A

Modelling and Stability Analysis of Hybrid Multiple Access in the IEEE 802.15.4 Protocol

Pangun Park, KTH, Royal Institute of Technology
Carlo Fischione, KTH, Royal Institute of Technology
Karl Henrik Johansson, KTH, Royal Institute of Technology

To offer flexible quality of service to several classes of applications, the medium access control (MAC) protocol of IEEE 802.15.4 wireless sensor networks (WSNs) combines the advantages of a random access with contention with a time division multiple access (TDMA) without contention. Understanding reliability, delay, and throughput is essential to characterize the fundamental limitations of the MAC and to optimize its parameters. Nevertheless, there is not yet a clear investigation of the achievable performance of hybrid MAC. In this paper, an analytical framework for modelling the behavior of the hybrid MAC protocol of the IEEE 802.15.4 standard is proposed. The main challenge for an accurate analysis is the coexistence of the stochastic behavior of the random access and the deterministic behavior of the TDMA scheme. The analysis is done in three steps. First, the contention access scheme of the IEEE 802.15.4 exponential backoff process is modeled through an extended Markov chain that takes into account channel, retry limits, acknowledgements, unsaturated traffic, and superframe period. Second, the behavior of the TDMA access scheme is modeled by another Markov chain. Finally, the two chains are coupled to obtain a complete model of the hybrid MAC. By using this model, the network performance in terms of reliability, average packet delay, average queuing delay, and throughput is evaluated through both theoretical analysis and experiments. The protocol has been implemented and evaluated on a test-bed with off-the-shelf wireless sensor devices to demonstrate the utility of the analysis in a practical set-up. It is established that the probability density function of the number of received packets per superframe follows a Poisson distribution. It is determined under which conditions the guaranteed time slot allocation mechanism of IEEE 802.15.4 is stable. It is shown that the mutual effect between throughput of the random access and the TDMA scheme for a fixed superframe length is critical to maximize the overall throughput of the hybrid MAC. In high traffic load, the throughput of the random access mechanism dominates over TDMA due to the constrained use of TDMA in the standard. Furthermore, it is shown that the effect of imperfect channel and carrier sensing on system performance heavily depends on the traffic load and limited range of the protocol parameters. Finally, it is argued that the traffic generation model established in this paper may be used to design an activation timer mechanism in a modified version of the CSMA/CA algorithm that guarantees a stable network performance.

Categories and Subject Descriptors: C.2.2 [**Computer-Communication Networks**]: Network Protocols

General Terms: Performance, Standardization, Theory, Verification

Additional Key Words and Phrases: IEEE 802.15.4, Hybrid MAC, Markov chain model

ACM Reference Format:
Park, P., Fischione, C., Johansson, K. H. 2011. Modelling and Stability Analysis of Hybrid Multiple Access in IEEE 802.15.4 Protocol. ACM Transactions on Sensor Networks V, N, Article A (January YYYY), 50 pages.
DOI = 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

1. INTRODUCTION

The IEEE 802.15.4 communication standard is arguably becoming the most popular protocol for low data rate and low power wireless sensor networks (WSNs) in many application domains, such

This work was supported by the EU project FeedNetBack and HYCON2, the Swedish Research Council, the Swedish Governmental Agency for Innovation Systems and the Knut and Alice Wallenberg Foundation. A preliminary version of this work has appeared as part of the PhD thesis [Park 2011]. The authors are with the ACCESS Linnaeus Center, Electrical Engineering, Royal Institute of Technology, Stockholm, Sweden. E-mails: {pgpark,carlofi,kallej}@ee.kth.se.

DOI 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

as industrial control, home automation, health care, and smart grids [Willig et al. 2005; IEEE 2006; Wheeler 2007]. One of the important aspects of IEEE 802.15.4 is the combination of the contention access and the time division multiple access (TDMA) medium access control (MAC) [IEEE 2006; 2010]. It enables the protocol to serve a variety of applications, such as remote control and automation.

It is instrumental to have a mathematical model of the IEEE 802.15.4 MAC for networked control systems. A major challenge in developing networked control systems using WSNs is the inconsistency in sensor measurements due to packet loss, delay, and false detections [Hespanha et al. 2007]. A key requirement in designing distributed control protocols for multiple sensor systems is the ability to verify that these protocols correctly implement the control specifications, and keep the system within given safety, stability, and performance limits. Control systems are particularly challenging because they must support the right decision even in the presence of unexpected network congestion, network failures, or external manipulations of the environment. Therefore, analysing the performance of communication protocols is essential to understand the fundamental limitations of control systems using WSNs. For example, a critical tool for safety verification of such control system is to provide a predictable and reliable mathematical description of the behavior of the underlying technology by modeling the interactions among the layers using simple expressions with adequate accuracy, which are then used for the design of the overall wireless networked control systems.

The idea of a hybrid MAC is not new. The IEEE 802.15.4 MAC has been inspired by the adaptive MAC protocol proposed in the late 70's by [Kleinrock and Yemini 1978] to maximize the throughput, which combined slotted ALOHA and TDMA. Many allocation schemes were designed to combine the advantages of both the ALOHA and the TDMA approaches. One of the extensions is to use a so called reservation scheme with contention, where users contend during a reservation period, and those who succeed in this contention transmit without experiencing interference. Such a scheme derives its efficiency from that reservation periods are several orders of magnitude shorter than transmission periods. The works proposed in [Crowther et al. 1973; Roberts 1973; Binder 1975] fall in this category of reservation schemes. Additional reservation protocols and their analysis can be found in [Tobagi and Kleinrock 1976; Tasaka and Ishibashi 1984; Tsai and Chang 1986]. For example, the demand assignment multiple access is successfully used for satellite and military communications [Fine and Tobagi 1984] and for the IEEE 802.15.3c standard. The throughput of IEEE 802.15.3c is studied in [Pyo and Harada 2009]. The essential difference between these reservation schemes and the IEEE 802.15.4 standard is that the contention access period supports not only the reservation scheme but also the data packet communication. Therefore, the IEEE 802.15.4 hybrid MAC is a more general protocol than the simpler hybrid MAC based on reservation schemes. In this paper, we give an analysis of the hybrid MAC of IEEE 802.15.4. We remark that our target is to model the hybrid MAC of IEEE 802.15.4 without proposing any modifications. Rather, we aim at understanding how to improve the network performance within the standard. To the best of our knowledge, this is the first paper simultaneously considering the random access and the TDMA of IEEE 802.15.4. We compare our analytical model to the results obtained by an experimental implementation, and thereby present the first comparison of an analytical model of the IEEE 802.15.4 protocol to real experiments. The contribution of the implementation and the experiments is to demonstrate the feasibility of the analytical model to understand the fundamental limitation of the protocol in a practical system.

The remainder of this paper is organized as follows. In Section 2, we give a brief overview of the IEEE 802.15.4 MAC, where we summarize the characteristic of the contention access period (CAP) and the TDMA period, (or contention free period (CFP)). In Section 3, we summarize previous work for the slotted carrier sense multiple access/collision avoidance (CSMA/CA) mechanism of the CAP and the guaranteed time slot (GTS) allocation mechanism of the CFP for the IEEE 802.15.4 MAC. Then, we follow three steps to analyze the performance of the hybrid MAC. First, we propose an extended Markov chain model of CSMA/CA with channel, retry limits, acknowledgements (ACKs), unsaturated traffic and superframe period of the CAP in Section 5. Then, we

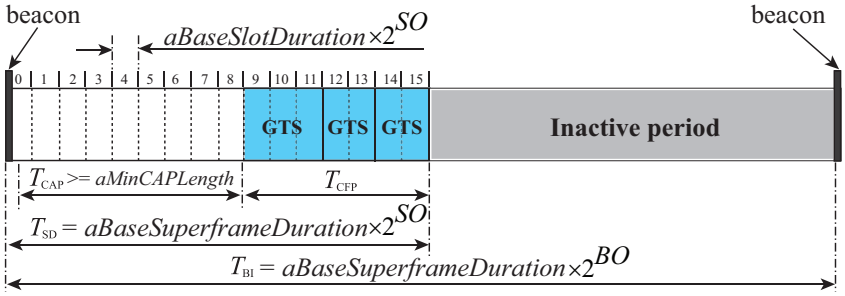


Fig. 1. Superframe structure of IEEE 802.15.4 MAC.

analyze the performance of the GTS allocation of the CFP based on a Markov chain model in Section 6. Third, we present a model for the hybrid MAC by connecting these two chains in Section 7. In Section 8, we first validate our analysis by experimental results, and show the performance of the hybrid MAC in terms of reliability, packet delay, queuing delay, and throughput. Then, we discuss the effect of imperfect channel and carrier sensing capability by using the simulation results. We discuss the motivation and possible applications of the model in Section 9. In Section 10, we conclude the paper.

2. OVERVIEW OF THE IEEE 802.15.4

In this section we give an overview of the key points of IEEE 802.15.4 MAC. The standard specifies the physical layer and the MAC sublayer for low-rate wireless networks. The standard defines two channel access modalities: the beacon-enabled modality, which uses a slotted CSMA/CA and the optional GTS allocation mechanism, and a simpler unslotted CSMA/CA without beacons. The communication is organized in temporal windows denoted superframes. Because we focus on the beacon-enabled modality featuring the hybrid MAC, we consider a star network with a coordinator. Many control applications rely on a star topology, or a topology close to a star, rather than on a general multi-hop network. In fact, multi-hop networks pose critical challenges fro real-time monitoring and control applications [ISA 2009; IEEE 2010; Willig 2008]. Fig. 1 shows the superframe structure of the beacon-enabled mode. In the following, we give the details of this figure.

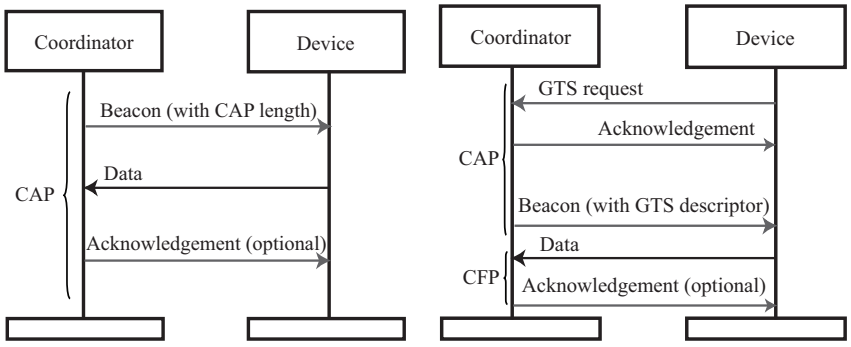
The network coordinator periodically sends beacon frames in every beacon interval T_{BI} to identify its personal area network and to synchronize devices that communicate with it. The main symbols used in the paper is provided in Table II. The coordinator and devices can communicate during the active period, called the superframe duration T_{SD}, and enter the low-power mode during the inactive period. The structure of the superframe is defined by two parameters, the beacon order (BO) and the superframe order (SO), which determine the length of the superframe and its active period, respectively, they are

$$T_{\text{BI}} = aBaseSuperframeDuration \times 2^{\text{BO}}, \tag{1}$$

$$T_{\text{SD}} = aBaseSuperframeDuration \times 2^{\text{SO}}, \tag{2}$$

where $0 \leq \text{SO} \leq \text{BO} \leq 14$ and *aBaseSuperframeDuration*[1] is the number of symbols forming a superframe when SO is equal to 0. In addition, the superframe is divided into 16 equally sized superframe slots of length *aBaseSlotDuration*. Each active period can be further divided into a CAP and an optional CFP, composed of GTSs. A slotted CSMA/CA mechanism is used to access the channel of non time-critical data frames and GTS requests during the CAP. In the CFP, the dedicated bandwidth is used for time-critical data frames. Figure 2 illustrates the date transfer mechanism of

[1]The italic letter is used for the parameters defined in the IEEE 802.15.4 standard in the manuscript.



(a) Non time-critical data packet or GTS request transmission.

(b) Time-critical data packet transmission.

Fig. 2. Data transfers of beacon-enabled mode during the CAP and CFP.

the beacon-enabled mode for the CAP and CFP. In the following section, we describe the data transmission mechanism for both CAP and CFP.

2.1. CSMA/CA algorithm of CAP

Consider a device trying to transmit a data packet during CAP. In slotted CSMA/CA of IEEE 802.15.4, first the MAC sublayer initializes four variables, i.e., the number of backoffs (NB=0), contention window (CW=2), backoff exponent (BE=*macMinBE*) and retransmission times (RT=0). Then the MAC sublayer delays for a random number of complete backoff periods in the range $[0, 2^{\mathrm{BE}} - 1]$ units. If the number of backoff periods is greater than the remaining number of backoff periods in the CAP, the MAC sublayer pauses the backoff countdown at the end of the CAP and resumes it at the start of the CAP in the next superframe. Otherwise the MAC sublayer counts its backoff delay. When the backoff period is zero, the device needs to perform the first clear channel assessment (CCA). The MAC sublayer proceeds if the remaining CSMA/CA algorithm steps (i.e., two CCAs), the frame transmission, and any ACK can be completed before the end of the CAP. If the MAC sublayer cannot proceed, it waits until the start of the CAP in the next superframe and apply a further random backoff delay in the range $[0, 2^{\mathrm{BE}} - 1]$ units before evaluating whether it can proceed again. Otherwise the MAC sublayer proceeds the CCA in the current superframe. If two consecutive CCAs are idle, then the device commences the packet transmission. If either of the CCA fails due to busy channel, the MAC sublayer increases the value of both NB and BE by one, up to a maximum value *macMaxCSMABackoffs* and *macMaxBE*, respectively. Hence, the values of NB and BE depend on the number of CCA failures of a packet. Once BE reaches *macMaxBE*, it remains at the value *macMaxBE* until it is reset. If NB exceeds *macMaxCSMABackoffs*, then the packet is discarded due to channel access failure. Otherwise, the CSMA/CA algorithm generates a random number of complete backoff periods and repeats the process. Here, the variable *macMaxCSMABackoffs* represents the maximum number of times the CSMA/CA algorithm is required to backoff. If channel access is successful, the device starts transmitting packets and waits for an ACK. The reception of the corresponding ACK is interpreted as successful packet transmission. If the device fails to receive ACK due to collision, bad channel or ACK timeout, the variable RT is increased by one up to *macMaxFrameRetries*. If RT is less than *macMaxFrameRetries*, the MAC sublayer initializes two variables CW=0, BE=*macMinBE* and follows the CSMA/CA mechanism to re-access the channel. Otherwise the packet is discarded due to the retry limit. Note that the default MAC parameters are $\mathit{macMinBE} = 3, \mathit{macMaxBE} = 5, \mathit{macMaxCSMABackoffs} = 4, \mathit{macMaxFrameRetries} = 3$. See [IEEE 2006] for further details.

Table I. Comparison between IEEE 802.15.4 MAC model features, analyzed performance indicators, and simulators or practical implementation tools from the literature and this paper. The letters $\mathbf{T}$, $\mathbf{R}$, $\mathbf{D}$, and $\mathbf{E}$ denote throughput, reliability, delay, and energy consumption. "Retry" in the feature column means the retransmission of the CSMA/CA algorithm. The circle ◯ denotes that the corresponding features or performance indicators are considered analytically. A white space denotes that the analytical model does not include the corresponding features, performance indicators or validation. The triangle △ denotes that the analytical model requires a significant modification of the IEEE 802.15.4 standard to model the corresponding features or to derive the corresponding performance indicators.

Analytical model	**Features**					**Indicators**				**Tool**
	CSMA/CA	ACK	Retry	GTS	Inactive	$\mathbf{T}$	$\mathbf{R}$	$\mathbf{D}$	$\mathbf{E}$	
[Misic and Misic 2005]	◯	◯	◯		◯	◯		◯		
[Misic et al. 2006]	◯	◯	◯		◯	◯	◯			Artifex [Boggio et al. 2001]
[Koubaa et al. 2006]				◯	◯			◯	◯	
[Ramachandran et al. 2007]	◯				◯	◯			◯	ns-2
[Pollin et al. 2008]	◯	◯				◯	◯		◯	Matlab
[Sahoo and Sheu 2008]	◯	◯	◯			◯	◯		◯	ns-2
[Jung et al. 2009]	◯	◯	◯		◯	◯				Monte Carlo
[Martalo et al. 2009]	◯				◯	◯		◯		ns-2
[Faridi et al. 2010]	◯	◯	◯			◯	◯	◯	◯	Matlab
[Buratti 2010]	◯			△	◯	◯	◯	△		Monte Carlo
This paper	◯	◯	◯	◯	◯	◯	◯	◯		TelosB, Contiki OS, COOJA

2.2. GTS allocation of CFP

The coordinator is responsible for the GTS allocation and determines the length of the CFP in a superframe. To request the allocation of a new GTS, the device sends the GTS request command to the coordinator. The coordinator confirms its receipt by sending an ACK frame within CAP. Upon receiving a GTS allocation request, the coordinator checks whether there are sufficient resources and, if possible, allocates the requested GTS. We recall that Figure 2(b) illustrates the GTS allocation mechanism. The GTS capacity in a superframe satisfies the following requirements:

(1) The maximum number of GTSs to be allocated to devices is seven, provided there is sufficient capacity in the superframe.
(2) The minimum length of a CAP $T_{\min}$ is *aMinCAPLength* .

