empty and the node is waiting for new packet arrivals. Note that the idle states $(Q_0, \dots, Q_{L_0-1})$ take into account the unsaturated traffic regime. States $(i, 0, j)$ and $(i, -1, j)$ represent CCA1 and CCA2, respectively. By knowing the duration of an ACK frame, ACK timeout, IFS, data packet length, and header duration, we define the packet successful transmission time L_s and the packet collision time L_c as

$$\begin{aligned} L_s &= L + t_{\text{ack}} + L_{\text{ack}} + IFS\,, \\ L_c &= L + t_{\text{m,ack}}, \end{aligned} \qquad (2)$$

where L is the total length of packet including overhead and payload, t_{ack} is ACK waiting time, L_{ack} is the length of ACK frame, IFS is the Inter-Frame Spacing and $t_{\text{m,ack}}$ is the timeout of the ACK, see the details [1,26]. States $(-1, k, j)$ and $(-2, k, j)$ consider the successful transmission and packet collision. Let α be the probability that CCA1 is busy, and β the probability that CCA2 is busy. We have the following results:

LEMMA 1 *Let* $b_{i,k,j} = \lim_{t\to\infty} P(s(t) = i, c(t) = k, r(t) = j), i \in (-2, m), k \in (-1, \max(W_i - 1, L_s - 1, L_c - 1)), j \in (0, n)$, *be the state stationary probability of the Markov chain in Fig. 1. Then, for* $i \leq m$

$$b_{i,k,j} = \frac{W_i - k}{W_i} b_{i,0,j}\,, \qquad k > 0 \qquad (3)$$

where

$$W_i = \begin{cases} 2^i W_0 & i \leq m_b - m_0\,, \\ 2^{m_b - m_0} W_0 & i > m_b - m_0\,, \end{cases}$$

$$
b_{0,0,0} = \begin{cases}
\Big[\frac{1}{2}\Big(\frac{1-(2x)^{m+1}}{1-2x}W_0 + \frac{1-x^{m+1}}{1-x}\Big)\frac{1-y^{n+1}}{1-y} + (1-\alpha)\frac{(1-x^{m+1})(1-y^{n+1})}{(1-x)(1-y)} \\
+(L_s(1-P_c)+L_cP_c)(1-x^{m+1})\frac{1-y^{n+1}}{1-y} + L_0\frac{q}{1-q}\Big(\frac{x^{m+1}(1-y^{n+1})}{1-y} \\
+P_c(1-x^{m+1})y^n + (1-P_c)\frac{(1-x^{m+1})(1-y^{n+1})}{1-y}\Big)\Big]^{-1} \\
\text{if} \quad m \le m_b - m_0 \\
\\
\Big[\frac{1}{2}\Big(\frac{1-(2x)^{m_b-m_0+1}}{1-2x}W_0 + \frac{1-x^{m_b-m_0+1}}{1-x} + (2^{m_b}+1)x^{m_b-m_0+1} \\
\times\frac{1-x^{m-m_b+m_0}}{1-x}\Big)\frac{1-y^{n+1}}{1-y} + (1-\alpha)\frac{(1-x^{m+1})(1-y^{n+1})}{(1-x)(1-y)} + (L_s(1-P_c) \\
+L_cP_c)(1-x^{m+1})\frac{1-y^{n+1}}{1-y} + L_0\frac{q}{1-q}\Big(\frac{x^{m+1}(1-y^{n+1})}{1-y} + P_c \\
\times(1-x^{m+1})y^n + (1-P_c)\frac{(1-x^{m+1})(1-y^{n+1})}{1-y}\Big)\Big]^{-1} \\
\text{otherwise}
\end{cases} \tag{5}
$$

and

$$
b_{i,0,j} = \left[(1-\alpha)(1-\beta)P_c\sum_{i=0}^{m}(\alpha+(1-\alpha)\beta)^i\right]^j (\alpha+(1-\alpha)\beta)^i b_{0,0,0}\,, \tag{4}
$$

and where $b_{0,0,0}$ *is given in Eq.* (5).

Proof: See Appendix A.1. ∎

We remark here that the term $b_{0,0,0}$, which plays a key role in the analysis, is different from the corresponding term given in [4], [13]– [16] due to our accurate modelling of the retransmissions, unsaturated traffic, and packet size. In the next section, we demonstrate the validity of the Markov chain model by Monte Carlo simulations.

Now, starting from the previous Lemma, we derive the busy channel probabilities α and β and the channel sensing probability τ. The probability τ that a node attempts a first carrier sensing (CCA1) in a randomly chosen time slot is

$$
\tau = \sum_{i=0}^{m}\sum_{j=0}^{n} b_{i,0,j} = \frac{1-x^{m+1}}{1-x}\frac{1-y^{n+1}}{1-y}b_{0,0,0}. \tag{6}
$$

This probability depends on the probability P_c that a transmitted packet encounters a collision and the probabilities α and β, which we give in the following.

The term P_c is the probability that at least one of the $N-1$ remaining nodes transmit in the same time slot. If all nodes transmit with probability τ, P_c is

$$
P_c = 1-(1-\tau)^{N-1}\,,
$$

where N is the number of nodes. Similarly to [4], we derive the busy channel probabilities α and β as follows. Since

$$\alpha = \alpha_1 + \alpha_2 \,, \tag{7}$$

where α_1 is the probability of finding channel busy during first CCA due to data transmission, namely

$$\alpha_1 = L(1-(1-\tau)^{N-1})(1-\alpha)(1-\beta)\,,$$

and α_2 is the probability of finding the channel busy during first CCA due to ACK transmission, which is

$$\alpha_2 = L_{\text{ack}}\frac{N\tau(1-\tau)^{N-1}}{1-(1-\tau)^N}(1-(1-\tau)^{N-1})(1-\alpha)(1-\beta)\,,$$

where L_{ack} is the length of the ACK. Finally,

$$\beta = \frac{1-(1-\tau)^{N-1}+N\tau(1-\tau)^{N-1}}{2-(1-\tau)^N+N\tau(1-\tau)^{N-1}}\,. \tag{8}$$

Now, we are in the position to derive the carrier sensing probability τ and the busy channel probabilities α and β by solving the system of non-linear equations (6), (7), and (8) for these probabilities. From these probabilities then one could derive the expressions of the reliability, delay for successful packet delivery, and power consumption that are needed in (1). The drawback of such an approach is that there is no closed form expression for these probabilities, and the system of equations that gives τ, α and β must be solved by numerical methods. This may be computationally demanding and inadequate for use in simple sensor devices. Therefore, in the following, we present a simple analytical model of the reliability, delay for successful packet delivery, and power consumption. The key idea is that sensor nodes can estimate the busy channel probabilities α and β and the channel sensing probability τ. Therefore, nodes exploit local measurements to evaluate the performance metrics, rather than solving nonlinear equations. Details follow in the sequel.

4.2 Reliability

The main contributions of this section are the derivation of both precise and approximated expression of the reliability (1b) of the optimization problem (1). Recall that the reliability is the probability of successful packet reception.

PROPOSITION 1 *The reliability is*

$$R(\mathbf{V}) = 1-\frac{x^{m+1}(1-y^{n+1})}{1-y}-y^{n+1}\,. \tag{9}$$

where $x=\alpha+(1-\alpha)\beta$, *and* $y=P_c(1-x^{m+1})$.

Proof: In slotted CSMA/CA, packets are unsuccessfully received due to two reasons: channel access failure and retry limits. Channel access failure happens when a packet fails to obtain idle channel in two consecutive CCAs within $m + 1$ backoffs. Furthermore, a packet is discarded if the transmission fails due to repeated collisions after $n+1$ attempts. Following the Markov model presented in Fig. 1, the probability that the packet is discarded due to channel access failure is

$$P_{cf} = \sum_{j=0}^{n} x\, b_{m,0,j} = \frac{x^{m+1}(1-y^{n+1})}{1-y} \,. \tag{10}$$

The probability of a packet being discarded due to retry limits, here denoted by P_{cr}, is

$$P_{cr} = \sum_{i=0}^{m} P_c(1-\beta) b_{i,-1,n} = y^{n+1} \,. \tag{11}$$

By putting Eq. (10) and (11), the reliability is given by

$$R(\mathbf{V}) = 1 - P_{cf} - P_{cr} \,,$$

from which the proposition follows. ■

CLAIM 1 *An approximation of the reliability is*

$$\widetilde{R}(\mathbf{V}) = 1 - x^{m+1}(1+\widetilde{y}) - \widetilde{y}^{n+1} \tag{12}$$

where

$$\begin{aligned}\widetilde{y} =& (1-(1-(1+x)(1+\hat{y})\widetilde{b}_{0,0,0})^{N-1})(1-x^2)\,,\\ \widetilde{b}_{0,0,0} =& 2/(W_0(1+2x)(1+\hat{y}) + 2L_s(1-x^2)(1+\hat{y})\\ &+ L_0 q/(1-q)(1+\hat{y}^2+\hat{y}^{n+1}))\,,\end{aligned}$$

and $\hat{y} = (1-(1-\tau)^{N-1})(1-x^2)$.

Proof: The expression of the state probability $b_{0,0,0}$ is the main responsible for the non-linear equations that give α, β and τ. Therefore, we approximate $b_{0,0,0}$. Let the approximation be $\widetilde{b}_{0,0,0}$. Given $z \geq 0$, we use that

$$\frac{1-z^{m+1}}{1-z} \approx 1 + z \qquad \text{if} \quad z \ll 1 \tag{13}$$

By using this approximation, Eq. (40) is approximated by

$$\sum_{i=0}^{m} \sum_{k=0}^{W_i-1} \sum_{j=0}^{n} b_{i,k,j} \approx \frac{b_{0,0,0}}{2} \left[(1+2x)W_0 + 1 + x\right](1+y) \tag{14}$$

Similarly, Eq. (41) is approximated by

$$\sum_{i=0}^{m}\sum_{j=0}^{n} b_{i,-1,j} \approx b_{0,0,0}(1-\alpha)(1+x)(1+y). \tag{15}$$

and Eq. (42) is approximated by

$$\sum_{j=0}^{n}\left(\sum_{k=0}^{L_s-1} b_{-1,k,j} + \sum_{k=0}^{L_c-1} b_{-2,k,j}\right) \approx b_{0,0,0}L_s(1-x^{m+1})(1+y), \tag{16}$$

where we assume that the packet collision time is approximated to the packet successful transmission time, namely $L_s \approx L_c$. Finally, using $K_0 = L_0 q/(1-q)$, the approximated idle stages of Eq. (43) is

$$\sum_{l=0}^{L_0-1} Q_l \approx b_{0,0,0}K_0\left[1+y+P_c(1-x^{m+1})(y^n-y-1)\right]. \tag{17}$$

By summing together Eqs. (14)–(17), the approximated state probability is

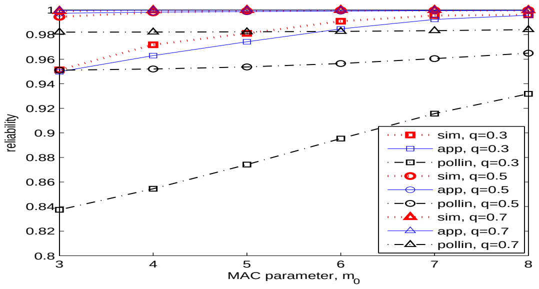
$$\widetilde{b}_{0,0,0} \approx \frac{2}{W_0 r_1 + 2r_2} \tag{18}$$

where $r_1 = (1+2x)(1+\hat{y})$, $r_2 = L_s(1-x^2)(1+\hat{y}) + K_0(1+\hat{y}^2+\hat{y}^{n+1})$, $\hat{y} = (1-(1-\tau)^{N-1})(1-x^2)$, and we neglect the term in Eq. (15) and use $1-x^{m+1} \approx 1-x^2$.
Now, we put the approximated state probability $\widetilde{b}_{0,0,0}$ into Eq. (9) to obtain the approximated reliability $\widetilde{R}(\mathbf{V})$ as follow

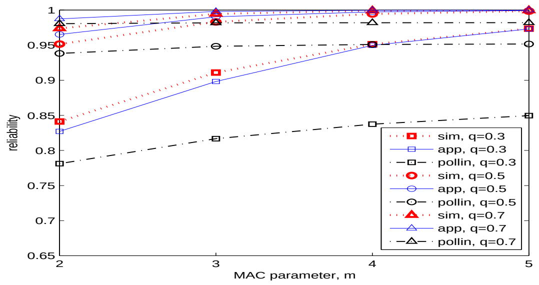
$$\widetilde{R}(\mathbf{V}) = 1 - x^{m+1}(1+\widetilde{y}) - \widetilde{y}^{n+1},$$

where $\widetilde{y} = (1-(1-\widetilde{\tau})^{N-1})(1-x^2)$ and the approximated carrier sensing probability $\widetilde{\tau} = (1+x)(1+\hat{y})\widetilde{b}_{0,0,0}$. ■

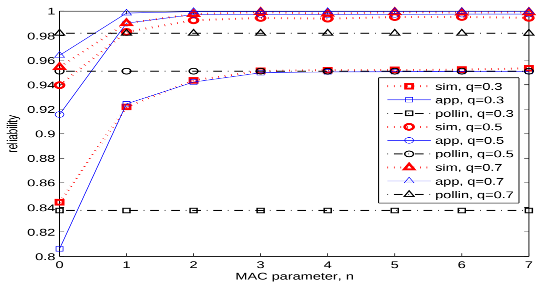
We remark that $\widetilde{R}(\mathbf{V})$ is a function of the measurable busy channel probabilities α and β, the channel access probability τ and the MAC parameters m_0, m_b, m, n. The approximation uses x and τ that are estimated.
We use Monte Carlo simulations to validate the approximated model of the reliability given by Eq. (12). The simulations are based on the specifications of the IEEE 802.15.4 [1] with several values of the traffic regime and MAC parameters. Fig. 2 compares Eq. (12), the analytical model in [4], and Monte Carlo simulations as a function of the traffic regimes $q = 0.3, 0.5, 0.7$ with $N = 10$ nodes and different MAC parameters m_0, m, n. In the figure, note that "Pollin" refers to the reliability model derived in [4]. Our analytical expression matches quite well the simulation results. The expression is closer to simulation results under low traffic



(a) $m_0 = 3, \ldots, 8$, $m_b = 8$, $m = 4$, $n = 3$



(b) $m = 2, \ldots, 5$, $m_0 = 3$, $m_b = 8$, $n = 3$



(c) $n = 0, \ldots, 7$, $m_0 = 3$, $m_b = 8$, $m = 4$

Figure 2: Reliability as a function of the traffic regimes $q = 0.3, 0.5, 0.7$, and MAC parameters $m_0 = 3, \ldots, 8$, $m_b = 8$, $m = 2, \ldots, 5$, $n = 0, \ldots, 7$ with Pollin's model [4]. The length of the packet is $L = 3$ and the number of nodes is $N = 20$.

regime $q = 0.5, 0.7$ than high traffic regime $q = 0.3$ because the approximation given by Eq. (13) holds if $x = \alpha + (1-\alpha)\beta \ll 1$, but x increases as the traffic and the number of nodes increases. The reliability approaches 1 under very low traffic regime $q = 0.7$. In Fig. 2(a), 2(b), the reliability increases as MAC parameters m_0, m increase, respectively. In Fig. 2(c), we observe that the improvement of reliability is small as the retry limits n increases if $n \geq 3$. Notice that the reliability saturates to 0.95 for traffic regime $q = 0.3$ if $n \geq 3$. Hence, the retransmissions are necessary but not sufficient to obtain high reliability under high traffic regimes.

4.3 Delay

In this section, we derive the constraint of average delay (1c) of the optimization problem (1). The average delay for a successfully received packet is defined as the time interval from the instant the packet is at the head of its MAC queue and ready to be transmitted, until the transmission is successful and the ACK is received. In this section, we develop an approximation for such an average delay, which is given by Claim 2. In order to prove such an important result, we need some intermediate technical steps. In particular, we need to characterize 1) the exact expression of the delay for a successful transmission at time $j + 1$ after jth events of unsuccessful transmission due to collision and 2) the expected value of the approximated backoff delay due to busy channel. We address these issues in the following.

Let D_j be the random time associated to the successful transmission of a packet at the jth backoff stage. Denote with $\mathcal{A}_j$ the event of a successful transmission at time $j+1$ after jth events of unsuccessful transmission. Let $\mathcal{A}_t$ be the event of successful transmission within the total attempts n. Then, the delay for a successful transmission after jth unsuccessful attempts is

$$D = \sum_{j=0}^{n} \mathbb{1}_{\mathcal{A}_j|\mathcal{A}_t} D_j \,,$$

where $D_j = L_s + j\, L_c + \sum_{h=0}^{j} T_h$, with T_h being the backoff stage delay, and recall that L_s is the packet successful transmission time and L_c is the packet collision transmission time as defined in Eq. (2).

LEMMA 2 *The probability of successful transmission at time $j+1$ after jth events of unsuccessful transmission due to collision is*

$$\Pr(\mathcal{A}_j|\mathcal{A}_t) = \frac{(1-y)\, y^j}{1-y^{n+1}} \tag{19}$$

where $y = P_c(1 - x^{m+1})$.

Proof: See Appendix A.2. ∎

In the following, we give the total backoff delay T_h. Let $T_{h,i}$ be the random time needed to obtain two successful CCAs from the selected backoff counter value in backoff level i. Recall that a node transmits the packet when the backoff counter is 0 and two successful CCAs are detected [1]. Denote with $\mathcal{B}_i$ the event occurring when the channel is busy for i times, and then idle at the $i+1$th time. Let $\mathcal{B}_t$ be the event of having a successful sensing within the total number of m sensing attempts. If the node accesses an idle channel after its i th busy CCA, then

$$T_h = \sum_{i=0}^{m} \mathbb{1}_{\mathcal{B}_i|\mathcal{B}_t} T_{h,i} \,,$$

where

$$T_{h,i} = 2\,T_{sc} + \sum_{k=1}^{i} T_{h,k}^{sc} + \sum_{k=0}^{i} T_{h,k}^{b} \,, \tag{20}$$

and where $2T_{sc}$ is the successful sensing time, $\sum_{k=1}^{i} T_{h,k}^{sc}$ is the unsuccessful sensing time due to busy channel during CCA, and $\sum_{k=0}^{i} T_{h,k}^{b}$ is the backoff time.

LEMMA 3 *The expected value of the approximated backoff delay is*

$$\begin{aligned}\mathbb{E}[\widetilde{T}_h] =& 2S_b \left(1 + \frac{1}{4}\left(\frac{1-b_l}{1-b_l^{m+1}}\left(2W_0 \frac{1-(2b_l)^{m+1}}{1-2b_l}\right.\right.\right.\\ &\left.\left.\left. - \frac{3(m+1)b_l^{m+1}}{1-b_l}\right) + \frac{3b_l}{1-b_l} - (W_0+1)\right)\right) \,,\end{aligned} \tag{21}$$

where $b_l = \max(\alpha, (1-\alpha)\beta)$.

Proof: See Appendix A.3. ∎

Now, we are in the position to derive an approximation of the average delay for successfully received packets.

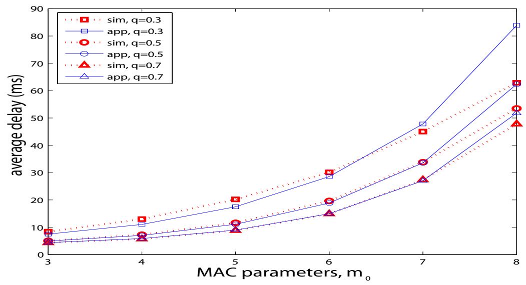
CLAIM 2 *The expected value of the approximated delay is*

$$\widetilde{D}(\mathbf{V}) = T_s + \mathbb{E}[\widetilde{T}_h] + \left(\frac{y}{1-y} - \frac{(n+1)\,y^{n+1}}{1-y^{n+1}}\right)(T_c + \mathbb{E}[\widetilde{T}_h]) \,. \tag{22}$$

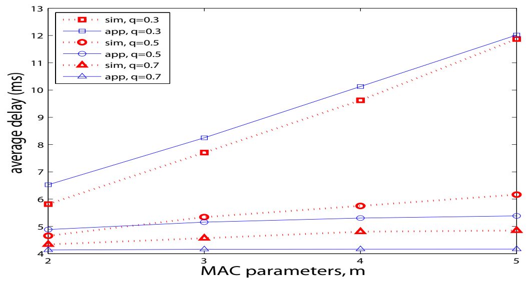
Proof: By considering the Lemma 2, we derive $\widetilde{D}(\mathbf{V})$

$$\widetilde{D}(\mathbf{V}) = \sum_{j=0}^{n} \Pr(\mathcal{A}_j|\mathcal{A}_t)\, \mathbb{E}[\widetilde{D}_j]$$

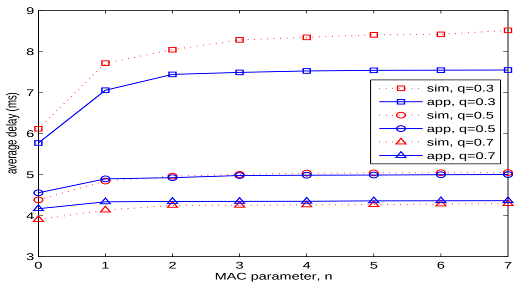
where $\mathbb{E}[\widetilde{D}_j] = T_s + j\,T_c + \sum_{h=0}^{j} \mathbb{E}[\widetilde{T}_h]$ and $\mathbb{E}[\widetilde{T}_h]$ is given in Eq. (21) from the Lemma 3. ∎



(a) $m_0 = 3, \ldots, 8$, $m_b = 8$, $m = 4$, $n = 3$



(b) $m = 2, \ldots, 5$, $m_0 = 3$, $m_b = 8$, $n = 3$



(c) $n = 0, \ldots, 7$, $m_0 = 3$, $m_b = 8$, $m = 4$

Figure 3: Average delay as a function of the traffic regimes $q = 0.3, 0.5, 0.7$ and MAC parameters $m_0 = 3, \ldots, 8$, $m_b = 8$, $m = 2, \ldots, 5$, $n = 0, \ldots, 7$. The length of the packet is $L = 3$ and the number of nodes is $N = 20$.

$$
\begin{aligned}
E_{\text{tot,i}}(\mathbf{V}) =& \frac{P_i \tau}{2}\left[\frac{(1-x)(1-(2x)^{m+1})}{(1-2x)(1-x^{m+1})}W_0 - 1\right] + P_{sc}(2-\alpha)\tau + (1-\alpha) \\
& \times (1-\beta)\tau\left(P_t L + P_i + L_{\text{ack}}\left(P_r(1-P_c) + P_i P_c\right)\right) + P_w q \left(x^{m+1}\right. \\
& \left.\times (1+y) + P_c(1-x^2)y^n + (1-P_c)(1-x^2)(1+y)\right)\widetilde{b}_{0,0,0} \quad (23) \\
E_{\text{tot,s}}(\mathbf{V}) =& P_{sc}(2-\alpha)\tau + (1-\alpha)(1-\beta)\tau\left(P_t L + P_i + L_{\text{ack}}\left(P_r(1-P_c)\right.\right. \\
& \left.\left. + P_i P_c\right)\right) + P_w\left(\tau - \frac{\widetilde{b}_{0,0,0}(1-(0.5x)^{m+1})}{W_0(1-0.5x)}\frac{1-y^{n+1}}{1-y}\right) \quad (24)
\end{aligned}
$$

Fig. 3 plots the average delay as obtained by Eq. (22) as a function of different traffic regimes $q = 0.3, 0.5, 0.7$ with a given number of nodes $N = 10$ and different MAC parameters m_0, m, n. The analytical model predict well the simulation results. The accuracy is reduced under high traffic regime $q = 0.3$ due to the approximation given by Eq. (13). Observe that the average delay increases as traffic regime increases due to high busy channel probability and collision probability. Fig. 3(a) shows that the average delay increases exponentially as m_0 increases. Hence, we conclude that m_0 is the key parameter on average delay with respect to m, n.

4.4 Power Consumption

Here, we derive the objective function, power consumption of the node (1a) of the optimization problem (1). We propose two models for the average power consumption, depending on the radio state during backoff mechanism specified by the IEEE 802.15.4 standard. Let us denote by *I-mode* and *S-mode* the situation when the radio is set in idle mode or in sleep mode during backoff period, respectively.

CLAIM 3 *The energy consumption of the I-mode* $E_{\text{tot,i}}(\mathbf{V})$ *is given by Eq.* (23) *and the S-mode* $E_{\text{tot,s}}(\mathbf{V})$ *is given by Eq.* (24) *where state probability* $\widetilde{b}_{0,0,0}$ *is given in Eq.* (18), $P_i, P_{sc}, P_{sp}, P_w, P_t, P_r$ *are the average power consumption in idle-listen, channel sensing, sleep states, wake-up state, transmit and receiving states, respectively.*

Proof: By considering the Markov chain given in Fig. 1, we see that the average power consumption of *I-mode* $E_{\text{tot,i}}(\mathbf{V})$ is

$$E_{\text{tot,i}}(\mathbf{V}) = E_{b,i} + E_{sc} + E_t + E_q + E_{w,i}\,.$$

In the following, we give the terms that concur in this power.

The idle backoff power consumption is

$$\begin{aligned} E_{b,i} =& P_i \sum_{i=0}^{m} \sum_{k=1}^{W_i-1} \sum_{j=0}^{n} b_{i,k,j} \\ =& \frac{P_i \tau}{2} \left[\frac{(1-x)(1-(2x)^{m+1})}{(1-2x)(1-x^{m+1})} W_0 - 1 \right] , \end{aligned} \tag{25}$$

where the carrier sensing probability τ is measured by each node and P_i is the average power consumption in idle-listen.
By putting together Eqs. (40), (41) and (6), the average power consumption of the sensing state is

$$E_{sc} = P_{sc} \sum_{i=0}^{m} \sum_{j=0}^{n} (b_{i,0,j} + b_{i,-1,j}) = P_{sc}(2-\alpha)\tau , \tag{26}$$

where P_{sc} is the average power consumption in channel sensing. Similarly, by substituting Eq. (42) and Eq. (6), the average power consumption for packet transmission including both successful transmission and packet collision E_t is

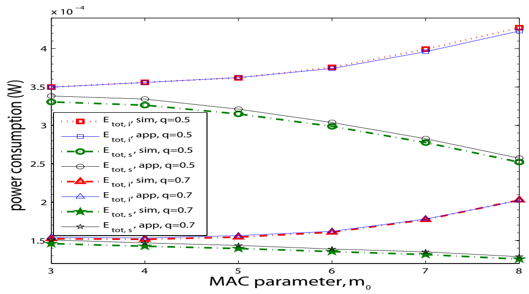
$$\begin{aligned} E_t =& P_t \sum_{j=0}^{n} \sum_{k=0}^{L-1} (b_{-1,k,j} + b_{-2,k,j}) + P_i \sum_{j=0}^{n} (b_{-1,L,j} + b_{-2,L,j}) \\ &+ \sum_{j=0}^{n} \sum_{k=L+1}^{L+L_{\text{ack}}+1} (P_r \, b_{-1,k,j} + P_i \, b_{-2,k,j}) \\ =& (1-\alpha)(1-\beta)\tau \left(P_t L + P_i + L_{\text{ack}} \left(P_r(1-P_c) + P_i P_c \right) \right) , \end{aligned} \tag{27}$$

where P_t, P_r are the average power consumption in transmit and receiving states, respectively. Analogously, E_q is the power consumption of idle stage without packet generation:

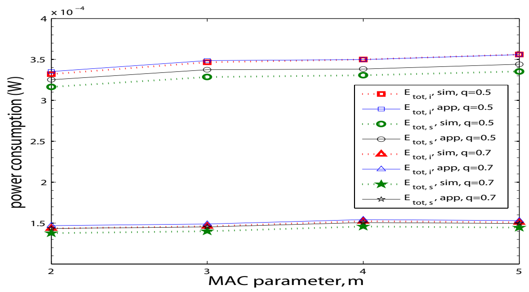
$$E_q = P_{sp} \sum_{l=0}^{L_0-1} Q_l \approx 0 , \tag{28}$$

where P_{sp} is the average power consumption in sleep states, respectively. We assume that the power consumption at sleeping state is negligible, namely $P_{sp} \approx 0$. Since a node wakes up only after generating packet, the wake-up power consumption $E_{w,i}$ is

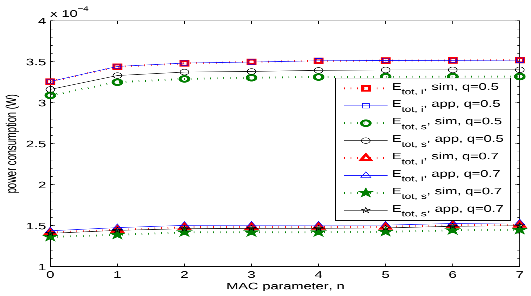
$$\begin{aligned} E_{w,i} =& P_w(1-q) Q_{L_0-1} \\ =& P_w q \left(x^{m+1}(1+y) + P_c(1-x^2) y^n + (1-P_c)(1-x^2)(1+y) \right) \widetilde{b}_{0,0,0} , \end{aligned} \tag{29}$$



(a) $m_0 = 3, \dots, 8, m_b = 8, m = 4, n = 3$



(b) $m = 2, \dots, 5, m_0 = 3, m_b = 8, n = 3$



(c) $n = 0, \dots, 7, m_0 = 3, m_b = 8, m = 4$

Figure 4: Average power consumption of *I-mode* and *S-mode* as a function of the traffic regimes $q = 0.3, 0.5, 0.7$ and MAC parameters $m_0 = 3, \dots, 8$, $m_b = 8$, $m = 2, \dots, 5$, $n = 0, \dots, 7$. The length of the packet is $L = 3$ and the number of nodes is $N = 20$.

where P_w is the average power consumption in wake-up state and the state probability $\widetilde{b}_{0,0,0}$ given in Eq. (18). By summing Eqs. (25)–(29), we obtain the average power consumption of *I-mode* in closed form.
The average power consumption of *S-mode* $E_{\mathrm{tot,s}}(\mathbf{V})$ during backoff states can be derived by following an approach similar to the *I-mode*:

$$E_{\mathrm{tot,s}}(\mathbf{V}) = E_{b,s} + E_{sc} + E_t + E_q + E_{w,s} \,,$$

where the sleep backoff power consumption is

$$E_{b,s} = P_{sp} \sum_{i=0}^{m} \sum_{k=1}^{W_i-1} \sum_{j=0}^{n} b_{i,k,j} \approx 0 \,,$$

the wake-up power consumption is

$$\begin{aligned} E_{w,s} =& P_w \sum_{i=0}^{m} \sum_{j=0}^{n} b_{i,1,j} \\ \approx& P_w \left(\tau - \frac{\widetilde{b}_{0,0,0}}{W_0} \frac{1-(0.5x)^{m+1}}{1-0.5x} \frac{1-y^{n+1}}{1-y} \right) , \end{aligned} \quad (30)$$

and E_{sc}, E_t, E_q is given in Eqs. (26), (27), (28), respectively. Note that since the radio is set in sleep mode during backoff period, node wakes up for each CCA1 state. ■

Fig. 4 compares the analytical model and simulation results of power consumption for both *I-mode* and *S-mode* as a function of different traffic regimes $q = 0.5, 0.7$ with a given number of nodes $N = 10$ and different MAC parameters m_0, m, n. We observe that the power consumption of *I-mode* increases as MAC parameters (m_0, m, n) increase under low traffic regime $q = 0.5, 0.7$ since the node needs to stay more time in idle sleep stage without packet generation under low traffic regime $q = 0.5, 0.7$, the main component of average power consumption is the idle backoff time rather than transmit or receiving power consumption i.e., $P_r > P_i > P_{sp}$ and $P_t > P_i > P_{sp}$. However, the power consumption of *S-mode* decreases as m_0 increases because of sleep mode during the backoff time. It is interesting to observe that the power consumption has a weaker dependence on m, n than m_0.

5 IEEE 802.15.4 Optimization

In the previous sections we developed the expressions of the performance metrics based on a generalized Markov chain model of the CSMA/CA mechanism. Here, we use these expressions and we present a novel approach where each node has to solve locally the optimization problem (1). In such a problem, the objective function is the total power consumption of the node, subject to reliability and delay constraints to transmit packets. The solution of the optimization problem gives the optimal MAC parameter m_0^*, m^*, n^* that each node has to use to minimize its energy expenditure. The average power consumption $E_{\rm tot}(\mathbf{V})$ depends on the radio mode during the backoff mechanism. It is given by Eq. (23) if the idle mode is selected, in which case we denote it $E_{\rm tot,i}(\mathbf{V})$, and it is given by Eq. (24), if the sleep mode is selected, in which case we denote it with $E_{\rm tot,s}(\mathbf{V})$ of Claim 3. The reliability $\widetilde{R}(\mathbf{V})$ and average delay $\widetilde{D}(\mathbf{V})$ are given by Eq. (12) of Claim 1 and (22) of Claim 2, respectively. Notice that the problem is combinatorial because the decision variables take on discrete values.

A vector of decision variables **V** is feasible if the reliability and delay constraints are satisfied. The optimal solution may be obtained by checking every combination of the elements of **V** that gives feasibility, and then checking the combination that gives the minimum objective function. Clearly, this approach may have a high computational complexity, since there are $6 \times 4 \times 8 = 192$ combinations of MAC parameters to check [1]. Therefore, in the following we propose an algorithm that gives the optimal solution by checking just a reduced number of combinations.