Therefore the CFP length depends on the GTS requests and the current available capacity in the superframe. If there is sufficient bandwidth in the next superframe, the coordinator determines a device list for GTS allocation based on a first-come-first-served (FCFS) policy. Then, the coordinator includes the GTS descriptor which is the device list that obtains GTSs in the following beacon to announce the allocation information. The coordinator makes this decision within *aGTSDescPersistenceTime* superframes. Note that on receipt of the ACK to the GTS request command, the device continues to track beacons and waits for at most *aGTSDescPersistenceTime* superframes. A device uses the dedicated bandwidth to transmit the packet within the CFP. In addition, a transmitting device ensures that its transaction is complete one IFS period before the end of its GTS.

3. RELATED WORK

In this section, we present related existing literature and we argue that it does not adequately investigate the performance of the hybrid MAC of IEEE 802.15.4. In particular, we first discuss the literature concerning the analysis of the CSMA/CA algorithm of the CAP, and then we consider previous work about the GTS allocation mechanism of the CFP. The modelling of IEEE 802.15.4 CAP is related to the IEEE 802.11 MAC [IEEE 1999]. Both IEEE 802.11 and 802.15.4 are based on a MAC that uses a binary exponential backoff scheme. Therefore, we also consider the literature related to the IEEE 802.11 MAC. Table I summarizes the features modeled, the performance indicators analyzed, and the simulators or practical implementations used in the relevant works from the literature and this paper. In the table, we have evidenced the performance indications as throughput $\mathbf{T}$, reliability $\mathbf{R}$, delay $\mathbf{D}$, and energy $\mathbf{E}$ and the analytical model validation. We discuss these analytical models in the following.

3.1. Analytical model of CAP

The basic functionalities of the IEEE 802.11 MAC has been modeled by Bianchi with a Markov chain under saturated traffic and ideal channel conditions [Bianchi 2000]. Extensions of this model have been used to analyze the MAC layer service time [Tickioo and Sikdar 2004] and throughput [Wu et al. 2002] of IEEE 802.11.

The analysis of the packet delay, throughput, and power consumption of the IEEE 802.15.4 MAC has been the focus of several simulation-based studies, e.g., [Zheng and Lee 2004], and some recent theoretical works, e.g., [Misic and Misic 2005; Misic et al. 2006; Pollin et al. 2008; Sahoo and Sheu 2008; Jung et al. 2009]. Inspired by Bianchi's work, a Markov model for IEEE 802.15.4 and an extension with ACK mechanism have been proposed in [Pollin et al. 2008] without retransmissions. In [Pollin et al. 2008], the authors provide an analytical evaluation of the performance including throughput and energy consumption in a star topology network. The probabilities of channel access and busy channel derived from a Markov chain model validate the proposed model. A modified Markov model including retransmissions with finite retry limits has been studied in [Sahoo and Sheu 2008] as an attempt to model the CSMA/CA mechanism. However, simulations show that this analysis gives inaccurate results in terms of power consumption and throughput under unsaturated traffic with finite retry limits. In [Faridi et al. 2010], the authors extend the work [Pollin et al. 2008] by considering the retransmission mechanism and investigate the reliability, delay, throughput, and power consumption of the slotted CSMA/CA algorithm of IEEE 802.15.4. It is shown that the probability of carrier sensing depends on the number of nodes and backoff stage at the expense of a more computationally intensive approach.

The works [Pollin et al. 2008; Sahoo and Sheu 2008; Faridi et al. 2010] do not consider the active and inactive periods of the superframe structure of the IEEE 802.15.4 MAC. This aspect is investigated in [Misic and Misic 2005; Misic et al. 2006; Jung et al. 2009]. In [Misic and Misic 2005], an unsaturated traffic and devices with a finite buffer in uplink communications are assumed. In [Misic et al. 2006], instead, downlink communications are taken into account, but devices are assumed to have infinite size buffers. Note that the power consumption, reliability, and delay performance are not investigated therein. Furthermore, the analytical models proposed in [Misic and Misic 2005; Misic et al. 2006] are not validated with the simulation results. In [Jung et al. 2009], a throughput analysis has been performed by an extension of the Markov chain proposed in [Misic et al. 2006], where the superframe structure, ACK, and retransmissions are considered. In [Jung et al. 2009], it is also shown that the models proposed in [Misic and Misic 2005; Misic et al. 2006], although very detailed, fail to agree with simulation results. The Markov chain proposed in [Jung et al. 2009] does not model the length of data and ACK packets, which is crucial for analyzing the performance metrics for IEEE 802.15.4 networks with low data rate.

In [Ramachandran et al. 2007], the authors propose a different approach with respect to Bianchi's work by approximating the CAP as the simple nonpersistent CSMA, which is a similar approach to [Cali et al. 2000]. They also propose the initialization of the contention window length to one to improve throughput and energy consumption when ACKs are not used. However, the authors assume that the entire superframe duration is active, that is, $\mathrm{SO} = \mathrm{BO}$ without considering the inactive period, which is not realistic. In [Martalo et al. 2009], the authors extend the framework proposed in [Ramachandran et al. 2007] to two-hop network scenarios, where sensors communicate with the coordinator through an intermediate relay device, which simply forwards the data packets without generating traffic on its own. However, the assumption of no ACK and beacon-enabled mode for the cluster tree scenario is not realistic. We remark that the GTS allocation mechanism of CFP is not considered in [Misic and Misic 2005; Misic et al. 2006; Ramachandran et al. 2007; Pollin et al. 2008; Sahoo and Sheu 2008; Jung et al. 2009; Martalo et al. 2009; Faridi et al. 2010]. Furthermore, none of these studies shows a comparison with experimental results as obtained by a practical implementation. The simulation results of [Misic and Misic 2005; Misic et al. 2006; Ramachandran et al. 2007; Pollin et al. 2008; Sahoo and Sheu 2008; Jung et al. 2009; Martalo et al.

2009; Faridi et al. 2010] are only derived by Monte Carlo simulation using Matlab, ns-2, or similar simulations.

3.2. Analytical model of CFP

Most of the literature does not consider satisfactorily the CFP, where the GTS mechanism operates. An interesting theoretical performance evaluation of the GTS allocation has been proposed by Koubaa et al. [Koubaa et al. 2006] by using network calculus. This paper focuses on the impact of the IEEE 802.15.4 MAC parameters $(\mathrm{SO}, \mathrm{BO})$, the delay, throughput, and energy consumption of the GTS allocation. Network calculus assumes a continuous flow model, whereas communication happens through low data rate packets in reality. Hence, network calculus gives the worst-case traffic flows, which leads to severe under-utilization of time slots in actual implementations. In [Park et al. 2009], a Markov chain model is used to analyze the stability, queuing delay, and throughput of the GTS allocation mechanism. This model requires the exact probability density function (PDF) of the number of received GTS requests per superframe, which abstracts the CSMA/CA mechanism of the CAP. It is shown that different PDFs of the number of GTS requests are critical in terms of queuing delay and throughput even if PDFs have same mean.

Some interesting algorithms have been proposed to improve the performance of the GTS allocation mechanism. To maximize the bandwidth utilization, a smaller slot size mechanism is proposed in [Cheng et al. 2007]. In [Na et al. 2008], the delay constraint and bandwidth utilization are considered for the design of a GTS scheduling algorithm. In [Huang et al. 2008], authors propose an adaptive GTS allocation scheme by considering low delay and fairness. Despite these studies, there is no explicit consideration of theoretical study of both CAP and CFP.

In [Buratti 2010], a query-based approach to analyze the performance of CAP and CFP of IEEE 802.15.4 MAC is considered. Although this approach is interesting, this mechanism is not explicitly considered in the IEEE 802.15.4 standard, i.e., the GTS allocation mechanism in [IEEE 2006] is not query-based. In IEEE 802.15.4 MAC, devices are allowed to use both CSMA/CA mechanism of CAP and GTS allocation mechanism of CFP in a superframe, which is not the case in [Buratti 2010]. The ACK mechanism and retransmission mechanism are not considered either. Moreover, it is assumed that if a device does not succeed in accessing the channel by the end of the superframe, the packet is lost. However, in IEEE 802.15.4 the packet transmission is deferred if the remaining time of the superframe is not sufficient to transmit a packet [IEEE 2006].

Finally, we remark that despite the theoretical promises of the analytical models for analyzing both CAP and CFP we surveyed in Sections 3.1 and 3.2, to date these theoretical results have not been compared to experimental results.

In the following section, we present the system model and assumption that we adopt to analyze the performance of the IEEE 802.15.4 MAC.

4. SYSTEM MODEL

We consider a star network with a coordinator and N devices. Every device contends to send data packet[2] to the coordinator. The coordinator acts as a data sink and we assume it does not experience the hidden device problem. Throughout this paper we consider applications where devices asynchronously generate packets for transmission. We consider the underlying minimum time unit corresponding to *aUnitBackoffPeriod*, as defined in the IEEE 802.15.4 standard and we denote it T_b. In the standard, T_b corresponds to 20 symbols in the physical layer (i.e., $320\,\mu s$ for $2.45\,\mathrm{GHz}$). When a device just has sent a packet successfully or just discarded a packet, we assume that a new packet is generated with probability η_t. If a new packet is not generated, then the device tries to generate a new packet after $h\,T_b$ s, where h is a positive integer. This packet is generated with probability η_p. Note that we pause the packet generation timer at the end of CAP and continue to count it at the beginning of the superframe. We discuss the motivation and possible applications of the traffic model in Sections 8.1 and 9. We consider two different types of data packets: non time-critical data

[2]Throughout this paper, we use the term "packet" to denote a protocol data unit or frame at the MAC layer.

packets to be transmitted during the CAP, and time-critical data packets to be transmitted during the CFP using the GTS allocation mechanism as described in [IEEE 2006].

When a device decides to generate a data packet, it generates a non time-critical data packet with probability η_d and time-critical data packet with probability $1 - \eta_d$ in our model. A device uses a beacon-enabled slotted CSMA/CA algorithm to send a non time-critical data packet and a GTS request to the coordinator during the CAP. Note that the packet transmission is successful if an ACK packet is received. For a time-critical data packet, the device informs the need of GTS resources by sending the request during the CAP. The coordinator allocates a number of GTSs by considering the received GTS requests. Each device may need to send a multiple number of time-critical packets wherein each packet has a fixed length due to the maximum length of a packet defined in the standard. The requests are stored in a queue of the coordinator, and wait to be served in the next superframes, where the related GTS may be allocated. If too many requests arrive with respect to the coordinator queue size, then we have a queue overflow. We assume an ideal channel condition of physical layer and perfect channel channel sensing capability of devices. Furthermore, we make the natural assumption that each device forwards a non time-critical packet or a GTS request within $2T_{\text{BI}}$ i.e., the maximum packet delay of the CAP is $2T_{\text{BI}}$.

Based on the introduced model and assumptions, we propose an analytical model of the slotted CSMA/CA algorithm of the CAP and the GTS allocation of the CFP based on two Markov chain models in Sections 5 and 6. Then, in Section 7, we connect these two models that allow us to investigate the performance of the hybrid MAC of IEEE 802.15.4 in terms of the reliability of the CAP, the average delay of the CAP, the queuing delay of the CFP, and the throughput of the network.

5. PERFORMANCE ANALYSIS OF CAP

In this section, we analyze the performance of the CSMA/CA algorithm of the CAP. In Section 5.1, we propose the model of such an algorithm to send a non time-critical data packet and GTS requests. The core contribution of the section is the proposal of an extended Markov chain model. Then, in Section 5.2, we derive the reliability and average packet delay of CSMA/CA mechanism of the CAP based on this model. In contrast to existing literature, all the key characteristics of IEEE 802.15.4 are considered, such as the channel, limited number of retransmissions, ACKs, unsaturated traffic, packet size, and superframe structure.

5.1. Modelling of CAP

Here, we develop a generalized Markov chain model of the slotted CSMA/CA algorithm of the beacon-enabled IEEE 802.15.4 MAC. The core contribution of the analysis is the derivation of the stationary probability distribution of the chain, which is summarized by Proposition 5.1. Compared to previous results, e.g., [Misic and Misic 2005; Misic et al. 2006; Ramachandran et al. 2007; Pollin et al. 2008; Sahoo and Sheu 2008; Jung et al. 2009; Buratti 2010], the novelty of this chain consists in the modelling of the channel, retry limits for each packet transmission, the ACK, the inclusion of unsaturated traffic regimes, packet size, and superframe structure. We will also discuss the strength of the proposed Markov chain model with respect to previous studies [Misic and Misic 2005; Misic et al. 2006; Jung et al. 2009; Buratti 2010].

Let $b(t), c(t), e(t)$ and $f(t)$ be the stochastic processes representing the backoff stage, the state of the backoff counter, the state of retransmission counter, and the state of deferred transmission at time t experienced by a device. The binary variable $f(t)$ indicates if a transmission has been deferred ($f(t) = 1$) or not ($f(t) = 0$), which is due to the limited size of superframe duration to transmit a packet. By making the natural assumption that devices start sensing independently, the stationary probability τ that a device attempts a first carrier sensing in a randomly chosen time slot is constant and independent of other devices. The quadruple $(b(t), c(t), e(t), f(t))$ is the state evolution of the Markov chain. We use (i, j, k, l) to denote a particular state. We assume the following notation for the MAC parameters: $m_0 \triangleq$ *macMinBE*, $m \triangleq$ *macMaxCSMABackoffs*, $n \triangleq$ *macMaxFrameRetries*, $m_b \triangleq$ *macMaxBE*, $W_0 \triangleq 2^{m_0}$, $W_m \triangleq 2^{\min(m_0+m, m_b)}$, where *macMinBE*

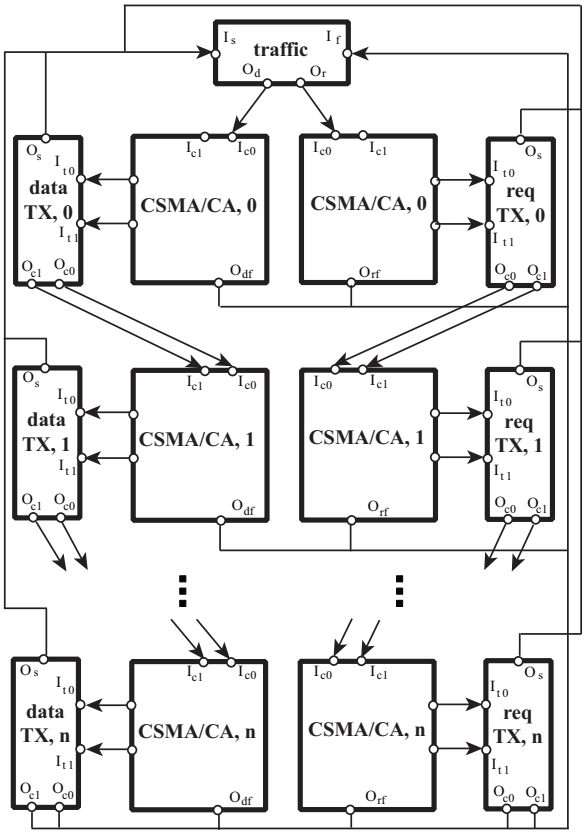
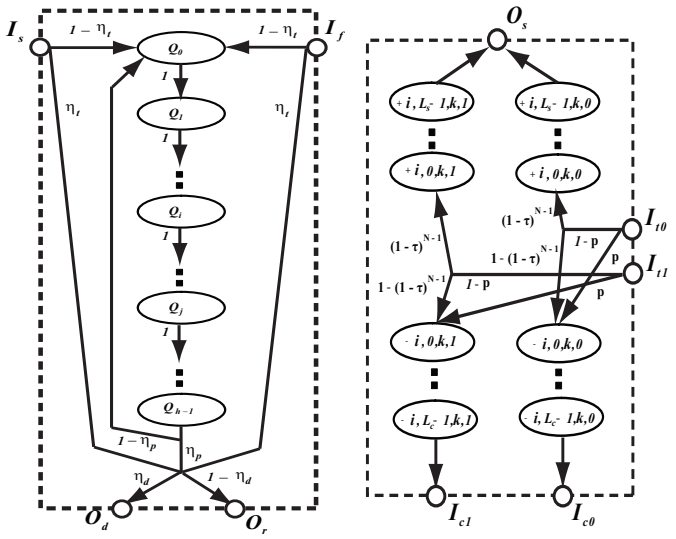