From Figs. 2, 3 and 4, we remark here that the reliability and power consumption of both *I-mode* and *S-mode* are increasing function as the parameter n increases. This properties are quite useful to solve (1) by an algorithm 1 with reduced computational complexity, as we see next.

The search of optimal MAC parameters uses an iterative procedure according to the component-based method [27]. In particular, the probabilities α, β, and τ are estimated periodically by each node. If a node detects a change, then the node solves the local optimization problem (1) using these estimated values.

The solution is achieved in two steps: first m_0 and m are fixed, and the value of n that minimizes the energy consumption is derived. Second, all the combinations of m_0 and m are checked, and the triple m_0^*, m^*, n^* that minimizes the energy consumption is selected. In particular, since the power consumption is increasing with n, it follows that the minimum is attained at the lowest value of n that satisfies the constraints. Given that the reliability is increasing with n, simple algebraic passages give that such a value is $n = f(m_0, m)$, with

$$f(m_0, m) = \left\lceil \frac{\ln(1 - x^{m+1}(1+\widetilde{y}) - R_{\min})}{\ln(\widetilde{y})} - 1 \right\rceil , \qquad (31)$$

where $\widetilde{y} = (1 - (1-\widetilde{\tau})^{N-1})(1-x^2)$ and

$$\widetilde{\tau} = \frac{2r_3}{W_0 r_1 + 2r_2} ,$$

with

$$\begin{aligned} r_1 &= (1+2x)(1+\hat{y}) , \\ r_2 &= L_s(1-x^2)(1+\hat{y}) + K_0(1+\hat{y}^2+\hat{y}^{n+1}) , \\ r_3 &= (1+x)(1+\hat{y}) , \end{aligned}$$

Algorithm 1: **Optimal solution to problem** (1)

```
Input: Feasible range of MAC parameters (m_0, m, n)
Output: m_0^*, m^*, n^*
begin
    Estimate α, β, τ
    currentObj ← ∞ ;
    V^* ← V_0 ;
    for m_0 ← 3 to 8 do
        for m ← 2 to 5 do
            n ← f(m_0, m) ;
            V ← [m_0, m, n] ;
            if D̃(V) ≤ D_max and isrg(n) then
                /* isrg(n): validity of n.                  */

                if currentObj > E_tot(V) then
                    currentObj ← E_tot(V);
                    V^* ← V;
                    /* else; not optimal.                   */

                /* else; not feasible.                      */

end
```

and $\hat{y} = (1-(1-\tau)^{N-1})(1-x^2)$. Notice that x and $\hat{y}$ are measurable since node estimates α, β, and τ.

This optimization procedure is summarized in Algorithm 1.

In the algorithm, Eq. (31) returns the optimal retry limits given a pair m_0, m. The function $isrg(n)$ checks the feasibility of given n by considering the allowed range. By using Algorithm 1, a node checks just $6 \times 4 = 24$ combinations of the MAC parameters m_0, m instead of $6 \times 4 \times 8 = 192$ combinations that would be required by an exhaustive search.

6 Practical Considerations

We have seen in Subsections 4.2, 4.3 and 4.4 that the performance metrics are function of the busy channel probabilities α and β and the channel access probability τ. Once these probabilities are known at each node, the optimal MAC parameters can be readily computed by the algorithm given in Subsection 5. In the algorithm, the number of nodes and packet generation rates are assumed to be known, whereas the busy channel probability and channel access probability are periodically estimated in each node during the sensing states of MAC layer, and they do not require an ACK mechanism, as we describe the details in the follow-

Algorithm 2: Algorithm for the node's MAC optimization.

INITIAL STATE:
1: Nodes communicate with coordinator at fixed MAC parameters.
2: The node estimates the busy channel probabilities α, β and channel access probability τ.
OPT STATE:
3: The node solves the optimization problem (1) by Algorithm. 1.
4: The optimal MAC solutions (m_0^*, m^*, n^*) are updated

ing. The robustness of the algorithm to possible errors in the estimation of the number of nodes and traffic load is then investigated in Section 7.3.

The average busy channel probabilities α and β are estimated at each node while sending a data packet to the coordinator. These probabilities are initialized at the beginning of the node's operation. Then, when the node senses the channel at CCA1 or CCA2, these probabilities are updated by $\alpha = \delta_b\alpha + (1-\delta_b)\hat{\alpha}, \beta = \delta_b\beta + (1-\delta_b)\hat{\beta}$ for some $\delta_b \in (0,1)$, respectively. The estimations of the busy channel probabilities $\hat{\alpha}$ and $\hat{\beta}$, and the channel access probability, use a sliding window with size W_b. Therefore, a node does not require any extra communication and sensing state to estimate these probabilities compared to the IEEE standard. By contrast, the estimation algorithms for IEEE 802.11 proposed in [18] and [28] are not energy efficient since a node needs to sense the channel state during the backoff stage. This allows one to estimate the average length of idle period. Hence, these schemes are implementable only in *I-mode*. By contrast, our scheme is applied both in *I-mode* and in *S-mode* and does not require any computation load during the backoff stage.

During an initialization phase, a node communicates with the default MAC parameters $m_0 = 3, m_b = 8, m = 4, n = 3$. Then, the busy channel probabilities α and β and the channel access probability τ are periodically estimated in each node during the channel sensing state. The application requirements are communicated by the coordinator to the node by piggybacking them in the beacon message from coordinator to sensor nodes if there are changes.

7 Numerical Results

In the following, we present Monte Carlo simulations to analyze the performance of the adaptive tuning algorithm of the MAC parameters we proposed, under both stationary and transient conditions. In the stationary conditions, the application requirements and network scenario are constant, whereas in transient condition there are variations. The simulations are based on the specifications of the IEEE 802.15.4 and the practical implementation aspects described in Section 6. In the simulations, the network considers the *I-mode* and *S-mode* of the node to compare the performance on the reliability, average packet delay and power consumption.

Furthermore, we investigate the fairness of resource allocation, robustness to network changes and sensitivity to inaccurate parameter estimations. Details follow in the sequel.

7.1 Protocol Behavior in Stationary Conditions

In this subsection, we are interested to the improvement of performance metrics of the proposed scheme at stationary conditions of the network, namely without changing application requirements and network scenarios. We also present a fairness analysis of the adaptive protocol.

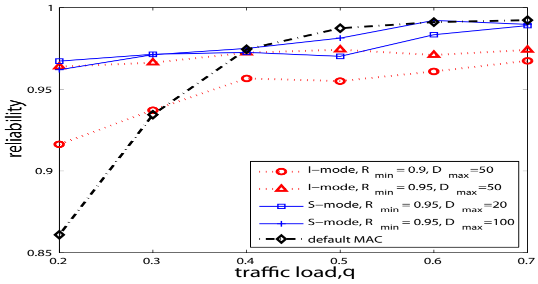
Figs. 5 compare the reliability, average delay and power gain values of the protocol as obtained by our algorithm and with default MAC parameters. Both the *I-mode* and *S-mode* for various traffic configurations and constraints are considered. The requirements for both the *I-mode* and *S-mode* are $R_{\min} = 0.9, 0.95, D_{\max} = 50$ and $R_{\min} = 0.95, D_{\max} = 20, 100\,\text{ms}$, respectively. Fig. 5(a) shows that both *I-mode* and *S-mode* satisfy the reliability constraint for different traffic regime. We observe the strong dependence of the reliability of default MAC with different traffic regime due to the fixed MAC parameters. At the high traffic regime $q = 0.2$, the reliability of default MAC is 0.861. In Fig. 5(b), the delay constraint is fulfilled both *I-mode* and *S-mode*. Observe that average delay of *I-mode* decreases when traffic regime is low $q \geq 0.5$. This is due to that the optimal MAC parameters at higher traffic regime increase more than the ones at lower traffic regime to satisfy the reliability constraint.

Recall that the target of our proposed adaptive algorithm is to use the tradeoff between application constraints and energy consumption instead of just maximization of reliability or minimization of delay. Therefore, to characterize quantitatively the power consumption, we define the power gain as

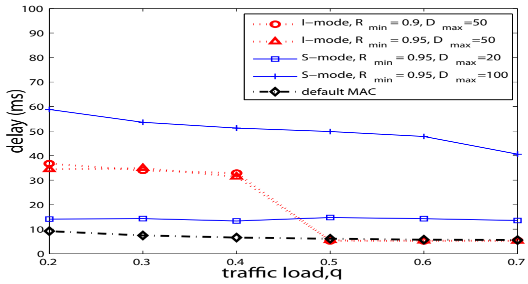
$$\rho = \frac{E_{\text{def}} - E_{\text{tot}}(\mathbf{V})}{E_{\text{def}}}$$

where E_{def} and $E_{\text{tot}}(\mathbf{V})$ are the average power consumption of *I-mode* or *S-mode* for default MAC and proposed scheme, respectively. The closer ρ to 1, the better the power efficiency. Fig. 5(c) shows that the power gain increases as traffic regime increases. This improvement is higher for *S-mode* than *I-mode*, e.g., power gain $\rho \approx 0.49$ for *S-mode* with $R_{\min} = 0.95, D_{\max} = 100$. Although there is a strong dependence of the power gain on the traffic regime, our proposed algorithm gives a better energy efficiency than the default MAC. Therefore, the numerical results show clearly the effectiveness of our adaptive IEEE 802.15.4 protocol while guaranteeing the constraints.

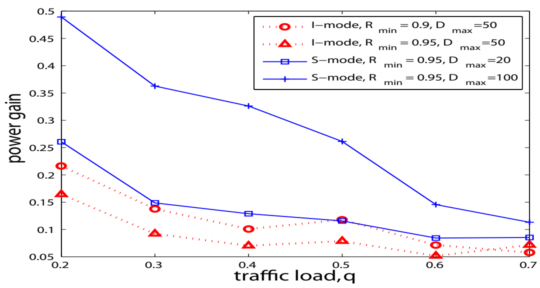
Next, we observe the tradeoff between the power consumption, reliability and delay constraints. Fig. 6 show the dependence of the power consumption in *S-mode* with reliability and delay constraints for a given traffic load, length of packets, and number of nodes. Observe that as the delay constraint becomes strict the power



(a) Reliability



(b) Average delay



(c) Power gain

Figure 5: Stationary condition: reliability, average delay and power gain of the *I-mode*, *S-mode* of proposed scheme and IEEE 802.15.4 with default parameter ($\textit{macMinBE} = 3, \textit{macMaxBE} = 5, \textit{macMaxCSMABackoffs} = 4, \textit{macMaxFrameRetries} = 3$) as a function of the traffic load $q = 0.2, \ldots, 0.7$, the reliability requirement $R_{\min} = 0.9, 0.95$ and delay requirement $D_{\max} = 20, 50, 100$ ms for a length of the packet $L = 7$ and $N = 10$ nodes. Note that "default MAC" refers to IEEE 802.15.4 with default MAC parameters.

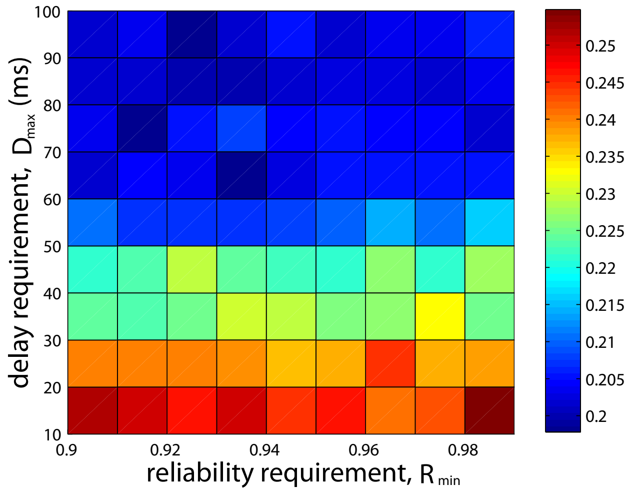


Figure 6: Stationary condition: power consumption of *S-mode* as a function of reliability constraint $R_{\min} = 0.9, \ldots, 0.99$ and delay requirement $D_{\max} = 10, \ldots, 100$ ms for the traffic load $q = 0.5$, the length of packet $L = 3$ and a number of nodes $N = 10$.

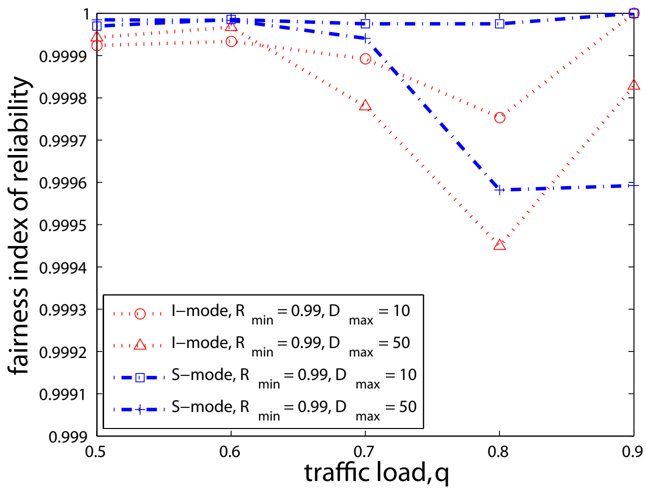


Figure 7: Fairness index of the reliability as a function of the traffic load $q = 0.5, \ldots, 0.9$, reliability requirement $R_{\min} = 0.99$ and delay requirement $D_{\max} = 10, 50$ ms for a length of the packet $L = 3$ and a number of nodes $N = 10$.

consumption increases. In other words, the reliability constraint of *S-mode* is less critical than delay constraint, see more results in [29].

The fairness of resource management is one of the most important concerns when implementing the tuning algorithm of the MAC parameters. We use Jain's fairness index [30] to show the fairness of our proposed scheme for both *I-mode* and *S-mode*. We compute the fairness index from a total number of 10 nodes in a stable network. The closer fairness index to 1, the better the achieved fairness. Fig. 7 compares the fairness index on reliability for the different requirements and traffic configurations with a given length of the packet and number of nodes. Fig. 7 reports the very high fairness achievement on reliability greater than 0.999, similar behavior for delay and power consumption. In other words, the MAC parameters of each node converge to similar parameter values. For the adaptive IEEE 802.15.4 protocol, we conclude that most of the node can share equally the common medium.

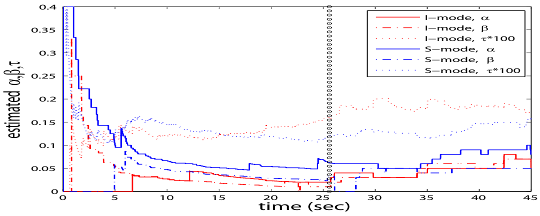
7.2 Protocol Behavior in Transient Conditions

The adaptive IEEE 802.15.4 protocol is based on the estimation of the busy channel probabilities α and β and the channel access probability τ. In this section, we investigate the convergence time of dynamic adaptation to the correct MAC parameters when the reliability and delay constraints change.

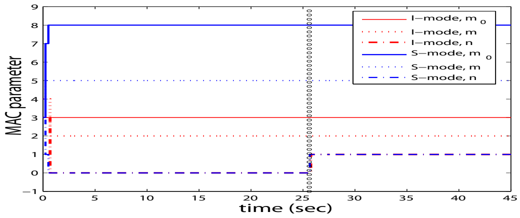
Figs. 8 and 9 show the dynamic adaptation of the MAC parameters of our proposed algorithm for both *I-mode* and *S-mode* on channel state, reliability and delay when the requirements change with a given traffic load and length of packets and number of nodes. Figs. 8(a), 8(b), 8(c), 8(d) and 9(a), 9(b), 9(c), 9(d) compare the behavior of busy channel, channel access probability, MAC parameters, reliability and delay when the reliability and delay requirements change.

Figs. 8(a) and 9(a) report the busy channel probabilities α and β and channel access probability τ over time. In Section 6, we noticed that the update frequency of α, β, τ is different. τ is updated in each *aUnitBackoffPeriod* and α and β are updated when a node stay in CCA1 and CCA2, respectively. Hence, the update frequency order of α, β, and τ is τ first, then α, and finally β.

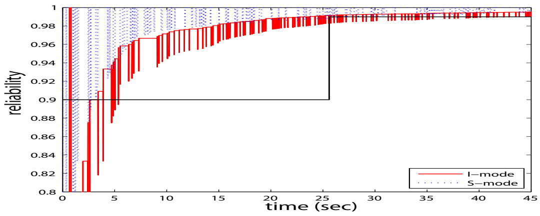
Figs. 8(b) and 9(b) show the adaptation of the MAC parameters (m_0, m, n) when the requirements change. Observe that the optimization algorithm returns different parameters for *S-mode* and *I-mode* due to the different power consumption model (see details in Section 4). Note that, as m_0 increases, the power consumption of *I-mode* increases, which is the opposite of *S-mode*. The optimal m_0, m, n of *I-mode* and *S-mode* adapts to $8, 5, 0$ and $3, 2, 0$, respectively. Furthermore, we observe the convergence of the MAC parameters of proposed scheme is very fast since our algorithm is based on analytical model instead of heuristic considerations as in link-based adaptation, where the algorithms adapt the contention window size by the ACK transmission [21]– [25]. By contrast, recall that our adaptive IEEE 802.15.4 is based on the physical sensing information before transmitting packets.



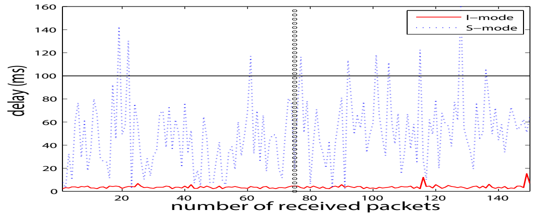
(a) α, β, τ behavior



(b) MAC parameter (m_0, m, n) behavior

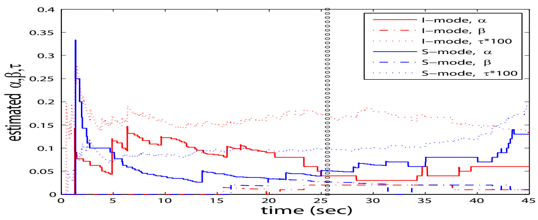


(c) Reliability behavior

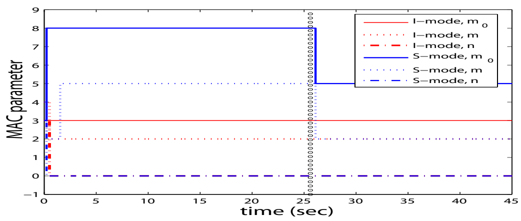


(d) Delay behavior

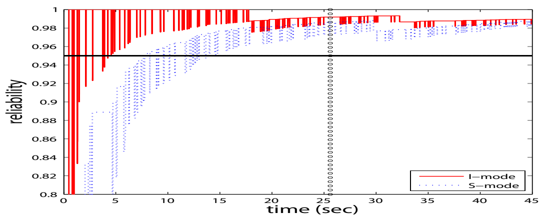
Figure 8: Transient condition: busy channel probabilities, channel access probability, MAC parameters, reliability and delay for *I-mode* and *S-mode* for traffic load $q = 0.6$, length of the packet $L = 3$ and number of nodes $N = 10$ when, at 26 s from the beginning, the reliability requirement varies from $R_{\min} = 0.9$ to $R_{\min} = 0.99$.



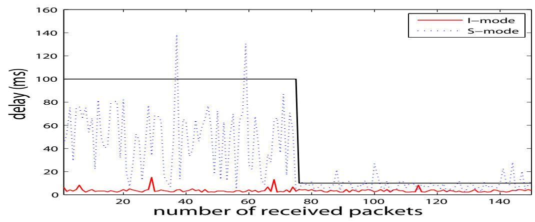
(a) α, β, τ behavior



(b) MAC parameter (m_0, m, n) behavior



(c) Reliability behavior



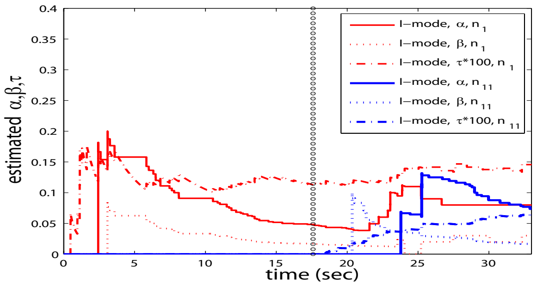
(d) Delay behavior

Figure 9: Transient condition: busy channel probabilities, channel access probability, MAC parameters, reliability and delay for *I-mode* and *S-mode* for traffic load $q = 0.6$, length of the packet $L = 3$ and number of nodes $N = 10$ when, at 26 s from the beginning, the delay requirement from $D_{\max} = 100$ ms to $D_{\max} = 10$ ms.

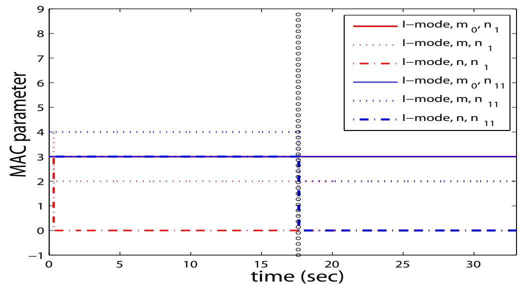
By comparing Figs. 8(b) and 8(c), the parameter n increases to 1 of both *I-mode* and *S-mode* when the reliability constraint increases to 0.99 at time 26 s. In addition, we observe the impact on the channel condition of increasing the MAC parameters. From Fig. 8(a), we see that the busy channel and channel access probabilities increase with the increasing of n. From Fig. 8(c), we see that both in the *I-mode* and *S-mode* the reliability requirement of 0.99 is fulfilled. The delay of *S-mode* is higher than *I-mode* due to larger MAC parameters m_0, m. The reliability of *S-mode* is larger than *I-mode* since the MAC parameters m_0, m are larger than the ones of *I-mode*. By the same argument, we observe that the packet delay of *S-mode* is about four times the one measured in *I-mode*. The packet delay is much more variable in *S-mode* than in *I-mode* one. Specifically, with *I-mode*, we have a reduction in the average MAC delay and a shorter tail for the MAC delay distribution with respect to the *S-mode*. Hence, there is a tradeoff between reliability and packet delay in the proposed scheme. The behavior of the adaptive MAC parameters measured by our proposed algorithm converges dramatically to the optimal values because it considers the busy channel and channel access probabilities.
Figs. 9(b) and 9(d) describe the dynamic adaptation of our algorithm to the delay requirement change. Fig. 9(b) reports that m_0 and m of *S-mode* adapt to $5, 2$ when $D_{\max}$ changes at time 26 s. In Fig. 9(d), we observe that the packet delay converges to 10 ms. In addition, the reliability decreases due to the decreasing parameters m_0, m at time 29 s. From the analysis, we can assess that the proposed scheme achieves the target of minimizing the power consumption while guaranteeing a quality of the service experienced by the network nodes (i.e., reliability and packet delay).

7.3 Robustness and Sensitivity Analysis

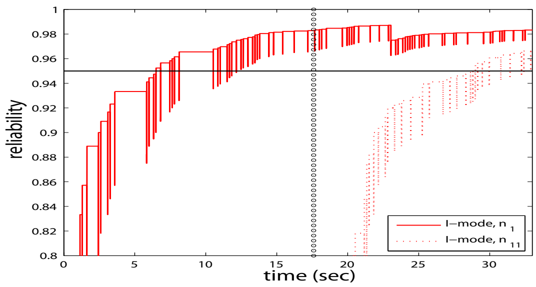
The performance analysis carried out so far assumed a network in stationary conditions. In particular, we have assumed that the number of nodes and traffic configuration are fixed. This assumption has allowed us to verify the effectiveness of our adaptive algorithm for IEEE 802.15.4 in steady state conditions. However, one of the critical issues in the design of wireless networks is dynamic topology. Therefore, in the following analysis, we will investigate our algorithm to react to changes in the number of nodes and traffic load when each node has an erroneous estimation of these parameters.
Figs. 10 considers the *I-mode* and show the dynamical behavior of the node when the number of nodes changes from $N = 10$ to $N = 20$ with an erroneous estimation of the number of nodes. At time 17.6 s, the number of nodes sharply increases to 20, when it was estimated to 10. This causes significant increase of the contention level. Note that n_1 is one of existing nodes before the network change and n_{11} is one of the new nodes that enters the network at time 17.6 s using its default MAC parameters. The already existing and new node adapt the MAC parameters by estimating α, β, τ. We assume that the wrong estimation happens due to some errors in the estimation or the biasing induced by the hidden-node phenomenon.



(a) α, β, τ behavior of *I-mode*

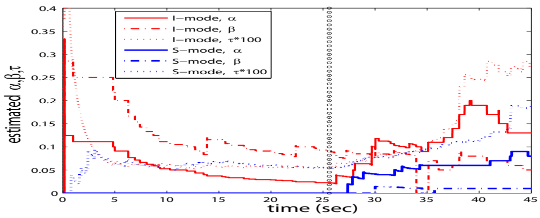


(b) MAC parameter (m_0, m, n) behavior of *I-mode*

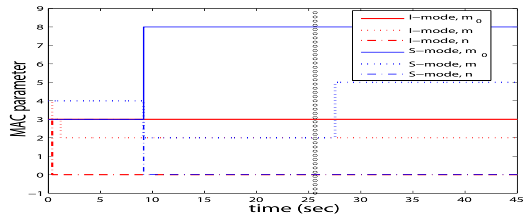


(c) Reliability behavior of *I-mode*

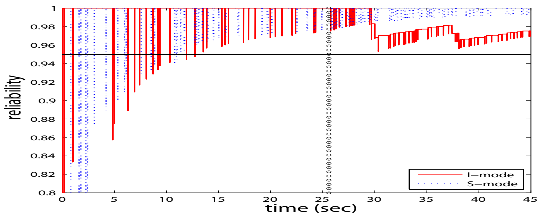
Figure 10: Robustness to the change of the number of nodes: busy channel probabilities, channel access probability, MAC parameters and reliability behavior of *I-mode* when the number of nodes changes sharply from $N = 10$ to $N = 20$ at time 17.6 s. Note that n_1 and n_{11} represent the behavior of one of $N = 10$ nodes plus new nodes after time 17.6 s, respectively. Traffic load $q = 0.6$, length of the packet $L = 3$ the reliability and delay constraint $R_{\min} = 0.95$ and $D_{\max} = 100$ ms, respectively.



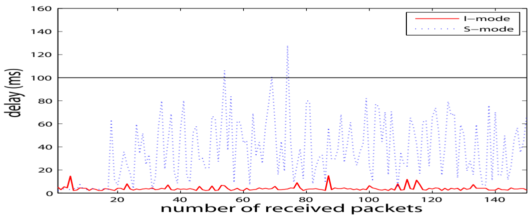
(a) α, β, τ behavior



(b) MAC parameter (m_0, m, n) behavior



(c) Reliability behavior



(d) Delay behavior

Figure 11: Robustness when the traffic load changes: busy channel probabilities, channel access probability, MAC parameters, reliability and delay behavior of *I-mode* and *S-mode* when the traffic load changes sharply from $q = 0.8$ to $q = 0.5$ at time 25.6 s for length of the packet $L = 3$ the reliability and delay constraint $R_{\min} = 0.95$ and $D_{\max} = 100$ ms, respectively.

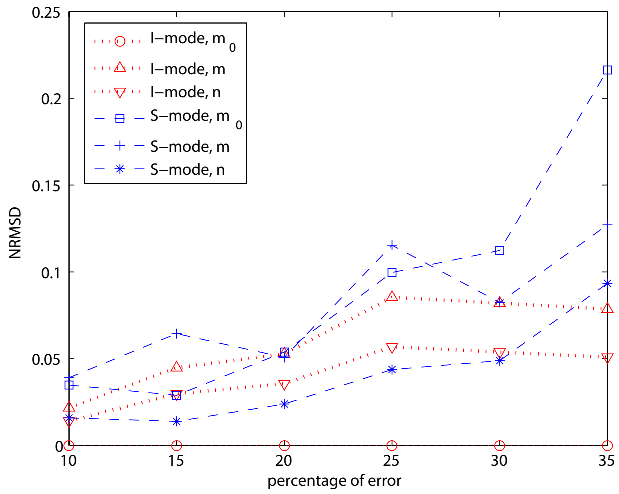


Figure 12: Sensitivity: normalized root mean square deviation (NRMSD) of *I-mode* and *S-mode* when traffic load $q = 0.6$, a reliability requirement $R_{\min} = 0.95$ and a delay requirement $D_{\max} = 100$ ms, length of the packet $L = 3$ and number of nodes $N = 10$ with different percentage error in busy channel probabilities α and β and channel access probability τ.

In Fig. 10(a), we observe that the busy channel and channel access probabilities of node n_{11} become stable after the network changes by updating the MAC parameters. Fig. 10(b) plots that the MAC parameters (m_0, m, n) converge to $3, 2, 0$ in *I-mode*. The figures indicate that the system reacts correctly to the erroneous estimation of the number of nodes after a few seconds. In Fig. 10(c), the reliability fulfills the requirement $R_{\min} = 0.95$ for both the existing and new nodes. Similar behaviors are observed for *S-mode*, see further details in [29].

Figs. 11 present the behavior of the node when the traffic load changes sharply from $q = 0.8$ to $q = 0.5$ at time 25.6s. Nodes uses a wrong estimation of the traffic load $q = 0.8$ after the traffic load changes. The results indicate that our algorithm is quite effective for the traffic configuration change. In Fig. 11(a), the busy channel and channel access probability increase as a result of higher traffic regime $q = 0.5$ for both *I-mode* and *S-mode*. Fig. 11(b) show that the parameter m of *S-mode* updates from 2 to 5 due to the increasing busy channel probability after the traffic load changes at time 28 s. The figure indicates that the system reacts correctly to the erroneous estimation of traffic configuration and, in few seconds, the estimate for α, β, τ allow to reach the optimal MAC parameters. In Figs. 11(c) and 11(d), the reliability and average delay constraint $(R_{\min} = 0.95, D_{\max} = 100\,\text{ms})$ are fulfilled for both *I-mode* and *S-mode*. Note that the reliability of *I-mode* is greater than 0.95 with some fluctuations after traffic load increases.

Fig. 12 illustrates the sensitivity of adaptive IEEE 802.15.4 with respect to the estimation errors to the busy channel probabilities α, β and the channel access probability τ. The normalized root mean squared deviation (NRMSD) between the optimal MAC parameters with exact estimation and the ones with erroneous estimation is used as the indicator of sensitivity. The normalization is taken over the range of MAC parameters (m_0, m, n). The NRMSD is approximately below 10% if the percentage of error is smaller than 20% for α, β, τ. It is interesting to observe that m_0 of ***I-mode*** is very robust to errors. This is due to the power consumption model, i.e., to the dominant factor m_0 of power consumption in *I-mode*. The robustness of MAC parameter is $m_0 > n > m$ and $n > m > m_0$ for *I-mode* and *S-mode*, respectively. We can show that errors below 20% in the estimation of α, β, τ give a performance degradation below 3% in terms on reliability, packet delay and energy gain for low traffic load.

8 Conclusions

In this paper we developed an analysis based on a generalized Markov chain model of IEEE 802.15.4, including retry limits, acknowledgements and unsaturated traffic regime. Then, we presented a novel adaptive MAC algorithm for minimizing the power consumption while guaranteeing reliability and delay constraints of IEEE 802.15.4 protocol. The algorithm does not require any modifications of the standard. The adaptive algorithm is grounded on an optimization problem where the objective function is the total power consumption, subject to constraints of reliability and delay of the packet delivery and the decision variables are the MAC parameters (*macMinBE*, *macMaxCSMABackoffs*, *macMaxFrameRetries*) of the standard. The proposed adaptive MAC algorithm is easily implementable on sensor nodes by estimating the busy channel and channel access probability.

We investigated the performance of our algorithm under both stationary and transient conditions. Numerical results showed that our optimization is efficient and ensures a longer lifetime of the network. In addition, we show that, even if the number of active nodes, traffic configuration and application constrains change sharply, our algorithm allow the system to recover quickly and operate at its optimal parameter by estimating just the busy channel and channel access probability. We also investigated the robustness of the protocol to possible errors during the estimation process on number of nodes and traffic load. Results indicated that the protocol reacts promptly to erroneous estimations.

Future investigations include the use of the aforementioned achievements to the practical implementation on sensor nodes based on specific application constraints.

Appendix A

A.1 Proof of Lemma 1

The proof has two steps. First, we derive the state transition probability of Markov chain. Second, the normalization condition is applied to compute the probability $b_{0,0,0}$.

The state transition probabilities associated with the Markov chain of Fig. 1 are

$$P(i,k,j|i,k+1,j) = 1, \text{ for } k \geq 0\,, \tag{32}$$

$$P(i,k,j|i-1,0,j) = \frac{\alpha + (1-\alpha)\beta}{W_i}, \text{ for } i \leq m\,, \tag{33}$$

$$P(0,k,j|i,0,j-1) = \frac{(1-\alpha)(1-\beta)P_c}{W_0}, \text{ for } j \leq n\,, \tag{34}$$

$$P(Q_0|m,0,j) = q\,(\alpha + (1-\alpha)\beta), \text{ for } j < n\,, \tag{35}$$

$$P(Q_0|i,0,n) = q\,(1-\alpha)(1-\beta), \text{ for } i < m\,, \tag{36}$$

$$P(Q_0|m,0,n) = q, \tag{37}$$

$$P(0,k,0|Q_0) = \frac{1-q}{W_0}, \text{ for } k \leq W_0 - 1\,. \tag{38}$$

Eq. (32) is the decrement of backoff counter, which happens with probability 1. Eq. (33) represents the probability of finding busy channel in CCA1 or CCA2 and a node selects uniformly a state in the next backoff stage. Eq. (34) gives the unsuccessful transmission probability after finding an idle channel in both CCA1 and CCA2, and a node picks uniformly a state in the next retransmission stage. Eq. (35) and (36) represent the probability of going back to the idle stage due to the channel access failure and retry limits, respectively. Eq. (37) accounts for the traffic regime and is the probability of going back to the idle stage at backoff counter m and retransmission stage n, which is given by q. Eq. (38) models the probability of going back to the first backoff stage from the idle stage. In the following, we use Eqs. (32)–(38) to compute the stationary distribution of the Markov chain.