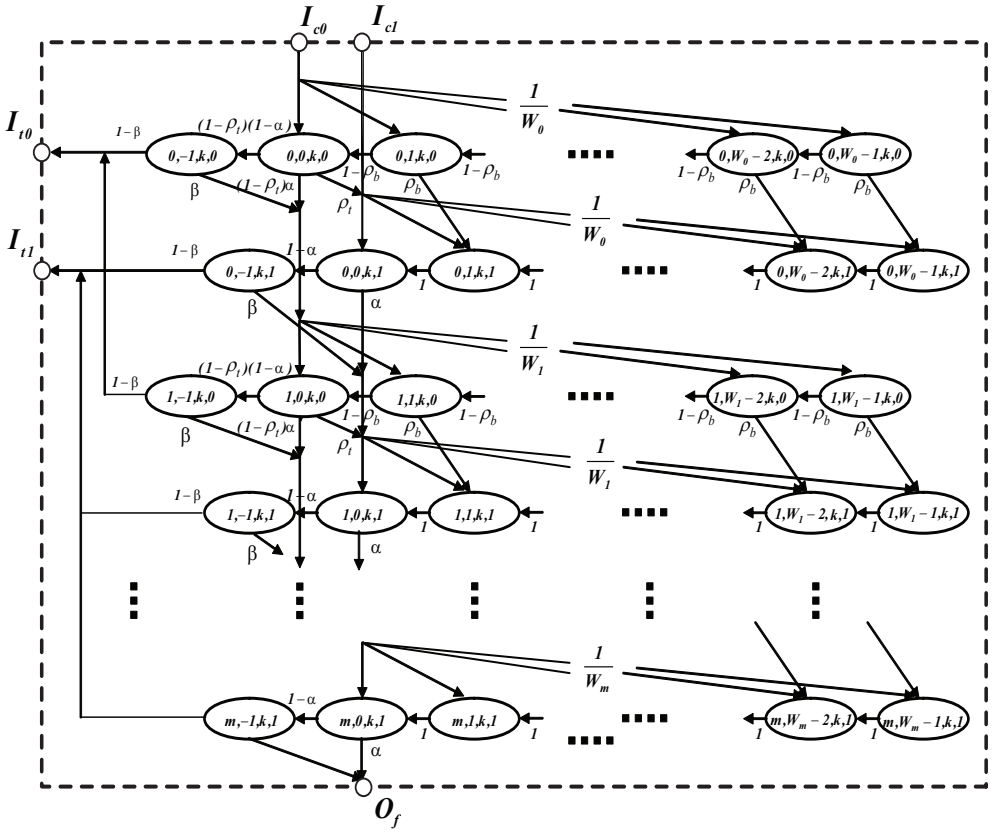
Fig. 3. Generalized Markov chain building blocks modelling the CSMA/CA algorithm of the IEEE 802.15.4 MAC for a single device. Details of traffic, TX, and CSMA/CA blocks are described in Figs 4(a)–(c), respectively.

is the minimum value of the backoff exponent, *macMaxCSMABackoffs* is the maximum number of backoffs allowed, *macMaxFrameRetries* is the maximum number of retries allowed, and *macMaxBE* is the maximum value of the backoff exponent in the CSMA/CA algorithm.

The Markov chain consists of three parts corresponding to the traffic generation block, CSMA/CA algorithm blocks, and packet transmission blocks in Fig. 3. The states $Q_0, \ldots, Q_{h-1}$ of Fig. 4(a) correspond to the idle-queue states when the packet queue is empty and the device is waiting for the next packet generation time. Note that the idle-queue states $Q_0, \ldots, Q_{h-1}$ take into account the sampling interval of the device. Each device generates a non time-critical data packet with probability η_d and a time-critical packet with probability $1 - \eta_d$. Then, the device performs the CSMA/CA algorithm to check a clean channel to send the non time-critical data packet or the GTS request for the time-critical packet. The Markov chain of the CSMA/CA algorithm consists of two parts corresponding to the backoff states and clear channel assessment (CCA) states reported in Fig. 4(c). The states from $(i, W_m - 1, k, l)$ to $(i, W_0 - 1, k, l)$ represent the backoff states. The states $(i, 0, k, l)$ and $(i, -1, k, l)$ represent the first CCA (CCA_1) and second CCA (CCA_2), respectively. Let α be the probability that CCA_1 is busy, and β the probability that CCA_2 is busy. If a device fails to obtain a clear channel due to repeated busy channel, then the packet is discarded. If the channel sensing is successful, then the device goes to the packet transmission block. In Fig. 4(b), the states $(+i, j, k, l)$ and $(-i, j, k, l)$ correspond to successful transmission and collision, respectively. Note that the states $i = 1$ and $i = 2$ denote the non time-critical data packet and the GTS request of time-critical data packet, respectively.



(a) **traffic**: Traffic generation block. (b) **TX**: Packet transmission block.



(c) **CSMA/CA**: CSMA/CA algorithm block of the k-th retransmission state.

Fig. 4. Detailed description of the components of every block of Fig. 3.

Before deriving the stationary probability of the Markov chain of Fig. 3, we need some definitions. Let P_c be the retransmission probability. If the packet collides due to other transmitted packets or is lost due to a bad channel, then the device repeats the CSMA/CA algorithm until a maximum number of retransmissions n. We model the bad channel by independent Bernoulli trials with the bad channel probability p where $0 \leq p \leq 1$. The independence of the trial results is assumed to be over all links and time slots. We assume that a transmitted packet in the bad channel does not affect other transmitted packets in the good channel. The slow variation of a fading channel is assumed so that the received power is constant throughout an entire transmission. If the packet transmission is successful, then it goes back to the traffic generation block. By knowing the duration of an ACK frame, ACK timeout, inter-frame spacing (IFS), data packet length, and header duration, we define the successful packet transmission time L_s and the packet collision time L_c with ACK and the successful packet transmission time L_g without ACK as

$$L_s = L_p + L_{\mathrm{w,ack}} + L_{\mathrm{ack}} + L_{\mathrm{IFS}}\,, \tag{3}$$

$$L_c = L_p + L_{\mathrm{m,ack}}\,, \tag{4}$$

$$L_g = L_p + L_{\mathrm{IFS}}\,, \tag{5}$$

where L_p is the total packet length including overhead and payload, $L_{\mathrm{w,ack}}$ is ACK waiting time, L_{ack} is the length of the ACK frame, L_{IFS} is the IFS time, and $L_{\mathrm{m,ack}}$ is the timeout of the ACK, see details in the standard [IEEE 2006]. Note that the successful packet transmission time $L_{s,d}$ and the packet collision time $L_{c,d}$ of a non time-critical data packet with packet length $L_{p,d}$ are obtained by substituting $L_{p,d}$ with L_p in Eqs. (3) and (4), respectively. Similarly, we derive the successful GTS request transmission time $L_{s,r}$ and the GTS request collision time $L_{c,r}$ for a length of GTS requests $L_{p,r}$.

To compute the stationary probability of the Markov chain, we first derive the probability ρ_t that the transmission is deferred due to the lack of the remaining time slots in a CAP as follows. The total number of time slots that are needed for a single transmission is $2L_{\mathrm{sc}} + L_p + L_{\mathrm{w,ack}} + L_{\mathrm{ack}}$ where two slots $2L_{\mathrm{sc}}$ are included due to the number of time slots for performing two CCAs and other components take into account the packet transmission with an ACK frame in Eqs. (3) and (4). Hence, ρ_t is approximated by

$$\rho_t = \frac{L_{\mathrm{tx}}}{T_{\mathrm{CAP}}} = \frac{2L_{\mathrm{sc}} + L_p + L_{\mathrm{w,ack}} + L_{\mathrm{ack}}}{T_{\mathrm{CAP}}}\,, \tag{6}$$

where T_{CAP} is the total number of time slots in a CAP. Similarly, the probabilities $\rho_{p,d}$ and $\rho_{p,r}$ of the events of deferred attempts due to the lack of remaining slot times for a non time-critical data packet and for a GTS request are obtained by replacing L_p with $L_{p,d}$ and $L_{p,r}$ in Eq. (6), respectively. We remark here that the previous literature [Misic and Misic 2005; Misic et al. 2006; Jung et al. 2009; Buratti 2010] does not consider the extra backoff mechanism for the event of deferred attempts of the transmission which is explicitly described in the standard [IEEE 2006]. Next, ρ_b is the probability that the MAC sublayer pauses the backoff countdown at the end of the CAP due to the limited length. Analogously, the approximated probability is

$$\rho_b = \frac{1}{T_{\mathrm{CAP}}}\,. \tag{7}$$

Note that the previous literature [Misic and Misic 2005; Misic et al. 2006; Jung et al. 2009; Buratti 2010] does not take into account this event although it is a very critical aspect for delay analysis, as we show in Section 5.2.2.

We have the following results:

PROPOSITION 5.1. *Let the stationary probability of the Markov chain in Fig. 3 be*

$$S_{i,j,k,l} = \lim_{t\to\infty} \Pr\left[b(t) = i, c(t) = j, e(t) = k, f(t) = l\right],$$

where $i \in (+2,-2,+1,-1) \cup (0,m)$, $j \in (-1, \max(W_i-1, L_s-1, L_c-1))$, $k \in (0,n)$, $l \in (0,1)$. *Then,*

$$S_{i,j,k,0} = \varpi_{i,j} g^k S_{0,0,0,0}\,,$$

with $\varpi_{i,j} = \left[1-(1-\rho_b)^{W_i-j}\right]\left[1-(1-\rho_b)^{W_i}\right]^{-1}$, *and* $g = P_c(1-\rho_t)(1-x)\sum_{i=0}^{m}\xi_i \sum_{j=0}^{W_0-1}(1-\rho_b)^j/W_0$, *where* $\xi_i = \prod_{r=1}^{i}(1-\rho_t)x/W_r \sum_{j=0}^{W_r-1}(1-\rho_b)^j$, $x = \alpha + (1-\alpha)\beta$ *and* $\xi_0 = 1$. *Moreover,*

$$S_{i,j,k,1} = \left[\sum_{r=j}^{W_i-1}\rho_b \varpi_{i,r}\xi_i g^k + \rho_t \frac{W_i-j}{W_i}\xi_i g^k + x\frac{W_i-j}{W_i}v_{i-1,k}u(i-1) + P_c(1-x)\frac{W_0-j}{W_0}\sum_{i=0}^{m} v_{i,k-1}\delta(i)u(k-1)\right] S_{0,0,0,0}\,,$$

where $u(i)$ *is the unit step function,* $\delta(i)$ *is the unit discrete delta function and*

$$v_{i,k} = x^i c + x^i \sum_{r=1}^{i} \frac{\rho_t + \rho_b \sum_{j=1}^{W_r-1}\varpi_{r,j}}{x^r} g^k + x^i a_k\,,$$

$$a_k = P_c(1-x)\sum_{i=0}^{m} v_{i,k-1}, \qquad k \geq 1\,,$$

$$v_{0,0} = \rho_t + \rho_b\, c \sum_{j=1}^{W_0-1}\varpi_{0,j}\,,$$

and $c = \rho_t + \rho_b \sum_{j=1}^{W_0-1}\varpi_{0,j}$ *and* $a_0 = 0$.

PROOF. See Appendix A.1. □

We remark here that the probability $S_{0,0,0,0}$, which plays a key role in the analysis, is different from the corresponding term given in [Misic and Misic 2005; Misic et al. 2006; Ramachandran et al. 2007; Pollin et al. 2008; Sahoo and Sheu 2008; Jung et al. 2009; Buratti 2010] due to our different modelling of channel, retry limits, ACK, unsaturated traffic, packet size, and superframe period. In Section 8, we demonstrate the validity of this Markov chain model by experimental results, and show how the performance analysis is affected by the extended model we derive in this paper.

5.2. Performance indicators of CAP

We now use the Markov chain model developed in the previous section to derive the performance indicators of the CAP in terms of reliability and average delay in Section 5.2.1 and 5.2.2, respectively. The main contributions of this section are given by Propositions 5.2 and 5.3 below.

5.2.1. Reliability. The main contributions of this section is the derivation of the probability of successful packet reception, reliability. With this goal in mind, we derive first the probability that a device attempts CCA_1 in a randomly chosen time slot:

$$\tau = \sum_{i=0}^{m}\sum_{k=0}^{n}\sum_{l=0}^{1} S_{i,0,k,l} = \left(\frac{1-g^{n+1}}{1-g}\sum_{i=0}^{m}\xi_i + \sum_{i=0}^{m}\sum_{k=0}^{n} v_{i,k}\right) S_{0,0,0,0}\,. \tag{8}$$

This probability depends on the probability that a transmitted packet encounters a collision or a bad channel P_c, the probability that CCA_1 is busy α, and the probability that CCA_2 is also busy β. These probabilities are developed in the following.

The packets are discarded due to either of two reasons: (i) collision, or (ii) bad channel. A transmitted packet in the good channel collides if at least one of the $N-1$ remaining devices transmits in the good channel and in same time slot. In addition, a transmitted packet is lost due to the bad channel. If all devices transmit with probability τ, then

$$P_c = \left(1-(1-\tau(1-p))^{N-1}\right)(1-p)+p\,,$$

where N is the number of total devices present in the network.

Similarly to [Pollin et al. 2008], we derive the busy channel probabilities α and β by considering the probability of the bad channel as follows. The busy channel probability of CCA_1 is

$$\alpha = \alpha_1 + \alpha_2\,, \tag{9}$$

where α_1 is the probability of finding channel busy during CCA_1 due to data transmission, namely

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

where the average length of packet is $\overline{L}_p = \eta_d L_{p,d} + (1-\eta_d)L_{p,r}$, and α_2 is the probability of finding the channel busy during CCA_1 due to ACK transmission, which is

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

where L_{ack} is the length of the ACK. The busy channel probability of CCA_2 is

$$\beta = \frac{1-(1-\tau)^{N-1}+N\tau(1-p)(1-\tau(1-p))^{N-1}}{2-(1-\tau)^N+N(1-p)\tau(1-\tau(1-p))^{N-1}}\,, \tag{10}$$

see details in [Pollin et al. 2008]. The expressions of the carrier sensing probability τ and the busy channel probabilities α and β form a system of non-linear equations that can be solved via numerical methods.

PROPOSITION 5.2. *Consider the definitions given in Proposition 5.1. Then, the reliability is*

$$P_s = 1-(1-\rho_t)x\xi_m\frac{1-g^{n+1}}{1-g} - x\sum_{k=0}^{n}v_{m,k} - P_c(1-x)\left((1-\rho_t)\sum_{i=0}^{m}\xi_i g^n + \sum_{i=0}^{m}v_{i,n}\right)\,. \tag{11}$$

PROOF. In slotted CSMA/CA, packets are discarded due to two reasons: (i) channel access failure and (ii) 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 or bad channel conditions after $n+1$ attempts. Following the Markov model presented in Fig. 3, the probability that the packet is discarded due to channel access failure is

$$P_{dc} = \sum_{k=0}^{n}((1-\rho_t)xS_{m,0,k,0} + xS_{m,0,k,1})S^{-1}_{0,0,0,0} = (1-\rho_t)x\xi_m\frac{1-g^{n+1}}{1-g} + x\sum_{k=0}^{n}v_{m,k}\,. \tag{12}$$

The probability of a packet being discarded due to retry limits is

$$P_{dr} = \sum_{i=0}^{m}\sum_{l=0}^{1}P_c(1-\beta)S_{i,-1,n,l}S^{-1}_{0,0,0,0} = P_c(1-x)\left((1-\rho_t)\sum_{i=0}^{m}\xi_i g^n + \sum_{i=0}^{m}v_{i,n}\right)\,. \tag{13}$$

Therefore, by considering Eqs. (12) and (13), the reliability is

$$P_s = 1-P_{dc}-P_{dr}\,. \tag{14}$$

□

5.2.2. Delay. In this section, we derive the analytical expression of the average delay for a successfully received packet. 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.

PROPOSITION 5.3. *The average delay to transmit a packet and successfully receive an ACK is*

$$\begin{aligned}\mathbb{E}[D] =& \sum_{k=0}^{n}(1-P_c)P_c^k\left[\Pr[\mathscr{B}_0]^k\left(\Pr[\mathscr{B}_0](L_s+kL_c+(k+1)\mu_{\mathscr{B}_0})\right.\right. \\ &+\Pr[\mathscr{T}](L_s+kL_c+k\mu_{\mathscr{B}_0})+\mu_{\mathscr{T}})+\sum_{r=0}^{k-1}\Pr[\mathscr{B}_0]^{k-1-r}\Pr[\mathscr{B}_1]^{r+1} \\ &\left.\left.\times(\Pr[\mathscr{T}](L_s+kL_c+(k-1-r)\mu_{\mathscr{B}_0}+(r+1)\mu_{\mathscr{B}_1})+\mu_{\mathscr{T}})u(k-1)\right]P_{\text{tot}}^{-1}\right.,\end{aligned} \tag{15}$$

where

$$\begin{aligned}\Pr[\mathscr{T}] =& \Pr[\mathscr{C}_0]+\Pr[\mathscr{C}_1]+\Pr[\mathscr{F}_0]+\Pr[\mathscr{F}_1]\,, \\ \mu_{\mathscr{T}} =& \Pr[\mathscr{C}_0]\mu_{\mathscr{C}_0}+\Pr[\mathscr{C}_1]\mu_{\mathscr{C}_1}+\Pr[\mathscr{F}_0]\mu_{\mathscr{F}_0}+\Pr[\mathscr{F}_1]\mu_{\mathscr{F}_1}\,, \\ P_{\text{tot}} =& \sum_{k=0}^{n}(1-P_c)P_c^k\left[\Pr[\mathscr{B}_0]^k\left(\Pr[\mathscr{B}_0]+\Pr[\mathscr{T}]\right)\right. \\ &\left.+\sum_{r=0}^{k-1}\Pr[\mathscr{B}_0]^{k-1-r}\Pr[\mathscr{B}_1]^{r+1}\Pr[\mathscr{T}]u(k-1)\right]\,,\end{aligned}$$

where $\mu_* = \sum_{i=0}^{m}\rho_{*,i}/\Pr[*]$ *, and where* $*$ *is one of the events* $\{\mathscr{B}_0,\mathscr{B}_1,\mathscr{C}_0,\mathscr{C}_1,\mathscr{F}_0,\mathscr{F}_1\}$*, with*

$$\begin{aligned}\rho_{\mathscr{B}_{0,i}} =& \varepsilon_i\sum_{k=0}^{i}(1-\rho_t)\Psi_i(1-x)\mu_{D_{k,i}}\,, \\ \rho_{\mathscr{B}_{1,i}} =& \sum_{k=0}^{i}\alpha^{i-k}((1-\alpha)\beta)^k(1-x)\mu_{D_{k,i}}\,, \\ \rho_{\mathscr{C}_{0,i}} =& \varepsilon_i\sum_{k=0}^{i}\rho_t\Psi_i(1-x)\left(\mu_{D_{k,i}}+T_{SP}+L_{\text{tx}}+\mu_{\Phi_i}\right)\,, \\ \rho_{\mathscr{C}_{1,i}} =& \varepsilon_i\sum_{k=0}^{i}(1-\Psi_i)(1-x)\left(\mu_{D_{k,i}}+T_{SP}+\mu_{\Phi_i}\right)\,, \\ \rho_{\mathscr{F}_{0,i}} =& \rho_t x\sum_{f=0}^{i-2}\sum_{h=0}^{i-f-1}\varepsilon_f\sum_{k=0}^{f}\Psi_f\alpha^{i-f-h-1}((1-\alpha)\beta)^h(1-x)\left(\mu_{D_{k,i}}+T_{SP}+L_{\text{tx}}+\mu_{\Phi_f}\right)\,, \\ \rho_{\mathscr{F}_{1,i}} =& x\sum_{f=0}^{i-2}\sum_{h=0}^{i-f-1}\varepsilon_f\sum_{k=0}^{f}(1-\Psi_f)\alpha^{i-f-h-1}((1-\alpha)\beta)^h(1-x)\left(\mu_{D_{k,i}}+T_{SP}+\mu_{\Phi_f}\right)\,,\end{aligned}$$

with

$$\begin{aligned}\Pr[\mathscr{B}_0] =& \sum_{i=0}^{m}\varepsilon_i\sum_{k=0}^{i}(1-\rho_t)\Psi_i(1-x)\,, \\ \Pr[\mathscr{B}_1] =& \sum_{i=0}^{m}\sum_{k=0}^{i}\alpha^{i-k}((1-\alpha)\beta)^k(1-x)\,, \\ \Pr[\mathscr{C}_0] =& \sum_{i=0}^{m}\varepsilon_i\sum_{k=0}^{i}\rho_t\Psi_i(1-x)\,,\end{aligned}$$

$$\Pr[\mathscr{C}_1] = \sum_{i=0}^{m} \varepsilon_i \sum_{k=0}^{i} (1-\Psi_i)(1-x),$$

$$\Pr[\mathscr{F}_0] = \rho_t x \sum_{i=0}^{m} \sum_{f=0}^{i-2} \sum_{h=0}^{i-f-1} \varepsilon_f \sum_{k=0}^{f} \Psi_f \alpha^{i-f-h-1} ((1-\alpha)\beta)^h (1-x) u(i-2),$$

$$\Pr[\mathscr{F}_1] = x \sum_{i=0}^{m} \sum_{f=0}^{i-2} \sum_{h=0}^{i-f-1} \varepsilon_f \sum_{k=0}^{f} (1-\Psi_f)\, \alpha^{i-f-h-1} ((1-\alpha)\beta)^h (1-x) u(i-2),$$

and $\mu_{\Phi_k} = \sum_{l=0}^{W_k-1} T_b\, l/W_k$, $\Psi_i = \sum_{j=0}^{W_i-1} (1-\rho_b)^j / W_i$,

$$\mu_{D_{k,i}} = \mu_{\Phi_k} + (i+1) 2 L_{\text{sc}},$$

$$\epsilon_r = (1-\rho_t)\Psi_r \max(\alpha, (1-\alpha)\beta),$$

$$\varepsilon_i = \begin{cases} \prod_{r=0}^{i-1} \epsilon_r & \textit{if } i \geq 1, \\ 1 & \textit{otherwise}, \end{cases}$$

T_{SP} *is the inactive period.*

PROOF. See Appendix A.2. □

We remark that $\mathbb{E}[D]$ is a function of the MAC parameters W_i, m, n, L_{sc} of CSMA/CA mechanism as well as the MAC parameters $T_{\text{SP}}, \rho_t, \rho_b$ related to superframe period. Furthermore, the busy channel probabilities α and β, and retransmission probability P_c of the network affect the average delay.

6. PERFORMANCE ANALYSIS OF CFP

We now turn our attention from the CAP to the CFP. We present the modelling of the GTS allocation mechanism based on a Markov chain model in Section 6.1. Then, in Section 6.2 we build on this modelling to characterize the average queuing delay of the GTS allocation.

6.1. Modelling of CFP

The modelling of the GTS allocation is given in two steps. First, we derive the constraints on the number of time slots to allocate from the IEEE 802.15.4 specification. Then, we model the behavior of the GTS allocation by a Markov chain.

First, we derive the number of GTSs Δ_u that can be allocated as a function of the IEEE 802.15.4 MAC parameters BO, SO and the number of time-critical data packets φ_n for each GTS request that can be served. Recall that during the CFP each packet has a fixed length L_g given by Eq. (5).

We assume that all GTS transmissions are successful and every request asks for at most one GTS. Each device may use a multiple number of GTSs. Furthermore, we assume that there is no reallocation of the GTS. Let the duration of a superframe slot be T_{SS}, then

$$T_{\text{SS}} = \frac{T_{\text{SD}}}{N_{\text{SS}}} = T_0 \times 2^{\text{SO}-4},$$

where T_{SD} is the number of symbols forming a superframe, N_{SS} is the number of slots contained in a superframe, and T_0 is the number of symbols forming a superframe when $\text{SO} = 0$. Consider that a single GTS may extend over a number of superframe slots θ_n. Since a given GTS needs to be larger than the total forward delay $\varphi_n L_g$, it follows that

$$\theta_n \geq \underline{\theta} = \left\lceil \frac{\varphi_n L_g}{T_{\text{SS}}} \right\rceil, \tag{16}$$

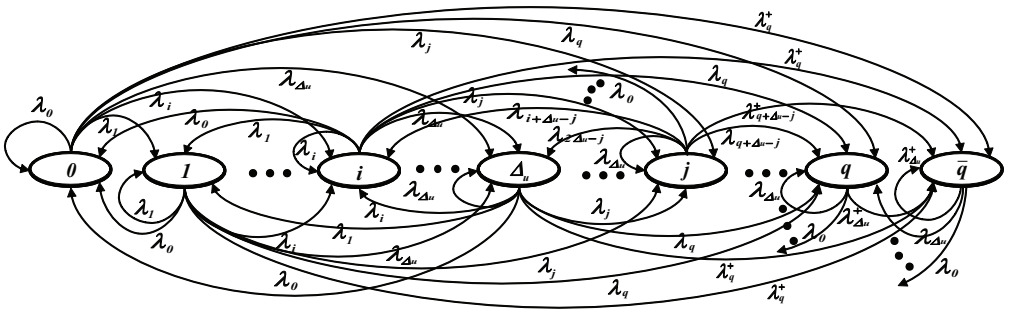


Fig. 5. Markov chain model for the GTS allocation of the CFP. The state defines the number of waiting requests to be served.

where $\underline{\theta}$ is the minimum number of superframe slots for a single GTS to serve the data frames $\varphi_n L_g$. Because the minimum CAP length, $T_{\text{CAP}} \geq T_{\min}$, the constraint of the maximum number of GTSs that the coordinator can allocate in a superframe is

$$T_{\text{SD}} - T_{\text{CFP}} = T_{\text{SD}} - T_{\text{SS}}\, k\, \underline{\theta} \geq T_{\min}\,,$$

where k is the number of waiting requests and T_{CFP} is the number of symbols forming a CFP. The maximum number of GTSs to be allocated to devices is

$$\Delta_u = \min\left(\left\lfloor \frac{N_{\text{SS}}\left(1-\frac{T_{\min}}{T_{\text{SD}}}\right)}{\underline{\theta}} \right\rfloor,\, N_{\text{GTS}}\right), \tag{17}$$

where N_{GTS} is the maximum number of GTS descriptors limited to seven in the IEEE 802.15.4 [IEEE 2006]. Note that Δ_u is a function of the parameters BO, SO and application constraints, i.e., the number of time-critical data packets φ_n wherein each packet has a fixed length L_g.

Now, we develop a Markov model for the GTS allocation mechanism. Let $r(t)$ be the stochastic processes representing the number of waiting requests of the coordinator at the beginning of the superframe t. This stochastic process defines a one-dimensional Markov chain of Fig. 5. If a number of received requests are over the limited queue size q after the new arrivals, then some of the new requests will be dropped on the base of a FCFS policy. We denote by $\bar{q}$ the state of the Markov chain indicating that the new arrived requests are dropped due to the limited queue size. Let λ_i be the probability of i successful requests during the CAP, given that the maximum number of requests is $\bar{\lambda}$. If the arriving requests observe that the queue size is over a threshold q, then these arriving requests are dropped because of a time limitation requirement N_{DPT}. Note that q is fixed and it is given by $\Delta_u\,(N_{\text{DPT}}+1)$ where Δ_u is the maximum number of GTSs for each superframe and $N_{\text{DPT}} = \textit{aGTSDescPersistenceTime}$. After N_{DPT} superframes, the GTS description of the beacon is removed. When there are k requests, $\Delta_u \leq k \leq q$, then some requests $k - \Delta_u$ will be delayed to obtain GTS in the next superframe. The transition probabilities of the chain are

$$P(j|i) = \lambda_j,\ \text{for } 0 \leq i < \Delta_u, j \neq \bar{q}\,, \tag{18}$$

$$P(j|i) = \lambda_{j+\Delta_u - i},\ \text{for } i \geq \Delta_u, j \geq i - \Delta_u\,, i \neq \bar{q} \tag{19}$$

$$P(j|i) = 0,\ \text{for } i \geq \Delta_u, j < i - \Delta_u, j \neq \bar{q}\,, \tag{20}$$

$$P(\bar{q}|i) = \lambda_q^+,\ \text{for } 0 \leq i < \Delta_u\,, \tag{21}$$

$$P(\bar{q}|i) = \lambda_{q+\Delta_u - i}^+,\ \text{for } i \geq \Delta_u\,, \tag{22}$$

where $\lambda_k^+ = 1 - \sum_{i=0}^{k} \lambda_i = \sum_{i=k+1}^{\bar{\lambda}} \lambda_i$. Eq. (18) gives the probability that j requests arrive when there are less than Δ_u requests, the maximum number of GTSs. Note that GTSs will be allocated to previously arrived requests before the CAP of the superframe, i.e., all waiting requests at the

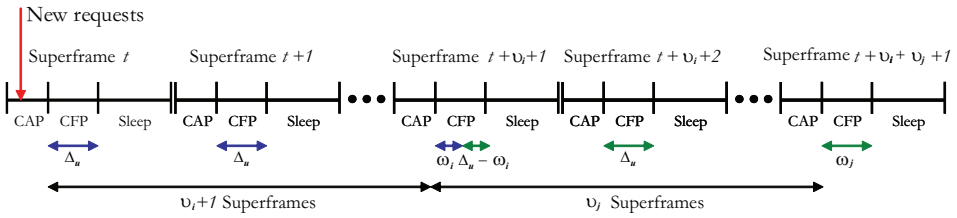


Fig. 6. Waiting time of j new requests that observe i old requests at superframe t.

beginning of the current superframe arrive during the CAP of the previous superframe. Eq. (19) gives the probability that j requests remain at the current superframe when there are some requests $i \geq \Delta_u$ at the previous superframe. The coordinator allocates a maximum number of GTS Δ_u to serve the requests $i \geq \Delta_u$. Furthermore, since the number of waiting requests is greater than Δ_u, some requests will fail to obtain GTS at the superframe. Hence, there are two different groups of requests at the current superframe, one group of requests arriving before the current superframe and another group of requests arriving at the current superframe. Let us call the first group "old requests" and the second group "new requests" throughout this paper. Eq. (20) shows the probability that there are less requests in the current superframe than $i - \Delta_u$ if there are $i \geq \Delta_u$ requests at previous superframe. Note that the transition probability of Eq. (20) is equal to 0 due to the maximum number of GTSs Δ_u. Eqs. (21) and (22) show the transition probabilities of the drop state $\bar{q}$ for the requests at the current superframe.

Let us denote the probability that the process described by the Markov chain is in state k at time t by $\pi_k^t = \Pr[r(t) = k]$. Let the state probability vector at time t be $\boldsymbol{\pi}^t = (\pi_0^t, \pi_1^t, \ldots, \pi_{\bar{q}}^t)$, for $t \in \{1, \cdots\}$ and $k \in \{0, \cdots, q, \bar{q}\}$. Then, we obtain $\boldsymbol{\pi}^t = \boldsymbol{\pi}^1 \mathbf{P}^t$ where $\mathbf{P}^t$ is the t-step transition probability matrix

$$\mathbf{P} = \begin{pmatrix} \lambda_0 & \lambda_1 & \cdots & \lambda_{\Delta_u} & \lambda_{\Delta_u+1} & \cdots & \lambda_q & \lambda_q^+ \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_0 & \lambda_1 & \cdots & \lambda_{\Delta_u} & \lambda_{\Delta_u+1} & \cdots & \lambda_q & \lambda_q^+ \\ 0 & \lambda_0 & \lambda_1 & \cdots & \lambda_{\Delta_u} & \cdots & \lambda_{q-1} & \lambda_{q-1}^+ \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_0 & \lambda_1 & \cdots & \lambda_{\Delta_u+1} & \lambda_{\Delta_u+1}^+ \\ 0 & 0 & \cdots & 0 & \lambda_0 & \cdots & \lambda_{\Delta_u} & \lambda_{\Delta_u}^+ \\ 0 & 0 & \cdots & 0 & \lambda_0 & \cdots & \lambda_{\Delta_u} & \lambda_{\Delta_u}^+ \end{pmatrix}. \tag{23}$$

The stationary distribution is

$$\boldsymbol{\pi} = \lim_{t \to \infty} \boldsymbol{\pi}^t = \mathbf{1} \left(\mathbf{P} - \mathbf{I} + \mathbf{1}^{\mathbf{T}} \mathbf{1} \right)^{-1}, \tag{24}$$

where $\boldsymbol{\pi} \in \mathbb{R}^{1 \times (q+2)}$, $\mathbf{P}$ is the transition probability matrix given in Eq. (23), $\mathbf{I} \in \mathbb{R}^{(q+2) \times (q+2)}$ is the identity matrix, $\mathbf{1} \in \mathbb{R}^{1 \times (q+2)}$ is the vector with all elements equal to one, and $\mathbf{1}^T \mathbf{1} \in \mathbb{R}^{(q+2) \times (q+2)}$ is the matrix with all the elements equal to one [Gross and Harris 1998].

6.2. Performance indicators of CFP

In this section, we analyze the expected queuing delay of the GTS allocation, namely the average delay between the arrival time of a GTS request to a coordinator and the actual transmission time of a time-critical packet after its allocation in some of the next superframes. The core contribution of this section is given by Proposition 6.1.