By the normalization condition, we know that

$$\sum_{i=0}^{m}\sum_{k=0}^{W_i-1}\sum_{j=0}^{n} b_{i,k,j} + \sum_{i=0}^{m}\sum_{j=0}^{n} b_{i,-1,j} + \sum_{j=0}^{n}\left(\sum_{k=0}^{L_s-1} b_{-1,k,j} + \sum_{k=0}^{L_c-1} b_{-2,k,j}\right) + \sum_{l=0}^{L_0-1} Q_l = 1\,. \tag{39}$$

We next derive the expressions of each term in Eq. (39).

From Eq. (3), (4), we have

$$\sum_{i=0}^{m}\sum_{k=0}^{W_i-1}\sum_{j=0}^{n} b_{i,k,j} \tag{40}$$

$$= \sum_{i=0}^{m}\sum_{j=0}^{n} \frac{W_i+1}{2}(\alpha+(1-\alpha)\beta)^i\, b_{0,0,j}$$

$$= \begin{cases} \frac{b_{0,0,0}}{2}\left(\frac{1-(2x)^{m+1}}{1-2x}W_0 + \frac{1-x^{m+1}}{1-x}\right)\frac{1-y^{n+1}}{1-y} \\ \quad \text{if } m \le m_b - m_0 \\ \\ \frac{b_{0,0,0}}{2}\left(\frac{1-(2x)^{m_b-m_0+1}}{1-2x}W_0 + \frac{1-x^{m_b-m_0+1}}{1-x} + \right. \\ \left.(2^{m_b}+1)x^{m_b-m_0+1}\frac{1-x^{m-m_b+m_0}}{1-x}\right)\frac{1-y^{n+1}}{1-y} \\ \quad \text{otherwise,} \end{cases}$$

where $x = \alpha + (1-\alpha)\beta$ and $y = P_c(1-x^{m+1})$. Similarly,

$$\sum_{i=0}^{m}\sum_{j=0}^{n} b_{i,-1,j} = \sum_{i=0}^{m}\sum_{j=0}^{n}(1-\alpha)(\alpha+(1-\alpha)\beta)^i\, b_{0,0,j}$$

$$= (1-\alpha)\frac{1-x^{m+1}}{1-x}\,\frac{1-y^{n+1}}{1-y}\, b_{0,0,0}\,, \tag{41}$$

and

$$\sum_{j=0}^{n}\left(\sum_{k=0}^{L_s-1} b_{-1,k,j} + \sum_{k=0}^{L_c-1} b_{-2,k,j}\right) \tag{42}$$

$$= (L_s(1-P_c)+L_cP_c)(1-x^{m+1})\frac{1-y^{n+1}}{1-y}\, b_{0,0,0}\,.$$

By considering that the successful transmission and the failure events are due to the limited number of backoff stages m and the retry limit n, the idle state probability is

$$Q_0 = q\,Q_{L_0-1} + q\left(\sum_{j=0}^{n}(\alpha+(1-\alpha)\beta)\; b_{m,0,j} + \sum_{i=0}^{m} P_c \right.$$

$$\left. \times (1-\beta)\, b_{i,-1,n} + \sum_{i=0}^{m}\sum_{j=0}^{n}(1-P_c)\,(1-\beta)\, b_{i,-1,j}\right)$$

$$= \frac{q}{1-q}\left(\frac{x^{m+1}(1-y^{n+1})}{1-y} + P_c(1-x^{m+1})y^n \right.$$

$$\left. + (1-P_c)\frac{(1-x^{m+1})(1-y^{n+1})}{1-y}\right) b_{0,0,0}\,, \tag{43}$$

where L_0 is the idle state length without generating packets and $\sum_{l=0}^{L_0-1} Q_l = L_0 Q_0$. Note that Eqs. (40)–(43) give the state values $b_{i,k,j}$ as a function of $b_{0,0,0}$. By replacing Eqs. (40)–(43) in the normalization condition given by Eq. (39), we obtain the expression for $b_{0,0,0}$.

A.2 Proof of Lemma 2

A transmission may be successful with probability $1 - P_c$, or collide with probability P_c. Then, the probability of the event $\mathcal{A}_j|\mathcal{A}_t$ is

$$\Pr(\mathcal{A}_j|\mathcal{A}_t) = \frac{P_c^j (1 - x^{m+1})^j}{\sum_{k=0}^{n} \left(P_c(1 - x^{m+1})\right)^k}$$

where the normalization comes by considering all the possible events of successful attempts $\mathcal{A}_t$. Note that $(1 - x^{m+1})$ is the probability of successful channel access within the maximum number of m backoff stages.

A.3 Proof of Lemma 3

By considering the busy channel during two CCAs, the probability of the event $\mathcal{B}_i|\mathcal{B}_t$ is approximated by

$$\widetilde{\Pr}(\mathcal{B}_i|\mathcal{B}_t) = \frac{b_l^i}{\sum_{k=0}^{m} b_l^k} , \tag{44}$$

where $b_l = \max(\alpha, (1 - \alpha)\beta)$ (note that this is the term that gives the approximation). The approximation of the average backoff period $\mathbb{E}[\widetilde{T}_h]$ is

$$\begin{aligned} \mathbb{E}[\widetilde{T}_h] =& \sum_{i=0}^{m} \widetilde{\Pr}(\mathcal{B}_i|\mathcal{B}_t)\, \mathbb{E}[\widetilde{T}_{h,i}] \\ =& 2T_{sc} + \sum_{i=0}^{m} \widetilde{\Pr}(\mathcal{B}_i|\mathcal{B}_t) \sum_{k=0}^{i} \left(\frac{W_0 2^k - 1}{2} S_b + 2T_{sc}\, k \right) \end{aligned} \tag{45}$$

where the approximated sensing time $\mathbb{E}[\widetilde{T}_{h,i}]$ considers the worst case, i.e., a failure of the second sensing (CCA2), which implies that $T_{sc} = S_b$ and that each sensing failure takes $2T_{sc}$ in Eq. (20).

References

[1] *IEEE Std 802.15.4-2996, September, Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)*, IEEE, 2006. [Online]. Available: http://www.ieee802.org/15

[2] A. Willig, K. Matheus, and A. Wolisz, “Wireless technology in industrial networks,” *Proceedings of the IEEE*, 2005.

[3] T. Abdelzaher, T. He, and J. Stankovic, “Feedback control of data aggregation in sensor networks,” in *IEEE CDC*, December 2004.

[4] S. Pollin, M. Ergen, S. C. Ergen, B. Bougard, F. Catthoor, A. Bahai, and P. Varaiya, “Performance analysis of slotted carrier sense IEEE 802.15.4 acknowledged uplink transmissions,” in *IEEE WCNC*, 2008.

[5] *IEEE Std 802.11 Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE, 1999. [Online]. Available: http://www.ieee802.org/11

[6] G. Bianchi, “Performance analysis of the IEEE 802.11 distributed cordination function,” *IEEE Journal on Selected Areas in Communications*, vol. 18, March 2000.

[7] P. Chatzimisios, A. C. Boucouvalas, and V. Vitsas, “IEEE 802.11 packet delay a finite retry limit analysis,” in *IEEE GLOBECOM*, 2003.

[8] Z. Hadzi-Velkov and B. Spasenovski, “Saturation throughput-delay analysis of IEEE 802.11 in fading channel,” in *IEEE ICC*, 2003.

[9] O. Tickioo and B. Sikdar, “Queueing analysis and delay mitigation in IEEE 802.11 random access MAC based wireless networks,” in *IEEE INFOCOM*, 2004.

[10] H. Wu, Y. Peng, K. Long, S. Cheng, and J. Ma, “Performance of reliable transport protocol over IEEE 802.11 wireless LAN: Analysis and enhancement,” in *IEEE INFOCOM*, 2002, pp. 599–607.

[11] J. Zheng and M. L. Lee, “A comprehensive performance study of IEEE 802.15.4,” in *IEEE Press Book*, 2004.

[12] A. Koubaa, M. Alves, and E. Tovar, “A comprehensive simulation study of slotted CSMA/CA for IEEE 802.15.4 wireless sensor networks,” in *IEEE IWFCS*, Jun 2006, pp. 183–192.

[13] S. Pollin, M. Ergen, S. C. Ergen, B. Bougard, L. V. D. Perre, F. Catthoor, I. Moerman, A. Bahai, and P. Varaiya, “Performance analysis of slotted carrier sense IEEE 802.15.4 medium access layer,” in *IEEE GLOBECOM*, 2006, pp. 1–6.

[14] J. Mišió, S. Shaf, and V. Mišió, “Performance of a beacon enabled IEEE 802.15.4 cluster with downlink and uplink traffic,” *IEEE Trans. Parallel and Distributed Systems*, pp. 361–376, 2006.

[15] P. K. Sahoo and J. P. Sheu, “Modeling IEEE 802.15.4 based wireless sensor network with packet retry limits,” in *ACM PE-WASUN*, 2008.

[16] C. Y. Jung, H. Y. Hwang, D. K. Sung, and G. U. Hwang, “Enhanced markov chain model and throughput analysis of the slotted CSMA/CA for IEEE 802.15.4 under unsaturated traffic conditions,” *IEEE Transactions on Vehicular Technology*, vol. 58, no. 1, pp. 473–478, 2009.

[17] A. Giridhar and P. R. Kumar, “Toward a Theory of In-network Computation in Wireless Sensor Networks,” *IEEE Communication Magazine*, pp. 97–107, April 2006.

[18] F. Cali, M. Conti, and E. Gregori, "IEEE 802.11 protocol: design and performance evaluation of an adaptive backoff mechanism," *IEEE Journal on Selected Areas in Communications*, 2000.

[19] K. Yedavalli and B. Krishnamachari, "Enhancement of the IEEE 802.15.4 MAC protocol for scalable data collection in dense sensor networks," in *ICST WiOPT*, 2008, pp. 152–161.

[20] R. Bruno, M. Conti, and E. Gregori, "Optimization of efficiency and energy consumption in p-persistent CSMA-based wireless LANs," *IEEE Transactions on Mobile Computing*, vol. 1, pp. 10 – 31, 2002.

[21] Q. Pang, S. C. Liew, J. Y. B. Lee, and V. C. M. Leung, "Performance evaluation of an adaptive backoff scheme for WLAN: Research articles," *Wirel. Commun. Mob. Comput.*, vol. 4, pp. 867–879, 2004.

[22] V. Bharghavan, A. J. Demers, S. Shenker, and L. Zhang, "MACAW: A media access protocol for wireless LAN's," *ACM SIGCOMM*, pp. 212–225, 1994.

[23] B. Bensaou, Y. Wang, and C. C. Ko, "Fair medium access in 802.11 based wireless ad-hoc networks," in *ACM MobiHoc*, 2000, pp. 99–106.

[24] J. G. Ko, Y. H. Cho, and H. Kim, "Performance evaluation of IEEE 802.15.4 MAC with different backoff ranges in wireless sensor networks," in *IEEE ICCS*, 2006, pp. 1–5.

[25] A. C. Pang and H. W. Tseng, "Dynamic backoff for wireless personal networks," in *IEEE GLOBECOM*, 2004, pp. 1580–1584.

[26] S. C. Ergen, "IEEE 802.15.4 summary," Advanced Technology Lab of National Semiconductor, Tech. Rep., 2004.

[27] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Athena Scientific, 1997.

[28] F. Cali, M. Conti, and E. Gregori, "Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit," *IEEE/ACM Transactions on Networking*, vol. 8, pp. 785 – 799, 2006.

[29] P. Park, P. D. Marco, C. Fischione, and K. H. Johansson, "Adaptive IEEE 802.15.4 protocol for reliable and timely constrined communnicaion," KTH, Tech. Rep., 2009.

[30] R. Jain, D. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," Digital Equipment Corporation, Tech. Rep., 1984.

Paper B

Analytical Modeling and Optimization of Duty-cycles in Preamble-based IEEE 802.15.4 Wireless Sensor Networks

C. Fischione, P. Park, S. Coleri Ergen,
K. H. Johansson and A. Sangiovanni-Vincentelli

Submitted to
IEEE/ACM Transactions on Networking

Analytical Modeling and Optimization of Duty-cycles in Preamble-based IEEE 802.15.4 Wireless Sensor Networks

C. Fischione, P. Park, S. Coleri Ergen,
K. H. Johansson and A. Sangiovanni-Vincentelli

Abstract

The efficient design and optimization of duty-cycled wireless sensor networks with preamble sampling random medium access control (MAC) greatly is based on accurate modeling of delay, packet reception probabilities, and energy consumption. The challenges for modeling are the random MAC and the sleep policy of the receivers that makes it impossible to determine the exact time of data packet transmission. A novel approach to the modeling of the delay, reliability, and energy consumption is proposed for a clustered network topology with unslotted IEEE 802.15.4 and preamble sampling MAC. The analysis developed in this paper gives expressions of the delay, reliability and energy consumption as a function of sleep time, listening time, traffic rate and MAC parameters. These expressions can then be effectively used to optimize the duty-cycle of the nodes. The optimization ensures a significant reduction of the energy consumption compared to existing solutions in the literature. Monte Carlo simulations using the ns-2 simulation tool demonstrate the validity of the analysis.

Index Terms–Wireless Sensor Networks, MAC, IEEE 802.15.4, Duty Cycle, Optimization.

1 Introduction

Energy-efficient IEEE 802.15.4 wireless sensor networks (WSNs) are providing new and affordable services for a variety of applications, including home and industrial automation, health-care monitoring, and smart grids [1, 2]. Ensuring energy efficiency is difficult for applications where WSNs must provide information for real-time action, because reliable and timely packet transmission may have a negative impact on energy consumption.

Idle listening of sensors is one of the major components of the energy budget. Duty-cycling has been proposed as an effective mechanism for reducing idle listening (see, e.g., GAF [3], SPAN [4] and S-MAC [5]). The idea is to periodically cycle between a sleep and a listening state. Sleep time and listening time are to be chosen so that energy is minimized while satisfying all the communication constraints. Duty-cycling MAC protocols are of two types: synchronous and asynchronous. Asynchronous duty-cycling Medium Access Control (MAC) protocols such as B-MAC [6], (the standard protocol for TinyOS [7]) and X-MAC [8] reduce idle listening in random access networks. In these protocols, the receiver wakes up periodically to check whether there is a transmission, and the sender transmits preambles to check if the receiver is awake. The main advantage of these protocols is that there is no complex control mechanisms, as in time division multiple access (TDMA) schemes, for discovering the network topology, keeping the nodes synchronized [9] and running the schedules efficiently [10]. Compared to synchronous duty-cycling protocols (e.g., [5, 11–13] and references therein), asynchronous protocols have the advantage of not requiring negotiation of the schedule among neighboring nodes to specify when the nodes are awake and asleep (see [8] for an extensive description of the advantages of asynchronous versus synchronous duty-cycling). However, the intrinsic simplicity of the asynchronous mechanism has the drawback of smaller energy saving potential as compared to the more complex solutions listed above, unless optimization of listening and sleep times is adapted to data traffic and network conditions.

In this paper, we consider the design of an energy efficient asynchronous duty-cycling based on the IEEE 802.15.4 communication standard [1]. Energy modeling and its use in listening and sleep time optimization was considered in B-MAC [6] and X-MAC [8], which can work on top of IEEE 802.15.4 and do not require any modification of the standard. However, these protocols on top of IEEE 802.15.4 do not take into account the effect of random access, which is a function of data traffic, MAC parameters and topology. This is a crucial aspect, since the duration of random access is much larger than the actual packet transmission: In IEEE 802.15.4 [1] radios with default parameter settings, the maximum back-off before packet transmission is 27.4 ms whereas the transmission time of a 56 byte packet is 1.79 ms at 250 kbps. Therefore, the random access may consume significant energy.

The amount of random access should be included in the energy minimization problem, because random access determines the time interval between the transmissions of two consecutive preamble packets. It determines listening time, since the receiver node should receive at least one preamble packet during the listening time. Furthermore, the amount of random access is affected by sleep time, since increasing sleep time increases the number of preambles. Consequently, if random access is not taken explicitly into account, asynchronous duty-cycling protocols experience long delays in packet transmission and may waste substantial energy. Furthermore, in [6, 8] and references therein, no delay or reliability constraint on packet delivery is considered. MAC protocols for sensor networks must have certain latency and reliability requirements in addition to low energy consumption. Since many applications require guaranteed arrival of sensor data to the collection center (e.g. security monitoring) and others require a certain degree of reliability in delivering sensor data (e.g., control and automation applications), latency and reliability must be considered in MAC design.

The original contributions of this paper are two: First, we provide accurate expressions of delay, reliability, and energy consumption as a function of random access, sleep time, listening time, traffic rate and MAC parameters. We demonstrate its validity by both analysis and Monte Carlo simulations using ns-2. Second, we illustrate the use of these formulations in the optimization of duty cycle of the nodes by minimizing energy consumption under latency and reliability constraints. To the best of our knowledge, this is the first approach to provide accurate analytical modeling and optimization of duty cycled networks with latency and reliability requirements.

The rest of the paper is organized as follows: Section 2 gives a description of our system model. Section 3 describes the preamble based protocol. Sections 4, 5 and 6 provide the analytical expressions for delay probability, reliability, and energy consumption of preamble sampling MAC on top of unslotted IEEE 802.15.4 networks, respectively. Section 7 deals with practical implementation aspects of our optimization. Section 8 illustrates the advantage of these models in the optimization of duty cycle parameters.

2 System Model

According to the IEEE 802.15.4 standard, we assume that the nodes of the WSN are organized into clusters (see Fig. 1). In a clustered topology, nodes organize themselves into clusters with a node acting as cluster head. All non-cluster head nodes transmit their data directly to the cluster head, while the cluster head receives data from all cluster members and transmits them to a remote base station. Throughout this paper we consider applications where nodes asynchronously generate packets with rate λ packets per second (see Table 2 for a list of main symbols used in the paper). The protocol we are investigating is referred to the typical low data rate applications using IEEE 802.15.4. Consequently, we assume that $\lambda \leq 1$.

We consider the unslotted IEEE 802.15.4 carrier sense multiple access with collision avoidance (CSMA/CA) mechanism, where each node in the network has two variables: NB and BE. NB is the number of times the CSMA/CA algorithm is required to back-off while attempting the current transmission and BE is the back-off exponent, which is related to how many back-off periods a device must wait before it attempts to assess the channel. The parameters that affect random back-off are $BE_{\min}$, $BE_{\max}$ and $NB_{\max}$, which correspond to the minimum and maximum of BE and the maximum of NB, respectively.

For the network topology and applications we are interested, the asynchronous

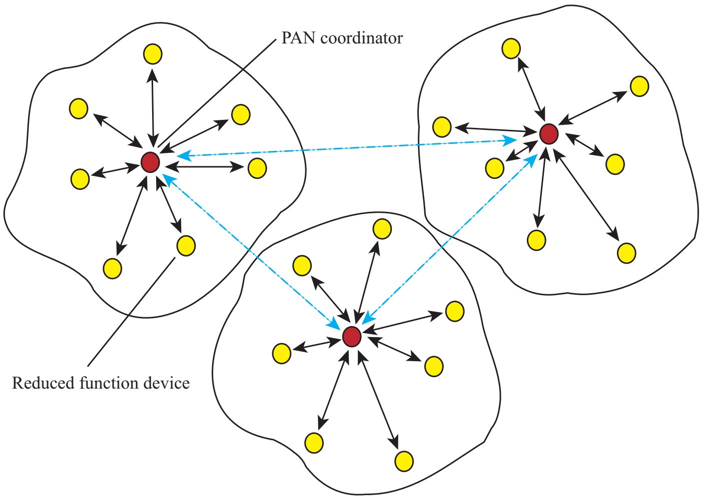


Figure 1: Clustered network topology. The packets generated by the yellow nodes are transmitted toward the cluster-head node depicted in the middle of each cluster.

duty-cycling MAC protocol based on preamble sampling with acknowledgment called X-MAC [8] offers good performance. In preamble sampling protocols, the receiver wakes up periodically for a short time to sample the medium. Such a time is defined as the listening time. When a sender has data, it transmits a series of short preamble packets, each containing the ID of the target node, until it either receives an acknowledgement packet (ACK) from the receiver or a maximum time is exceeded (see Fig. 2). We assume that such a maximum time is given by the sleep plus listening time of the receiver. Following the transmission of each preamble packet, the transmitter node goes in a listening state having a maximum timeout duration $T_{\mathrm{TX,out}}$. If the receiver is the target, it sends an acknowledgement (ACK) during the pause between the preamble packets. When the receiver node sends an ACK, it waits for data packets for a duration of at least T_{out} even after the end of the wake-up time. Consequently, the maximum listening time is $R_l + T_{\mathrm{out}}$. The extension of T_{out} to the regular listening time allows for the reception of the data packets whose ACK was sent near the expiration of R_l. Upon reception of the ACK, the sender transmits the data packet to the receiver. However, the transmission of such a packet occurs after sensing the channel idle. If the channel is busy, data transmission may be delayed too much. The transmitter gives up the transmission of the data packet if the delay from the first attempt to transmit a preamble is larger than $R_s + R_l$.

It is natural that preambles and acknowledgements in preamble sampling protocols are sent by using a random access to avoid collisions, as allowed by IEEE

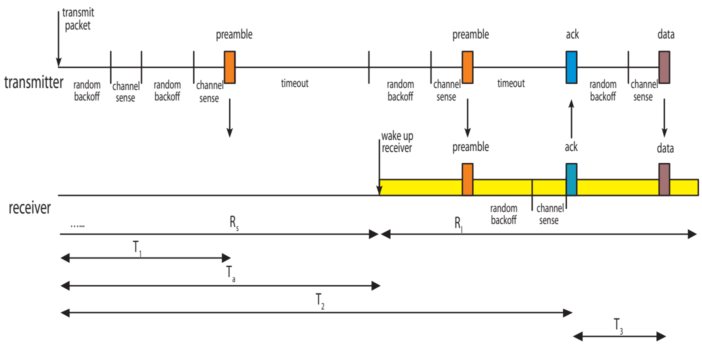


Figure 2: Communication states between a transmitter and a receiver. A random number of preambles are sent before that one falls in the listening period of the receiver. Afterwards, the receiver sends an ACK. When the transmitter hears the ACK, the data packet is sent.

802.15.4 [1]. Although this is not considered in [8], it gives obvious benefits in terms of delay and reliability, since preambles and acknowledgments may collide with any other packet.

3 Protocol Description

The variables established by our protocol are the listening time R_l and sleep time R_s of the receiver node given the channel condition, data traffic, topology of the network, and number of nodes. Consider a cluster of nodes that want to transmit packets to the cluster head. These variables can be obtained as a solution to optimization problems that consider the total energy consumption of the cluster, delay, and reliability in the packet delivery, namely

$$\min_{R_l, R_s} \quad E(R_l, R_s) \tag{1}$$

$$\text{s.t.} \quad D_{\max}(R_l, R_s, t_{\max}) \leq \tau_{\max}\,, \tag{2}$$

$$R_{\min}(R_l, R_s) \geq \phi_{\min}\,. \tag{3}$$

In this problem, $E(R_l, R_s)$ is the total energy consumption of the transmitters and the cluster head, $D_{\max}(R_l, R_s, t_{\max})$ is the delay probability, and $R_{\min}(R_l, R_s)$ is the reliability for the cluster head to receive successfully packets. The term $\tau_{\max}$ is the desired probability that the delay is less than $t_{\max}$, and $\phi_{\min}$ is the minimum desired probability with which a data packet should be received. We define $t_{\max}$, $\tau_{\max}$, and $\phi_{\min}$ as the application requirements. In other words, the application sets the desired $t_{\max}$, $\tau_{\max}$, and $\phi_{\min}$ and then the solution to the optimization problem gives the optimal sleeping and listening times that minimize the total energy. We assume that the two constraints of the optimization problem are associated to the transmitter node that experiences the worst delay and reliability, so that if such a node can transmit packets with the specified $\tau_{\max}$, $t_{\max}$, and $\phi_{\min}$, then all other nodes of the cluster will have the same specification [1]. The decision variables are the listening time R_l and the sleep time R_s of the cluster head. Notice that the optimal selection of these variables minimizes also the energy expenditure of the transmitters. This is because the cost function includes such an energy, which is an implicit function of the duty cycle of the cluster head. Observe that energy, delay, and reliability are also a function of the traffic, the number of nodes, busy channel probability, and the loss probability of preambles or ACK. We did not evidence such a dependency for simplicity of notation.

The optimization problem (1) is motivated by that in many applications, if latency and reliability requirements ($t_{\max}$, $\tau_{\max}$, and $\phi_{\min}$) are not met, the correct execution of decisions concerning the phenomena sensed may be severely compromised. High reliability and low delay may demand significant energy consumption, whereas many applications, e.g., control and actuation applications, can usually tolerate a certain degree of packet losses and delay [14, 15]. Clearly, maximizing reliability and minimizing delay are not in general the optimal design strategies: delay and reliability must be flexible design parameters that need to be just adequate for the application requirements, so that minimization of the total energy is possible to ensure long lifetime of the network while guaranteeing a desired delay and reliability.

We remark here that problem (1) is much more general than the optimization of the duty-cycle performed by X-MAC [8], where the two constraints were not considered, and the random access was not modelled.

To solve the optimization problem (1), we need the expressions of the cost function and constraints, which we develop in the following Sections 4, 5 and 6.

4 Delay Modeling

In this section we tackle the challenging problem of modeling the delay experienced for the successful packet transmission from a transmitter node to the receiver, which we need in Eq. (2). In the following analysis, first we restrict our attention to a transmitter-receiver pair, and then we generalize the analysis to the case of many transmitters in Subsection 4.4.

[1] We could have considered a pair of reliability and delay constraints per each transmitter, thus having a problem with $2N$ constraints and 2 variables, where N is the number of nodes per cluster. Such an optimization problem would have been over-constrained and therefore difficult to solve. We mention that it is easy to check the node experiencing the worst delay and reliability by just checking the packet loss probability from the transmitter to the receiver. We see that our simplification gives quite good results in Section 8

Let us denote by TX the transmitting node, and by RX the receiver node. In the modeling of the delay, we assume that the time is counted from the moment in which the TX node has a packet to send.
The delay to transmit a packet successfully is a function of three random components (see Fig. 2):

- T_1: random delay spent by the TX node to complete the transmission of a preamble packet. It includes also the processing time and the transmission time of the preamble.
- T_2: random delay spent by the TX node until the receiver node is in the listening state and an acknowledgment packet reaches the TX node;
- T_3: random delay spent by the TX node from the instant of acknowledgement reception until the transmission of a data packet. It includes also the processing time and the transmission time of the data packet.

Hence, the delay to transmit successfully a data packet is given $T_p = T_2 + T_3$. In the following, we characterize the three delay components T_1, T_2 and T_3.

4.1 Modeling of T_1

In this subsection, we provide the exact expressions of the average and variance of T_1. Then, we approximate the distribution of T_1 by a normal distribution, whose average and variance are obtained through a moment matching approach. Such an approximation is motivated by that a closed form expression for the distribution of T_1 cannot be achieved, as we will discuss later. We will show that the approximation is quite accurate.
The mechanism to transmit a preamble packet is the same as the one for data packets, for we are assuming to use IEEE 802.15.4. If the channel is busy, a random back off is spent before a further trial. Let $NB_{\max} \leq N_b \leq R_s/S_c$ be the maximum number of back-off of a preamble, namely the number of times that the TX node attempts to access the channel before giving up the transmission of a preamble, where S_c is the sensing time. By denoting with $S_{p,j}$ the random back-off time at the j-th trial, it follows that $S_{p,j}$ has a uniform distribution in the interval $[0, 2^{r(j)} - 1]$, for $j = 1, \ldots, N_b$, where $r(j) = \min(\text{rem}\,(j, NB_{\max}) + BE_{\min} - 1, BE_{\max})$, with $\text{rem}(\cdot,\cdot)$ being the remainder of the division of the first by the second argument.
Denote by $\mathcal{A}_k$ the event occurring when attempting to send a preamble the channel is busy for $k-1$ times, and then is free at the k-th time. The probability of such an event is

$$\Pr[\mathcal{A}_k] = \beta^{k-1}(1-\beta)\,,$$

where β is the busy channel probability. We assume that β is independent at each attempt. Such an assumption is quite accurate for saturated traffic in [16], and has been widely adopted in the literature also for unsaturated traffic (see, e.g., [17–19] and references therein). In Sections 4, 5, and 6 we show by Monte Carlo simulations that this approximation is accurate within the operational region of WSNs. Consider the attempt of transmission of the i-th preamble. Then, random delay T_1 spent by the TX node before transmitting a preamble packet within N_b attempts can be described as

$$T_1 = \begin{cases} S_{p,1} + S_c + T_{\text{hr}}, & \text{if } \mathcal{A}_1|\mathcal{A}; \\ S_{p,1} + S_c + S_{p,2} + S_c + T_{\text{hr}}, & \text{if } \mathcal{A}_2|\mathcal{A}\ ; \\ \vdots & \\ \sum_{j=1}^{N_b}(S_{p,j} + S_c) + T_{\text{hr}}, & \text{if } \mathcal{A}_{N_b}|\mathcal{A}. \end{cases}$$

where T_{hr} is the time employed by the hardware platform to process the packets and transmit them, and $\mathcal{A}$ is the event that a preamble is transmitted with at maximum N_b preambles:

$$\Pr[\mathcal{A}] = \Pr\left[\sum_{j=1}^{N_b} \mathcal{A}_j\right] = \sum_{j=1}^{N_b} \Pr[\mathcal{A}_j]\,,$$

where previous inequality comes from that the events $\mathcal{A}_j$, $j = 1, \ldots, N_b$ are mutually exclusive. It holds

$$\begin{aligned} \Pr[\mathcal{A}_k|\mathcal{A}] &= \frac{\Pr\left[\mathcal{A}_k \sum_{j=1}^{N_b} \mathcal{A}_j\right]}{\Pr[\mathcal{A}]} = \frac{\Pr\left[\mathcal{A}_k\right]}{\sum_{j=1}^{N_b} \Pr\left[\mathcal{A}_j\right]} \\ &= \frac{\beta^{k-1}}{\sum_{j=1}^{Nb} \beta^{j-1}}\,, \end{aligned}$$

We can rewrite T_1 as

$$T_1 = \sum_{k=1}^{N_b}\left[\sum_{j=1}^{k}(S_{p,j} + S_c) + T_{\text{hr}}\right]\mathbb{1}_{\mathcal{A}_k|\mathcal{A}} = \sum_{j=1}^{N_b} \Sigma_k \mathbb{1}_{\mathcal{A}_k|\mathcal{A}}\,, \tag{4}$$

where $\mathbb{1}_{(\cdot)}$ it the indicator function (its value is 1 if the argument is true, and 0 otherwise) and

$$\Sigma_k = \sum_{j=1}^{k}(S_{p,j} + S_c) + T_{\text{hr}}\,.$$

From previous equation, Σ_k is given by the sum of independent uniformly distributed random variables plus a constant. The expectation Σ_k can be computed by recalling the distribution of $S_{p,j}$, whose average is

$$\mu_{S_{p,j}} = \frac{\left(2^{r(j)} - 1\right) S_b}{2}\,,$$

where $S_b = \mathit{aUnitback\text{-}offPeriod}$ is the back-off period [1]. Hence

$$\mu_{\Sigma_k} = \mathbb{E}\left[\Sigma_k\right] = \sum_{j=1}^{k} \left[\mu_{S_{p,j}} + S_c\right] + T_{\mathrm{hr}}. \tag{5}$$

The variance of Σ_k is given by the sum of the variances of $S_{p,j}$, which is

$$\sigma^2_{S_{p,j}} = \frac{\left(2^{2r(j)} - 1\right) S_b^2}{12},$$

hence

$$\sigma^2_{\Sigma_k} = \mathbb{E}\left[\Sigma_k - \mathbb{E}\,\Sigma_k\right]^2 = \sum_{j=1}^{k} \sigma^2_{S_{p,j}}. \tag{6}$$

Using (5) and (6) it is possible to compute the exact expression of the average value and the correlation of T_1 as

$$\mu_{T_1} = \mathbb{E}\,T_1 = \sum_{k=1}^{N_b} \frac{\mu_{\Sigma_k}\beta^{k-1}}{\sum_{j=1}^{Nb}\beta^{j-1}},$$

$$\rho_{T_1} = \mathbb{E}\,T_1^2 = \sum_{k=1}^{N_b} \frac{\rho_{\Sigma_k}\beta^{k-1}}{\sum_{j=1}^{Nb}\beta^{j-1}},$$

where $\rho_{\Sigma_k} = \sigma^2_{\Sigma_k} + \mu^2_{\Sigma_k}$. From these moments, the variance of T_1 follows $\sigma^2_{T_1} \triangleq \rho_{T_1} - \mu^2_{T_1}$. The computation in a closed form of previous expressions can be obtained from the sum of powers [20, pag. 193].
Since T_1 is the weighted sum of uniform random variables having different mean and variance, no closed form expression is available for the probability mass function (PMF). However, we resort to a normal distribution to approximate the PMF of T_1, namely, we assume that

$$f_{T_1}(x) \sim \frac{1}{\sigma_{T_1}\sqrt{2\pi}} \exp\left(-\frac{(x-\mu_{T_1})^2}{2\sigma^2_{T_1}}\right). \tag{7}$$

In Subsection 4.4 we show that this approximation matches well the real one obtained via Monte Carlo simulations. Finally, from (7) we can compute the cumulative distribution function (CDF) of T_1 by the error function. Let $P_1(t)$ be such a CDF:

$$\Pr[T_1 \le t] = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{t - \mu_{T_1}}{\sigma_{T_1}\sqrt{2}}\right)\right) \triangleq P_1(t), \tag{8}$$

where $\mathrm{erf}()$ is the error function. We will use such a probability in the next sections.