We recall that each device sends the GTS requests using CSMA/CA mechanism of CAP. If the GTS request command cannot be sent due to a CSMA/CA algorithm failure, the device discards the GTS request. The coordinator determines a device list of the GTS allocation of the next super-

frame based on a FCFS policy. The number of GTS descriptors is limited by the remaining length of the CAP and the desired length of the requested GTS. Hence, the CFP length depends on the GTS requests and the current available capacity in the superframe. The coordinator includes the GTS descriptor which is the device list that obtains GTSs in the following beacon to announce the allocation information. The coordinator makes this decision within *aGTSDescPersistenceTime* superframes. Once the device receives an ACK corresponding to the GTS request, the device continues to track beacons and waits for at most *aGTSDescPersistenceTime* superframes. When new requests are received during the CAP, then the delay of the GTS allocation can be estimated by observing the queue size of the waiting requests. If there are not sufficient GTS resources for the request in the current superframe, the GTS allocation will be delayed to the next superframe. If the coordinator is not able to assign the dedicated slot within *aGTSDescPersistenceTime* superframes, then it discards the GTS requests. Assume that the arrival process of requests is uniformly distributed during the CAP. The average delay between the arrival time of the CAP and the end of the CAP at a superframe is then half of the CAP.

Suppose that j requests arrive at superframe t and that there are i requests already waiting in the queue. We can calculate the number of successful requests that obtain a GTS and delayed requests that fail to obtain a GTS at the next superframe, as explained in the following. Based on the FCFS policy, j new requests are able to obtain GTS after i old requests arriving before the current superframe t. In Fig. 6, we first compute the number of superframes to allocate GTSs for old requests. i old requests require υ_i+1 superframes, plus ω_i remainders that obtain the GTSs later, after $\upsilon_i + 1$ superframes have been served. Hence, i old requests require $\upsilon_i + 2$ superframes to obtain GTS if there are ω_i remainders. These remainders will share the CFP along with the new requests at superframe $t + \upsilon_i + 1$, see Fig. 6. It follows that

$$\upsilon_i = \left\lfloor \frac{i-\Delta_u}{\Delta_u} \right\rfloor , \quad \omega_i = \mathrm{rem}\left(\frac{i-\Delta_u}{\Delta_u}\right) ,$$

where υ_i is the integer quotient and the function $\mathbf{rem}$ returns an integer remainder. The nominator $i-\Delta_u$ takes into account the current GTS allocation mechanism of superframe t. Note that the new requests are not able to get GTS at the current superframe t since the beacon includes information of the GTS allocation at each superframe.

By considering the waiting time of old requests, it is possible to calculate the delay of the j new requests. The quotient υ_i of old requests can be considered as a delay offset in the computation of the delay of the new requests. We consider ω_i remainders by summing it with j new requests due to the shared CFP with the ω_i remained requests. In a similar way, the quotient and remainder of the new requests are given by

$$\upsilon_j = \left\lfloor \frac{j+\omega_i}{\Delta_u} \right\rfloor , \quad \omega_j = \mathrm{rem}\left(\frac{j+\omega_i}{\Delta_u}\right) .$$

The new requests require $\upsilon_j + 1$ superframes to allocate all GTS requests. In Fig. 6, ω_j waiting requests will remain waiting to obtain GTS at superframe $t + \upsilon_i + \upsilon_j + 1$.

To analyze the delay, we define the beacon interval as

$$T_{\mathrm{BI}}^{t,t+1} = T_{\mathrm{CFP}}^{t} + T_{\mathrm{SP}}^{t} + T_{\mathrm{CAP}}^{t+1} ,$$

where T_{CFP}^{t} is a CFP, T_{SP}^{t} is an inactive period, and T_{CAP}^{t} is a CAP at superframe t. Let $D_{i,j,t}$ denotes the expected delay of j new requests that observe a queue size of i waiting requests at superframe t. We distinguish the number of successful requests that obtain GTS out of j new requests for a different superframe. Note that j new requests of the superframe t will obtain GTSs from the superframe $t + \upsilon_i + 1$ to the superframe $t + \upsilon_i + \upsilon_j + 1$. The delay of allocating GTS is the sum of three components:

— Arrival delay of the CAP: We consider the arrival delay based on the FCFS policy of queue management. Note that that the arrival time of new requests is uniformly distributed in the CAP.

— Offset delay of old and new requests: The new requests need to wait a number of superframes before obtaining GTS since there are old and new requests arriving before the current request.
— Service delay of the CFP: If the new requests are able to obtain GTS at the current superframe, then the superframe slot size T_{SS} and the minimum number of superframe slots $\underline{\theta}$ need to be accounted for.

Hence, the following proposition holds:

PROPOSITION 6.1. *Let q be the queue size, then the expected delay of j new requests arriving at superframe t is*

$$\mathbb{E}[D_q(t)] = \sum_{i=0}^{\Delta_u - 1} G(i,q) + \sum_{i=\Delta_u}^{q} G(i, q-i+\Delta_u) + G(q,\Delta_u)\frac{\pi_{\bar{q}}^t}{\pi_q^t}\,, \tag{25}$$

where

$$G(a,b) = \sum_{j=1}^{\bar{\lambda}} P_{a,j}^t \left(D_{a,j,t}\, u(b-j) + D_{a,b,t}\, u(j-(b+1)) \right)\,,$$

with $P_{i,j}^t = \pi_i^t \lambda_j / \sum_{k=1}^{\bar{\lambda}} \lambda_k$, and $\bar{\lambda}$ is the maximum number of requests.

PROOF. See Appendix A.3. □

We remark here that the average queuing delay mainly depends on the traffic pattern λ_i of the number of GTS requests and protocol parameters BO, SO of Δ_u.

7. HYBRID MARKOV CHAIN MODEL

We are now in the position to give the core contribution of the overall paper. We propose a performance analysis of the hybrid MAC by connecting the two Markov chain models of CAP and CFP in Section 5 and Section 6. Then, we characterize the throughput of hybrid MAC by taking into account both the non time-critical data packet transmission of the CAP and the time-critical data packet transmission of the CFP.

7.1. Connection between CAP and CFP

In this section, we connect the two Markov chain models of the CAP of Section 5.1 and the CFP of Section 6.1. Furthermore, it is determined under which conditions the GTS allocation mechanism is stable. Recall that the number of GTS requests depends on the number of time-critical data packets that the devices want to send to the coordinator. With contention-based transmissions during the CAP, there will be loss and delay of GTS requests due to the busy channel, collisions, and bad channel conditions. The number of GTS requests affects the CAP length, i.e., increasing the number of GTS requests decreases the CAP length for a fixed superframe length. Hence, the performance of the GTS allocation of the CFP depends significantly on the number of successfully received GTS requests by the coordinator during the CAP. In other words, the stability analysis of the GTS allocation mechanism shows under which conditions the CAP and CFP portion of the superframe structure is stable. The stability analysis of the superframe structure is given by Lemma 7.1.

The coupling between the two chains is given by the PDF of the number of GTS requests per superframe, which abstracts the CSMA/CA mechanism of the CAP. Such a PDF is the input to analyze the GTS allocation mechanism of the CFP. We first derive the average number of successfully received GTS requests for a given CAP length T_{CAP}. Then, we propose the approximated PDF of the number of GTS requests per superframe based on the analysis of the CAP.

Suppose that the arrival process of GTS requests is independent with the reliability of GTS requests P_s of the CAP given in Eq. (11). The average service time for a successfully received packet

including the average packet generation time and packet delay is

$$\mathbb{E}[\psi_s] = \frac{1-\eta_t}{\eta_p} h + \mathbb{E}[D_s]\,,$$

where $\mathbb{E}[D_s]$ is the average delay for successfully received packets given in Eq. (15). By the same argument, the average service time for discarded packets due to channel access failure and retry limits are described by the following equations: $\mathbb{E}[\psi_{dc}] = (1-\eta_t)\,h/\eta_p + \mathbb{E}[D_{dc}]\,, \mathbb{E}[\psi_{dr}] = (1-\eta_t)\,h/\eta_p + \mathbb{E}[D_{dr}]$, where $\mathbb{E}[D_{dc}]$ and $\mathbb{E}[D_{dr}]$ are the average delay for discarded packets due to channel access failure and retry limits, respectively. To consider the limited CAP length T_{CAP}, we derive the feasible set of $(\mathbb{E}[\psi_s], \mathbb{E}[\psi_{dc}], \mathbb{E}[\psi_{dr}])$ which should satisfy the following constraint:

$$\mathscr{R} = \{\mathbf{r} | T_{\text{CAP}} - \min(\boldsymbol{\psi}_c) < \boldsymbol{\psi}_c^T \mathbf{r} \le T_{\text{CAP}}, \mathbf{r} \in \mathbb{Z}_+^3\}\,,$$

where $\boldsymbol{\psi}_c = (\,\mathbb{E}[\psi_s],\ \mathbb{E}[\psi_{dc}],\ \mathbb{E}[\psi_{dr}]\,)^T$, $\mathbf{r} = (\,r_1,\ r_2,\ r_3\,)^T$, and $\mathbb{Z}_+^3$ is triples of non-negative integers. Let us assume that the feasible set is a sequence of independent, identically distributed random variables which consists of V_{tot} trials, and each taking three possible outcomes, R_1, R_2, R_3. Notice that each possible outcome can occur with probability P_s given by Eq. (11), P_{dr} given by Eq. (12), and P_{dc} given by Eq. (13). Therefore, the probability mass function that R_1 occurs r_1 times, R_2 occurs r_2 times, and R_3 occurs r_3 times follows a multinomial distribution:

$$\Pr\left[\bigcap_{i=1}^{3} R_i = r_i\right] = \frac{\frac{V_{\text{tot}}!}{r_1!r_2!r_3!} P_s^{r_1} P_{dc}^{r_2} P_{dr}^{r_3}}{\sum\limits_{\mathbf{r}\in\mathscr{R}} \frac{V_{\text{tot}}!}{r_1!r_2!r_3!} P_s^{r_1} P_{dc}^{r_2} P_{dr}^{r_3}}\,,$$

where $V_{\text{tot}} = \sum\limits_{i=1}^{3} r_i$. From the multinomial distribution, the probability that a successfully received GTS request occurs r_1 times is

$$\Pr[R_1 = r_1] = \frac{\sum\limits_{r_2,r_3\in\mathscr{R}} \frac{V_{\text{tot}}!}{r_1!r_2!r_3!} P_s^{r_1} P_{dc}^{r_2} P_{dr}^{r_3}}{\sum\limits_{\mathbf{r}\in\mathscr{R}} \frac{V_{\text{tot}}!}{r_1!r_2!r_3!} P_s^{r_1} P_{dc}^{r_2} P_{dr}^{r_3}}\,.$$

Note that if $\mathbb{E}[\psi_s] = \mathbb{E}[\psi_{dc}] = \mathbb{E}[\psi_{dr}]$, then it follows that the distribution is binomial. Therefore, the expected number of GTS requests for a given CAP length T_{CAP} is

$$\mathbb{E}[R_1] = \sum_{r_1=0}^{\left\lfloor \frac{T_{\text{CAP},k}}{\mathbb{E}[\psi_s]} \right\rfloor} r_1\ \Pr[R_1 = r_1]\,. \tag{26}$$

Next, we derive the average number of successfully received non time-critical packet and GTS requests for a given CAP length for a hybrid MAC. Assume that the average number of successfully received packet is proportional to the superframe length. Let us denote the average number of successfully received packets N_{SD} including the non time-critical packet and GTS requests for a given superframe length T_{SD}. According to the standard [IEEE 2006], the CAP and CFP lengths are updated via the following equations:

$$\begin{aligned} T_{\text{CAP}}^{t+1} &= T_{\text{SD}} - T_{\text{CFP}}^{t+1}\,, \\ T_{\text{CFP}}^{t+1} &= \min\left(\frac{(1-\eta_d)N_{\text{SD}}}{T_{\text{SD}}} T_{\text{CAP}}^{t},\ \Delta_u\right) T_{\text{SS}}\,, \end{aligned} \tag{27}$$

where $T_{\text{CAP}}^0 = T_{\text{SD}}$ and $T_{\text{CFP}}^0 = 0$, and where the index t denotes the discrete time in the superframe unit. Recall that the probability to generate a time-critical packet is $1-\eta_d$. Then, an instantaneous average number of successfully received non time-critical packets N_{CAP}^t and GTS requests N_{CFP}^t for

a given superframe length T_{SD} are given by

$$N^t_{\text{CAP}} = \frac{\eta_d N_{\text{SD}} T^t_{\text{CAP}}}{T_{\text{SD}}}\,, \tag{28}$$
$$N^t_{\text{CFP}} = \frac{(1-\eta_d) N_{\text{SD}} T^t_{\text{CAP}}}{T_{\text{SD}}}\,.$$

LEMMA 7.1. *Iteration* (27) *converges to the fixed points* $(T^\star_{CAP}, T^\star_{CFP})$ *with*

$$T^\star_{CAP} = \begin{cases} \frac{T^2_{SD}}{T_{SD}+(1-\eta_d)N_{SD}T_{SS}} & \text{if} \quad N_{SD} < \frac{N_{SS}}{1-\eta_d}\,, \\ T_{SD} - \Delta_u T_{SS} & \text{if} \quad N_{SD} \geq \frac{N_{SS}\Delta_u}{(1-\eta_d)(N_{SS}-\Delta_u)}\,, \end{cases}$$

and $T^\star_{CFP} = T_{SD} - T^\star_{CAP}$.

PROOF. See Appendix A.4. □

Lemma 7.1 gives conditions on when the CAP and CFP lengths converge. Note that the case when $\frac{N_{\text{SS}}}{1-\eta_d} \leq N_{\text{SD}} < \frac{N_{\text{SS}}\Delta_u}{(1-\eta_d)(N_{\text{SS}}-\Delta_u)}$, the lemma is inconclusive. We may have an oscillatory behavior. When $N_{\text{SD}} \geq \frac{N_{\text{SS}}\Delta_u}{(1-\eta_d)(N_{\text{SS}}-\Delta_u)}$, since the average number of successfully received GTS requests is greater than the maximum number of GTSs to be allocated to devices Δ_u, many GTS requests will be dropped due to the limited queue size. When $N_{\text{SD}} < \frac{N_{\text{SS}}}{1-\eta_d}$, the portion of CAP and CFP of the superframe structure is stabilized without significant drop of GTS requests.

PROPOSITION 7.2. *The average number of successfully received non time-critical packets* N_{CAP} *and GTS requests* N_{CFP} *for a given superframe length* T_{SD} *are given by*

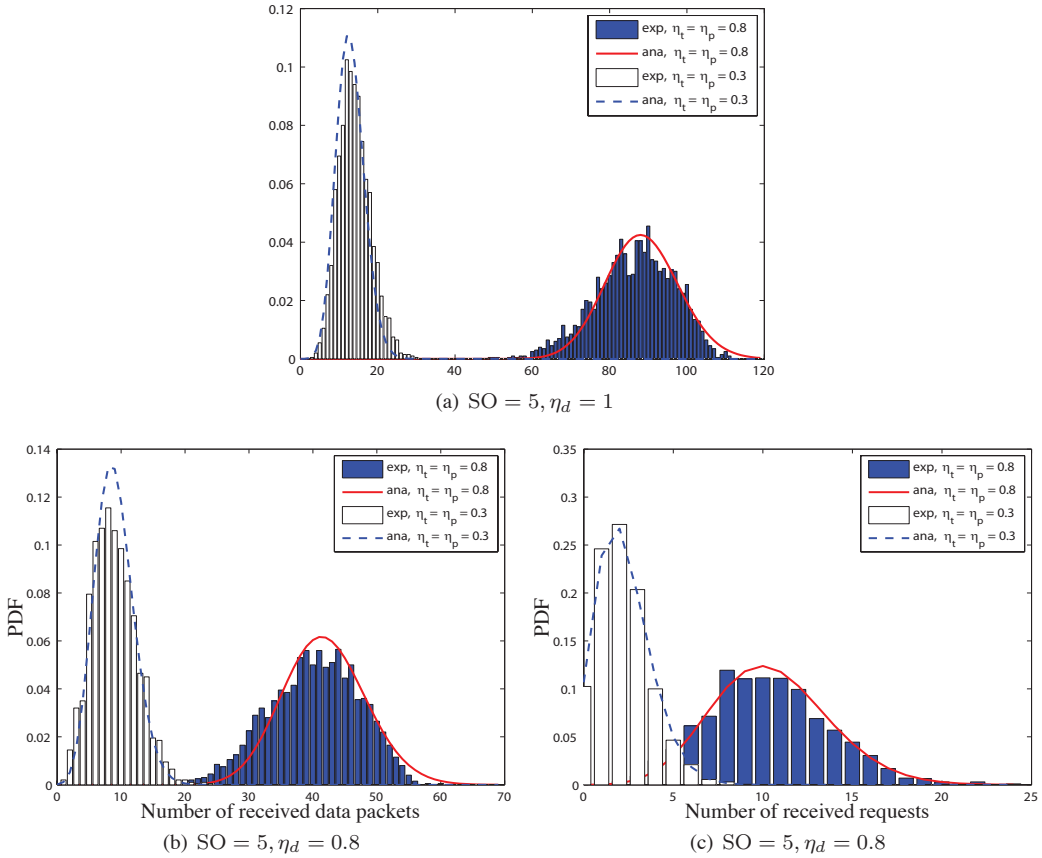
$$N_{CAP} = \frac{\eta_d N_{SD} T^\star_{CAP}}{T_{SD}}\,, \tag{29}$$
$$N_{CFP} = \frac{(1-\eta_d) N_{SD} T^\star_{CAP}}{T_{SD}}\,. \tag{30}$$

PROOF. By following Lemma 7.1 and Eq. (28), we compute the average number of successfully received non time-critical packets and GTS requests. Note that the oscillatory behavior predicted by Lemma 7.1 does not exist if $2\Delta_u \leq N_{\text{SS}}$. The IEEE 802.15.4 MAC fulfills the convergence condition since $\Delta_u \leq N_{\text{GTS}}$ and $N_{\text{SS}} = 16$. □

7.2. PDF of the number of received packets

We are now in the position to propose the PDF of the number of GTS requests per superframe, which has the input vector to analyze the GTS allocation mechanism of the CFP in Section 6. The challenge for the derivation of the analytical model is the stochastic behavior of all devices of the network, which makes it quite complicated and requires heavy computations to characterize the exact PDF of the number of GTS requests per superframe. In our earlier work [Park et al. 2009], it is shown that the exponential distribution is a good approximation of the packet delay distribution. Since the inter-arrival time of packet transmission can be approximated as exponentially distributed, the Poisson distribution with the mean value given in Eq. (26) approximates the arrival process of GTS requests in the CAP.

We validate the PDF of the number of received packets per superframe by using the Poisson approximation. Fig. 7 shows the PDF of the number of received data packets and requests per superframe as obtained by experiments and the Poisson distribution with the mean value given in Eqs. (29) and (30) as a function of the probabilities for generating non time-critical data packets with $\eta_d = 0.8, 1$, and traffic load with $\eta_t = \eta_p = 0.3, 0.8$ with the superframe order $\text{SO} = 5$, MAC parameters $m_0 = 3, m_b = 8, m = 4, n = 1$, number of time-critical data packets $\varphi_n = 2$ for each GTS request and packet length $L_p = 7$. The IEEE 802.15.4 MAC protocol is implemented



(a) SO $= 5, \eta_d = 1$

(b) SO $= 5, \eta_d = 0.8$

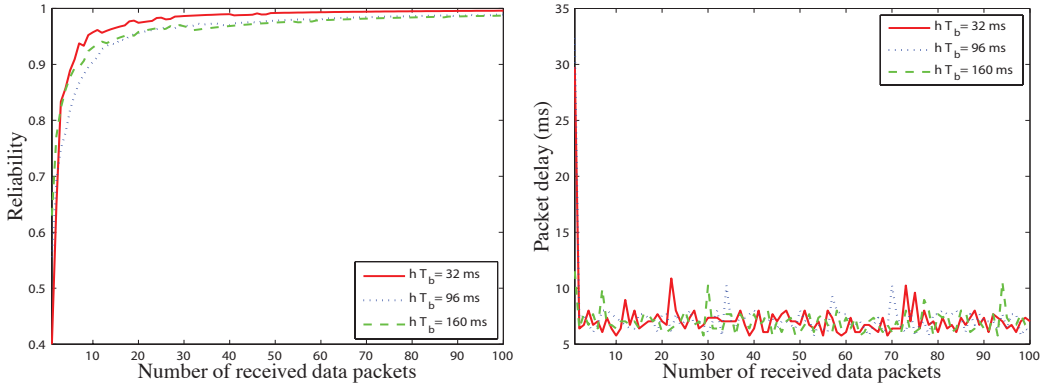
(c) SO $= 5, \eta_d = 0.8$

Fig. 7. Probability density function of the number of received data packets and requests per superframe as a function of the different probabilities for generating a non time-critical data packet $\eta_d = 0.8, 1$, and traffic load $\eta_t = \eta_p = 0.3, 0.8$ given a superframe order SO $= 5$, the MAC parameters $m_0 = 3, m_b = 8, m = 4, n = 1$, a number of time-critical data packets $\varphi_n = 2$ for each GTS request, and a packet length $L_p = 7$. The experiment runs for 96000×2^{10} time slots.

on a real test-bed using the IEEE 802.15.4-based TelosB platform [Polastre et al. 2005] and Contiki OS [Dunkels et al. 2004]. Fig. 7(a) reports the PDF of the number of received data packets for the CAP without having a GTS allocation mechanism. We observe that a Poisson distribution predicts well the PDF of the number of received data packets. The higher traffic load $\eta_t = \eta_p = 0.8$ gives a greater number of received data packets than lower traffic load $\eta_t = \eta_p = 0.3$. Figs. 7(b) and 7(c) show the PDF of the number of data packets and requests per superframe for the probability of generating a non time-critical data packet $\eta_d = 0.8$. Note that the mean number of data packets for the traffic load $\eta_t = \eta_p = 0.3, 0.8$ is closer than the case of the probability for generating a non time-critical data packet $\eta_d = 1$. This is due to the decreasing CAP length caused by the GTS allocation based on the number of received requests. If the GTS allocation mechanism is activated with $\eta_d < 1$, then the CAP length decreases since the GTS allocation takes a portion of the superframe resource.

7.3. Throughput of Hybrid MAC

Here we characterize the throughput of the hybrid MAC, namely, the average amount of both non time-critical packets and time-critical packets that can be transmitted during the beacon interval of



(a) Reliability (b) Packet delay

Fig. 8. Snapshot of reliability and packet delay as a function of traffic load $\eta_t = 0, \eta_p = 1$, the length of idle-queue $hT_b = 32, 96, 160$ ms, superframe order SO $= 10$ given a probability for generating a non-time critical data packet $\eta_d = 1$, and a packet length $L_p = 7$.

length T_{BI}. The normalized system throughput is defined as

$$\Theta = \frac{N_{\text{CAP}} L_{p,d} + \min(N_{\text{CFP}}, \Delta_u)\varphi_n L_{p,g}}{T_{\text{BI}}}, \tag{31}$$

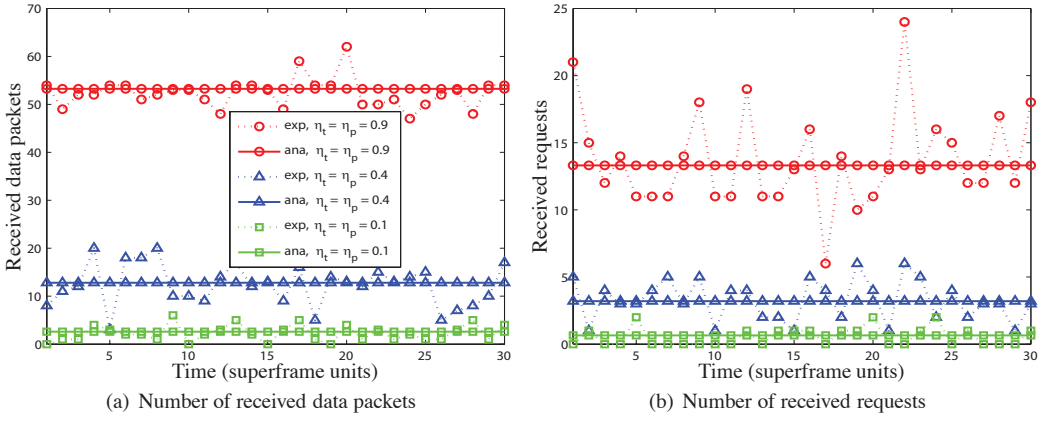
where an average number of successfully received non time-critical packets N_{CAP} and GTS requests N_{CFP} are given in Eqs. (29) and (30), respectively. The first and second terms of the nominator in Eq. (31) gives the average number of successful non time-critical packets of the CAP and time-critical packets of the CFP in a slot unit, respectively. The normalized system throughput Θ depends on the traffic pattern since an average number of successfully received non time-critical packets N_{CAP} and GTS requests N_{CAP} are related to the number of data packets φ_n, the frame size L_p, the mean and variance of requests.

8. PERFORMANCE EVALUATION

Here we present extensive experimental and simulation results of the hybrid IEEE 802.15.4 MAC to validate our analytical expressions in terms of the reliability given in Eq. (11), average packet delay given in Eq. (15), average queuing delay given in Eq. (25), and throughput given in Eq. (31).

We implement the IEEE 802.15.4 protocol to demonstrate the feasibility of the analytical model to understand the fundamental limitation of the protocol in a practical system. The IEEE 802.15.4 protocol was implemented on a test-bed using the TelosB platform [Polastre et al. 2005] running the Contiki OS [Dunkels et al. 2004] based on the specifications of the IEEE 802.15.4 [IEEE 2006]. TCP/IP communication is provided by the uIP stack [Dunkels et al. 2004]. The IEEE 802.15.4 defines one backoff as 20 symbols that correspond to 320 μs for 2.45 GHz. Since the hardware timer available for TelosB is based on a 32768 Hz clock, we use a backoff with duration of 305 μs instead of the 320 μs. The implementation is available for download [Qin and Park 2011]. We consider a typical indoor environment with concrete walls. $5, 10, 15, 25$ devices are placed to mimic a star topology. Each device is at a distance of around 3 m from the coordinator. Network devices generate packets with traffic load equal to $\eta_t = \eta_p = 0.1, \ldots, 1$ and send packets to the coordinator.

We first investigate the fundamental limitation of the IEEE 802.15.4 protocol by assuming perfect channel and carrier sensing capabilities. We present the transient behavior of the protocol to show the convergence of the network behavior. Then, we focus on the stationary behavior by using the experimental results. The experiments include several values of the traffic load, the probability for generating a non time-critical packet, and superframe length. Note that we fix the beacon order



(a) Number of received data packets

(b) Number of received requests

(c) Number of allocated GTSs

Fig. 9. Experimental snapshot of the number of received data packets, received requests, and allocated GTSs as a function of traffic load $\eta_t = \eta_p = 0.1, 0.4, 0.9$ given a probability for generating a non-time critical data packet $\eta_d = 0.8$, a superframe order $\mathrm{SO} = 5$, a number of time critical data packets $\tau_n = 2$ for each GTS request, and a packet length $L_p = 7$.

$\mathrm{BO} = 10$ and MAC parameters $m_0 = 3, m_b = 8, m = 4, n = 1$ in the experiments. Furthermore, we discuss the effect of imperfect channel and carrier sensing capabilities by using simulations.

8.1. Transient behavior

In this subsection, we investigate the transient behavior of the protocol. In particular, we first present the convergence of the network behavior in terms of reliability and packet delay of CAP without having the GTS allocation mechanisms for a given length of the superframe. Then, we describe the transient behavior of the hybrid MAC by considering both CAP and CFP.

Fig. 8 shows the convergence of the network behavior in terms of reliability and packet delay of CAP as a function of the length of idle-queue $hT_b = 32, 96, 160$ ms, superframe order $\mathrm{SO} = 10$ with the packet length $L_p = 7$ for $\eta_t = 0, \eta_p = 1$, which represents a similar traffic pattern with periodic packet generation time. The network contention level of the periodic traffic model significantly depends on the initial packet generation time. Hence, the initial traffic generation time is a critical factor to analyze the different performance indicators such as reliability, delay, and throughput. To investigate the effect of the traffic model, in the experiment all devices synchronize the packet generation time to maximize the contention in the CSMA/CA algorithm of CAP at the initial time. We recall that the traffic model of the network is described in Section 4. The packet generation time adopted in this paper varies depending on the backoff time of CSMA/CA and retransmission

time, which is different from the typical periodic traffic pattern commonly assumed in WSNs. The traffic model explicitly considers the contention level of the network. The basic idea is that as the contention level increases, the packet generation time increases so that the collisions are reduced. On the other hand, as the contention level decreases, the packet generation time decreases to increase the throughput of the network. This simple model allows us to make a statistical analysis of the performance of the network. Notice that we discuss the motivation and possible applications of traffic generation model in Section 9.

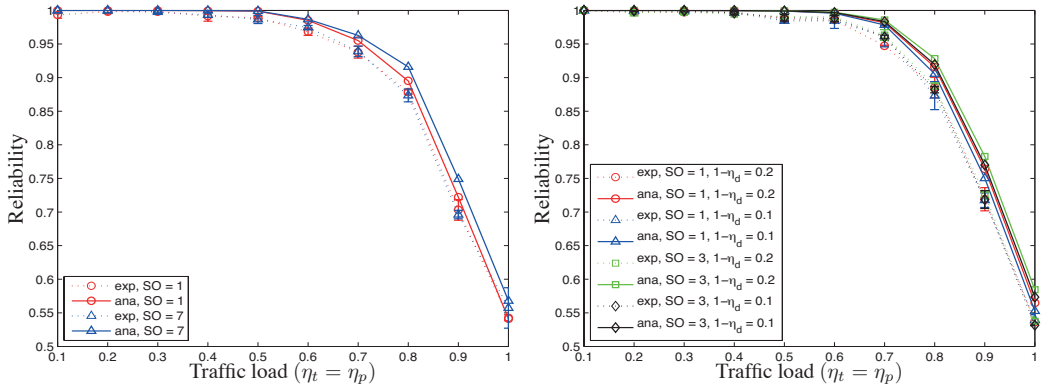
In Figs. 8(a) and 8(b), we observe a high packet loss and a high packet delay during the initial time due to the synchronized packet generation. However, this congestion is resolved after a few number of packet communications, as shown by a high reliability and low packet delay. In other words, all devices asynchronize the packet generation time to reduce the contention even though the packet generation time is synchronized at the initial time. It is interesting to observe that the reliability and packet delay are similar for different interval of packet generation by resolving the contention of the initial time. We remark that as the length of idle-queue increases, the proposed traffic model is similar to the periodic traffic model. The reason is that the length of idle-queue becomes dominant for packet generation interval rather than the packet delay.

Fig. 9 shows the analytical model of Proposition 7.2 and experimental snapshots of the number of received data packets, received requests, and allocated GTSs as a function of traffic load $\eta_t = \eta_p = 0.1, 0.4, 0.9$ given a probability for generating a non-time critical data packet $\eta_d = 0.8$, a superframe order $\mathrm{SO} = 5$, a number of time critical data packets $\tau_n = 2$ for each GTS request, and a packet length $L_p = 7$. The analytical model predicts well the experimental results. The number of allocated GTSs of Fig. 9(c) follows well the number of received requests of Fig. 9(b) for the traffic load $\eta_t = \eta_p = 0.4$ due to the GTS allocation mechanism based on received requests. In Fig. 9(a) and 9(b), the number of received data packets and requests decreases at the 2th superframe. Note that the length of CAP decreases due to the allocated GTSs after receiving the GTS requests, i.e., there are no allocated GTSs at the beginning of the network. In Fig. 9(b), a large number of requests are received at the 22th superframe for the traffic load $\eta_t = \eta_p = 0.4$, then the coordinator allocates the available number of GTSs, which is 6, of the superframe in Fig. 9(c). We observe that the number of data packets decreases since the length of CAP decreases due to the number of allocated GTSs in Fig. 9(a). Furthermore, we remark that the saturation condition of the portion of CAP and CFP of Lemma 7.1 is observed for the traffic load $\eta_t = \eta_p = 0.9$ due to the greater number of received requests. We note that although the number of received data packets and requests varies, the average value follows well the analytical model based on the convergence condition of Lemma 7.1.

8.2. Reliability

Fig. 10(a) compares the reliability of the CAP without having GTS allocations $\eta_d = 1$ of our proposed model and experimental results as a function of traffic load $\eta_t = \eta_p = 0.1, \ldots, 1$ and superframe order $\mathrm{SO} = 1, 3, 7$ with the packet length $L_p = 7$. The vertical bars indicate the standard deviation as obtained out of 5 experimental runs of 48000×2^{10} time slots each. The analytical model of the reliability as obtained by Eq. (11) follows well the experimental results. We observe that the smaller SO the worse is the reliability. The reason is that a small SO increases the event of deferred attempts due to the lack of remaining slot of the CAP, and then it increases the contention at the beginning of superframe. We remark that the effect of SO on the reliability of the CAP is negligible due to the extra backoff mechanism for the event of deferred attempts of the standard. In [Misic and Misic 2005; Misic et al. 2006; Jung et al. 2009], authors observe a high contention and high packet loss at the beginning of superframes since they do not consider the extra backoff mechanism for the event of deferred attempts which is explicitly described in the standard [IEEE 2006]. We remark that the number of deferred attempts increases as the traffic load increases and SO decreases. Therefore, the extra backoff mechanism for the event of deferred attempts at the beginning of superframe significantly improves the reliability of the CAP.