4.2 Modeling of T_{ack}

Since the acknowledgement packets are transmitted by following exactly the same mechanism of the preamble packets, the time employed by the RX node to send an acknowledgement packet upon the reception of a preamble packet can be modeled as done for T_1, so that T_{ack} is approximated by a normal distributionwith average

$$\mu_{T_{\text{ack}}} = \sum_{k=1}^{NB_{\max}} \mu_{\Sigma_k} \frac{\beta^{k-1}}{\sum_{k=1}^{NB_{\max}} \beta^{k-1}} ,$$

and correlation

$$\rho_{T_{\text{ack}}} = \sum_{k=1}^{NB_{\max}} \rho_{\Sigma_k} \frac{\beta^{k-1}}{\sum_{k=1}^{NB_{\max}} \beta^{k-1}} ,$$

and $\sigma^2_{T_{\text{ack}}} \triangleq \rho_{T_{\text{ack}}} - \mu^2_{T_{\text{ack}}}$.

4.3 Modeling of T_2

In this section we model T_2, the random delay the TX node waits until an ACK is sent by the RX node and reaches the TX node that sent the preamble. With this goal in mind, we need to define three random variables, T_a, T_l and N_p, which we present next.

First, let us denote by T_a the random time to wait from the beginning of the transmissions until the start of the listen time. The time T_a can be modeled by the following random variable:

$$T_a = \begin{cases} 0, & \text{if } \bar{\mathcal{S}} ; \\ T_s, & \text{if } \mathcal{S} . \end{cases} \tag{9}$$

where T_s is the random time to wait that the receiver wakes up. This time can be modeled as a uniform distribution in the range $[0, R_s]$, since such a time is computed from the beginning of the transmission of the TX node, which may uniformly fall in the interval $[0, R_s]$ (see Fig. 2). The event $\mathcal{S}$ occurs when the RX node is sleeping. Since a node sleeps for R_s seconds and is awake for R_l seconds, it follows that

$$\Pr[\mathcal{S}] = \frac{R_s}{R_s + R_l} , \qquad \Pr[\bar{\mathcal{S}}] = 1 - \Pr[\mathcal{S}] .$$

From the definition (9) we rewrite T_a as $T_a = 0\mathbb{1}_{\bar{\mathcal{S}}} + T_s\mathbb{1}_{\mathcal{S}} = R_s\mathbb{1}_{\mathcal{S}}$. It follows that the PMF of T_a is

$$\Pr[T_a] = \Pr[T_s] \Pr[\mathcal{S}] = \Pr[T_s] \frac{R_s}{R_s + R_l} .$$

We define by T_l the time interval from the moment wherein the preamble packet is received in the listening time of the receiver, until the listening time expires. By following the same approach as the one used for the characterization of T_s, it follows that T_l has a uniform distribution in the interval $[0, R_l]$.
Consider N_p, the random number of preambles that should be sent before one falls in the active time of the receiver and the acknowledgment is sent back by the RX node. Denote with T_{ack} the random time to complete the transmission of an acknowledgement sent by the RX node after a preamble packet is received. Notice that its statistical distribution is the same as T_1, since an acknowledgement is transmitted by following the same mechanism of a preamble, the only difference being that N_b must be replaced with $NB_{\max}$. Furthermore, let us define $\mathcal{B}_k$ as the event that a preamble has to be sent k times before being received in the active time of the RX node and the corresponding acknowledgement is sent by the RX node and received before the time out of the TX node. We assume that the event $\mathcal{B}_k$ is conditioned on the random active time T_a of the RX and on the random remaining listening time T_l of the RX. These times are random from the point of view of the transmitter, which does not know when the receiver wakes up and when it will go to sleep.
We are now in the position of defining the delay T_2:

$$T_2 = \begin{cases} T_{1,1} + T_{\mathrm{ack}}, & \text{if } \mathcal{B}_1 | \mathcal{B}; \\ T_{1,1} + T_{\mathrm{TX,out}} + T_{1,2} + T_{\mathrm{ack}}, & \text{if } \mathcal{B}_2 | \mathcal{B}\ ; \\ \vdots & \\ \sum_{j=1}^{N_p} T_{1,j} + (N_p - 1) T_{\mathrm{TX,out}} + T_{\mathrm{ack}}, & \text{if } \mathcal{B}_{N_p} | \mathcal{B}. \end{cases}$$

where $T_{1,j}$ is the delay for the transmission of the j-th preamble. The distribution of $T_{1,j}$ is given by (4). $\mathcal{B}$ is the probability that the TX node receives an ACK within N_p preambles:

$$\Pr[\mathcal{B}] = \Pr\left[\sum_{k=1}^{N_p} \mathcal{B}_k\right] = \sum_{k=1}^{N_p} \Pr[\mathcal{B}_k]\,,$$

where previous inequality comes from that the events $\mathcal{B}_j$, $j = 1, \ldots, N_p$ are mutually exclusive. It holds

$$\Pr[\mathcal{B}_l | \mathcal{B}] = \frac{\Pr\left[\mathcal{B}_l \sum_{k=1}^{N_b} \mathcal{B}_k\right]}{\Pr\left[\mathcal{B}\right]} = \frac{\Pr\left[\mathcal{B}_l\right]}{\sum_{j=1}^{N_p} \Pr\left[\mathcal{B}_k\right]}\,.$$

We describe $\mathcal{B}_k$ next. First, let α be the loss probability of preambles or ACK due to bad channel or collisions. Such a probability is different from the loss probability of data packets, which we denote by p, because the size of preambles and ACK is much smaller than data packets. We assume that these probabilities are independent at each attempt. Such an approximation has been widely adopted in the literature (see, e.g., [16–19] and references therein). In Sections 4, 5, and 6 we show by Monte Carlo simulations that this approximation is quite accurate within the operational region of WSNs.

PROPOSITION 1 *Let $\mathcal{B}_k$, with $k \in \mathbb{N}$, the event occurring when $k-1$ preambles are sent before the k-th is received in the active time of the RX node, and the acknowledgement is sent back by the RX node and received before the time out of the TX node. Let Ω be the certain event. Then*

$$\begin{aligned}\mathcal{B}_k = \big[&\mathcal{C}_k + \mathcal{D}_{k-1}\mathcal{E}_{k-1}\,\alpha + \mathcal{D}_{k-1}\mathcal{E}_{k-1}(1-\alpha)\bar{\mathcal{F}}_{k-1} \\ &+\mathcal{D}_{k-1}\mathcal{E}_{k-1}(1-\alpha)\mathcal{F}_{k-1}\,\alpha\big]\,\mathcal{D}_k\mathcal{E}_k\mathcal{F}_k(1-\alpha)^2\,, \end{aligned} \tag{10}$$

where

$$\begin{aligned} \mathcal{C}_k &= [(k-1)\mathbb{1}_{k-1\geq 0}T_1 + (k-2)\mathbb{1}_{k-2\geq 0}T_{\mathrm{TX,out}} \leq T_a]\,, \\ \mathcal{D}_k &= [kT_1 + (k-1)\mathbb{1}_{k-1\geq 0}T_{\mathrm{TX,out}} > T_a]\,, \\ \mathcal{E}_k &= [kT_1 + (k-1)\mathbb{1}_{k-1\geq 0}T_{\mathrm{TX,out}} \leq T_a + T_l]\,, \\ \bar{\mathcal{F}}_{k-1} &= [T_{\mathrm{ack}} > T_{\mathrm{TX,out}}|\mathcal{D}_{k-1}]\,, \\ \mathcal{F}_k &= [T_{\mathrm{ack}} \leq T_{\mathrm{TX,out}}|\mathcal{D}_k]\,, \\ \bar{\mathcal{F}}_0 &= [T_{\mathrm{ack}} > T_{\mathrm{TX,out}}]\,, \\ \mathcal{D}_0 &= \Omega\,. \end{aligned} \tag{11}$$

Proof: See Appendix A.1. ∎

PROPOSITION 2

$$\begin{aligned} \Pr[\mathcal{B}_k] = &\big(\Pr[\mathcal{C}_k\mathcal{E}_k] - \Pr[\bar{\mathcal{D}}_k]\big)\Pr[T_{\mathrm{ack}} \leq T_{\mathrm{TX,out}}] \\ &\times (1-\alpha)^2 + (\Pr[\mathcal{E}_k] - \Pr[\mathcal{C}_k\mathcal{E}_k]) \\ &\times \Pr[T_{\mathrm{ack}} \leq T_{\mathrm{TX,out}}]\alpha(1-\alpha)^2 \\ &+ (\Pr[\mathcal{E}_k] - \Pr[\mathcal{C}_k\mathcal{E}_k])\,(1 - \Pr[T_{\mathrm{ack}} \leq T_{\mathrm{TX,out}}]) \\ &\times \Pr[T_{\mathrm{ack}} \leq T_{\mathrm{TX,out}}](1-\alpha)^3 \\ &+ (\Pr[\mathcal{E}_k] - \Pr[\mathcal{C}_k\mathcal{E}_k]) \\ &\times (\Pr[T_{\mathrm{ack}} \leq T_{\mathrm{TX,out}}])^2\,\alpha(1-\alpha)^3\,, \end{aligned}$$

where

$$
\begin{aligned}
\Pr[\mathcal{C}_k] &= P_1\left(\frac{T_a-(k-2)T_{\mathrm{TX,out}}}{k-1}\right),\\
\Pr[\bar{\mathcal{D}}_k] &= P_1\left(\frac{T_a-(k-1)T_{\mathrm{TX,out}}}{k}\right),\\
\Pr[\mathcal{E}_k] &= P_1\left(\frac{T_a+T_l-(k-1)T_{\mathrm{TX,out}}}{k}\right),\\
\Pr[\mathcal{C}_k\mathcal{E}_k] &= \Pr[\mathcal{C}_k]\Pr[T_1\le T_l-T_{\mathrm{TX,out}}]\\
&\quad+\Pr[\mathcal{E}_k](1-\Pr[T_1\le T_l-T_{\mathrm{TX,out}}]).
\end{aligned}
$$

Proof: See Appendix A.2. ■

PROPOSITION 3 *The maximum number of preambles to send is*

$$N_p = 2+\frac{R_s}{T_{\mathrm{TX,out}}}.$$

Proof: The set (10) occurs with probability zero when k grows so large that either $\mathcal{C}_k$ or $\mathcal{E}_k$ equal the impossible event with probability 1. N_p is the maximum of such k.
Recalling that the probability of these events is defined over positive distributions, it follows that $\mathcal{C}_k$ and $\mathcal{E}_k$ occur with non-zero probability if $T_a-(k-2)T_{\mathrm{TX,out}}\ge 0$ and $T_a+T_l-(k-1)T_{\mathrm{TX,out}}\ge 0$, whereby

$$
\begin{aligned}
k \le& 2+\frac{T_a}{T_{\mathrm{TX,out}}}\le 2+\frac{R_s}{T_{\mathrm{TX,out}}},\\
k \le& 1+\frac{T_a}{T_{\mathrm{TX,out}}}+\frac{T_l}{T_{\mathrm{TX,out}}}\le 1+\frac{R_s}{T_{\mathrm{TX,out}}}+\frac{T_l}{T_{\mathrm{TX,out}}}.
\end{aligned}
$$

Since $T_l > T_{\mathrm{TX,out}}$, we obtain that

$$
\begin{aligned}
&\min\left(2+R_s/T_{\mathrm{TX,out}},1+R_s/T_{\mathrm{TX,out}}+T_l/T_{\mathrm{TX,out}}\right)\\
&\quad=2+\frac{R_s}{T_{\mathrm{TX,out}}},
\end{aligned}
$$

which concludes the proof. ■

From Propositions 1 and 3, the average and variance of T_2 is

$$
\begin{aligned}
\mu_{T_2} &= \sum_{k=1}^{N_p}\left[k\mu_{T_1}+(k-1)T_{\mathrm{TX,out}}+\mu_{T_{\mathrm{ack}}}\right]\Pr[\mathcal{B}_l|\mathcal{B}],\\
\sigma^2_{T_2} &= \sum_{k=1}^{N_p}\sigma^2_{T_{2,k}}\Pr[\mathcal{B}_l|\mathcal{B}],
\end{aligned}
$$

where $\sigma^2_{T_{2,k}}$ is the variance of $\sum_{j=1}^{k} T_{1,j} + (k-1)T_{\text{TX,out}} + T_{\text{ack}}$.
Since T_2 is given by the weighted sum of variables that we approximated in Subsection 4.1 as normal distributed, it follows that the PMF of T_2 can be approximated by a normal random variable. In Subsection 4.4 we show that this approximation is quite accurate.

4.4 Delay Probability

The delay to send successfully a data packet is given by $T_p = T_2 + T_3$, given T_a and T_l. In particular, T_2 is a function of T_a and T_l (see Subsection 4.3). Looking at Fig. 2, it is straightforward to see that T_3 can be characterized as T_{ack} except for a higher constant transmission time within T_{hr}, so that T_3 is approximated by a normal random variable. It follows that T_p is approximated by a normal distribution as well, with average $\mu_{T_p} = \mu_{T_2} + \mu_{T_3}$, and variance $\sigma^2_{T_p} = \sigma^2_{T_2} + \sigma^2_{T_3}$.
The distribution of the delay we have modeled so far is conditioned on the active time T_a of the receiver and the time interval T_l from preamble reception in the listening time of the receiver until the listening time expires, since the event $\mathcal{B}_k$ is conditioned on these times (see Subsection 4.1 and Eq. (11)). Therefore, the probability that a packet is delayed some $t_{\max}$ seconds and falls in the listening time of the RX node is given by

$$D_{\max}(R_l, R_s, t_{\max}) \triangleq \mathbb{E}_{T_a} \mathbb{E}_{T_l} \Pr[(T_p \leq t_{\max})] , \qquad (12)$$

where $\mathbb{E}_{T_a}$ and $\mathbb{E}_{T_l}$ denote the statistical average with respect to the distribution of T_a and T_l, respectively. Since the CDF of T_p is given by a cumulative Gaussian distribution, it is a highly nonlinear function of the random variables T_a and T_l. Therefore, the averages $\mathbb{E}_{T_a} \mathbb{E}_{T_l}$ are obtained by replacing T_a and T_l with their respective expectations, as proposed in [20, pag. 428]. This is equivalent to replacing μ_{T_p} with $\mathbb{E}_{T_a} \mathbb{E}_{T_l} \mu_{T_p}$ and $\sigma^2_{T_p}$ with $\mathbb{E}_{T_a} \mathbb{E}_{T_l} \sigma^2_{T_p}$ in the argument of the cumulative Gaussian distribution of T_p. In the notation adopted for (12), we remarked that the reliability depends on the listening and sleep times R_l and R_s and the maximum desired delay $t_{\max}$.
We validated the analysis of the delay by comparing the expectation and variance of (12) to extensive Monte Carlo simulations obtained by an ns-2 simulator. The simulator reproduced the system depicted in Fig. 1, where transmitter nodes send packets according to the preamble-based MAC. All the numerical values set for the simulations are taken coherently with the IEEE 802.15.4 standard [1] with default MAC parameters and the Tmote sensors [21]. Each simulation result was computed by running simulations to reproduce 20000 seconds of real time, and five simulations were run to remove the dependance on the initial seed of the random generators. The delay analysis was then evaluated by collecting results from these simulations, as we discuss below.
Although the expectation and variance of (12) have been derived for a single transmitter-receiver pair, we consider it as an approximation for the general case

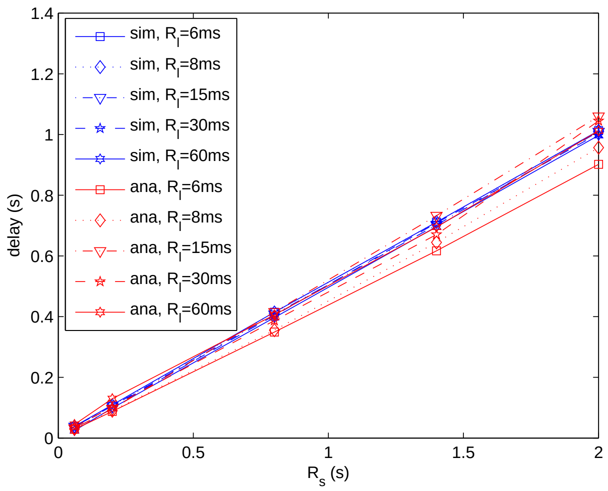


Figure 3: Average delay to send successfully a data packet $\mathbb{E}_{T_a}\mathbb{E}_{T_l}\mu_{T_p}$ as obtained by analysis and simulations for a network with $N = 8$ nodes and traffic period $1/\lambda = 30$ s. On the x axis, the sleep time R_s is reported.

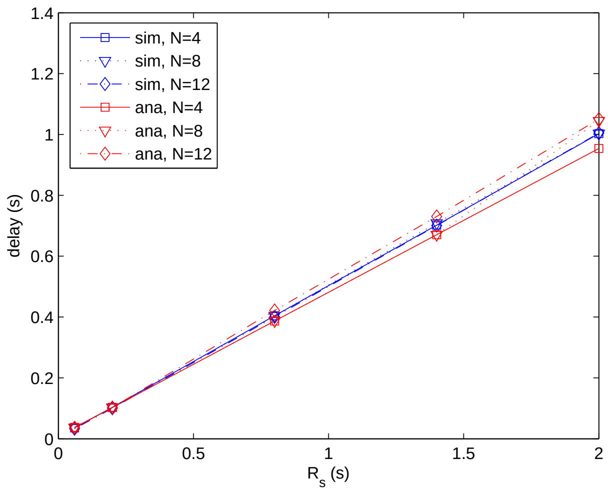


Figure 4: Average delay to send successfully a data packet $\mathbb{E}_{T_a}\mathbb{E}_{T_l}\mu_{T_p}$ as obtained by analysis and simulations for a network with fixed listen time $R_l =$ 60 ms, $N = 4, 8, 12$ nodes and traffic period $1/\lambda = 30$ s. The case with $N = 8$ considers the hidden terminal problem.

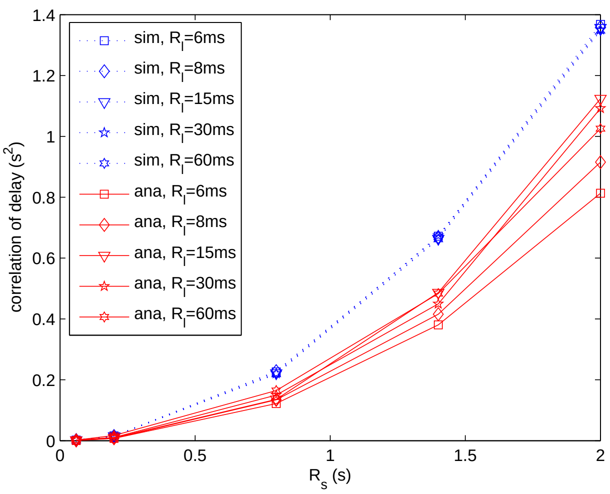


Figure 5: Variance of the delay to send successfully a data packet $\mathbb{E}_{T_a}\,\mathbb{E}_{T_l}\sigma^2_{T_p}$ as obtained by analysis and simulations for a network with different listen time $R_l \geq 6$ ms, $N = 8$ nodes and traffic period 30 s. On the x axis, the sleep time R_s is reported.

of several transmitters. This is motivated by that the analysis considers the loss and busy channel probabilities, which accounts for the case of multiple transmitters. As a matter of fact, we observed good matching analysis-simulations for all the cases of practical interest. Figs. 3 and 4 show the average delay for different listening and sleep times, different number of nodes and with hidden node terminals. We chose a traffic period larger than 10 s since higher traffic rates would exhibit packet loss probabilities larger than 50%, which is of no-interest. A good linear relationship between delay and sleep time can be inferred from the Monte Carlo simulations since the packet transmission time and wake time are very short compared to the sleep time. This approximation is valid only when $R_s \geq R_l$ for all the cases considered. However, this is not a limitation, because, to save energy, sensors have to use duty-cycles much smaller than 50%, which is perfectly compatible with $R_s \geq R_l$. Figs. 4 present that the impacts of the different number of nodes and hidden nodes are negligible in terms of delay for successfully received packets. This is due to the low data rate. In Fig. 5, report the variance of the delay. The analysis gives a matching less accurate then the average delay, but we will see in the next sections that the achieved accuracy is satisfactorily for optimization purposes. The same observation holds for a different number of nodes, with hidden node terminals, and different traffic generation rates. We conclude that the analysis of the delay is satisfactorily both for the single transmitter-receiver pair, and for multiple transmitters.

5 Reliability Analysis

In this section, we analyze the reliability, or probability that a data packet is successfully received.
The failure of a data packet transmission is owed to three possibilities: 1) a preamble is not successfully received, 2) the ACK is not successfully received, and 3) the data packet is not successfully received. In the following, we characterize these events.
In Proposition 1, we defined $\mathcal{B}_k$, with $k \in \mathbb{N}$, as the event occurring when $k-1$ preambles are sent before the k-th is received in the active time of the RX node, and the acknowledgement is sent back by the RX node and received before the time out of the TX node. For analytical tractability, $\mathcal{B}_k$ was derived for a single transmitter-receiver pair. However, we assume to use it also for the derivation of the reliability in the general case of several transmitters. Such an assumption is reasonable, as we will show by extensive Monte Carlo simulations presented at the end of this Section.
Let the event $\mathcal{G}$ occur when a preamble is successfully transmitted during the active time of the receiver within N_b trials (which occurs with probability $\Pr[\mathcal{J}_p] = 1-\beta^{N_b}$) and the corresponding ACK is successfully sent within $NB_{\max}$ trials (which occurs with probability $\Pr[\mathcal{J}_a] = 1-\beta^{NB_{\max}}$). Then

$$\mathcal{G}|\mathcal{J}_p\mathcal{J}_a = \sum_{k=1}^{N_p} \mathcal{B}_k \,. \tag{13}$$

By observing that $\mathcal{B}_k$ and $\mathcal{B}_j$ are mutually exclusive if $i \neq j$, it follows

$$\Pr[\mathcal{G}] = (1-\beta^{N_b})(1-\beta^{NB_{\max}}) \sum_{k=1}^{N_p} \Pr\left[\mathcal{B}_k\right] \,. \tag{14}$$

Define the event $\mathcal{I}|\mathcal{G}$, which occurs when the TX sends successfully a data packet, provided that a preamble is successfully received and the ACK is also successfully received, then

$$\Pr[\mathcal{I}|\mathcal{G}] = \left(1-\beta^{NB_{\max}}\right)(1-p) \,. \tag{15}$$

Finally, by putting together (13) and (15), and averaging with respect to the distribution of T_a and T_l, the reliability is given by

$$R_{\min}(R_l, R_s) = \mathbb{E}_{T_a}\,\mathbb{E}_{T_l} \Pr[\mathcal{G}] \Pr[\mathcal{I}|\mathcal{G}] \,. \tag{16}$$

These expectations are computed as done in Eq. (12). In the notation, we remarked that the reliability depends on the listening and sleep times R_l and R_s.

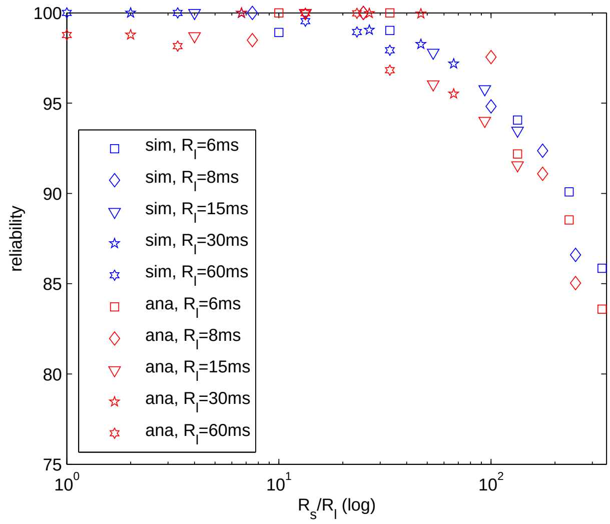


Figure 6: Reliability as obtained by Eq. (16) and simulations for a network with $N = 8$ nodes and traffic period $1/\lambda = 30$ s. On the x axis, the ratio R_s/R_l is reported.

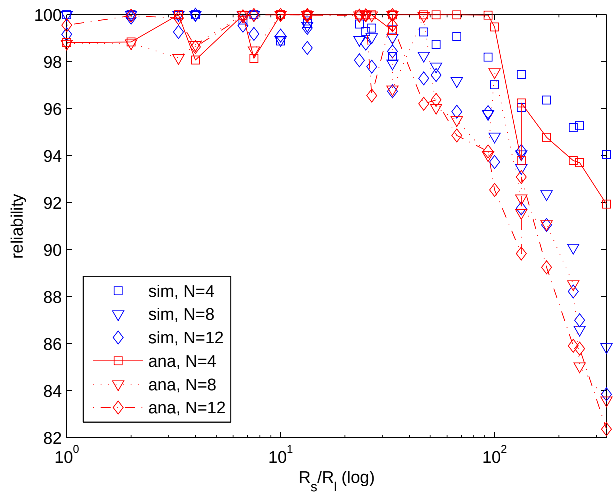


Figure 7: Reliability as obtained by by Eq. (16) and simulations for a network with $N = 4, 8, 12$ nodes and traffic period $1/\lambda = 30$ s.

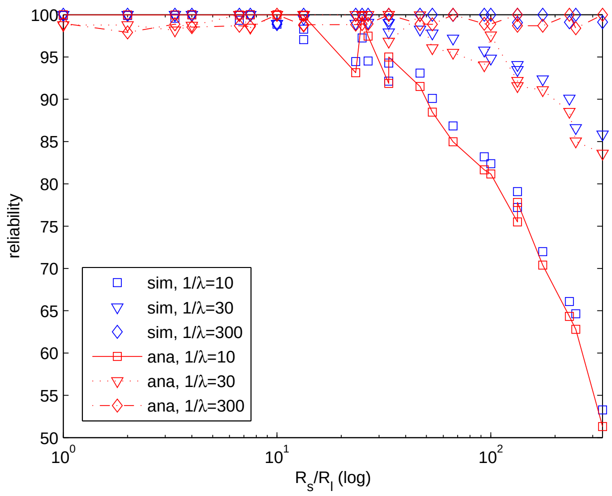


Figure 8: Reliability as obtained by Eq. (16) and simulations for a network with $N = 8$ nodes and traffic period $1/\lambda = 10, 30, 300$ s.

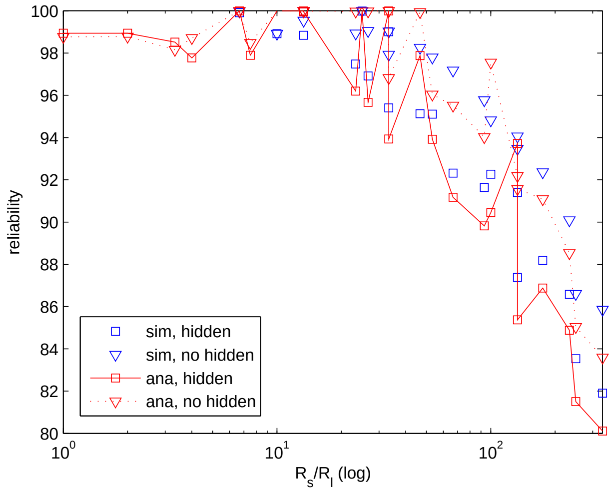


Figure 9: Reliability as obtained by Eq. (16) and simulations for a network with $N = 8$ nodes and traffic period $1/\lambda = 30$ s with and without hidden node terminal

We validated the analysis of the reliability by comparing Eq. (16) to extensive Monte Carlo simulations that were obtained by ns-2, as described in Subsection 4.4. In Fig. 6, we reported an example of such simulations as a function of different listen time for the case of 8 nodes. We see that the analysis follows well the simulations results. The sharp fluctuations are due to the PMF of the random back-off. The PMF is indeed discontinuous with sudden jumps due to the discrete increase of the exponential back-off and the magnitude of the busy channel probability. These sharp fluctuations of the PMF are also typical in IEEE 802.11 [22]. Figs. 7 and 8 show that the reliability decreases as the number of nodes and packet generation rate increase. In addition, Fig. 9 reports the impacts of hidden nodes on reliability that the reliability decreases much faster as the number of hidden nodes increases. We observed that the difference analysis-simulations is always below 5%, so we conclude that Eq. (16) is a good approximation. As the ratio of sleep time to wake time increases, the expected number of preambles increases, which in turn increases the total traffic. This decreases the network reliability by the resulting increase in collisions and random back-off.

6 Energy Consumption

In this section we characterize the energy consumption of the network. The total normalized energy consumption over a listening-sleeping time $R_s + R_l$ is calculated considering the energy spent by a TX node to send a data packet ($\mathbb{E}\, E_{\rm tx}$) and by the RX node to receive a data packet ($\mathbb{E}\, E_{\rm rx}$):

$$\mathbb{E}\, E_{\rm tot} = \frac{d_{\rm TX}\, \mathbb{E}\, E_{\rm tx} + \mathbb{E}\, E_{\rm rx}}{R_s + R_l}\,, \tag{17}$$

where $d_{\rm TX}$ is the probability that a TX node has at least one data packet to send during the time $R_l + R_s$:

$$d_{\rm TX} = 1 - e^{-\lambda(R_s + R_l)}\,.$$

The total energy consumption was normalized by $R_s + R_l$ so that it is intended as an energy per time unit. The remaining energy components in (17) are characterized in the following.

6.1 Energy Consumption at the Transmitters

Denote by $P_{\rm tx}$, $P_{\rm rx}$ and P_s the power required to transmit, receive and sleep, respectively. We have the following results:

CLAIM 4 *Let $S^n_{p,k}$ be the k-th random back-off of the n-th preamble. The instantaneous transmit energy for a TX node is upper bounded as*

$$\begin{aligned} E_{\mathrm{tx}} \leq & \sum_{i=1}^{N_p} \left[\sum_{n=1}^{i} E^{(n)}_{\mathrm{tx},T_1} + (i-1) E_{T_{\mathrm{TX,out}}} + E_{\mathrm{tx},T_{\mathrm{ack}}} + \right. \\ & \left. E_{\mathrm{tx},T_{\mathrm{data}}} \right] \mathbb{1}_{\mathcal{B}_i} \qquad (18) \\ & + \left[\sum_{n=1}^{N_p} E^{n}_{\mathrm{tx},T_1} + (N_p - 1) E_{T_{\mathrm{TX,out}}} + E_{\mathrm{tx},T_{\mathrm{ack}}} \right] \mathbb{1}_{\bar{\mathcal{B}}} \end{aligned}$$

where

$$E^{(n)}_{\mathrm{tx},T_1} = \sum_{j=1}^{N_b} \left[\sum_{k=1}^{j} \left(P_s S^n_{p,k} + P_{\mathrm{rx}} S_c \right) + P_{\mathrm{tx}} S_p \right] \mathbb{1}_{\mathcal{A}_j} \,, \qquad (19)$$

$$E_{T_{\mathrm{TX,out}}} = T_{\mathrm{TX,out}} P_{\mathrm{rx}} \,, \qquad (20)$$

$$E_{\mathrm{tx},T_{\mathrm{ack}}} = \sum_{j=1}^{NB_{\max}} \left[\sum_{k=1}^{j} (S_{p,k} + S_c) + S_a \right] P_{\mathrm{rx}} \mathbb{1}_{\mathcal{A}_j} \,, \qquad (21)$$

$$\begin{aligned} E_{\mathrm{tx},T_{\mathrm{data}}} = & \sum_{j=1}^{NB_{\max}} \left[\sum_{k=1}^{j} (P_s S_{p,k} + P_{\mathrm{rx}} S_c) + P_{\mathrm{tx}} S_d \right] \mathbb{1}_{\mathcal{A}_j} \\ & + \left[\sum_{k=1}^{NB_{\max}} (P_s S_{p,k} + P_{\mathrm{rx}} S_c) \right] \mathbb{1}_{\bar{\mathcal{A}}} \,. \qquad (22) \end{aligned}$$

Proof: See Appendix A.3. ∎

REMARK 1 We derived an upper bound for analytical tractability. The upper bound in Eq. (18) is given by considering the worst case in the number of preambles to be transmitted when no idle channel is found. We see in Subsection (6.3) that such a bound is reasonable.

Using previous claim, we can compute easily the average energy to transmit a data packet.

CLAIM 5 *The average transmit energy per TX node with respect to the random* *back off is upper-bounded by*

$$
\begin{aligned}
\mathbb{E}\,E_{\mathrm{tx}} \leq \sum_{i=1}^{N_p} & \left[i\,\mathbb{E}\,E_{\mathrm{tx},T_1} + (i-1)\,\mathbb{E}\,E_{T_{\mathrm{TX,out}}} + \mathbb{E}\,E_{\mathrm{tx},T_{\mathrm{ack}}}\right. \\
& \left. + \mathbb{E}\,E_{\mathrm{tx},T_{\mathrm{data}}}\right]\Pr[\mathcal{B}_i] + \left[N_p\,\mathbb{E}\,E_{\mathrm{tx},T_1}\right. \\
& \left. + (N_p-1)\,\mathbb{E}\,E_{T_{\mathrm{TX,out}}} + \mathbb{E}\,E_{\mathrm{tx},T_{\mathrm{ack}}}\right] \\
& \times \left(1 - \sum_{i=1}^{N_p}\Pr[\mathcal{B}_i]\right) . \qquad (23)
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbb{E}\,E_{\mathrm{tx},T_1} &= \mathbb{E}\,E^n_{\mathrm{tx},T_1} \qquad (24) \\
&= \sum_{j=1}^{N_b}\left[\sum_{k=1}^{j}\left(P_s\mu_{S_{p,k}} + P_{\mathrm{rx}}S_c\right) + P_{\mathrm{tx}}S_p\right]\Pr[\mathcal{A}_j] \, ,
\end{aligned}
$$

$$
E_{T_{\mathrm{TX,out}}} = T_{\mathrm{TX,out}}P_{\mathrm{rx}} \, , \qquad (25)
$$

$$
\mathbb{E}\,E_{\mathrm{tx},T_{\mathrm{ack}}} = \sum_{j=1}^{NB_{\max}}\left[\sum_{k=1}^{j}\left(\mu_{S_{p,k}} + S_c\right) + S_a\right]P_{\mathrm{rx}}\Pr[\mathcal{A}_j] \, , \qquad (26)
$$

$$
\begin{aligned}
&\mathbb{E}\,E_{\mathrm{tx},T_{\mathrm{data}}} \\
&\quad = \sum_{j=1}^{NB_{\max}}\left[\sum_{k=1}^{j}\left(P_s\mu_{S_{p,k}} + P_{\mathrm{rx}}S_c\right) + P_{\mathrm{tx}}S_d\right]\Pr[\mathcal{A}_j] \\
&\quad + \left[\sum_{k=1}^{NB_{\max}}\left(P_s\mu_{S_{p,k}} + P_{\mathrm{rx}}S_c\right)\right]\Pr\left[\bar{\mathcal{A}}\right] \, , \qquad (27)
\end{aligned}
$$

Proof: (Sketch) Eq. (23) is computed by applying the linear properties of the expectation operator to (18). ■

6.2 Energy Consumption at the Receiver

The energy consumed at the receiver is upper bounded by

$$
\mathbb{E}\,E_{\mathrm{rx}} \leq R_sP_s + (R_l + T_{\mathrm{out}})\max(P_{\mathrm{tx}}, P_{\mathrm{rx}}) \, . \qquad (28)
$$

where we considered that the RX can be listening for a time T_{out} after the end of the listening time if an acknowledgement was sent just before the end of the

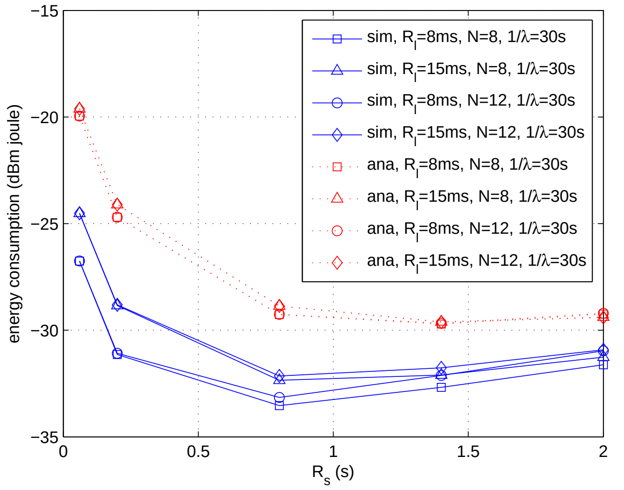


Figure 10: Average energy consumption as obtained by Eq. (29) and simulations for a traffic rate of $1/\lambda = 30\,\text{s}$ and $8, 12$ nodes with different listening time $R_l = 8, 15$ ms. On the x axis, the sleep time R_s is reported.

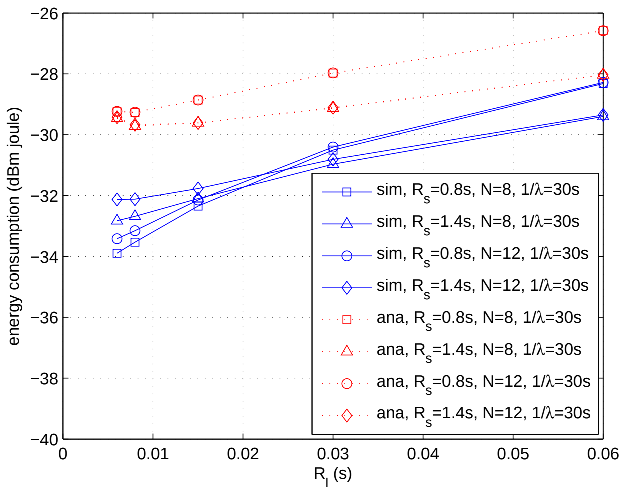


Figure 11: Average energy consumption as obtained by Eq. (29) and simulations for a traffic rate of $1/\lambda = 30$ s and $8, 12$ nodes with different sleep time $R_s = 0.8, 1.4$ s. On the x axis, the listen time R_l is reported.

listening time. The upper bound for $\mathbb{E}\,E_{\rm rx}$ is motivated by that such an energy is given by the idle listening, sending an acknowledgement and receiving data packets. Since these events are highly cross correlated among them and among different sensors, it is difficult to provide a closed form expression for the probabilities of these events. As a result, an accurate characterization would require modeling the probability that the RX node is busy with the reception of a data packet while some other node is trying to send another data packet, which is very difficult, if not impossible, to model. In the next Subsection, we will see that Eq. (28) is a satisfactory bound.

6.3 Average Energy Consumption

The dependency of $\mathbb{E}\,E_{\rm tot}$ on the random variables T_a and T_l can be removed by taking the expectation with respect to T_a and T_l. The total average energy consumption is denoted as

$$E(R_l, R_s) \triangleq \mathbb{E}_{T_a}\,\mathbb{E}_{T_l}\,\mathbb{E}\,E_{\rm tot}\,. \tag{29}$$

In this equation, we have evidenced that the average total energy consumption depends on the sleep time and the listen time. The expectations are computed as done in Eq. (12).

We validated the analysis of the average energy consumption by comparing the upper bound given by Eq. (29) to Monte Carlo ns-2 simulations, which were obtained as described in Subsection 4.4. We observed a good matching for all cases of practical interest. Figs. 10 and 11 reports the analytical model and simulation results of the energy as a function of R_s and R_l for packet generation periods of 30 s. The figures show a good matching of the upper bound with the simulations. The same conclusion holds for other choices of the network parameters (number of nodes and traffic generation rate). We remark that the upper bound is useful for optimization purposes.

7 Practical Considerations

In the previous sections, we have modeled the distributions of the delay to send a data packet from the transmitter to the receiver, the reliability and the energy consumption. Given a set of loss and busy channel probabilities, these expressions can be used off-line to select the optimal values of the sleep time R_s and listening time R_l that minimize the energy consumption given the latency and reliability constraints. Now, we are in the position to solve the optimization problem (1).

The problem (1) shows highly non linear functions in the decision variables. Solving such a problem is not a burden for processors without computational constraints. Hence, the optimal solution can be computed off-line and stored in a look-up table as function of the loss and collision probabilities. Observe that the optimal solution is a function of the network topology, channel condition, traffic and number of nodes. From the cluster-head point of view, these factors are summarized by the loss and busy channel probabilities α and β. Therefore, given α, β, the optimal solution of problem (1), denoted by $R_l^*(\alpha, \beta)$ and $R_s^*(\alpha, \beta)$, can be loaded in a light look up table to be stored in the cluster-head node. The table can be thought of as a matrix with rows associated to the set of values of α and columns associated to the values of β. The cluster-head node can easily do an estimation of the loss probabilities $\hat{\alpha}$ and busy channel $\hat{\beta}$, and read from the look-up table the entries $R_l^*(\alpha, \beta)$ and $R_s^*(\alpha, \beta)$ at location α, β closer to $\hat{\alpha}$ and $\hat{\beta}$. For instance, if we consider 10 values for $\hat{\alpha}$ and 10 for $\hat{\beta}$, the table would have 100 entries. By assuming that each entry takes 1 byte, the table has the size of just 0.1 Kbytes.

8 Simulation Results

In this section, we present an optimization of the duty cycle by using the modeling of the energy, delay and reliability that we have developed in the previous section. The simulator reproduced the system depicted in Fig. 1, where transmitter nodes send packets by a preamble-based IEEE 802.15.4 MAC. All the numerical values set for the simulations are taken coherently with the IEEE 802.15.4 standard and the Tmote sensors. The cluster head node employed a look-up table as described in Section 7 to optimize the duty cycle. In particular, during the simulations, the cluster head estimated the loss and busy channel probabilities, and then read the look-up table at a location with loss and busy probabilities closest in a Euclidean distance sense to the ones estimated. We considered 5 representative values of the loss probability and 5 representative values for the busy channel probability. Each simulation result was computed by running simulations to reproduce 20000 seconds of real time, and five simulations were run to remove the dependance on the initial seed of the random generators. The delay, reliability and energy consumption were then evaluated by collecting results from these simulations. In the following, we present the results of the optimization without delay and reliability constraints, and with these constraints.

8.1 Unconstrained Optimization

We compared the minimization of (1) to the one provided by X-MAC [8]. Recall that such a protocol does not take into account random back-off, delay and reliability constraints. Therefore, for the sake of comparison of the protocol proposed in this paper and X-MAC, we pose $t_{\max} = \infty$, $\tau_{\max} = 1$, and $\psi_{\min} = 0$, which implies neglecting the delay and reliability requirements, i.e., the energy is minimized without constraints, as done in X-MAC.

The energy consumption corresponding to the optimal protocol parameters within region $\{R_s \leq 2\,\mathrm{s}, R_l \geq 6\,\mathrm{ms}, R_s \geq R_l\}$ is shown in Fig. 12. The figure also

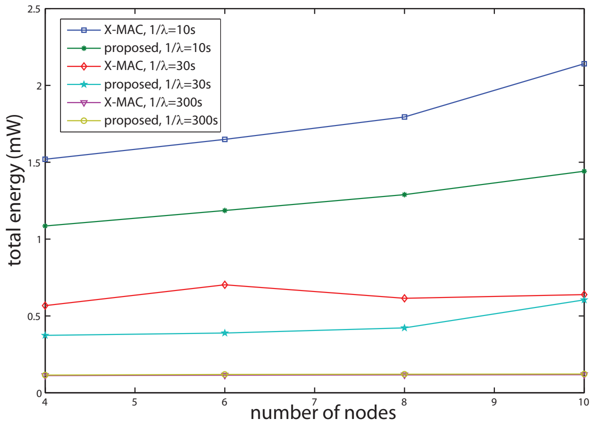


Figure 12: Comparison of energy consumption as obtained by the proposed protocol and X-MAC.

shows the energy consumption achieved by X-MAC. Our protocol outperforms X-MAC in all the scenarios considered. Specifically, when the packet generation period is high (300 s) the difference between X-MAC and our protocol is small (5% less than X-MAC), but as the packet generation period decreases the improvement is substantial, more than 50%. The main reason for this difference is that the nodes consume much less energy in packet transmission compared to the model in [8]. X-MAC is based on the assumption that the transmitter sends preamble packets back to back until the receiver wakes up, while actually there is random back-off before packet transmissions during which the transmitter puts its radio in sleep mode. Since the transmit energy dominates the receive energy much earlier according to the model in [8], the optimal wake time becomes considerably higher compared to the actual optimal wake time that we achieve.

8.2 Constrained Optimization

We validate our optimization by considering the delay and reliability requirements. In the following, we report the simulation results with optimal listening and sleep time as obtained by the solution of problem (1) for various cases of application requirements.

In Fig. 13, the average delay in the packet delivery probability achieved by simulations after employing the optimized listening and sleep times is plotted as function of the reliability requirement $\phi_{\min} = 93, 96, 99\,\%$ with delay requirement $t_{\max} = 0.2, 0.4, 0.6, 0.8\,\mathrm{s}$, $\tau_{\max} = 95\,\%$. Observe that the average delay is smaller than delay requirement because we consider the delay distribution as con-

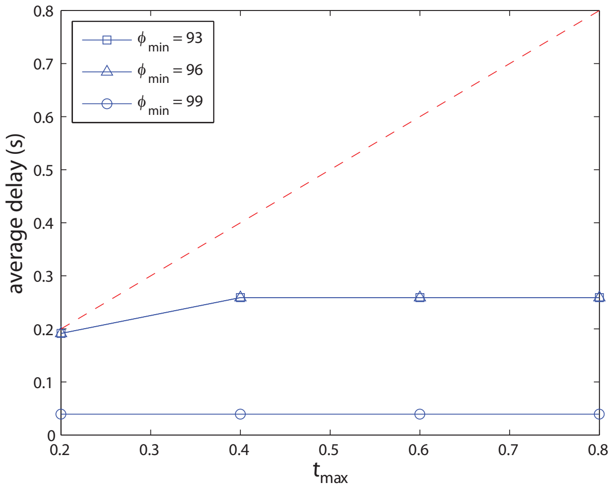


Figure 13: Average delay as function of reliability requirements $\phi_{\min} = 93, 96, 99\,\%$ with delay requirements $t_{\max} = 0.2, 0.4, 0.6, 0.8\,\mathrm{s}$, $\tau_{\max} = 95\,\%$ for a traffic rate of $1/\lambda = 30\,\mathrm{s}$ and 8 nodes.

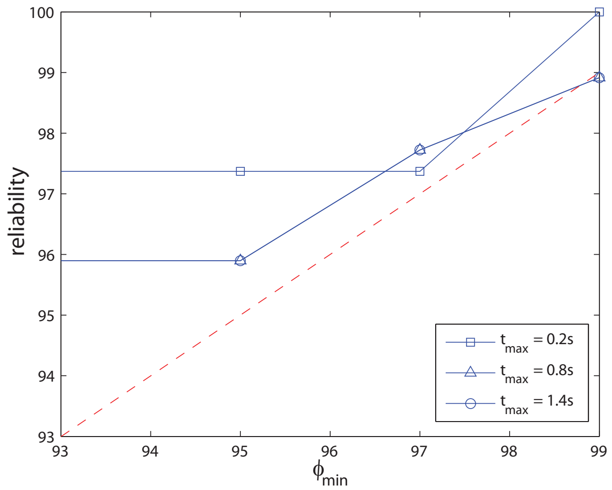


Figure 14: Reliability as function of delay requirements $t_{\max} = 0.2, 0.8, 1.4\,\mathrm{s}$, $\tau_{\max} = 95\,\%$ with reliability requirements $\phi_{\min} = 93, 95, 97, 99\,\%$ for a traffic rate of $1/\lambda = 30\,\mathrm{s}$ and 8 nodes.

straint. The average delay decreases as the reliability constraint becomes strict $\phi_{\min} = 99\,\%$ due to the fact that the sleep time decreases as the reliability constraint increases. Furthermore, the strict reliability requirement $\phi_{\min} = 99\,\%$ determines that the dominant constraint of the optimization problem is the reliability, whereas the delay requirement $t_{\max} = 0.2, 0.4, 0.6, 0.8\,\mathrm{s}$ makes inactive the delay constraint.

Fig. 14 shows the reliability as obtained by ns-2 simulations after employing the optimal listening and sleep times with the delay requirement given by $t_{\max} = 0.2, 0.8, 1.4\,\mathrm{s}$, $\tau_{\max} = 95\,\%$, and the reliability requirement given by $\phi_{\min} = 93, 95, 97, 99\,\%$. We see that the simulation results confirm the validity of the optimization because the requirement on the reliability are satisfied in most of cases. Note that the reliability requirement $\phi_{\min} = 99\,\%$ is slightly below $99\,\%$ when $t_{\max} = 0.8, 1.4\,\mathrm{s}$ due to the Euclidean distance method to approximate the estimated loss and busy channel probabilities. It is interesting to observe that the strict delay requirement $t_{\max} = 0.2\,\mathrm{s}$ determines that the dominant constraint of the optimization problem is the delay, whereas the reliability requirement $\phi_{\min} = 93, 95, 97\,\%$ makes inactive the reliability constraint.

We conclude that the listening and sleep time computed by our modeling and optimization allow packets to meet the delay and reliability requirements set by the application.

9 Conclusions

We developed a novel analytical characterization of the delay and packet loss probability distribution, and energy consumption for a clustered network topology with unslotted IEEE 802.15.4 and preamble sampling MAC. The analysis was based on the statistical modeling of the preamble, acknowledgement and data packet transmission. Monte Carlo simulations using ns-2 validated and illustrated our approach.

Our analysis can be used efficiently to provide a set of optimal listening and sleep times that minimize the energy consumption of the network while guaranteeing latency and reliability constraints. Compared to existing protocols that minimize only energy consumption, as B-MAC and X-MAC, our approach obtains much better results. Thus our method can be effectively employed to ensure a longer lifetime of the network.

In the future, we will focus on the extension of our theoretical analysis to the hybrid random-access/TDMA MAC, and on the investigation of the interaction of duty-cycle with routing.

Appendix A

A.1 Proof of Proposition 1

The proof is by iteration. First, consider the case of $k = 1$. Then, the event of reception of a preamble after the first attempt occurs when the preamble is sent during the listening time of the RX node, and the acknowledgement is sent before the timeout of the RX node:

$$\mathcal{B}_1 = (T_1 > T_a)(T_1 \leq T_a + T_l)(T_{\text{ack}} \leq T_{\text{TX,out}})\,.$$

Consider the case of $k = 2$. A preamble fails because 1) it was sent during the sleep time of the receiver, or 2) it was sent during the active time of the receiver but there was a loss due to bad channel or collisions, or 3) it was sent during the active time of the receiver without loss but the RX was not able to send back the acknowledgment before the time out of the TX, or 4) it was sent during the active time of the receiver without loss but the transmitted acknowledgment before the time out of the TX was collided. Then, a second preamble is sent during the listening time and an acknowledgement is sent back before the time out of the TX:

$$\begin{aligned}
\mathcal{B}_2 = &[(T_1 \leq T_a) + (T_1 > T_a)(T_1 \leq T_a + T_l)\,\alpha \\
&+ (T_1 > T_a)(T_1 \leq T_a + T_l)(1-\alpha) \\
&\times (T_{\text{ack}} > T_{\text{TX,out}} | T_1 > T_a) + (T_1 > T_a)(T_1 \leq T_a + T_l) \\
&\times (1-\alpha)(T_{\text{ack}} \leq T_{\text{TX,out}} | T_1 > T_a)\,\alpha] \\
&\times [T_1 + T_{\text{TX,out}} + T_1 > T_a] \\
&\times [T_1 + T_{\text{TX,out}} + T_1 \leq T_a + T_l] \\
&\times [T_{\text{ack}} \leq T_{\text{TX,out}} | T_1 + T_{\text{TX,out}} + T_1 > T_a]\,(1-\alpha)^2\,.
\end{aligned}$$

Consider the case $k = 3$. This happens when a preamble is sent a first time during the sleep time, then, after a time out, a second preamble is sent again during the sleep time, and, finally, after another time out, a preamble is sent during the active time, and the acknowledgement is transmitted before the time out:

$$\begin{aligned}
\mathcal{B}_3 = &[T_1 \leq T_a][T_1 + T_{\text{TX,out}} + T_1 \leq T_a] \\
&\times [T_1 + T_{\text{TX,out}} + T_1 + T_{\text{TX,out}} + T_1 > T_a] \\
&\times [T_1 + T_{\text{TX,out}} + T_1 + T_{\text{TX,out}} + T_1 \leq T_a + T_l] \\
&\times [T_{\text{ack}} \leq T_{\text{TX,out}}] \\
= &[2T_1 + T_{\text{TX,out}} \leq T_a][3T_1 + 2T_{\text{TX,out}} > T_a] \\
&\times [3T_1 + 2T_{\text{TX,out}} \leq T_a + T_l][T_{\text{ack}} \leq T_{\text{TX,out}}]\,. \qquad (30)
\end{aligned}$$

It is straightforward to generalize previous expression so to obtain the sought proof.

A.2 Proof of Proposition 2

From the proof of Proposition 1, a preamble fails because of four events. It follows that $\mathcal{B}_k = \mathcal{B}_{1,k} + \mathcal{B}_{2,k} + \mathcal{B}_{3,k} + \mathcal{B}_{4,k}$, where

$$
\begin{aligned}
\mathcal{B}_{1,k} &= \mathcal{C}_k\mathcal{D}_k\mathcal{E}_k\mathcal{F}_k(1-\alpha)^2\,,\\
\mathcal{B}_{2,k} &= \mathcal{D}_{k-1}\mathcal{E}_{k-1}\mathcal{D}_k\mathcal{E}_k\mathcal{F}_k\,\alpha(1-\alpha)^2\,,\\
\mathcal{B}_{3,k} &= \mathcal{D}_{k-1}\mathcal{E}_{k-1}\bar{\mathcal{F}}_{k-1}\mathcal{D}_k\mathcal{E}_k\mathcal{F}_k(1-\alpha)^3\,,\\
\mathcal{B}_{4,k} &= \mathcal{D}_{k-1}\mathcal{E}_{k-1}\mathcal{F}_{k-1}\mathcal{D}_k\mathcal{E}_k\mathcal{F}_k\,\alpha(1-\alpha)^3\,.
\end{aligned}
$$

Notice that $\mathcal{B}_{1,k}\mathcal{B}_{2,k} = \emptyset$, $\mathcal{B}_{1,k}\mathcal{B}_{3,k} = \emptyset$, $\mathcal{B}_{1,k}\mathcal{B}_{4,k} = \emptyset$, $\mathcal{B}_{2,k}\mathcal{B}_{3,k} = \emptyset$, $\mathcal{B}_{2,k}\mathcal{B}_{4,k} = \emptyset$, $\mathcal{B}_{3,k}\mathcal{B}_{4,k} = \emptyset$. It follows that

$$
\Pr[\mathcal{B}_k] = \Pr[\mathcal{B}_{1,k}] + \Pr[\mathcal{B}_{2,k}] + \Pr[\mathcal{B}_{3,k}] + \Pr[\mathcal{B}_{4,k}]\,. \tag{31}
$$

In the following, the probabilities of $\mathcal{B}_{1,k}$, $\mathcal{B}_{2,k}$, $\mathcal{B}_{3,k}$ and $\mathcal{B}_{4,k}$ are computed. The probability of $\mathcal{B}_{1,k}$ is given by considering that the event $\mathcal{F}_k$ is independent of the others, so that

$$
\Pr[\mathcal{B}_{1,k}] = \Pr[\mathcal{C}_k\mathcal{D}_k\mathcal{E}_k]\Pr[\mathcal{F}_k](1-\alpha)^2\,.
$$

From [23], we have

$$
\Pr[\mathcal{C}_k\mathcal{E}_k] = \Pr[\mathcal{C}_k\mathcal{D}_k\mathcal{E}_k] + \Pr[\mathcal{C}_k\bar{\mathcal{D}}_k\mathcal{E}_k]\,,
$$

and that $\mathcal{C}_k\bar{\mathcal{D}}_k\mathcal{E}_k = \bar{\mathcal{D}}_k\mathcal{E}_k = \bar{\mathcal{D}}_k$, from which it holds

$$
\Pr[\mathcal{C}_k\mathcal{D}_k\mathcal{E}_k] = \Pr[\mathcal{C}_k\mathcal{E}_k] - \Pr[\bar{\mathcal{D}}_k]\,. \tag{32}
$$

Rewriting $\mathcal{C}_k = [kT_1 + (k-1)T_{\mathrm{TX,out}} \le T_a + T_1 + T_{\mathrm{TX,out}}]$, it implies

$$
\mathcal{C}_k\mathcal{E}_k = \begin{cases} \mathcal{E}_k & \text{if } T_a + T_l \le T_a + T_1 + T_{\mathrm{TX,out}} \\ \mathcal{C}_k & \text{otherwise} \end{cases} \tag{33}
$$

whereby

$$
\begin{aligned}
\Pr[\mathcal{C}_k\mathcal{E}_k] = &\Pr[\mathcal{C}_k]\Pr[T_1 \le T_l - T_{\mathrm{TX,out}}]\\
&+ \Pr[\mathcal{E}_k](1 - \Pr[T_1 \le T_l - T_{\mathrm{TX,out}}])\,.
\end{aligned}
$$

This equation and (32) provide us

$$
\begin{aligned}
\Pr[\mathcal{B}_{1,k}] = &\big(\Pr[\mathcal{C}_k\mathcal{E}_k] - \Pr[\bar{\mathcal{D}}_k]\big)\Pr[T_{\mathrm{ack}} \le T_{\mathrm{TX,out}}]\\
&\times (1-\alpha)^2\,,
\end{aligned} \tag{34}
$$

where $\Pr[\mathcal{F}_k] = \Pr[T_{\mathrm{ack}} \le T_{\mathrm{TX,out}}]$.

To compute $\Pr[\mathcal{B}_{2,k}]$, observe that

$$\mathcal{D}_{k-1}\mathcal{E}_{k-1}\mathcal{D}_k\mathcal{E}_k\mathcal{F}_k\,\alpha(1-\alpha)^2 = \mathcal{D}_{k-1}\mathcal{E}_k\mathcal{F}_k\,\alpha(1-\alpha)^2\,,$$

because $\mathcal{D}_k\mathcal{D}_{k-1} = \mathcal{D}_{k-1} = \bar{\mathcal{C}}_k$ and $\mathcal{E}_k\mathcal{E}_{k-1} = \mathcal{E}_k$. It follows that

$$\begin{aligned}\Pr[\mathcal{B}_{2,k}] =& \Pr[\bar{\mathcal{C}}_k\mathcal{E}_k]\Pr[\mathcal{F}_k]\alpha(1-\alpha)^2\\ =&(\Pr[\mathcal{E}_k]-\Pr[\mathcal{C}_k\mathcal{E}_k])\Pr[T_{\text{ack}}\le T_{\text{TX,out}}]\\ &\times\alpha(1-\alpha)^2\,.\end{aligned} \tag{35}$$

To compute $\Pr[\mathcal{B}_{3,k}]$, observe that

$$\mathcal{D}_{k-1}\mathcal{E}_{k-1}\bar{\mathcal{F}}_{k-1}\mathcal{D}_k\mathcal{E}_k\mathcal{F}_k(1-\alpha)^3 = \bar{\mathcal{F}}_{k-1}\mathcal{D}_{k-1}\mathcal{E}_k\mathcal{F}_k(1-\alpha)^3\,,$$

because $\mathcal{D}_k\mathcal{D}_{k-1} = \mathcal{D}_{k-1} = \bar{\mathcal{C}}_k$ and $\mathcal{E}_k\mathcal{E}_{k-1} = \mathcal{E}_k$. Note that $\bar{\mathcal{F}}_{k-1}$ and $\mathcal{F}_k$ are independent, and that $\Pr[\bar{\mathcal{F}}_{k-1}] = 1-\Pr[T_{\text{ack}}\le T_{\text{TX,out}}]$, so

$$\begin{aligned}\Pr[\mathcal{B}_{3,k}] =& \Pr[\bar{\mathcal{C}}_k\mathcal{E}_k]\Pr[\bar{\mathcal{F}}_{k-1}]\Pr[\mathcal{F}_k](1-\alpha)^3\\ =&(\Pr[\mathcal{E}_k]-\Pr[\mathcal{C}_k\mathcal{E}_k])(1-\Pr[T_{\text{ack}}\le T_{\text{TX,out}}])\\ &\times\Pr[T_{\text{ack}}\le T_{\text{TX,out}}](1-\alpha)^3\,.\end{aligned} \tag{36}$$

To compute $\Pr[\mathcal{B}_{4,k}]$, observe that

$$\begin{aligned}\mathcal{D}_{k-1}\mathcal{E}_{k-1}\mathcal{F}_{k-1}\mathcal{D}_k\mathcal{E}_k\mathcal{F}_k\,\alpha(1-\alpha)^3 =& \mathcal{F}_{k-1}\mathcal{D}_{k-1}\mathcal{E}_k\mathcal{F}_k\\ &\times\alpha(1-\alpha)^3\,,\end{aligned}$$

because $\mathcal{D}_k\mathcal{D}_{k-1} = \mathcal{D}_{k-1} = \bar{\mathcal{C}}_k$ and $\mathcal{E}_k\mathcal{E}_{k-1} = \mathcal{E}_k$. Note that $\mathcal{F}_{k-1}$ and $\mathcal{F}_k$ are independent, so

$$\begin{aligned}\Pr[\mathcal{B}_{4,k}] =& \Pr[\bar{\mathcal{C}}_k\mathcal{E}_k]\Pr[\mathcal{F}_{k-1}]\Pr[\mathcal{F}_k]\alpha(1-\alpha)^3\\ =&(\Pr[\mathcal{E}_k]-\Pr[\mathcal{C}_k\mathcal{E}_k])\left(\Pr[T_{\text{ack}}\le T_{\text{TX,out}}]\right)^2\\ &\times\alpha(1-\alpha)^3\,.\end{aligned} \tag{37}$$

By putting together (31)– (37), the proof follows.

A.3 Proof of Claim 4

Eq. (18) is given by the sum of two main components: the energy spent in the case of successful data packet transmission, and the energy spent in the case of unsuccessful data packet transmission. To characterize these terms, let us start to observe that in Eq. (18), $\mathbb{1}_{\mathcal{B}_i} = 1$ if $\mathcal{B}_i$ occurs, namely the i-th preamble was successfully transmitted, then the corresponding acknowledgement was successful received. Analogously, $\mathbb{1}_{\mathcal{B}_i} = 0$ if $\mathcal{B}_i$ does not occur, namely either the i-th preamble was not transmitted, and/or the corresponding acknowledgement was not received. From Proposition 1, notice that if $\mathbb{1}_{\mathcal{B}_i} = 1$ for some i, then $\mathbb{1}_{\mathcal{B}_i} = 0$ for $k = 1, \dots, N_p$ and $k \neq i$. Furthermore, there is a non zero probability that $\mathbb{1}_{\mathcal{B}_i} = 0$ for $i = 1, \dots, N_p$, meaning that there is a non zero probability that a preamble is not sent within N_p trials. In the following, the various terms of Eq. (18) are explained.

The first sum of Eq. (18) is derived as follows. Suppose that the i-th preamble was successfully transmitted in the active time of the receiver, and the corresponding acknowledgement was received, i.e., $\mathbb{1}_{\mathcal{B}_i} = 1$. The energy spent for transmission is given by the sum of the energy spent trying to access the channel (E_{tx,T_1}) from the first preamble until the ith, plus the energy spent during the time out periods $E_{T_{\mathrm{TX,out}}}$ for the preambles $1, \dots, i-1$, plus the energy spent by the receiver to hear the coming ACK transmission ($E_{\mathrm{tx},T_{\mathrm{ack}}}$), and transmission of data packet ($E_{\mathrm{tx},T_{\mathrm{data}}}$).

The second sum of Eq. (18) is derived as follows. Suppose that preamble was not transmitted, and/or the corresponding acknowledgement was not received, i.e., $\mathbb{1}_{\mathcal{B}_i} = 0$ for $i = 1, \dots, N_p$. This is equivalent to $\mathbb{1}_{\bar{\mathcal{B}}} = 1$, where $\bar{\mathcal{B}}$ is the complementary of $\sum_{i=1}^{N_p} \mathcal{B}_i$. In the sum of the events, we considered the worst case, which occurs when N_p preambles have been sent. This justifies why Eq. (18) is an upper bound.

Eq. (19) is the energy spent to sense the channel and to transmit a preamble, provided that the channel is sensed free and that the acknowledgement is received ($\mathbb{1}_{\mathcal{A}_j}$). At back-off j of n-th preamble, the energy spent is

$$\sum_{k=1}^{j} \left(P_s S_{p,k}^{n} + P_{\mathrm{rx}} S_p \right) .$$

A preamble is transmitted spending the energy $P_{\mathrm{tx}} S_p$.

Eq. (20) is the energy spent by the transmitter during the time out.

Eq. (21) is the energy spent by the transmitter while waiting for the acknowledgement. It is derived by following the same approach used for (19).

Eq. (22) is the energy spent by the transmitter while trying to send the data packet, if a previous preamble was received. The first term in (22) is derived by following the same approach used for (19). The last term takes into account that a packet may not be sent if the back-off procedure after the reception of the acknowledgement lasts too long.

A.4 Proof of Claim 5

The average energy consumption at the transmitter can be computed by applying the linear properties of the expectation operator to (18). This computation involves terms as $\mathbb{1}_{\mathcal{A}_j} \mathbb{1}_{\mathcal{B}_i}$ in (19) and (21), for which we have the approximation $\mathbb{E}\left[\mathbb{1}_{\mathcal{A}_j} \mathbb{1}_{\mathcal{B}_i}\right] \simeq \Pr[\mathcal{A}_j] \Pr[\mathcal{B}_i]$, while in (22) the product $\mathbb{1}_{\mathcal{A}_j} \mathbb{1}_{\mathcal{B}_i}$ is approximated as given by two independent random variables, so that $\mathbb{E}\left[\mathbb{1}_{\mathcal{A}_j} \mathbb{1}_{\mathcal{B}_i}\right] \simeq \Pr[\mathcal{A}_i] \Pr[\mathcal{B}_i]$.