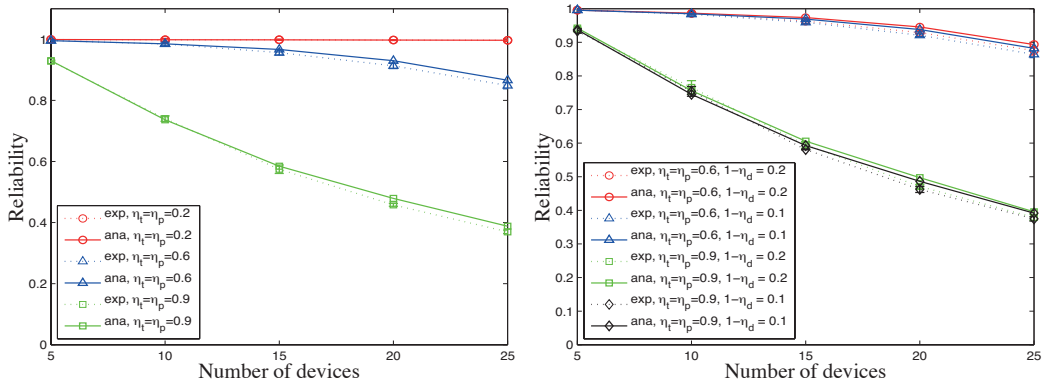
Fig. 10(b) shows the reliability of hybrid MAC with the CFP $\eta_d < 1$ as a function of traffic load, the probabilities for generating a time-critical data packet $1 - \eta_d = 0.1, 0.2$, and superframe



(a) Superframe order SO $= 1, 3, 7$ given a probability for generating a non time-critical data packet $\eta_d = 1$ and a packet length $L_p = 7$.

(b) Superframe order SO $= 1, 3$ and the probabilities for generating a time critical data packet $1 - \eta_d = 0.1, 0.2$, given a number of time-critical data packets $\varphi_n = 2$ for each GTS request and packet length $L_p = 7$.

Fig. 10. Reliability as a function of traffic load $\eta_t = \eta_p = 0.1, \ldots, 1$. The vertical bars indicate the standard deviation as obtained out of 5 experimental runs of 48000×2^{10} time slots each.



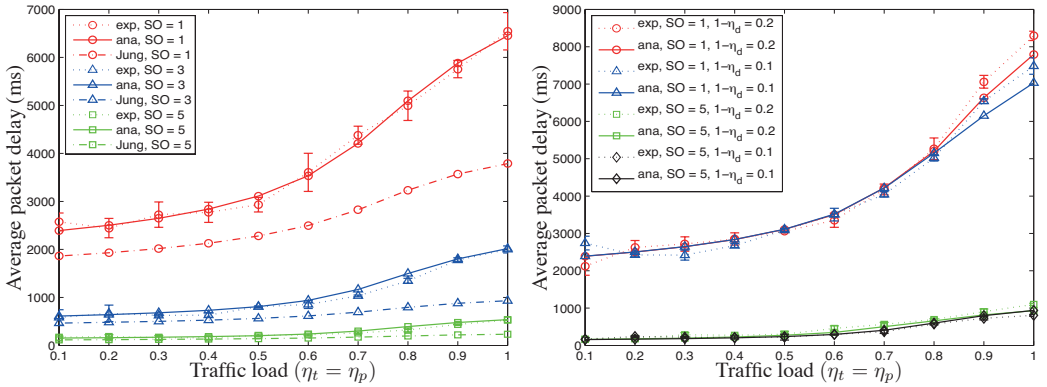
(a) Traffic load $\eta_t = \eta_p = 0.2, 0.6, 0.9$ given Superframe order SO $= 5$ and a probability for generating a non time-critical data packet $\eta_d = 1$ and a packet length $L_p = 7$.

(b) Traffic load $\eta_t = \eta_p = 0.6, 0.9$ and the probabilities for generating a time critical data packet $1 - \eta_d = 0.1, 0.2$, given superframe order SO $= 5$ and a number of time-critical data packets $\varphi_n = 2$ for each GTS request and packet length $L_p = 7$.

Fig. 11. Reliability as a function of the number of devices $N = 5, \ldots, 25$.

order SO $= 1, 3$ with a number of time-critical data packets $\varphi_n = 2$ for each GTS request and packet length $L_p = 7$. The analytical model predicts well the experimental results. Since the CFP length increases as the probability for generating time-critical packet η_d increases, the CAP length decreases for a fixed superframe order. As we observe in Fig. 10(a), the reliability decreases as the CAP length decreases due to the deferred attempts of packet transmission. Although the CAP length is even more decreasing as the probability for generating a time-critical packet $1 - \eta_d$ increases, the effect of different probabilities $1 - \eta_d$ is negligible.

Fig. 11(a) reports the reliability of the CAP without having GTS allocations $\eta_d = 1$ and experimental results as a function of the number of devices $N = 5, \ldots, 25$, traffic load $\eta_t = \eta_p = 0.2, 0.6, 0.9$ and superframe order SO $= 5$ with the packet length $L_p = 7$. The analytical model predicts well the experimental results. As the number of devices increases, the reliability decreases



(a) Superframe order SO $= 1, 3, 5$ given a probability for generating a non time-critical data packet $\eta_d = 1$ and a packet length $L_p = 7$.

(b) Superframe order SO $= 1, 5$ and the probabilities for generating a time-critical data packet $1 - \eta_d = 0.1, 0.2$, with a number of time-critical data packets $\varphi_n = 2$ for each GTS request and packet length $L_p = 7$.

Fig. 12. Average packet delay as a function of traffic load $\eta_t = \eta_p = 0.1, \ldots, 1$. Note that "Jung" refers to the average delay expression derived from Markov chain model in [Jung et al. 2009]. The reason that the model in [Jung et al. 2009] is not accurate is that it does not explicitly consider the deferred attempt of the backoff state.

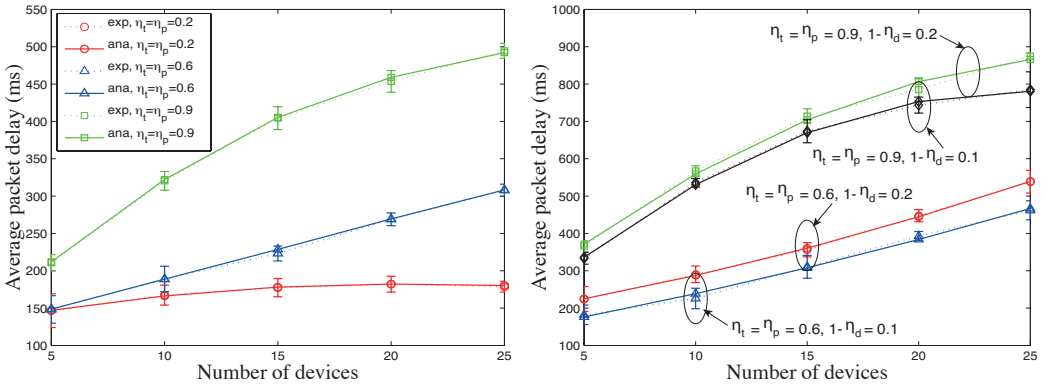
due to the high contention to send a packet. We validate the analytical model also for a large-scale network $N = 50, \ldots, 100$ in Appendix A.6.

Fig. 11(b) shows the reliability of hybrid MAC with the CFP as a function of the number of devices $N = 5, \ldots, 25$, traffic load $\eta_t = \eta_p = 0.2, 0.6, 0.9$, the probabilities for generating a time-critical data packet $1 - \eta_d = 0.1, 0.2$, and superframe order SO $= 5$ with a number of time-critical data packets $\varphi_n = 2$ for each GTS request and packet length $L_p = 7$. As the probability for generating time-critical packet $1 - \eta_d$ increases, the reliability increases for given traffic load and superframe order. As we observe in Figs. 10(a) and 10(b), the reliability decreases as the probability of deferred attempts of packet transmission increases. As the length of packets increases, the probability of deferred attempts increases in Eq. (6). Because the length of GTS requests $L_{p,r}$ is less than $L_{p,d}$, the probability of deferred attempts decreases as the probability for generating time-critical packet $1 - \eta_d$ increases.

8.3. Delay

In this section, we study the average packet delay of the CAP and the average queuing delay of the GTS allocation mechanism of the CFP. Recall that the measured packet delay of the experiments with a backoff time unit $305\ \mu s$ is tuned by considering the theoretically backoff time corresponding to $320\ \mu s$.

Fig. 12(a) compares the average packet delay of data packets of the pure CAP without having GTS allocations $\eta_d = 1$ given in Eq. (15), the analytical model derived from [Jung et al. 2009], and the experimental results as a function of traffic load $\eta_t = \eta_p = 0.1, \ldots, 1$ and superframe order SO $= 1, 3, 5$ with the packet length $L_p = 7$. Since in [Misic and Misic 2005; Misic et al. 2006; Jung et al. 2009] there is not direct investigation of the average packet delay with superframe structure, we derive it from the Markov chain model in [Jung et al. 2009]. Our analytical model of the average packet delay follows well the experimental results. The average packet delay increases as the traffic load increases due to the high contention. Furthermore, the average packet delay increases as the parameter SO decreases. This is due to that as the parameter SO decreases, the inactive period of the fixed superframe length $BO = 10$ increases. Note that if BO $=$ SO, then the behavior of the hybrid MAC is very similar to CSMA/CA algorithm of the CAP, see details in [Park et al. 2009]. The average packet delay derived from [Jung et al. 2009] does not predict well the experimental results. The main reason is that [Misic and Misic 2005; Misic et al. 2006; Jung et al. 2009; Buratti



(a) Traffic load $\eta_t = \eta_p = 0.2, 0.6, 0.9$ given superframe order $\mathrm{SO} = 5$ and a probability for generating a non time-critical data packet $\eta_d = 1$ and a packet length $L_p = 7$.

(b) Traffic load $\eta_t = \eta_p = 0.6, 0.9$ and the probabilities for generating a time-critical data packet $1 - \eta_d = 0.1, 0.2$, with superframe order $\mathrm{SO} = 5$ and a number of time-critical data packets $\varphi_n = 2$ for each GTS request and packet length $L_p = 7$.

Fig. 13. Average packet delay as a function of the number of devices $N = 5, \ldots, 25$.

2010] do not explicitly consider the deferred attempt of the backoff state. We remark that the deferred attempt of the backoff state becomes a critical factor as decreasing the parameter SO because more packets are delayed for the inactive period before actual packet transmission. Furthermore, the backoff mechanism for the deferred attempt is not considered as we discussed in Section 8.2.

Fig. 12(b) shows the average packet delay of the hybrid MAC with the CFP as a function of traffic load, the probabilities for generating a time-critical data packet $1 - \eta_d = 0.1, 0.2$, and superframe order $\mathrm{SO} = 1, 5$ with a number of time-critical data packets $\varphi_n = 2$ for each GTS request and packet length $L_p = 7$. We observe a similar behavior of the average packet delay of the hybrid MAC as the delay of the pure CAP reported in Fig. 12(a). The average packet delay increases as the probability for generating a non time-critical data packet increases. This is due to that as the GTS requests increases, the portion of the CFP increases for the fixed superframe length. Therefore, the deferred packet delay increases since a device waits more time for the inactive period and the CFP. By a similar argument, we see that the average packet delay of hybrid MAC in Fig. 12(b) is higher than the delay of the pure CAP in Fig. 12(a). Note the scale of the axis.

Fig. 13(a) reports the analytical model of the average packet delay of the pure CAP and the experimental results as a function of the number of devices $N = 5, \ldots, 25$, traffic load $\eta_t = \eta_p = 0.2, 0.6, 0.9$ and superframe order $\mathrm{SO} = 5$ with the packet length $L_p = 7$. The analytical model of the average packet delay predicts well the experimental results. As the number of devices increases, the average packet delay increases due to the high busy channel and collision probabilities.

Fig. 13(b) shows the average packet delay of the hybrid MAC with the CFP as a function of the number of devices, traffic load $\eta_t = \eta_p = 0.6, 0.9$, the probabilities for generating a time-critical data packet $1 - \eta_d = 0.1, 0.2$, and superframe order $\mathrm{SO} = 5$ with a number of time-critical data packets $\varphi_n = 2$ for each GTS request, and the packet length $L_p = 7$. Similar to Fig. 12(b), the average packet delay increases as the probability for generating a time-critical data packet increases because the portion of the CFP increases for the fixed superframe length. It is interesting to observe that the difference of the average packet delay for the lower traffic load $\eta_t = \eta_p = 0.6$ is almost constant for different probabilities for generating a time-critical data packet $1 - \eta_d = 0.1, 0.2$.

Next, we discuss the queuing delay of the GTS allocation mechanism. Fig. 14 shows the average queuing delay of the CFP as a function of traffic load, the probabilities for generating a time-critical data packet $1 - \eta_d = 0.1, 0.2$, and superframe order $\mathrm{SO} = 5, 7, 9$ with a number of time-critical data packets $\varphi_n = 2$ for each GTS request and packet length $L_p = 7$. The analytical model of the

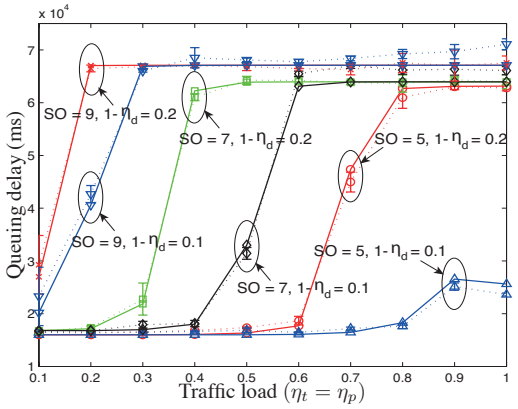


Fig. 14. Average queuing delay of GTS request as a function of traffic load $\eta_t = \eta_p = 0.1, \ldots, 1$, the probabilities for generating a time-critical data packet $1 - \eta_d = 0.1, 0.2$, and superframe order $\mathrm{SO} = 5, 7, 9$ with a number of time-critical data packets $\varphi_n = 2$ for each GTS request and packet length $L_p = 7$.

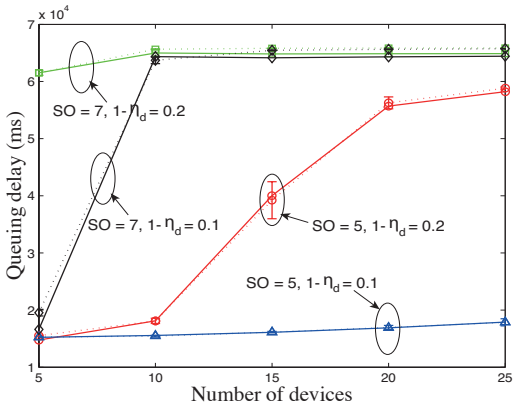


Fig. 15. Average queuing delay of GTS request as a function of the number of devices $N = 5, \ldots, 25$, traffic load $\eta_t = \eta_p = 0.6$, the probabilities for generating a time-critical data packet $1 - \eta_d = 0.1, 0.2$, and superframe order $\mathrm{SO} = 5, 7$ with a number of time-critical data packets $\varphi_n = 2$ for each GTS request and packet length $L_p = 7$.

queuing delay given by Eq. (25) predicts well the experimental results. As the traffic load $\eta_t = \eta_p$ and the probability for generating a time-critical data packet $1 - \eta_d$ increase, the queuing delay increases due to a higher number of received GTS requests at the coordinator. By a similar argument, as the parameter SO increases, the queuing delay increases due to a higher number of received GTS requests. Observe that the queuing delay is saturated when the number of received GTS requests increases for higher traffic load and higher superframe order because of the maximum number of GTSs to be allocated to devices, see details in Section 6.1.

Fig. 15 reports the average queuing delay of the CFP as a function of the number of devices $N = 5, \ldots, 25$, traffic load $\eta_t = \eta_p = 0.6$, the probabilities for generating a time-critical data packet $1 - \eta_d = 0.1, 0.2$, and superframe order $\mathrm{SO} = 5, 7$ with a number of time-critical data packets $\varphi_n = 2$ for each GTS request and packet length $L_p = 7$. By a similar argument of Fig. 15, as the number of devices increases, the queuing delay increases due to a higher number of received GTS requests at the coordinator. Hence, the queuing delay of the GTS allocation mechanism is not only dependent on the traffic load, probability for generating a time-critical data packet, and superframe order but also the number of devices of the network.

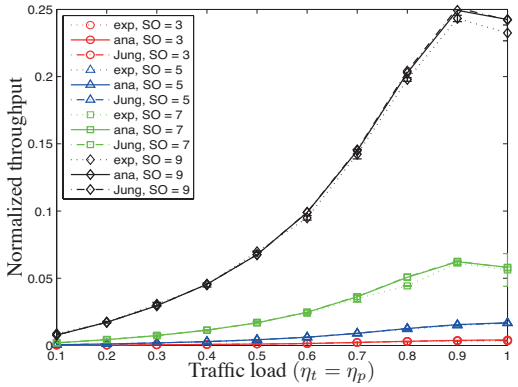


Fig. 16. Throughput as a function of traffic load $\eta_t = \eta_p = 0.1, \ldots, 1$ and superframe order $\mathrm{SO} = 3, 5, 7, 9$ with a probability for generating a non time-critical data packet $\eta_d = 1$ and packet length $L_p = 7$. We observe a similar results as the one in [Jung et al. 2009]. However, recall that our model is much more general because we consider the CFP.