References

[1] *IEEE Std 802.15.4-2996, September, Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)*, IEEE, 2006. [Online]. Available: http://www.ieee802.org/15

[2] *Routing Over Low power and Lossy networks (roll)*. [Online]. Available: http://www.ietf.org/dyn/wg/charter/roll-charter.html

[3] Y. Xu, J. Heidemann, and D. Estrin, “Geography-informed energy conservation for ad hoc routing,” in *ACM MobiCom*, vol. 70-84, 2001.

[4] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris, “Span: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks,” in *ACM MobiCom*, 2001.

[5] W. Ye, J. Heidemann, and D. Estrin, “Medium access control with coordinated sleeping for wireless sensor networks,” *IEEE/ACM Transactions on Networking*, vol. 12, no. 3, June 2004.

[6] J. Polastre, J. Hill, and D.Culler, “Versatile low power media access for wireless sensor networks,” in *2nd International Conference on Embedded networked sensor systems*, November 2004, pp. 95–107.

[7] *The TinyOS community forum*. [Online]. Available: http://www.tinyos.net

[8] M. Buettner, G. Yee, E. Anderson, and R. Han, “X-MAC: A short preamble MAC protocol for duty-cycled wireless sensor networks,” in *4th ACM Conference on Embedded Sensor Systems (SenSys)*, November 2006.

[9] S. C. Ergen and P. Varaiya, “Pedamacs: Power efficient and delay aware medium access protocol for sensor networks,” *IEEE Transactions on Mobile Computing*, vol. 5, no. 7, 2006.

[10] E. Uysal-Biyikoglu, B. Prabhakar, and A. E. Gamal, “Energy-efficient packet transmission over a wireless link,” *IEEE/ACM Transactions on Networking*, vol. 10, no. 12, August 2002.

[11] T. V. Dam and K. Langendoen, “An adaptive energy-efficient MAC protocol for wireless sensornetworks,” in *1st ACM Conf. on Embedded Networked Sensor Systems*, November 2003, pp. 171–180.

[12] S. Ganeriwal, I. Tsikogiannis, H. Shim, V. Tsiatsis, M. Srivastava, and D. Ganesan, “Estimating clock uncertainty for efficient duty-cycling in sensor networks,” *IEEE/ACM Transactions on Networking*, September 2008.

[13] X. Shi and G. Stromber, “Syncwuf: An ultra low-power MAC protocol for wireless sensor networks,” *IEEE Transactions on Mobile Computing*, pp. 115–125, January 2007.

[14] W. Zhang, M. S. Braniky, S. M. Phillips, “Stability of networked control systems,” *IEEE Control Systems Magazine*, February 2001.

[15] A. Willig, “Recent and emerging topics in wireless industrial communication,” *IEEE Transactions on Industrial Informatics*, vol. 4, no. 2, pp. 102–124, 2008.

[16] G. Bianchi, “Performance analysis of the ieee 802.11 distributed cordination function,” in *IEEE Journal on Selected Areas in Communications*, vol. 18, March 2000.

[17] S. Pollin, M. Ergen, S. C. Ergen, B. Bougard, L. Perre, I. Moerman, A. Bahai, P. Varaiya, and F. Catthoor, “Performance analysis of slotted carrier sense IEEE 802.15.4 medium access layer,” in *IEEE Transactions on Wireless Communication*, vol. 7, 2008, pp. 3359–3371.

[18] H. Zhai, Y. Kwon, and Y. Fang, “Performance analysis of IEEE 802.11 MAC protocols

in wireless LANs: Research articles," in *Wirel. Commun. Mob. Comput.*, 2004, pp. 917–931.

[19] H. Wu, Y. Peng, K. Long, S. Cheng, and J. Ma, "Performance of reliable transport protocol over IEEE 802.11 wireless LAN: Analysis and enhancement," in *Proc. of IEEE INFOCOM*, 2002.

[20] L. Råde and B. Westergren, *Mathematics Handbook for Science and Engineering*. Studentlitteratur, 2004.

[21] *Tmote Sky Data Sheet*, Moteiv, San Francisco, CA, 2006. [Online]. Available: http://www.moteiv.com/products/docs/tmote-sky-datasheet.pdf

[22] M. Ergen and P. Varaiya, "Delay analysis of distributed coordination function in IEEE 802.11," in *Proc. of IEEE GLOBECOM*, 2006.

[23] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. Mc Graw Hill, Third Edition, 1991.

Symbol	Meaning
TX node	a transmitter node
RX node	a receiver node (cluster head)
T_1	random delay spent by the TX node before transmitting a preamble packet
T_2	random delay spent by the TX node from the beginning of a transmission until the reception of the acknowledgement
T_3	random delay spent by the TX node from the acknowledgement reception until the transmission of a data packet
T_p	random delay to wait before a data packet is successfully received
T_s	random sleep time of the RX as seen from the TX (it is uniformly distributed over $[0, R_s]$)
T_l	random listening time of the RX as computed upon the reception of a preamble (it is uniformly distributed over $[0, R_l]$)
T_{hr}	time employed by the hardware platform to process the packets and transmit them
T_{ack}	random time before the RX node can access the channel and send an acknowledgement
$T_{\mathrm{TX,out}}$	maximum time that a TX node waits for an ACK after having sent a preamble.
T_{out}	maximum time that a TX node waits from the moment of the reception of an ACK before giving up the data packet transmission.
N_p	maximum number of preambles that can be sent
N_b	maximum number of back-off to sense the channel for sending a preamble packet
NB	number of back-off of CSMA/CA
BE	back-off exponent
$NB_{\max}$	maximum number of back-offs before declaring a channel access failure
N	number of nodes in a cluster
λ	packet generation rate per node
d_{TX}	probability that a TX node has a packet to send in the interval $R_s + R_l$

Table 1: Main symbols used in the paper.

α	probability of preamble or ACK loss
β	probability of busy channel
p	probability of data packet loss
$\phi_{\min}$	minimum probability of successful packet transmission (reliability requirement)
$\tau_{\max}$	maximum probability of maximum delay (latency requirement)
E_{tot}	total energy consumption
$S_{p,j}$	j-th random back-off time of a preamble
$\mu_{S_{p,j}}$	average of $S_{p,j}$
S_c	channel sensing duration for clear channel assessment
S_p	preamble packet duration
S_a	acknowledgement packet duration
S_d	data packet duration
S_b	unit time used by the CSMA/CA algorithm
P_{tx}	transmit power
P_{rx}	receive power
P_s	sleep power
R_s	sleep time of the receiver node (cluster head)
R_l	active time of the receiver node (cluster head)
$\mathcal{A}_k$	event occurring when the channel is busy for $k-1$ times
$\mathcal{B}_k$	event occurring when a preamble has to be sent k times before being received in the active time of the RX node and the corresponding acknowledgement is sent by the RX node and received before the time out of the TX node
$\mathcal{G}$	event occurring when a preamble is successfully received during the listening state of the receiver
$\mathcal{H}\vert\mathcal{G}$	event occurring when the ACK is successfully sent before the time out of the RX expires provided that a preamble is successfully received
$\mathcal{I}\vert\mathcal{G},\mathcal{H}$	event occurring when the TX sends successfully a data packet provided that a preamble is successfully received and the ACK is also successfully received

Table 2: Main symbols used in the paper (continuation).

Paper C

Breath: an Adaptive Protocol for Industrial Control Applications using Wireless Sensor Networks

P. Park, C. Fischione, A. Bonivento,
K. H. Johansson and A. Sangiovanni-Vincentelli

Submitted to
IEEE Transactions on Mobile Computing

Breath: an Adaptive Protocol for Industrial Control Applications using Wireless Sensor Networks

P. Park, C. Fischione, A. Bonivento,
K. H. Johansson and A. Sangiovanni-Vincentelli

Abstract

Energy-efficient, reliable and timely data transmission is essential for wireless sensor networks (WSNs) employed in scenarios where plant information must be available for control applications. To reach a maximum efficiency, cross layer interaction is a major design paradigm to exploit the complex interaction among the layers of the protocol stack. This is challenging because latency, reliability, and energy are at odds, and resource constrained nodes support only simple algorithms. In this paper, the novel protocol Breath is proposed for control applications. Breath is designed for WSNs where source nodes attached to a plant must transmit information via multi-hop routing to a sink. Breath ensures a desired packet delivery and delay probabilities while minimizing the energy consumption of the network. The protocol is based on randomized routing, medium access control, and duty-cycling jointly optimized for energy efficiency. The design approach relies on a constrained optimization problem, whereby the objective function is the energy consumption and the constraints are the packet reliability and delay. The challenging part is the modelling of the interactions among the layers by simple expressions of adequate accuracy, which are then used for the optimization by in-network processing. The optimal working point of the protocol is achieved by a simple algorithm, which adapts to traffic variations and channel conditions with negligible overhead. The protocol has been implemented and experimentally evaluated on a test-bed with off-the-shelf wireless sensor nodes, and it has been compared with a standard IEEE 802.15.4 solution. Analytical and experimental results show that Breath is tunable and meets reliability and delay requirements. Breath exhibits a good distribution of the working load, thus ensuring a long lifetime of the network. Therefore, Breath is a good candidate for efficient, reliable, *and timely data gathering for control applications.*

Index Terms–Wireless Sensor Networks, Networked Control System, Control over Multi-hop WSNs, Routing, MAC, Radio Power Control, Duty Cycle, Optimization.

1 Introduction

Wireless sensor networks (WSNs) are networks of tiny sensing devices for wireless communication, monitoring, control, and actuation. Given the potential benefits offered by these networks, as, e.g., simple deployment, low installation cost, lack of cabling, and high mobility, they are specially appealing for control and industrial applications [1, 2]. The variety of application domains and theoretical challenges for WSNs has attracted research efforts for more than one decade. Nevertheless, a lively research and standardization activity is ongoing [2, 3].

Although WSNs provide a great advantage for process, manufacturing and industry, they are not yet widely deployed. This is due to that the software for these applications is usually written by process and software engineers that are expert in process control technology, but know little of the network and sensing infrastructure that has to be deployed to support control applications. On the other side, the communication infrastructure is designed by communication engineers that know little about process control technology. Moreover, the adoption of wireless technology further complicates the design of these networks. Being able to satisfy high requirements on communication performance over unreliable communication channels is a difficult task.

Standard practice for control system design over communication networks is as follows: First, deploy the networked embedded system on a predefined distributed architecture, chosen on the basis of experience and heuristic considerations. Then, tweak the software implementation of the control algorithm to meet latency, bandwidth, and reliability offered by the network. In many control designs, the network imperfections are completely disregarded, assuming instead that sensor and control data instantaneously reach the controller and actuator node, respectively.

This is far from ideal, because many control systems are highly cost sensitive, and using a non-optimized network is clearly expensive. Moreover, the complexity of large networked embedded systems continues to increase, making heuristic and experience-based design practices inadequate at best. To bridge this gap and derive a correct and efficient implementation, a system-level approach has been proposed in [4, 5]. By a system-level design for WSNs, the control algorithm designers impose a set of requirements on reliability, packet delay and energy consumption that the communication infrastructure must satisfy.

An efficient system-level design process for operations of WSNs in industrial control applications poses extra challenges compared to more traditional communication networks, namely:

- Reliability: Sensor information must be sent to the sink of the network with a given probability of success, because missing these data could prevent the correct execution of control actions or decisions concerning the phenomena sensed. However, maximizing the reliability may increase substantially the network energy consumption [2]. Hence, the network designers need to consider the tradeoff between reliability and energy consumption.

- Delay: Sensor information must reach the sink within some deadline. A probabilistic delay requirement must be considered instead of using average packet delay since the delay jitter can be too difficult to compensate for, especially if the delay variability is large [6]. Retransmission of old data to maximize the reliability may increase the delay and is generally not useful for control application [7].

- Energy efficiency: The lack of battery replacement, which is essential for affordable WSN deployment, requires energy-efficient operations. Since high reliability and low delay may demand a significant energy consumption of the network, thus reducing the WSN lifetime, the reliability and delay must be flexible design parameters that need to be adequate for the requirements. Note that controllers can usually tolerate a certain degree of packet losses and delay [6], [8], [9], [10]. Hence, the maximization of the reliability and minimization of the delay are not the optimal design strategies for the control applications we are concerned within this paper.

- Adaptation: The network operation should adapt to application requirement changes, varying wireless channel and network topology. For instance, the set of application requirements may change dynamically and the communication protocol must adapt its design parameters according to the specific requests of the control actions. To support changing requirements, it is essential to have an analytical model describing the relation between the protocol parameters and performance indicators (reliability, delay, and energy consumption).

- Scalability: Since the processing resources are limited, the protocol procedures must be computationally light. These operations should be performed within the network, to avoid the burden of too much communication with a central coordinator. This is, particularly important for large networks. The protocol should also be able to adapt to size variation of the network, as, for example, caused by moving obstacles, or addition of new nodes.

In this paper, we offer a complete design approach that embraces all the factors mentioned above. We propose the Breath protocol, a self-adapting efficient solution for reliable and timely data transmission. Since the protocol adapts to the network variations by enlarging or shrinking next-hop distance, sleep time of the nodes, and transmit radio power, we think that it behaves like a breathing organism.

The rest of the paper is organized as follows: In Section 1.1, we motivate our study and summarize existing work. Section 1.2 presents the main contributions of the paper. In Section 2 we define the system scenario. In Section 3, we introduce Breath in detail. In Section 4 an optimization problem is posed to optimize the protocol, whereas in Section 5 the constraints and cost function of the protocol are modelled. In Section 6, we derive the optimal solution and in Section 7 we present an adaptive algorithm to obtain the working point of the protocol. Some fundamental working limits of Breath are given in Section 8. A complete experimental implementation of the protocol is presented in Section 9. Finally, in Section 10 concluding remarks and future perspectives are given.

1.1 Related Works

There have been many contributions to the problem of protocol design for WSNs, both in academia (e.g., [2, 11]) and industry (e.g., [12–14]). New protocols have been built around standardized low-power protocols such as IEEE 802.15.4 [3], Zigbee [15] and WirelessHART [16]. WirelessHART is a promising solution for the replacement of the wired HART protocol in industrial contexts. However, the power consumption is not a main concern in WirelessHART, whereas the data link layer is based on Time Division Multiple Access (TDMA), which requires time synchronization and pre-scheduled fixed length time-slots by a centralized network manager. Such a manager should update the schedule frequently to consider reliability and delay requirements and dynamic changes of the network, which demands complex hardware equipments. WirelessHART is thus in contrast with the necessity of simple protocols able to work with limited energy and computing resources. In Tab. 1, we summarize the characteristics of the protocols that are relevant for the category of applications we are concerned within this paper. In the table, we have evidenced wether indications as energy $\mathbf{E}$, reliability $\mathbf{R}$, and delay $\mathbf{D}$ have been included in the protocol design and validation, and whether a cross-layer approach has been adopted. We discuss these protocols in the following.
GAF [17], SPAN [18] and X-MAC [19] consider the energy efficiency as a performance indicator, which is attained by algorithms under the routing layer and above the MAC layer so called bridge layer. Simulation results of reliability and delay are reported in [17, 18]. These protocols have not been designed out of an analytical modelling of reliability and delay, so there is not systematic control of them. One of the first protocol for WSNs designed to offer a high reliability is RMST [20], but energy consumption of the network or delay have not been accounted for in this protocol. The same lack of energy efficiency and delay requirements can be found in the reliable solutions presented in [21–23]. Dozer [24] comprises the MAC and routing layer to minimize the energy consumption while maximizing the reliability of the network, but an analytical approach has not been followed. Specially, Fetch [23] and Dozer [24] are designed for monitoring application, which mainly deals with lower traffic load than control applications. The latency of Fetch [23] is significantly dependent on the depth of the routing tree and

Table 1: Protocol comparison. The letters **E**, **R**, and **D** denote energy, reliability and communication delay. The circle denotes that a protocol is designed by considering the indication of the column, but it has not been validated experimentally. The circle with plus denotes that the protocol is designed by considering the indication and experimentally validated. The dot denotes that the protocol design does not include indication and hence cannot control it, but simulation or experiment results include it. The term "bridge" means that the protocol is designed by bridging MAC and routing layers.

Protocol	**E**	**R**	**D**	Layer
GAF [17]	$\bigcirc$	$\cdot$	$\cdot$	bridge
SPAN [18]	$\bigcirc$	$\cdot$	$\cdot$	bridge
XMAC [19]	$\oplus$	$\cdot$	$\cdot$	MAC
Flush [21]		$\oplus$		MAC
Fetch [23]	$\cdot$	$\oplus$	$\cdot$	phy, MAC, routing
GERAF [25]	$\bigcirc$		$\bigcirc$	MAC, routing
Dozer [24]	$\oplus$	$\oplus$		MAC, routing
MMSPEED [26]		$\bigcirc$	$\bigcirc$	routing
Breath	$\oplus$	$\oplus$	$\oplus$	phy, MAC, routing

is around some hundred seconds. In addition, experimental results of Dozer [24] show good energy efficiency and reliability under very low traffic intensity (with data sampling interval of 120 s,) but the delay in the packet delivery is not considered, which is essential for control applications [7], [8]. Energy efficiency with delay requirement for MAC and randomized routing is considered in GERAF [25], without simulation or experimental validation.

The focuses of the protocols mentioned above [17]– [25] are the maximization of the energy efficiency or reliability, or just minimization of the delay, without considering simultaneously application requirements in terms of reliability and delay in the packet delivery. In other words, these protocols are mostly designed for monitoring applications and does not support typical control application requirements. Control and industrial applications are able to cope with a certain degree of packet losses and delay [6, 8, 9], which implies that the approaches followed in the protocols mentioned above are not the ideal solution for these applications. The maximization of the energy efficiency and reliability may give a long delay, which are bad for the stability of the closed-loop control system. Analogously, the maximization of the reliability may be energy demanding and may give long delay, all of which are not tolerable for control applications. In addition, the protocols mentioned above do not support an adaptation to the changes of the reliability and delay, which may be required by the controller.

The protocols MMSPEED [26] and SERAN [27] are appealing for control and industrial applications. However, MMSPEED is not energy efficient because it considers a routing technique with an optimization of reliability and delay without energy constraints. The protocol satisfies a high reliability requirement by using duplicated packets over multi-path routing. Duplicated packets increase the traffic load with negative effect on the stability and energy efficiency of the network. In SERAN, a system-level design methodology has been presented for industrial applications, but even though SERAN allows the network to operate with low energy consumption subject to delay requirements, it does not consider tunable reliability requirements nor duty-cycling policies, which are essential to reduce energy consumption. Furthermore, SERAN focuses on low traffic networks. These characteristics limit the performance of SERAN both in term of energy and reliability in our application setup.

Given the availability of numerous techniques to reduce energy consumption and ensure reliability and low delays, a cross-layer optimization is a natural approach to integrate the protocol layers. Some cross-layer design challenges of the physical, MAC and network layers to minimize the energy consumption of WSNs have been surveyed in [28], [29] and [30]. Many of the cross-layer solutions proposed in the literature are hardly useful for the application domain we are targeting, because they require sophisticated processing resources, or instantaneous global network knowledge, which are out of reach of the capabilities of real nodes. Network design can be formulated as an optimization problem. However, as it was noted in [31], the complex interdependence of the decision variables (sleep disciplines, clustering, MAC, routing, power control, etc.) lead to difficult problems even in simple network topologies, where the analytical relations describing packet reception rate, delay and energy consumption may be highly nonlinear expressions. Such a difficulty is further exacerbated when considering non-TDMA scheme [32]. We propose next a design approach that offers a computationally attractive solution by simplifications of adequate accuracy.

1.2 Original Contribution

In this paper we present Breath, an adaptive protocol for WSNs for reliable and timely data gathering. Our system model considers source nodes that have to send packets to the sink via multi-hop routing under tunable reliability and delay requirements. We present a solution based on randomized routing, CSMA/CA MAC and randomized sleep discipline that are jointly optimized for energy consumption. To the best of our knowledge, no efficient and simple cross-layer protocol that includes all the relevant characteristics of the physical layer, MAC, routing, duty cycling, load balancing that minimizes the energy consumption of the network under reliability and delay requirements has been proposed. No protocol in the literature guarantees adaptation to reliability and delay requirements over multi-hop communication, with optimizing the energy consumption. Especially, our original contribution is as follows:

1. We provide explicit analytical relations of the reliability, delay and total energy consumption as a function of MAC, routing, physical layer, duty-cycle and

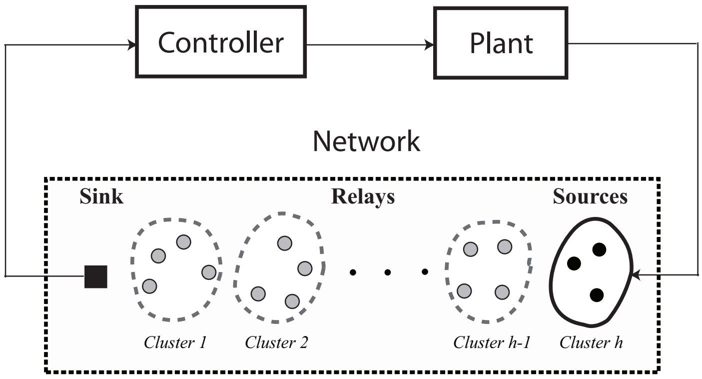


Figure 1: Wireless control loop. An wireless network closes the loop from sensors to controller. The network includes source nodes (black dots) attached to the plant, $h-1$ relay clusters (grey dots), and a sink (black rectangular) attached to the controller.

radio power. The approach is based on simple yet good approximations whose accuracy is systematically verified.

2. The analytical relations allow us to pose and solve a mixed integer-real optimization problem where the energy minimization is achieved under tunable reliability and delay requirements.

3. Based on this optimization, we develop a novel algorithm that allows for rapid deployment and self-adaptation of the network to traffic variations and channel conditions, and guarantee the application requirements without heavy computation or communication overhead.

4. The protocol is implemented on a test-bed using Tmote Sky sensors [12]. We show by analysis and experimental evaluation the benefits of our solution.

2 System Scenario

We consider the scenario depicted in Fig. 1, where a plant is remotely controlled over a WSN [7, 9]. Outputs of the plant are sampled at periodic intervals by the sensors with total packet generation rate of λ pckt/s (see Table 2 for main symbols used in the paper). We assume that source packets are transmitted to a sink, which is connected to the controller, over a multi-hop network of uniformly and randomly distributed relaying nodes. No direct communication is possible between source and sink. Relay nodes add their own information to incoming packet and relay these packets. When the controller receives the measurements, they are used in a control algorithm to compensate the control output. The control law is open based on an estimation of the delay induced by the WSN, which induces constraints on the communication delay and the packet loss probability. Packets must reach the sink within some minimum reliability and maximum delay. These boundaries are denoted as *application requirements* throughout this paper. The application requirements are chosen by the control algorithm designers. Since they can change from one control algorithm to another, or a control algorithm can ask to change them from time to time, we allow them to vary. We assume that nodes of the network cannot be recharged, so the operations must conserve energy. The system scenario is quite general, because it applies to any interconnection of a plant by a multi-hop WSN to a controller tolerating a certain degree of data loss and delay.

3 The Breath Protocol

The Breath protocol groups all N nodes between the sources and the sink with $h-1$ relay clusters. Data packets can be transmitted only from a cluster to the next cluster closer to the sink. Clustered network topology is supported in networks that require energy efficiency, since transmitting data through relays consumes less energy than routing directly to the sink [33]. In [34], a dynamic clustering method adapts the network parameters. In [33] and [35], a cluster header is selected based on the residual energy levels for clustered environments. However, the periodic selection of clustering may not be energy-efficient, and does not ensure the flexibility of the network to a time-varying wireless channel environment. A simpler geographic clustering is instead used in Breath. Nodes in the forwarding region send short beacon messages when they are available to receive data packets. Beacon messages are exploited to carry information related to the control parameters of the protocol.

In the following sections, we will describe the protocol stack and state machine of Breath in Section 3.1 and 3.2, respectively.

3.1 The Breath Protocol Stack

Breath uses a randomized routing, a hybrid TDMA at the MAC, radio power control at the physical layer, and sleeping disciplines. We give details in the following.

In many industrial environments, the wireless conditions vary heavily because of moving metal obstacles and other radio disturbances. In such situations, routing schemes that use fixed routing tables are not able to provide the flexibility over mobile equipments, physical design limitation and reconfiguration typical of an industrial control application. Fixed routing is inefficient in WSNs due to the cost of building and maintaining routing tables. To overcome this limitation, routing through a random sequence of hops has been introduced in [25]. The Breath protocol is built on an optimized random routing, where next hop route is efficiently selected at random. Randomized routing allows us to reduce overhead because no node coordination or routing state needs to be maintained by the network. Robustness to node failures is also considerably increased by randomized routing. Therefore, nodes route data packets to next-hop nodes randomly selected in a forwarding region.

Each node, either transmitter or receiver, does not stay in an active state all time, but goes to sleep for a random amount of time, which depends on the traffic and channel conditions. Since traffic, wireless channel, and network topology may be time-varying, the Breath protocol uses a randomized duty-cycling algorithm. Sleep disciplines turn off a node whenever its presence is not required for the correct operation of the network. GAF [17], SPAN [18] and S-MAC [36] focus on controlling the effective network topology by selecting a connected set of nodes to be active and turning off the rest of the nodes. These approaches require extra communication, since nodes maintain partial knowledge of the state of their individual neighbors. In Breath, each node goes to sleep for an amount of time that is a random variable dependent on traffic and network conditions. Let μ_c be the cumulative wake-up rate of each cluster, i.e., the sum of the wake-up rates that a node sees from all nodes of the next cluster. The cumulative wake-up rate of each cluster must be the same for each cluster to avoid congestions and bottlenecks.

The MAC of Breath is based on a CSMA/CA mechanism similar to IEEE 802.15.4. Both data packets and beacon packets are transmitted using the same MAC. Specifically, the CSMA/CA checks the channel activity by performing clear channel assessment (CCA) before the transmission can commence. Each node maintains a variable NB for each transmission attempt, which is initialized to 0 and counts the number of additional backoffs the algorithm does while attempting the current transmission of a packet. Each backoff unit has duration T_{ca} ms. Before performing CCAs a node takes a backoff of $\mathrm{random}(0, W-1)$ backoff units i.e., a random number of backoffs with uniformly distributed over $0, 1, \ldots, W-1$. If the CCA fails, i.e., the channel is busy, NB is increased by one and the transmission is delayed of $\mathrm{random}(0, W-1)$ backoff periods. This operation is repeated at most M_{ca} times, after which a packet is discarded.

The Breath protocol assumes that each node has a rough knowledge of its location. This information, which is commonly required for the applications we are targeting [2], can be obtained running a coarse positioning algorithm, or using the Received Signal Strength Indicator (RSSI), which is typically provided by off-the-shelf sensor nodes [37]. Some radio chip already provide a location engine based on RSSI [38]. Location information is needed for tuning the transmit radio power and to change the number of hops, as we will see later. The energy spent for radio transmission plays an important role in the energy budget and for the interference in the network. Breath, therefore, includes an effective radio power control algorithm.

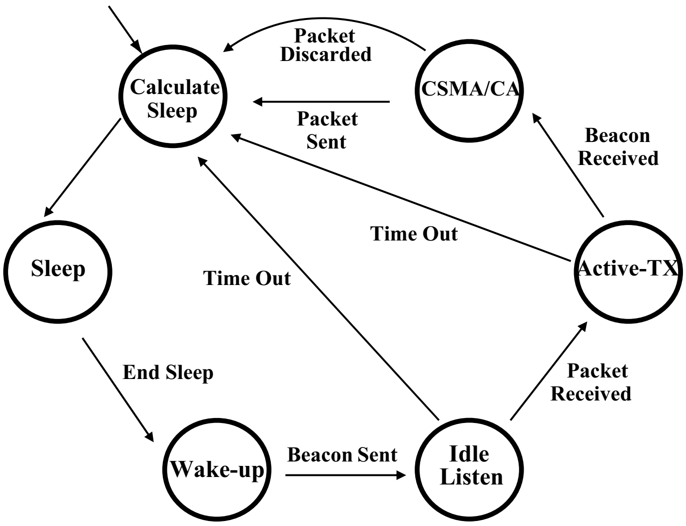


Figure 2: State machine description of a relay node executing the Breath protocol.

3.2 State Machine Description

Breath distinguishes between three node classes: source nodes, relays, and the sink.

The source nodes wake-up as soon as they sense packets. Before sending packets, a source node waits for a beacon message from the cluster of nodes closer to the source. Upon the reception of a beacon, the node sends the packet.

Consider a relay node k. Its detailed behavior is illustrated by the state machine of Fig. 2, as we describe in the following:

- **Calculate Sleep State**: the node calculates the parameter μ_k for the next sleeping time and generates an exponentially distributed random variable having average $1/\mu_k$. After this the node goes back to the Sleep State. μ_k is computed such that the cumulative wake-up rate of the cluster μ_c in ensured.
- **Sleep State**: the node turns off its radio and starts a timer whose duration is an exponentially distributed random variable with average $1/\mu_k$. When the timer expires, the node goes to the Wake-up State.
- **Wake-up State**: the node turns its beacon channel on, and broadcasts a beacon indicating its location. Then, it switches to listen to the data channel, and it goes to the Idle Listen State.
- **Idle Listen State**: the node starts a timer of a fixed duration that must be long enough to receive a packet. If a data packet is received, the timer is discarded,

the node goes to the Active-TX State, and its radio is switched from the data channel to the beacon channel. If the timer expires before any data packet is received, the node goes to the Calculate Sleep State.

- **Active-TX State**: the node starts a waiting timer of a fixed duration. If the node receives the first beacon coming from a node in the forwarding region within the waiting time, it retrieves the node ID and goes to the CSMA/CA State. Otherwise if the waiting timer is expired before receiving a beacon, the node goes to the Calculate Sleep State.
- **CSMA/CA State**: the node switches its radio to hear the data channel, and it tries to send a data packet to a node in the next cluster by the CSMA/CA MAC. If the channel is not clean within the maximum number of tries, the node discards the data packet and goes to the Calculate Sleep State. If the channel is clear within the maximum number of attempts, the node transmits the data packet using an appropriate level of radio power and goes to the Calculate Sleep State.

The sink node sends periodically beacon messages to the last cluster of the network to receive data packets. Such a node estimates periodically the traffic rate and the wireless channel conditions. By using this information, the sink runs an algorithm to optimize the protocol parameters, as we describe in Section 4. Once the results of the optimization are achieved, they are communicated to the relays by beacons. According to the protocol given above, the packet delivery depends on the traffic rate, the channel conditions, number of forwarding regions, and the cumulative wake-up time. In the next sections, we show how to model and optimize online these parameters.

4 Protocol Optimization

The protocol is optimized dynamically by a constrained optimization problem. The objective function, denoted by $E_{\text{tot}}(h, \mu_c)$, is the total energy consumption for transmitting and receiving packets from the sources to the sink. The constraint are given by the packet reception probability and delay probability. The optimization problem is

$$\min_{h,\mu_c} \quad E_{\text{tot}}(h, \mu_c) \tag{1a}$$

$$\text{s.t.} \quad R(h, \mu_c) \geq \Omega \,, \tag{1b}$$

$$\Pr[D(h, \mu_c) \leq \tau] \geq \Delta \,, \tag{1c}$$

$$h \geq 2 \,, \tag{1d}$$

$$\mu_{\min} \leq \mu_c \leq \mu_{\max} \,. \tag{1e}$$

The decision variables are the cumulative wake-up rate μ_c of each cluster and the number of relay clusters, $h - 1$. $R(h, \mu_c)$ is the probability of successful packet delivery (reliability) from the source cluster to the sink, and Ω is the minimum desired probability. $D(h, \mu_c)$ is a random variable describing the delay to transmit a packet from the source cluster to the sink. τ is the desired maximum delay, and Δ is the minimum probability with which such a maximum delay should be achieved. The next-to-last constraint is due to that there is at least two hops from the sources to the sink. The constraint $\mu_{\min} \leq \mu_c \leq \mu_{\max}$ is due to that the wake-up rate cannot be less than a minimum value $\mu_{\min}$, and larger than a maximum value $\mu_{\max}$ due to hardware reasons. Note that Problem (1) is a mixed integer-real optimization problem, because μ_c is real and h is integer. We need to have Δ and Ω close to one. We let $\Delta \geq 0.95$ and $\Omega \geq 0.9$, namely we assume that the delay τ must be achieved at least with a probability of 95%, and the reliability must be larger than 90%. We remark that τ, Δ, and Ω are application requirements, and h, μ_c and nodes' radio transmit power are *protocol parameters* that must be adapted to the traffic rate λ, the wireless channel conditions, and the application requirements for an efficient network operation.
In the following sections, we will propose an approach to model the quantities of Problem (1), along with a strategy to achieve the optimal solution, namely the values of h^* and μ_c^* that minimize the cost function and satisfy the application requirements. As we will see later, the system complexity prevents us to derive the exact expressions for the analytical relations of the optimization problem. An approximation of the requirements and an upper bound of the energy consumption will be used.

5 Modelling of the Protocol

In this section, we model the reliability, packet delay distribution and total energy consumption of the network.

5.1 Reliability Constraint

In this subsection, we provide an analytical expression for the reliability constraint (1b) in Problem (1).
A data packet can be lost at a hop because of a bad wireless channel or packet collisions. The collision probability is determined by the CSMA/CA MAC. Therefore, to analyze such a behavior, we use a Markov chain. The approach is similar to the one proposed in [39] and [40] (see also [41–43]). Let m be the maximum backoff stage, and W be the maximum backoff time of CSMA/CA. Let $s(t) \in \{0, \ldots, m\}$ and $b(t) \in \{0, \ldots, W - 1\}$ be the stochastic process representing the backoff stage and the backoff time counter, respectively. The delay spent before a node senses the channel idle is modelled by the Markov chain depicted in Fig. 3. The Markov chain state is $(s(t), b(t))$, where $b(t) = -1$ refers to the assessment of the

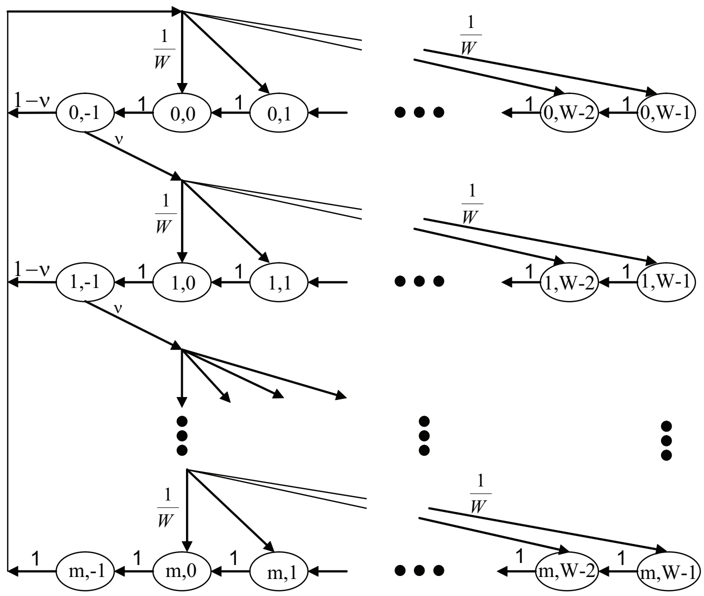


Figure 3: Markov chain model for CSMA/CA state evolution of Breath.

channel state during CCA. Denote the Markov chain's steady-state probabilities by $b_{i,k} = \Pr\{(s(t), b(t)) = (i,k)\}$. They allow us to compute the probability of successful transmission in CSMA/CA as the probability that exactly one node transmits and $n-1$ are silent:

$$\psi_{\rm sc}(n) = \frac{\zeta(1-\zeta)^{n-1}}{1-(1-\zeta)^n},$$

where

$$\zeta = \sum_{i=0}^{m} b_{i,0} = \frac{2}{W+3}.$$

From the Markov chain, we derive also the busy channel probability $\nu(n)$, which is

$$\nu(n) = \frac{1-(1-\tau)^{n-1}}{2-(1-\tau)^{n-1}}. \tag{2}$$

We will use this probability in Section 5.3.

The probability of successful transmission in CSMA/CA $\psi_{\rm sc}(n)$ depends on the number of nodes n that are contending to transmit packets. We need therefore to compute the probability $\psi_{\rm sb}(\mu_c, n)$ that a generic number n of contending nodes compete within a period $1/\lambda$ to transmit a data packet. By recalling that the cumulative wake-up rate is exponentially distributed random variable with intensity μ_c,

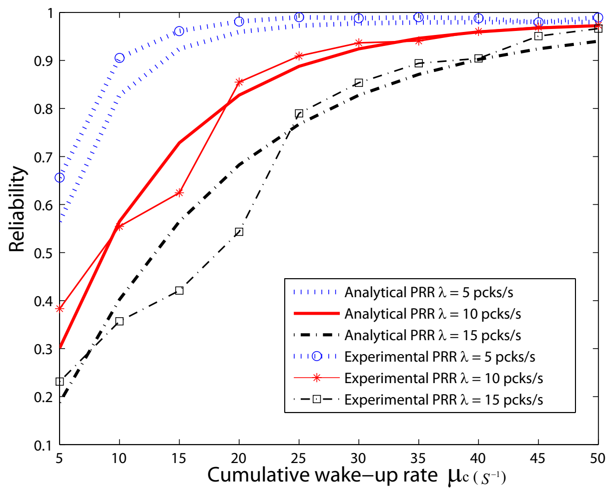


Figure 4: Reliability as obtained by Eq. (4) and experimental results as a function of μ_c. Curves refer to traffic rates $\lambda = 5, 10, 15$ pckt/s for $h = 2$ hops and $N = 15$ nodes.

and noting that $e^{-\mu_c n/\lambda}$ is the probability to have more than n contending nodes, we conclude that

$$\psi_{\rm sb}(\mu_c, n) = e^{-\mu_c(n-1)/\lambda} - e^{-\mu_c n/\lambda} \,. \tag{3}$$

Hence, $\psi_{\rm sb}(n)$ and $\psi_{\rm sc}(n)$, we the reliability is

$$R(h, \mu_c) = \prod_{i=1}^{h} p_i \sum_{n=1}^{\infty} \psi_{\rm sb}(n)\psi_{\rm sc}(n) \,, \tag{4}$$

where p_i denotes the probability of successful packet reception during a single-hop transmission from cluster i to cluster $i-1$.

Since the components of the sum in (4) with $n \geq 2$ give a small contribution, we set $n = 2$ and validate (4) with experimental results. Fig. 4 reports the reliability vs μ_c, as obtained by Eq. (4) with $n = 2$ and experiments for a two-hop network. We see that (4) provides a good approximation of the experimental results because it is always around 5% of the experiments for reliability values of practical interest (larger than 0.7). The same behavior is found for h up to 4.

We can rewrite the reliability constraint $R(h, \mu_c) \geq \Omega$ by using (4) with $n = 2$, thus obtaining

$$\mu_c \geq f_r(h, \Omega) \triangleq \lambda \ln(2C_r) - \lambda \ln \left[C_r - 1 + \sqrt{(C_r - 1)^2 - 4C_r \left(\Omega^{1/h} / p_{\min} - 1 \right)} \right], \quad (5)$$

where $C_r = \zeta(1-\zeta)/(1-(1-\zeta)^2)$, and $p_{\min} = \min(p_1, \ldots, p_h)$. Note that we used the worst channel condition of the network $p_{\min}$, which is acceptable for optimization purpose because in doing so we consider the minimum of (4). Since the argument of the square root in (5) must be positive, an additional constraint is introduced:

$$h \leq h_r \triangleq \frac{\ln(\Omega)}{\ln(p_{\min})}. \quad (6)$$

We will use (5) and (6) in Section 6 to find the solution of Problem (1). Now, we turn our attention to the delay constraint.

5.2 Delay Constraint

The delay $D(h, \mu_c)$ between source to sink is given by the sum of the delays experienced by a packet at each hop. There are two sources of delay:

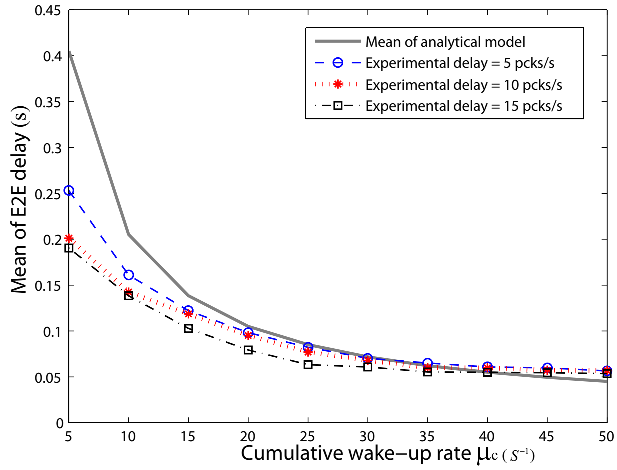
- Time to wait before the first wake-up of a node in the next cluster: Let such a time be denoted with α_i for cluster i.
- Time to wait for clean channel: Since the Breath protocol uses CSMA/CA, a node spends a random time before sensing idle channel. Denote with ε_i such a time for cluster i.

By summing these delays per each hop, we obtain the delay model

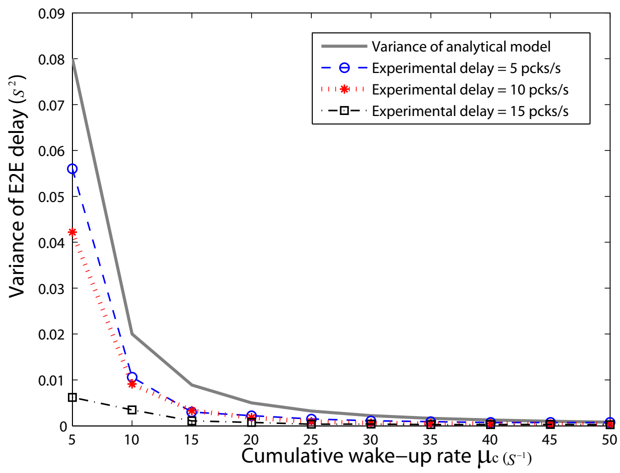
$$D(h, \mu_c) = \sum_{i=1}^{h} (\alpha_i + \varepsilon_i). \quad (7)$$

In this equation, α_i is an exponentially distributed random variable whose intensity μ_c is the sum of the wake-up intensities of the nodes in the next cluster. Characterization of ε_i is more difficult, owing to the backoff mechanism of the CSMA/CA algorithm. However, we assume that the backoff time can be approximated by a Gaussian distribution whose average is matched with the average and standard deviation of a uniformly distributed random variable between 0 and $(M_{\text{ca}} + 1)(W - 1)$. Namely,

$$\varepsilon_i \in \mathcal{N} \left(\frac{(M_{\text{ca}} + 1)(W - 1)T_{\text{ca}}}{2}, \frac{(M_{\text{ca}} + 1)(W^2 - 1)T_{\text{ca}}^2}{12} \right).$$



(a) average delay



(b) variance of delay

Figure 5: Validation of average and variance of delay given by Eq. (8), (9) by experimental results, respectively. The traffic rates $\lambda = 5, 10, 15$ pckt/s are considered w.r.t. wake-up rates μ_c from 5 to 50, $h = 2$ hops and $N = 15$ nodes.

According to such an assumption, the delay $D(h, \mu_c)$ is approximated by a Gaussian random variable $\mathcal{N}(\mu_D, \sigma_D^2)$, where

$$\mu_D = \frac{h}{\mu_c} + \frac{h(M_{\text{ca}}+1)(W-1)T_{\text{ca}}}{2}\,, \tag{8}$$

$$\sigma_D^2 = \frac{h}{\mu_c^2} + \frac{h(M_{\text{ca}}+1)(W^2-1)T_{\text{ca}}^2}{12}\,. \tag{9}$$

We validated these approximations by comparing the analysis with experimental results. Figs. 5(a), 5(b) show the mean and variance of the delay given by Eq. (8), (9) and the experimental results, respectively. The analytical model describes well the experimental data because it gives an upper bound for wake-up rates up to $35\ \text{s}^{-1}$, and then the model underestimates the experimental result less than 5%. These properties are quite useful for optimization purposes. Same dependence is found on h up to 4.

We are now in the position to express the delay constraint in Problem (1) by using Eqs. (8) and (9) that we just derived:

$$\Pr[D \le \tau] \approx 1 - Q\left(\frac{\tau - \mu_D}{\sigma_D}\right) \ge \Delta\,, \tag{10}$$

where $Q(x) = 1/\sqrt{2\pi}\int_x^\infty e^{-t^2/2}dt$ is the complementary standard Gaussian distribution. After some manipulations, it follows that (10) can be rewritten as

$$\mu_c \ge \frac{12\,C_{d1}\,h + 2\sqrt{3\,C_{d3}\,h\,\left[12\,C_{d1}^2 + C_{d2}\,(h - C_{d3})\right]}}{12\,C_{d1}^2 - C_{d2}\,C_{d3}}\,,$$

where

$$C_{d1} = \tau - \frac{h(M_{\text{ca}}+1)(W-1)T_{\text{ca}}}{2}\,,$$
$$C_{d2} = h(M_{\text{ca}}+1)(W^2-1)T_{\text{ca}}^2\,,$$
$$C_{d3} = \left(Q^{-1}(1-\Delta)\right)^2.$$

Since $T_{\text{ca}}^2 = 0.1024 \times 10^{-6}$ [12], and h, M_{ca}, W are positive integers, it follows that $T_{\text{ca}}^2 \ll h(M_{\text{ca}}+1)(W^2-1)$. Then $C_{d2} \ll C_{d1}$ and (10) is approximated by

$$\mu_c \ge f_d(h,\tau,\Delta) \triangleq \frac{2\left[h + Q^{-1}(1-\Delta)\sqrt{h}\right]}{2\tau - h(M_{\text{ca}}+1)(W-1)T_{\text{ca}}}\,. \tag{11}$$

Inequality (11) has been derived under the additional constraint

$$h \le h_d \triangleq \frac{2\tau}{(M_{\text{ca}}+1)(W-1)T_{\text{ca}}}\,. \tag{12}$$

We will use (11) and (12) in Section 6 to find the solution of the optimization problem (1). Now, we investigate the total energy consumption.

5.3 Energy Consumption

The total energy consumption is

$$E_{\text{tot}}(h, \mu_c) = E_{\text{pck}}(h, \mu_c) + E_{wu}(h, \mu_c)\,, \tag{13}$$

where $E_{\text{pck}}(h, \mu_c)$ is the total energy for transmission and reception of data packets and $E_{wu}(h, \mu_c)$ is the energy consumption for wake-up, listening and beaconing during a time T, which we characterize in Section 5.3, 5.3, respectively.

Data Packet Communication Energy

Assuming h hops, and recalling that sources emits λ pckt/s,

$$E_{\text{pck}}(h, \mu_c) = T\lambda \sum_{i=1}^{h} \left[Q_m(d_i) + \frac{A_{\text{rx}}}{\mu_c} + E_{\text{ca}}(\mu_c) + E_r \right]\,, \tag{14}$$

where E_r accounts for the fixed cost of the RF circuit for the reception of a data packet. The term $Q_m(d_i)$ is the energy consumption for radio transmission, where d_i is the transmission distance to which a data packet has to be sent. The term $E_{\text{ca}}(\mu_c)$ is the energy spent during the CSMA/CA state.
The energy model given by Eq. (14) is derived under the assumption that all packets generated at the sources reach the sink. Obviously, some packet may be lost before reaching the sink, therefore (14) gives a upper bound on the energy consumption. This is reasonable, since our goal is the minimization of the cost function.
The energy spent for radio transmission is a function of the radio power used to transmit packets:

$$Q_m(d_i) = V\, I(P_t(d_i))\, t_m\,, \tag{15}$$

where V is the voltage consumption of the RF circuit at the node, t_m is the transmission time of a data packet, $I(P_t(d_i))$ is the current consumption of the electronic circuit needed to transmit packets at radio power $P_t(d_i)$, and d_i is the distance from the transmitter which a packet must reach to with some desired probability. The relation between the current consumption and radio power depends on the hardware platform. For Tmote Sky sensors, it holds that [44]

$$\begin{aligned} I(P_t(d_i)) \approx & -19P_t(d_i)^4 + 53P_t(d_i)^3 - 53P_t(d_i)^2 \\ & + 29P_t(d_i) + 8.7\,. \end{aligned}$$

Given this approximation, minimization of $Q_m(d_i)$ is achieved by minimizing $P_t(d_i)$. $P_t(d_i)$ can be minimized by computing the minimum radio power that ensures packets to reach a given distance with a given probability, as we see next.
The optimal transmit power is derived by considering the distribution of the Signal to Interference plus Noise Ratio (SINR). By imposing a requirement p_{con} on the

probability of successful packet reception at a distance d_k from node k, we can translate the requirement on the average SINR, thus obtaining a bound $\overline{\gamma}_c$ on such an average SINR. From this we can then derive the transmit radio power necessary to successfully receive packets at a distance d_k with probability $p_{\rm con}$. It follows that the minimum transmit power is [45]

$$P_t(d_k)_{\rm ~dB} = \overline{\gamma}_{c\rm ~dB} + {\rm PL}(d_0)_{\rm ~dB} + 10\,\beta_k\,\log_{10}\frac{d_k}{d_0} + P_{n\rm ~dB} - \frac{\ln 10}{20}\sigma^2_{\gamma_k\rm ~dB}\,,$$

where ${\rm PL}(d_0)_{\rm ~dB}$ is the path loss at a reference distance d_0, β_k is the path loss decay constant, P_n is the noise floor, and σ_{γ_k} is the variance of the SINR (see [45] for details). We remark that the power $P_t(d_k)_{\rm ~dB}$ minimizes the energy spent for radio transmission in Eq. (15). Notice that the actual packet reception probability p_i may fluctuate around $p_{\rm con}$ due to the delay of this power control and the limited maximum transmit power.

In the following, we characterize $E_{\rm ca}$. Consider the energy spent for transmission of a data packet in the i th cluster. Let $E_{\rm ca}$ is the energy spent by a node to check the channel status by the CSMA/CA algorithm upon the reception of a beacon. This energy, which is due to CCA, is dependent on the maximum number of tries $M_{\rm ca}$. We have two situations: the number of contending nodes n attempting to transmit a data packet is less then $M_{\rm ca}$, or the number of contending nodes is larger than $M_{\rm ca}$. If $n < M_{\rm ca} + 1$, all nodes will succeed to sense a clean channel with the energy $E_{\rm ca1}(n)$, otherwise we need to consider the transmission success and failure probabilities to perform CCA with the energy $E_{\rm ca2}(n)$, which is the function of the busy channel probability $\nu(k)$ conditioned on k contending nodes defined by Eq. (2), see the details in [45]

By summing two the energy components $E_{\rm ca1}(n), E_{\rm ca2}(n)$, the average energy consumption spent by the CSMA/CA is

$$E_{\rm ca}(\mu_c) = \sum_{n=1}^{\infty} \psi_{\rm sb}(\mu_c, n)\,[E_{\rm ca1}(n)\,u(M_{\rm ca} - n) + E_{\rm ca2}(n)\,u(n - M_{\rm ca} - 1)]\ , \qquad (16)$$

where $\psi_{\rm sb}(\mu_c, n)$ is the probability to have n contending nodes in a cluster given by Eq. (3) and $u(x) = 1$ if $x \geq 0$, whereas $u(x) = 0$ otherwise.

Control Signalling Energy

A node randomly cycles between an awake state and a sleep state. Each time a node wakes up, it spends an energy given by the power needed to wake-up A_w during the wake-up time T_w, plus the energy to listen for the reception of a data packet within a maximum time $T_{\rm ac}$. After a node wakes up, it transmits a beacon

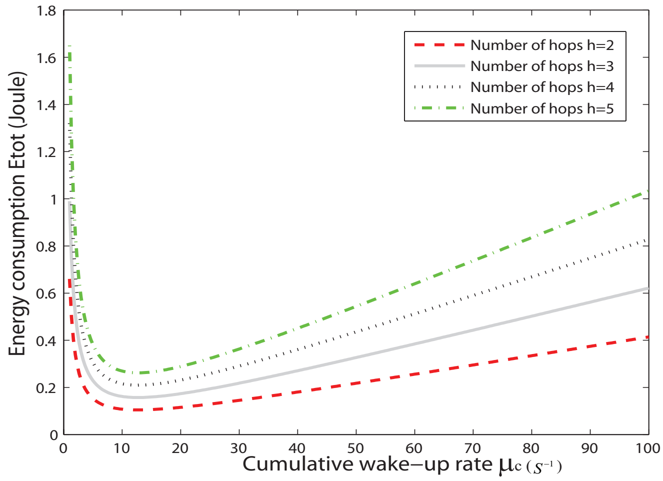


Figure 6: Total energy consumption given by Eq. (18) for a different number of hops ($h = 2, 3, 4, 5$) over wake-up rates μ_c from 1 to 100 in traffic rate ($\lambda = 5$ pcks/s) and $N = 15$ nodes.

to the next cluster. Let the wireless channel loss probability be $1 - p_i$ of i cluster, then nodes of $i - 1$ cluster have to wake-up on average $1/p_i$ times to create the effect of a single wake-up so that a transmitter node successfully receives a beacon. Recalling that there are h hops and a cumulative wake-up rate per cluster μ_c, the total cost in a time T for wake-ups and beaconing is

$$E_{wu}(h, \mu_c) = T\mu_c \sum_{i=1}^{h} \frac{1}{p_i} \left[Q_b(d_i) + A_w T_w + A_{\text{rx}}(T_{\text{ac}} - T_w)\right] , \qquad (17)$$

where $Q_b(d_i)$ is the expected energy consumption to transmit a beacon message at the distance d_i.

Total Energy Consumption

Here we put together the energy analysis developed in the previous two sections. The total energy consumption is

$$\begin{aligned} E_{\text{tot}}(h, \mu_c) = & T\lambda \left[Q_m\left(\frac{S}{h-1}\right) + Q_m\left(\frac{S}{h-1}\right)(h-1)u(h-1)\right. \qquad (18) \\ & \left. + \; h\left(\frac{A_{\text{rx}}}{\mu_c} + E_{\text{ca}}(\mu_c) + E_r\right)\right] + \frac{T\mu_c}{p_{\min}}\left[2Q_b\left(\frac{S}{h-1}\right)\right. \\ & \left. + Q_b\left(\frac{2S}{h-1}\right)(h-2)\, u(h-2) + h\left(A_w T_w + A_{\text{rx}}(T_{\text{ac}} - T_w)\right)\right] . \end{aligned}$$

where we upper bounded (14) and (17) by considering the worst distance to which data and beacon packets must be sent, which are $S/(h-1)$ and $2S/(h-1)$, and the worst reception probability $p_{\min}$.

Fig. 6 shows the energy given by Eq. (18) as a function of the number of hops h over different wake-up rates μ_c. The total energy consumption increases with h given μ_c. This is due to that, for a given number of total nodes present in the network, increasing h implies higher wake-up rates per node. In other words, increasing the number of hops is energy inefficient. Observe also that a low wake-up rate does not minimize the total energy consumption, because of the longer waiting time to receive a beacon message that such a rate causes. Hence, there is a tradeoff between the energy consumption for wake-up and waiting to get a beacon message. We explore this tradeoff for optimization problem in the following section.

6 Optimal Protocol Parameters

In this section we give the optimal protocol parameters used by Breath. Consider the reliability and delay constraints, and the total energy consumption as investigated in Sections 5.1, 5.2, and 5.3. The optimization problem (1) becomes

$$\begin{aligned} \min_{h,\mu_c} \quad & E_{\text{tot}}(h,\mu_c) \\ \text{s.t.} \quad & \mu_c \geq \max(f_r(h,\Omega), f_d(h,\tau,\Delta))\,, \\ & 2 \leq h \leq \min(h_r, h_d)\ , \\ & \mu_{\min} \leq \mu_c \leq \mu_{\max}\,, \end{aligned} \tag{19}$$

where the first constraint comes from (5) and (6), and the second from (11) and (12). We assume that this problem is feasible. Infeasibility means that for any $h = 2, \ldots, \min(h_r, h_d)$, then $\mu_c \geq \max(f_r(h,\Omega), f_d(h,\tau,\Delta)) > \mu_{\max}$, namely it is not possible to guarantee the satisfaction of the reliability and delay constraint given the application requirements. This means that the application requirements must be relaxed, so that feasibility is ensured and the problem can be solved. The solution of this optimization problem, h^* and μ_c^*, is derived in the following.

By using the numerical values given for the Tmote Sky sensors [12] for all the constants in the optimization problem, the cost function of Problem (19) is increasing in h and convex in μ_c. This allows us to derive the optimal solution in two steps: for each value of $h = 2, \ldots, \min(h_r, h_d)$, the cost function is minimized for μ_c, achieving $\mu_c^*(h)$. Then, the optimal solution if found in the pair $h, \mu_c^*(h)$ that gives the minimum energy consumption. We describe this procedure next.

Let h be fixed. From the properties the cost function of Problem (19), the optimal solution $\mu_c^*(h)$ is attained either at the minimum of the cost function or at the boundaries of the feasibility region given by the requirements on μ_c. The minimum of the cost function can be achieved by taking its derivative with respect

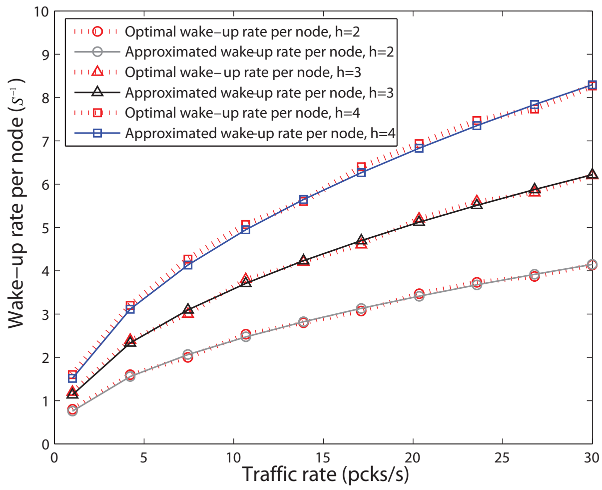


Figure 7: Wake-up rate that minimizes the total energy consumption and approximated wake-up rate as obtained by (20) for different number of hops ($h = 2, 3, 4$), traffic rates λ from 1 to 30 pcks/s and $N = 15$. The y-axis was normalized by 15.

to μ_c. To obtain this derivative in an explicit form, we assume that CSMA/CA energy consumption can be approximated by a constant value since the numerical value is smaller than other factors. Under this assumption, the minimization by the derivative is approximated by

$$\mu_e(h) = (p_{\min}\,\lambda\,A_{\rm tx})^{\frac{1}{2}} \left[\frac{h-2}{h} Q_b\left(\frac{2S}{h-1}\right) u(h-2) + \frac{2}{h} Q_b\left(\frac{S}{h-1}\right) + A_w T_w + A_{\rm rx}(T_{\rm ac} - T_w)\right]^{-\frac{1}{2}} . \qquad (20)$$

In Fig. 7 we check the validity of this approximation. The figure reports the approximated minimum as obtained by Eq. (20) compared to the wake-up rate that minimizes the actual energy consumption as obtained by a numerical minimization algorithm. The approximation is very tight because the error is less than 2%.

By using Eq. (20), we see that an optimal solution μ_c^* provided h is given by $\mu_e(h)$ if $\mu_{\min} \le \mu_e(h) \le \mu_{\max}$ and $\mu_e(h) \ge \max(f_r(h,\Omega), f_d(h,\tau,\Delta))$, otherwise an optimal solution is given by the value between $\mu_{\max}$ and $\max(f_r(h,\Omega), f_d(h,\tau,\Delta))$ that minimizes $E_{\rm tot}(h,\mu_c)$. Therefore, for any $h = 2,\ldots,\min(h_r,h_d)$, we compute μ_c^*. Then, the optimal solution h^* and μ_c^* is given by the pair μ_c^*, h that minimizes the cost function. This procedure to compute the optimal solution is illustrated by Algorithm 3.

7 Adaptation Mechanisms

In the previous sections, we showed how to determine the optimal number of clusters and cumulative wake-up rate by solving an optimization problem. Here, we present in detail some adaptation algorithms that the sink must run to determine correctly h^* and μ_c^* as the traffic rate and channel conditions changes. These algorithms allow us to adapt the protocol parameters to the traffic rate and channel condition without high message overhead.

7.1 Traffic Rate and Channel Estimation

The sink node estimates the traffic rate λ and the worst channel probability $p_{\min}$ of the network. To estimate the global minimum of the worst channel condition, each p_i should be estimated at a local node and sent to the sink for each link of the path $i = 1, \dots, h$. This might increase considerably the packet size. To avoid this, we propose the following strategy. Consider a relay node of the i th cluster. It estimates p_i by the signal of the beacon packet. Then the nodes compares p_i with the channel condition information carried by the received data packet and selects the minimum. This minimum is then encoded in the data packet and sent with it to the next-hop node. After the sink node retrieves the channel condition of the route by receiving a data packet, it computes an average of the worst channel conditions among the last received data packets. Using this estimate, the sink solves the optimization problem running Algorithm 3. Afterwards, the return value of the algorithm, h^* and μ_c^*, can be piggybacked on beacons that the sink sends toward the relays closer to the sink. Then, these protocol parameters are forwarded when the nodes wake-up and send beacons to the next cluster toward the sources. During the initial state, nodes set $h = 2$ before receiving a beacon.

7.2 Wake-up rate and Radio Power Adaptation

Once a cluster received μ_c^*, each node in the cluster must adapt its wake-up rate so that the cluster generates such a cumulative wake-up rate. We consider the natural solution of distributing μ_c^* equally between all nodes of the cluster. Let μ_k be the wake-up rate of node k, and suppose that there are l nodes in a cluster. The fair solution is $\mu_k = \mu_c^*/l$ for any node. However, a node does not know and cannot estimate efficiently the number of nodes in its cluster.

To overcome this problem, we follow the same approach proposed in [30], where an Additive Increase and Multiplicative Decrease (AIMD) algorithm leads to a fair distribution of the wake-up duties within a single cluster. Specifically, each node that is waiting to forward a data packet observes the time before the first wake-up in the forwarding region. Starting from this observation, it estimates the cumulative

Algorithm 3: **Algorithm for the computation of the optimal solution of Problem** (19) **.**

Input: Requirements Ω, τ, Δ, feasible range $\mu_{\min}, \mu_{\max}$
Output: h^*, μ_c^*
begin
  $h^* \leftarrow 2$;
  if $\left(\mu_e(h^*) \geq \max(f_r(h^*,\Omega), f_d(h^*,\tau,\Delta))\right) \cap \left(\mu_{\min} \leq \mu_e(h^*) \leq \mu_{\max}\right)$ **then**
    $\mu_c^* \leftarrow \mu_e(h^*)$;
    $\mathrm{SAT} \leftarrow \mathrm{true}$;
  else if $\left[(\mu_e(h^*) \geq \mu_{\max}) \cup (\mu_e(h^*) \leq \mu_{\min}) \cup (\mu_e(h^*) \leq \max(f_r(h^*,\Omega), f_d(h^*,\tau,\Delta)))\right] \cap \left(\mu_{\min} \leq \max(f_r(h^*,\Omega), f_d(h^*,\tau,\Delta)) \leq \mu_{\max}\right)$ **then**
    $\mu_c^* \leftarrow \max(f_r(h^*,\Omega), f_d(h^*,\tau,\Delta))$;
    $\mathrm{SAT} \leftarrow \mathrm{true}$;
  else
    $\mu_c^* \leftarrow \mu_{\max}$;
    $\mathrm{SAT} \leftarrow \mathrm{fail}$;
  $E \leftarrow E_{\mathrm{tot}}(h^*, \mu_c^*)$;
  for $h \leftarrow 3$ **to** $\min(h_r, h_d)$ **do**
    if $(\mu_e(h) \geq \max(f_r(h,\Omega), f_d(h,\tau,\Delta))) \cap (\mu_{\min} \leq \mu_e(h) \leq \mu_{\max})$ **then**
      $\mu_{\mathrm{tmp}} \leftarrow \mu_e(h)$;
      $\mathrm{SAT} \leftarrow \mathrm{true}$;
    else if $\left[(\mu_e(h) \geq \mu_{\max}) \cup (\mu_e(h) \leq \mu_{\min}) \cup (\mu_e(h) \leq \max(f_r(h,\Omega), f_d(h,\tau,\Delta)))\right] \cap \left(\mu_{\min} \leq \max(f_r(h,\Omega), f_d(h,\tau,\Delta)) \leq \mu_{\max}\right)$ **then**
      $\mu_{\mathrm{tmp}} \leftarrow \max(f_r(h,\Omega), f_d(h,\tau,\Delta))$;
      $\mathrm{SAT} \leftarrow \mathrm{true}$;
    else
      $\mu_{\mathrm{tmp}} \leftarrow \mu_{\max}$;
      $\mathrm{SAT} \leftarrow \mathrm{fail}$;
    $E_{\mathrm{tmp}} \leftarrow E_{\mathrm{tot}}(h, \mu_{\mathrm{tmp}})$;
    if $(E > E_{\mathrm{tmp}}) \cap \mathrm{SAT}$ **then**
      $h^* \leftarrow h$;
      $\mu_c^* \leftarrow \mu_{\mathrm{tmp}}$;
      $E \leftarrow E_{\mathrm{tmp}}$;
end

wake-up rate $\tilde{\mu}_c$ of the forwarding region and it compares it with the optimal value of the wake-up rate μ_c^* when a node receives a beacon. Note that the node retrieves information on h^*, μ_c^* and location information of the beacon node. If $\tilde{\mu}_c < \mu_c^*$ the node sends by the data packet an Additive Increase (AI) command for the wake-up rate of next-hop cluster, else it sends a Multiplicative Decrease (MD) command. Furthermore, the node updates the probability of successful transmission p_i based on the channel information using the RSSI and distance information d_k between its own location and beacon node. After the node updates the channel condition estimation, it sets the data packet transmission power to $P_t(d_k)$, and encodes the channel estimation in the packet as described in Section 7.2.

If a data packet is received, the node retrieves information on wake-up rate update: if AI then $\mu_k = \mu_k + \theta$, else $\mu_k = \mu_k/\phi$, where θ and ϕ are control parameters. From experimental results, we obtained that $\theta = 3$ and $\phi = 1.05$ achieve good performance. Furthermore, the node runs the reset mechanism for load balancing of wake-up rate as discussed in Subsection 7.2. The command on the wake-up rate variation is piggybacked on data packets and does not require any additional message.

However, this approach may generate a load balancing problem because of different wake-up rates among relays within a short period. Load balancing is a critical issue, since some nodes may wake-up at higher rate than desired rate of other nodes, thus wasting energy. To overcome this situation, each relay node runs a simple reset mechanism. We assign an upper and lower bound to the wake-up rate for each node. If the wake-up rate of a node is larger than the upper bound $(1+\xi)\mu_c^*(h^*-1)/N$ or is smaller than the lower bound $(1-\xi)\mu_c^*(h^*-1)/N$, then a node resets its wake-up rate to $\mu_c^*(h^*-1)/N$, where ξ assumes a small value and $(h^*-1)/N$ is an estimation of number of nodes per cluster.

8 Fundamental limits

Understanding the fundamental limits of Breath is critical for its appropriate use. This section focuses on the minimum number of relays required to support the protocol, and the minimum delay that can be set by the application.

8.1 Minimum number of nodes per cluster

The minimum number $n_{\min}$ of nodes per cluster to support the protocol with given reliability and delay requirements is

$$n_{\min} \geq \frac{\mu_c^*}{\mu_{k,\max}} ,$$

where $\mu_{k,\max}$ is the maximum wake-up rate of node k. By considering the worst active time for the duty cycle, we have

$$\mu_{k,\max} = (T_{\text{ac}} + T_{\text{be}})^{-1} ,$$

where $T_{\text{ac}} = 30$ ms is the maximum listening time to receive a data packet and $T_{\text{be}} = 500$ ms is the maximum waiting time before receiving a beacon [12].

8.2 Minimum delay

The minimum delay that the application can set is achieved by considering a very high wake-up rate per cluster. This minimizes the waiting time before receiving a beacon. Hence, by summing the delays of the CSMA/CA state and physical limits of the wireless channel, the minimum delay is

$$\tau \geq h\ [2(M_{\text{ca}} + 1)(W - 1)T_{\text{ca}} + 2\,T_{\text{prop}} + t_m + t_b]\ ,$$

where the fist term is the maximum delay of CSMA/CA state, T_{prop} is propagation delay, and t_m and t_b are, respectively, the transmission delay of data and beacon packets. Since $T_{\text{prop}} = 0.875$ ms, $t_m = 1.5$ ms and $t_b = 0.64$ ms [12], they can be basically ignored because they are negligible with respect to other delays.

9 Experimental Implementation

In this section we provide an extensive set of experiments to validate the Breath protocol. The experiments enable us to assess Breath in terms of reliability and delay in the packet transmission, and energy consumption of the network both in stationary and transitionary condition. The protocol was implemented on a test bed of Tmote Sky sensors [12], and was compared with a standard implementation of IEEE 802.15.4 [3], as we discuss next.
We consider a typical indoor environment with concrete walls. The experiments were performed in a static AWGN propagation and time-varying Rayleigh fading environment, respectively:

- AWGN environment: nodes and surrounding objects were static, with minimal time-varying changes in the wireless channel due to multi-path fading effects. In this case, the wireless channel is well described by an Additive White Gaussian Noise (AWGN) model.

- Rayleigh fading environment: obstacles were moved within the network, along a line of 20 m. Furthermore, a metal object was put in front of the source node, so the source node and the relays were not in line-of-sight. The source was moved on a distance of few tens of centimeters.

A node acted as source and generated packets periodically at different rates ($\lambda =$ 5, 10, and 15 pckt/s). 15 relays were placed to mimic the topology in Fig. 1. The sources was at a distance of 20 m far from the sink. The sink node collected packets and then computed the optimal solution using Algorithm 3. The delay requirement was set to $\tau = 1$s and the reliability to $\Omega = 0.9$ and 0.95. In other

words, we imposed that packet must reach destination within 1s with a probability of Ω. These requirements were chosen as representative for control applications.
We compared Breath against an implementation of the unslotted IEEE 802.15.4 [3] standard, which is similar to the randomized MAC that we use in this paper. In such an IEEE 802.15.4 implementation, we set nodes to a fixed sleep schedule, defined by CT_{ac} where C is integer number (recall that T_{ac} is the maximum node listening time in Breath). We defined the case $\mathcal{L}$ (low sleep), were the IEEE 802.15.4 implementation is set with $C = 1$, whereas we defined the case $\mathcal{H}$ (high sleep) by setting $C = 4$. The case $\mathcal{H}$ represents a fair comparison between Breath and IEEE 802.15.4, while in the case $\mathcal{L}$ nodes are let to listen much longer time than nodes in Breath. The power level in the IEEE 802.15.4 implementation where set to -5dBm. We set the IEEE 802.15.4 protocol parameters to default values *macMinBE* $= 3$, *aMaxBE* $= 5$, *macMaxCSMABackoffs* $= 4$. Details follows in the sequel.

9.1 Protocol Behavior for Stationary Requirements

In this subsection, we investigate the performance of Breath about the reliability, average delay and energy consumption that can be achieved in a stationary configuration of the requirements, i.e., during the experiment there was not change of application requirements. Data was collected out of 10 experiments, each lasting 1 hour.

Reliability

Fig. 8 indicates that the network converges by Breath to a stable error rate lower than $1-\Omega$ and hence satisfies the required reliability with traffic rate $\lambda = 10$ pckt/s, the delay requirement $\tau = 1$s, and $\Omega = 0.9$, 0.95. IEEE 802.15.4 $\mathcal{H}$ in AWGN channel provides the worse performance than the other protocols because of lower wake-up rate. Observe that $\Omega = 0.9$ in Rayleigh fading environment gives the better reliability than $\Omega = 0.95$ in AWGN channel due to higher wake-up rate to compensate the fading channel condition. Notice that the higher fluctuation of reliability between the number of received packets 2500 and 2800 for Rayleigh fading environment with $\Omega = 0.95$ is due to deep attenuations in the wireless channel.
Fig. 9 shows the reliability of Breath and IEEE 802.15.4 $\mathcal{L}$, $\mathcal{H}$ as a function of the reliability requirement $\Omega = 0.9, 0.95$ and traffic rate $\lambda = 5, 10, 15$ pckt/s in AWGN and Rayleigh fading environments, with the vertical bars indicating the standard deviation as obtained out of 10 experimental runs of 1 hour each. Observe that the reliability is stable around the required value for Breath, and this holds for different traffic rates and environments. However, IEEE 802.15.4 $\mathcal{L}$ and $\mathcal{H}$ do not ensure the reliability satisfaction for large traffic rates. Specifically, IEEE 802.15.4 $\mathcal{H}$ shows poor reliability in any case, and performance worsen as the the environment moves from the AWGN to the Rayleigh fading. Furthermore,

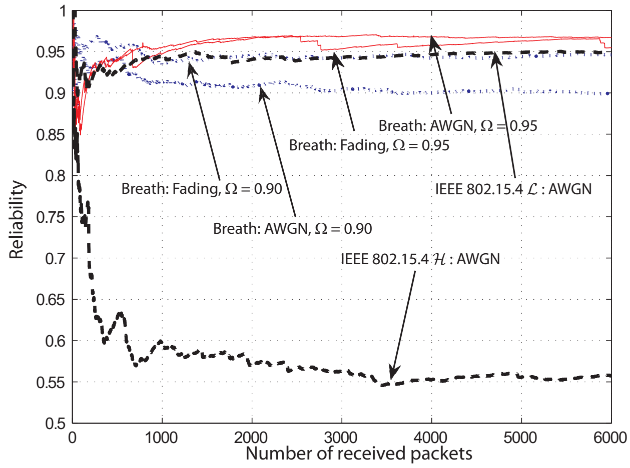


Figure 8: Convergence over time of the reliability for IEEE 802.15.4 $\mathcal{L}$, $\mathcal{H}$ in AWGN, and Breath with reliability requirements $\Omega = 0.9, 0.95$ and traffic rates $\lambda = 10$ pckt/s in AWGN and Rayleigh fading environments.

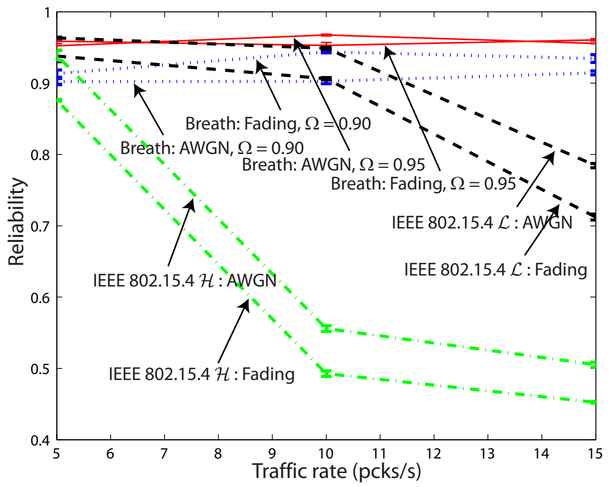


Figure 9: Reliability in IEEE 802.15.4 $\mathcal{L}$, $\mathcal{H}$, and Breath with requirement $\Omega = 0.9, 0.95$ for traffic rates $\lambda = 5, 10, 15$ pckt/s in AWGN and Rayleigh fading environments.

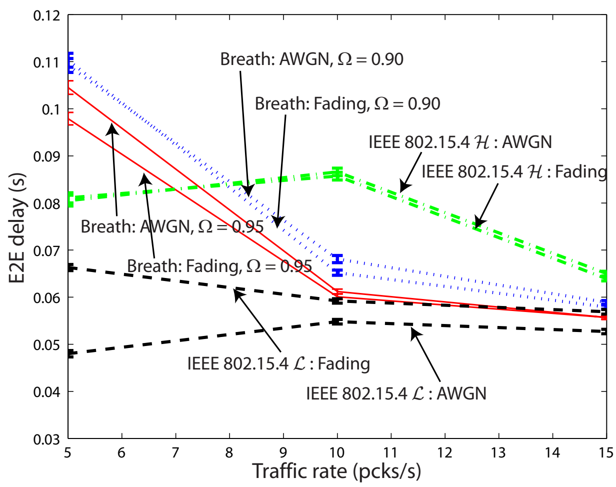


Figure 10: Temporal average of the delay of Breath and IEEE 802.15.4 $\mathcal{L}$, $\mathcal{H}$ with reliability requirement $\Omega = 0.9, 0.95$ and delay requirement $\tau = 1$ s over traffic rates $\lambda = 5, 10, 15$ pckt/s in AWGN and Rayleigh fading environment.

even though IEEE 802.15.4 $\mathcal{L}$ imposes that nodes wakes up more often, it does not guarantee a good reliability in higher traffic rates. The reason is found in the sleep schedule of the IEEE 802.15.4 case, which is independent on traffic rate and wireless channel conditions. The result is that the fixed sleep schedule is not feasible to support high traffic and time-varying wireless channels. Moreover, the fixed sleep schedule does not guarantee a uniform distribution of cumulative wake-up rate within certain time in a cluster, which means that there may be congestions. On the contrary, Breath presents an excellent behavior in any situation of traffic load and channel condition.

Delay

In Fig. 10, the sample average of the delay for packet delivery of Breath, IEEE 802.15.4 $\mathcal{L}$ and $\mathcal{H}$ are plotted as a function of the reliability requirement Ω and traffic rate λ in AWGN and Rayleigh fading environments, with the vertical bars indicating the standard deviation of the samples around the average. The sample variance of the delay exhibits similar behavior as the average. The delay meets quite well the constrains. Observe that delay decreases as the traffic rate rises. This is due to that Breath increases linearly the wake-up rate of nodes when the traffic rate increases (see Eq. (5)). The delay is larger for worse reliability requirements. Note that Eq. (5) increases as the reliability requirement Ω increases. IEEE 802.15.4 $\mathcal{L}$ has lower delay than IEEE 802.15.4 $\mathcal{H}$ because nodes have higher

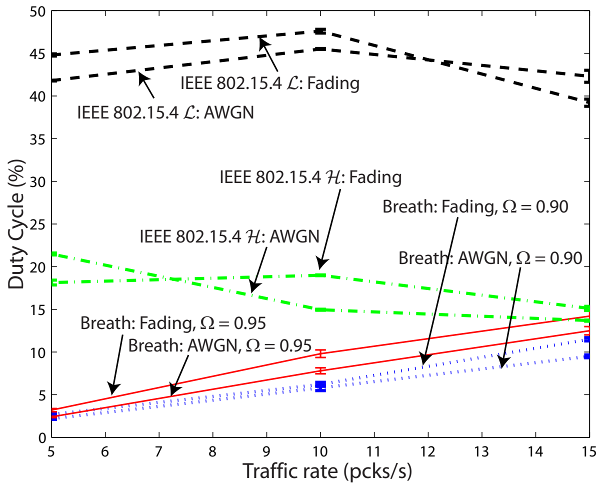


Figure 11: Sample average of the node's duty cycle in IEEE 802.15.4 $\mathcal{L}$, $\mathcal{H}$ and Breath with reliability requirement $\Omega = 0.9, 0.95$ for traffic rates $\lambda = 5, 10, 15$ pckt/s in AWGN and Rayleigh fading environments.

wake-up time. Breath has an intermediate behavior with respect to IEEE 802.15.4 $\mathcal{L}$ and $\mathcal{H}$ after $\lambda = 7$. From these experimental results, we conclude that both Breath and IEEE 802.15.4 meet the delay requirement. However, notice that the delay for IEEE 802.15.4 is related to only packet successfully received, which may be quite few.

Duty Cycle

In this section we study the energy consumption of the nodes.

As energy performance indicator, we measured the node's duty cycle, which is the ratio of the active time of the node to the total experimental time. Obviously, the lower is the duty cycle, the better is the performance of the protocol on energy consumption.

Fig. 11 shows the sample average of duty cycle of Breath, IEEE 802.15.4 $\mathcal{L}$ and $\mathcal{H}$ with respect to the traffic rates $\lambda = 5$, 10, 15 pckt/s and $\Omega = 0.9, 0.95$, both in AWGN and Rayleigh fading environments, with the vertical bars indicating the standard deviation of the samples. Note that IEEE 802.15.4 $\mathcal{L}$ and $\mathcal{H}$ do not exhibit a clear relationship with respect to traffic rate and have almost flat duty cycle around 42% and 18%, respectively, because of fixed sleep time. Considering Breath, observe that the duty cycle increases linearly with the traffic rate and reliability requirement. As for the delay, this is explained by Eq. (5). Since Breath minimizes the total energy consumption on the base of a tradeoff between wake-

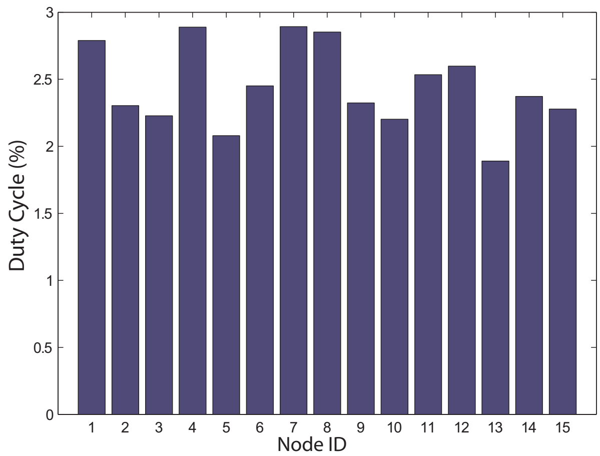


Figure 12: Distribution of the duty cycle in each node with $N = 15$ relays. The reliability requirement is $\Omega = 0.95$ and traffic rate is $\lambda = 5$ pckt/s.

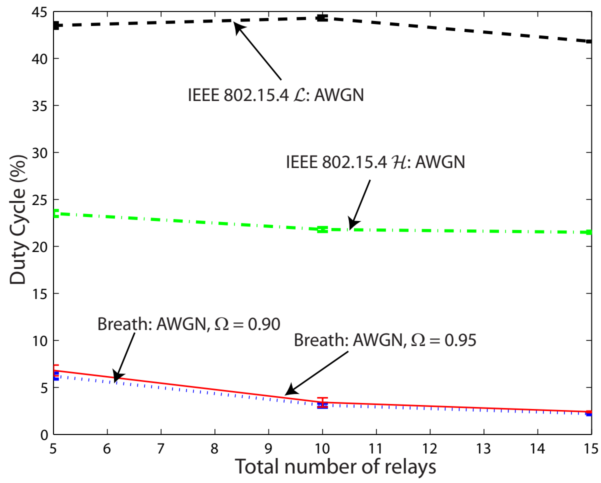


Figure 13: Sample average of the duty cycle in IEEE 802.15.4 $\mathcal{L}$, $\mathcal{H}$ and Breath with reliability requirement $\Omega = 0.9, 0.95$ and traffic rates $\lambda = 5, 10, 15$ pckt/s in AWGN environment for different networks, each with a different number of relaying nodes.

up rate and waiting time of beacon messages (recall the analysis in Section 5.3), lower wake-up rates do not guarantee lower duty cycle. Observe that choosing a lower active time for the nodes of the IEEE 802.15.4 implementation would obviously obtain energy savings comparable with Breath, however, the reliability of the IEEE 802.15.4 implementation would be heavily affected (recall Fig. 9). In other words, ensuring a duty cycle for the IEEE 802.15.4 implementation comparable with Breath would be very detrimental with respect to the reliability.

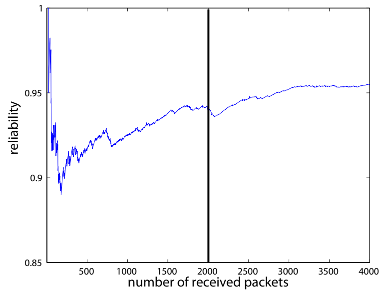
Fig. 12 shows the experimental results for the duty cycle of each relay node for $\lambda = 5$ pckt/s and $\Omega = 0.95$. A fair uniform distribution of the duty cycles among all nodes of the network is achieved. This is an important result, because the small variance of the wake-up rate among nodes signifies that duty cycle and load are uniformly distributed, with obvious advantages for the network lifetime.

Fig. 13 reports the case of several networks, where each network corresponds to a number of relays between the source and the sink in an AWGN environment. From the figure it is possible to evaluate how much Breath extends the network lifetime compared to IEEE 802.15.4 $\mathcal{L}$ and $\mathcal{H}$. Observe that the duty cycle is proportional to the density of nodes. Hence, the network lifetime is extended fairly by adding more nodes without creating load balancing problems.

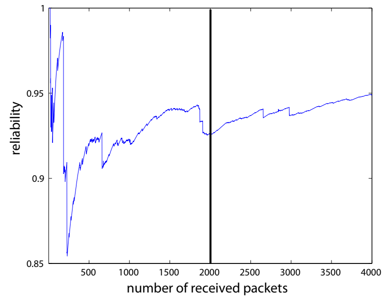
Finally, recall that Breath uses a radio power control (Subsection 5.3), so that further energy savings are actually obtained with respect to the IEEE 802.15.4 implementation.

9.2 Protocol Behavior for Time-Varying Requirements

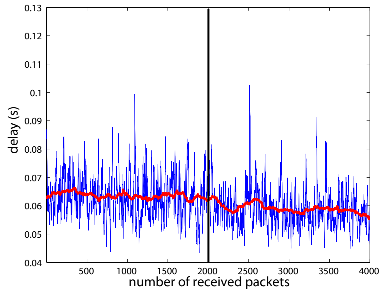
Performance of Breath protocol is based on the application requirements and estimation of the channel condition. In this subsection, we investigate the dynamic adaptation of Breath when the reliability Ω and delay τ requirements change. Figs. 14 show the dynamic adaptability of reliability, packet delay and energy consumption when the requirements are changed for given traffic rate and number of nodes. Figs. 14(a), 14(c), 14(e) and 14(b), 14(d), 14(f) present the behavior of the reliability, packet delay and average active time when the reliability and delay requirements change, respectively. More specifically, the average active time is defined as the average time nodes are active. We observe performance in terms of reliability, delay, and average active time in Figs. 14(a), 14(c), 14(e) for a reliability requirement variation. When Ω increases from 0.9 to 0.95 at a time corresponding to the number of received packet 2000 the reliability converges to 0.95. At the same time, packet delay decreases and average active time increase since optimal wake-up rate increases to guarantee the higher reliability requirement. Analogously, Breath adapts the network by considering the delay requirement variation in Figs. 14(b), 14(d), and 14(f). It is clear that the average delay is under 60 ms since we consider the distribution of the delay probability. The average active time increases when the delay requirement changes due to a higher optimal wake-up rate. Hence, in the optimization problem, the delay requirement 60 ms gives a stricter constraint than reliability constraint computed at a requirement of



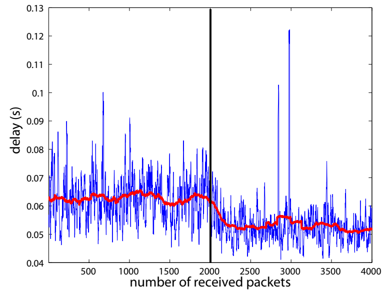
(a) Reliability behavior when Ω changes from 0.9 to 0.95



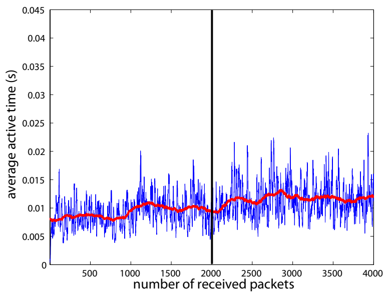
(b) Reliability behavior when τ changes from 1 s to 60 ms



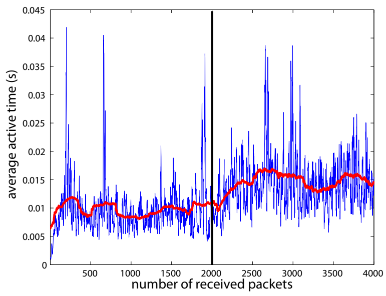
(c) Delay behavior when Ω changes from 0.9 to 0.95



(d) Delay behavior when τ changes from 1 s to 60 ms



(e) Active time behavior when Ω changes from 0.9 to 0.95



(f) Active time behavior when τ changes from 1 s to 60 ms

Figure 14: Reliability, packet delay and active time behavior for $\lambda = 10$ pckt/s, $N = 15$ when reliability and delay requirements vary from $\Omega = 0.9$ to $\Omega = 0.95$ and from $\tau = 1$ s to $\tau = 60$ ms at a time instant corresponding to the number of received packets 2000 in AWGN environment.

0.9. From this analysis, we can conclude that Breath adaptively achieves its target (i.e., minimization of power consumption) while guaranteeing the reliability and delay requirements. Furthermore, we observe clearly the tradeoff between the application requirements and energy consumption, i.e., as application requirements become strict, energy consumption increases.

10 Conclusions

We designed and implemented Breath, a protocol that consider a system-level approach to guarantee explicitly reliability and delay requirements in wireless sensor networks for control and actuation application. The protocol considers duty-cycle, routing, MAC, and physical layers all together to maximize the network lifetime by taking into account the tradeoff between energy consumption and application requirements for control applications.

We developed an analytical expression of the total energy consumption of the network, as well as reliability and delay for the packet delivery. These relations allowed us to pose a mixed real-integer constrained optimization problem to optimize the number of hops in the multi-hop routing, the wake-up rates of the nodes, and the transmit radio power as a function of the routing, MAC, physical layer, traffic, and hardware platform. An algorithm for the dynamic and continuous adaptation of the network operations to the traffic and channel conditions, and application requirements, was proposed.

We provided a complete test-bed implementation of the protocol, building a wireless sensor network with TinyOS and Tmote Sky sensors. An experimental campaign was conducted to test the validity of Breath in an indoor environment with both AWGN and Rayleigh fading channels. Experimental results showed that the protocol achieves the reliability and delay requirements, while minimizing the energy consumption. It outperformed a standard IEEE 802.15.4 implementation in terms of both energy efficiency and reliability. In addition, Breath showed good load balancing performance, and is scalable with the number of nodes. Given its good performance, Breath is a good candidate for many control and industrial applications, since these applications ask for both reliability and delay requirements in the packet delivery. A practical application of the protocol was reported in [10].

We are currently investigating the extension of the design methodology to consider mesh networks (ad-hoc and wireless sensor networks) and test it for other control applications.

References

[1] A. Willig, K. Matheus, and A. Wolisz, “Wireless technology in industrial networks,” *Proceedings of the IEEE*, 2005.

[2] A. Willig, "Recent and emerging topics in wireless industrial communication," *IEEE Transactions on Industrial Informatics*, vol. 4, no. 2, pp. 102–124, 2008.

[3] *IEEE Std 802.15.4-2996, September, Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)*, IEEE, 2006. [Online]. Available: http://www.ieee802.org/15

[4] A. Bonivento, L. Carloni, and A. Sangiovanni-Vincentelli, "Platform-based design for wireless sensor networks," *Mobile Netw. Appl.*, 2006.

[5] A. Sangiovanni-Vincentelli, "Quo vadis, sld? reasoning about the trends and challenges of system level design," *Proceedings of the IEEE*, vol. 95, no. 3, 2007.

[6] J. R. Moyne and D. M. Tilbury, "The emergence of industrial control networks for manufacturing control, diagnostics, and safety data," *Proceedings of the IEEE*, 2007.

[7] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, 2007.

[8] S. M. P. W. Zhang, M. S. Braniky, "Stability of networked control systems," *IEEE Control Systems Magazine*, 2001.

[9] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, "Foundations of control and estimation over lossy networks," *Proceedings of the IEEE*, 2007.

[10] E. Witrant, P. Park, M. Johansson, C. Fischione, and K. H. Johansson, "Predictive control over wireless multi-hop networks," in *IEEE MSC*, 2007.

[11] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: A scalable and robust communication paradigm for sensor networks," in *ACM MobiCOM*, 2000.

[12] *Tmote Sky Data Sheet*, Moteiv, San Francisco, CA, 2006. [Online]. Available: http://www.moteiv.com/products/docs/tmote-sky-datasheet.pdf

[13] P. Buonadonna, D. Gay, J. Hellerstein, W. Hong, and S. Madden, "TASK: Sensor network in a box," Intel Research Lab Report, Tech. Rep., 2007.

[14] R. Steigman, and J. Endresen, "Introduction to WISA and WPS, WISA-wireless interface for sensors and actuators and WPS-wireless proximity switches," *White paper*, 2004. [Online]. Available: http://www.eit.uni-kl.de/litz/WISA.pdf

[15] *The ZigBee Alliance* . [Online]. Available: http://www.zigbee.org

[16] *Wirelesshart data sheet*, IEEE, 2007. [Online]. Available: http://www.hartcomm2.org/hartprotocol/wirelesshart/wirelesshartmain.html

[17] Y. Xu, J. Heidemann, and D. Estrin, "Geography-informed energy conservation for ad hoc routing," in *ACM MobiCom*, vol. pp. 70-84, 2001.

[18] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris, “Span: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks,” in *ACM MobiCom*, 2001.

[19] M. Buettner, G. Yee, E. Anderson, and R. Han, “X-MAC: A short preamble MAC protocol for duty-cycled wireless sensor networks,” in *ACM SenSys*, 2006.

[20] F. Stann and J. Heidemann, “RMST: Reliable Data Transport in Sensor Networks,” in *IEEE SNPA*, 2003.

[21] S. Kim, R. Fonseca, P. Dutta, A. Tavakoli, D. Culler, P. Levis, S. Shenker, and I. Stoica, “Flush: a reliable bulk transport protocol for multihop wireless networks,” in *ACM SenSys*, 2007.

[22] O. B. Akan and F. Akyildiz, “Event-to-sink reliable transport in wireless sensor networks,” *IEEE Transactions on Networking*, vol. 13, no. 5, pp. 1003–1016, 2005.

[23] G. Werner-Allen, K. Lorincz, J. Johnson, J. Lees, and M. Welsh, “Fidelity and yield in a volcano monitoring sensor network,” in *USENIX OSDI*, 2006.

[24] N. Burri, P. von Rickenbach, and R. Wattenhofer, “Dozer: ultra-low power data gathering in sensor networks,” in *ACM/IEEE IPSN*, 2007.

[25] M. Zorzi and R. R. Rao, “Energy and latency performance of geographic random forwarding for ad hoc and sensor networks,” in *IEEE WCNC*, 2003.

[26] E. Felemban, C. G. Lee, and E. Eylem, “MMSPEED: Multipath multi-speed protocol for QoS guarantee of reliability and timeliness in wireless sensor networks,” *IEEE Transactions on Mobile Computing*, vol. 5, no. 6, pp. 738–754, 2006.

[27] A. Bonivento, C. Fischione, L. Necchi, F. Pianegiani, A. Sangiovanni-Vincentelli, “System level design for clustered wireless sensor networks,” *IEEE Transactions on Industrial Informatics*, 2007.

[28] J. Misic, S. Shafi, and V. Misic, “Cross-layer activity management in an 802.15.4 sensor network,” *IEEE Communications Magazine*, vol. 44, no. 1, pp. 131–136, 2006.

[29] L. van Hoesel, T. Nieberg, J. Wu, and P. Havinga, “Prolonging the lifetime of wireless sensor networks by cross-layer interaction,” *IEEE Wireless Communications*, vol. 11, no. 6, pp. 78–86, 2004.

[30] J. Van Greuen, D. Petrovic, A. Bonivento, J. Rabaey, K. Ramchandran, and A. Sangiovanni-Vincentelli, “Adaptive sleep discipline for energy conservation and robustness in dense sensor networks,” in *IEEE ICC*, 2004.

[31] R. Cristescu, B. Beferull-Lozano, M. Vetterli, and R. Wattenhofer, “Network correlated data gathering with explicit communication: NP-completeness and algorithms,” *ACM/IEEE Transactions on Networking*, vol. 14, no. 1, 2006.

[32] P. Chen and S. Sastry, "Latency and connectivity analysis tools for wireless mesh networks," in *ACM/IEEE ROBOCOMM*, 2007.

[33] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660–670, 2002.

[34] Y. Ma and J. H. Aylor, "System lifetime optimization for heterogeneous sensor networks with a hub-spoke topology," *IEEE Transactions on Mobile Computing*, vol. 3, no. 3, pp. 286–294, 2004.

[35] O. Younis and S. Fahmy, "HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 366–379, 2004.

[36] W. Ye, J. Heidemann and D. Estrin, "Medium access control with coordinated adaptive sleeping for wireless sensor networks," *ACM/IEEE Transactions on Networking*, 2004.

[37] *CC2420 Data Sheet*, Chipcon, Oslo, Norway, 2005. [Online]. Available: http://www.chipcon.com/files/CC2420-Data-Sheet-1-3.pdf

[38] *CC2431 Data Sheet*, Texas Instruments. [Online]. Available: http://focus.tij.co.jp/jp/lit/ds/symlink/cc2431.pdf

[39] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal of Selected Areas in Telecommunications*, vol. 18, no. 3, pp. 535–547, 2000.

[40] S. Pollin, M. Ergen, S. C. Ergen, B. Bougard, L. V. D. Perre, F. Catthoor, I. Moerman, A. Bahai, and P. Varaiya, "Performance analysis of slotted carrier sense IEEE 802.15.4 medium access layer," in *IEEE GLOBECOM*, 2006.

[41] Y. Xiao and J. Rosdahl, "Throughput and delay limits of IEEE 802.11," *IEEE Communication Letters*, 2002.

[42] H. Zhai, Y. Kwon, and Y. Fang, "Performance analysis of IEEE 802.11 mac protocols in wireless lans," *Wireless Communications and Mobile Computing*, 2004.

[43] S. Shafi, J. Misic, and V. B. Misic, "Performance of a beacon enabled IEEE 802.15.4 cluster with downlink and uplink traffic," *IEEE Transactions on Vehicular Technology*, 2006.

[44] A. Lymberopoulos, D. Savvides, "XYZ: a motion-enabled, power aware sensor node platform for distributed sensor network applications," in *ACM/IEEE IPSN*, 2005.

[45] P. Park, C. Fischione, A. Bonivento, K. H. Johansson, and A. Sangiovanni-Vincentelli, "Breath: a self-adapting protocol for wireless sensor networks in control and automation," in *IEEE SECON*, 2008.

Symbol	**Meaning**
$D(h, \mu_c)$	Distribution of delay
E_{ca}	Average energy consumption in CSMA/CA
E_{tot}	Total energy consumption of the network
E_{pck}	Data packet transmission Energy consumption
E_r	Energy consumption for receiving a data packet
E_{wu}	Energy consumption for wake-up and beconing
h	Number of hops from source to sink
M_{ca}	Maximum number of CSMA/CA TX tries
N	Number of relays of the network
p_{min}	Minimum successful packet reception probability
P_t	Radio transmission power
Q_b	Energy consumption for a beacon transmission
Q_m	Energy consumption for a packet transmission
$R(h, \mu_c)$	Probability of successful packet delivery
S	Distance from the source to the sink
T	Total time
T_{ac}	Maximum listening time to receive a data packet
T_{ca}	A unit of backoff period
T_W	Wake-up time from sleep mode
W	Maximum number of random backoff time
A_{rx}	Power consumption at RX mode
A_s	Power consumption to scan a channel
A_{tx}	Power consumption at TX mode
A_w	Power consumption to wake-up
α_i	Exponentially distributed time of intensity μ_c
ε_i	Uniformly distributed backoff time of CSMA/CA
ζ	Percentage of slots in the backoff time counter 0
λ	Traffic rate
μ_c	Cumulative cluster wake-up rate
μ_k	Wake-up rate of node k
ν	Busy channel probability in CSMA/CA
τ	Required delay
$\psi_{sc}(n)$	Probability of successful TX in CSMA/CA
$\psi_{sb}(\mu_c, n)$	Competition probability of n nodes in $1/\lambda$ s
Δ	Required delay probability
Ω	Required reliability

Table 2: Main symbols used in the paper.