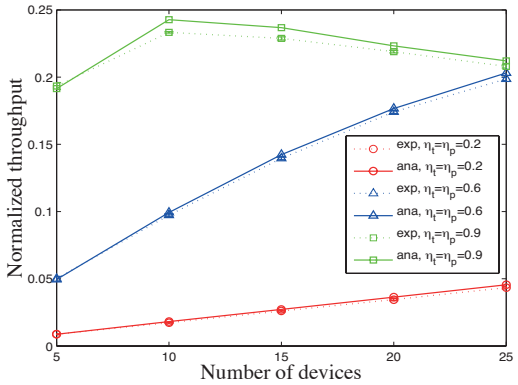
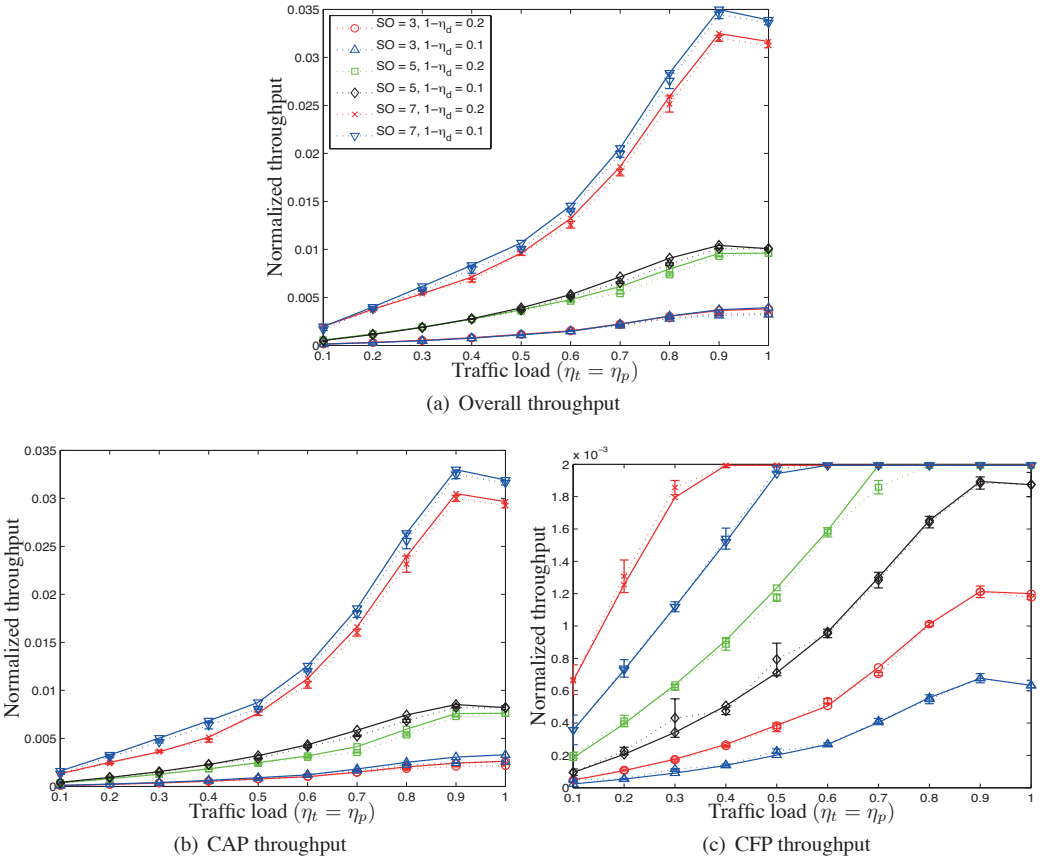


Fig. 17. Throughput as a function of the number of devices $N = 5, \ldots, 25$, traffic load $\eta_t = \eta_p = 0.2, 0.6, 0.9$ and superframe order $\mathrm{SO} = 9$ with a probability for generating a non time-critical data packet $\eta_d = 1$ and packet length $L_p = 7$.

8.4. Throughput

Fig. 16 compares the throughput of the pure CAP without having GTS allocation given in Eq. (31), the throughput given in [Jung et al. 2009], and experimental results as a function of traffic load and superframe order $\mathrm{SO} = 3, 5, 7, 9$ with the packet length $L_p = 7$. The analytical model of throughput predicts well the experimental results. The throughput of the CAP increases as the traffic load $\eta_t = \eta_p = 0.1, \ldots, 1$ increases. Furthermore, as the parameter SO increases, the throughput of the CAP increases due to the longer CAP length. As the traffic load decreases and SO increases, the throughput in [Jung et al. 2009] is very close to the throughput given in Eq. (31) because the number of deferred attempts decreases.

Fig. 17 reports the throughput of the pure CAP without having GTS allocation and experimental results as a function of a number of devices $N = 5, \ldots, 25$, traffic load $\eta_t = \eta_p = 0.2, 0.6, 0.9$, and superframe order $\mathrm{SO} = 9$ with the packet length $L_p = 7$. The throughput of the CAP increases as the number of devices $N = 5, \ldots, 25$ increases for the lower traffic load $\eta_t = \eta_p = 0.2, 0.6$. However, the throughput of the CAP decreases as the number of devices $N = 10, \ldots, 25$ increases for the higher traffic load $\eta_t = \eta_p = 0.9$. Furthermore, it is interesting to observe that the throughput of the CAP for the different traffic load $\eta_t = \eta_p = 0.6, 0.9$ is similar for given number of devices



(a) Overall throughput

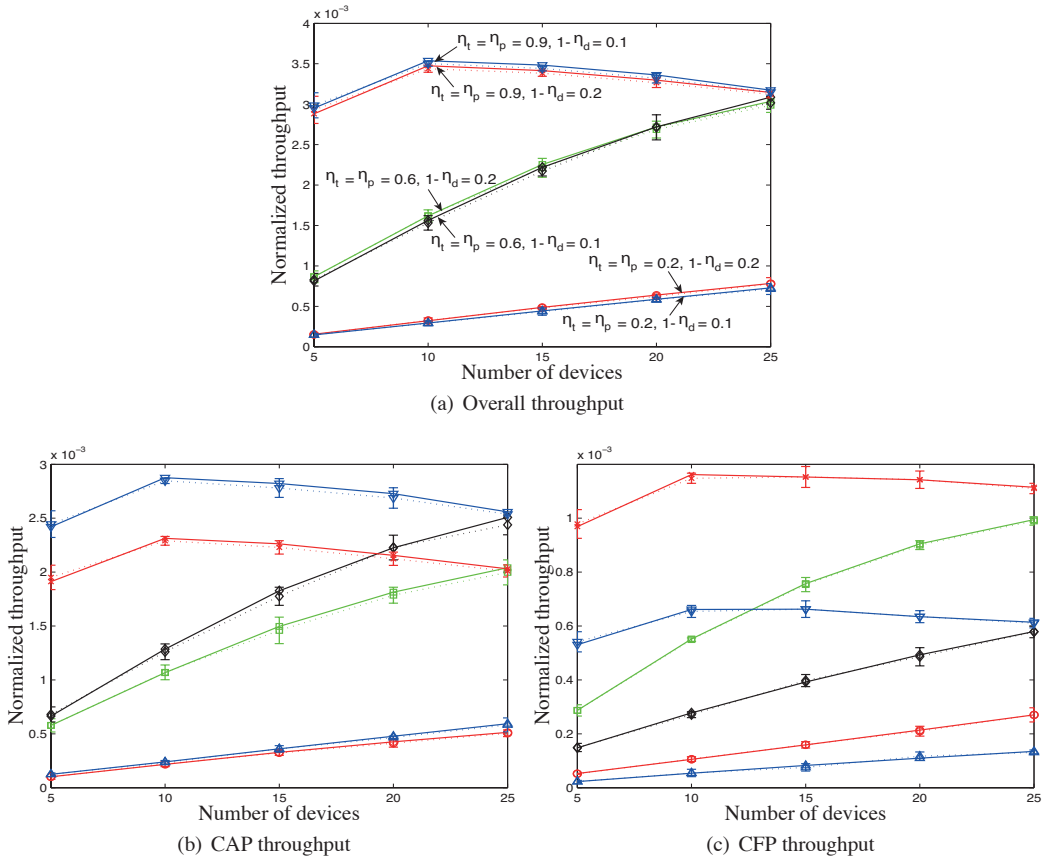
(b) CAP throughput

(c) CFP throughput

Fig. 18. Throughput as a function of traffic load $\eta_t = \eta_p = 0.1, \ldots, 1$, the probabilities for generating a time-critical data packet $1 - \eta_d = 0.1, 0.2$, and superframe order $\mathrm{SO} = 3, 5, 7$ with a number of time-critical data packets $\varphi_n = 2$ for each GTS request and packet length $L_p = 7$.

$N = 25$. We note that the the network is stable for the lower traffic load $\eta_t = \eta_p = 0.6$, whereas it is not for the higher traffic load $\eta_t = \eta_p = 0.9$.

Fig. 18 shows the throughput of hybrid MAC as a function of traffic load, the probabilities for generating a time-critical data packet $1 - \eta_d = 0.1, 0.2$, and superframe order $\mathrm{SO} = 3, 5, 7$ with a number of time-critical data packets $\varphi_n = 2$ for each GTS request and packet length $L_p = 7$. We remark that the model in [Jung et al. 2009] does not take into account the CFP of the superframe. Fig. 18(a) presents the overall throughput, namely the sum of throughput of the CAP in Fig. 18(b) and the CFP in Fig. 18(c). We observe a similar trend to the throughput of the CAP as in Fig. 16. The throughput of the CAP and CFP increases as the traffic load $\eta_t = \eta_p$ and SO increase. It is interesting to observe that as the probability for generating a time-critical data packet $1 - \eta_d$ increases, the throughput of CFP increases, however the throughput of the CAP decreases. Hence, there is a tradeoff between throughput of the CAP and CFP for a fixed superframe length. In our setup, the throughput of the CAP is more dominant for the overall throughput of the network. Main reason is that the throughput of the CFP is strictly limited by the maximum number of GTSs. For the lower parameter $\mathrm{SO} = 3, 5$, the overall throughput of different probabilities for generating a time-critical data packets is closer with respect to the throughput of the CAP because the throughput of the CFP compensates the lower throughput of the CAP. Therefore, the throughput of the hybrid MAC depends on the probability for generating time-critical packets.



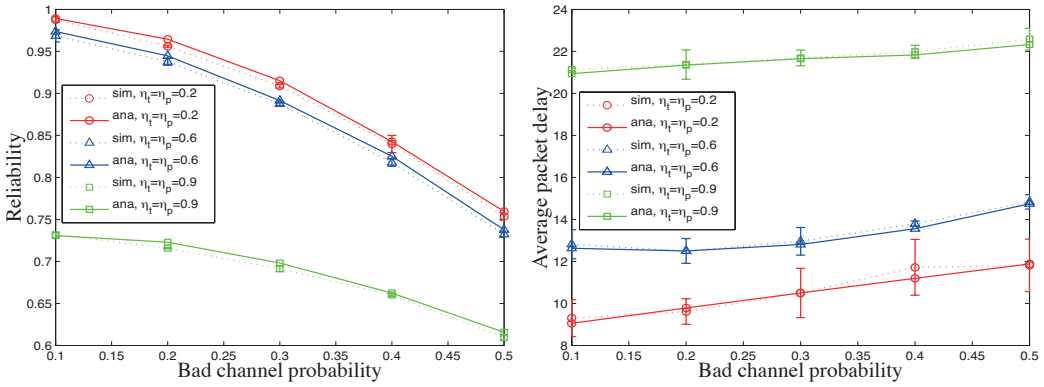
(a) Overall throughput

(b) CAP throughput

(c) CFP throughput

Fig. 19. Throughput as a function of the number of devices $N = 5, \ldots, 25$, traffic load $\eta_t = \eta_p = 0.2, 0.6, 0.9$, the probabilities for generating a time-critical data packet $1 - \eta_d = 0.1, 0.2$, and superframe order $\mathrm{SO} = 3$ with a number of time-critical data packets $\varphi_n = 2$ for each GTS request and packet length $L_p = 7$.

In Fig. 19, we show that the portion of CAP and CFP is critical to maximize the throughput of the hybrid MAC. Fig. 19 reports the throughput of hybrid MAC as a function of the number of devices $N = 5, \ldots, 25$, traffic load $\eta_t = \eta_p = 0.2, 0.6, 0.9$, the probabilities for generating a time-critical data packet $1 - \eta_d = 0.1, 0.2$, and superframe order $\mathrm{SO} = 3$ with a number of time-critical data packets $\varphi_n = 2$ for each GTS request and packet length $L_p = 7$. With a similar argument as in Fig. 18, we clearly observe a tradeoff between throughput of the CAP and CFP for a fixed superframe length. For the higher traffic load $\eta_t = \eta_p = 0.9$, the throughput of the CAP is dominant for the overall throughput. In other words, it is better to decrease the probability for generating a time-critical data packet $1 - \eta_d$ to maximize the overall throughput for the higher traffic load. However, by comparing Figs. 18(a) and 18(b) for the lower traffic load $\eta_t = \eta_p = 0.2$, we clearly observe that the compensation of the CAP becomes dominant for the overall throughput. The overall throughput increases as the probability for generating a time-critical data packet $1 - \eta_d$ increases for the lower traffic load $\eta_t = \eta_p = 0.2$. For the traffic load $\eta_t = \eta_p = 0.6$, we remark that the throughput of the CAP is dominant for the greater number of devices $N = 20, 25$, however, the throughput of the CAP becomes dominant for the less number of devices $N = 5, \ldots, 15$. Therefore, the optimal portion of the CFP to maximize the overall throughput depends on the traffic load of the network. We note that this observation of the IEEE 802.15.4 protocol is not general because the TDMA mechanism is in general better to maximize the throughput for the higher traffic load.



(a) Reliability (b) Average packet delay

Fig. 20. Effect of imperfect channel as a function of loss probabilities $p = 0.1, \ldots, 0.5$, traffic load $\eta_t = \eta_p = 0.2, 0.6, 0.9$ given beacon order BO $= 8$, superframe order SO $= 7$, number of devices $N = 10$, probability for generating a non time-critical data packet $\eta_d = 1$ and packet length $L_p = 7$.

It is because of the limited usability of the GTS allocation mechanism as described in Sections 2.2 and 6.2.

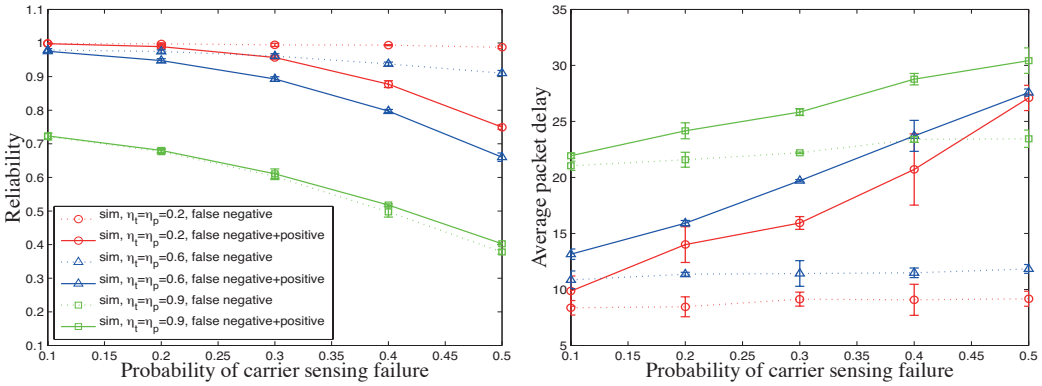
8.5. Effect of imperfect channel and carrier sensing

Even though the assumption of the perfect wireless channel and carrier sensing plays a critical role to understand the fundamental limitations of the IEEE 802.15.4 protocol, wireless channel and carrier sensing in practice are not perfect. In this section, we first analyze the effect of imperfect wireless channel and then discuss the system performance with the imperfect carrier sensing through simulations. The main reason to use the simulation is that it is difficult to characterize the wireless channel and carrier sensing failure in real world experiments. We compare the computation time of the analytical model and simulation in Appendix A.5.

Fig. 20 compares the reliability and average packet delay of the CAP without having GTS allocations $\eta_d = 1$ of our proposed model and simulation results as a function of the probabilities of the bad channel $p = 0.1, \ldots, 0.5$, traffic load $\eta_t = \eta_p = 0.2, 0.6, 0.9$, beacon order BO $= 8$, superframe order SO $= 7$ with the packet length $L_p = 7$. We use the COOJA simulator [Osterlind et al. 2006] to analyze the effect of the imperfect wireless channel modeled by a Bernoulli random process. The analytical model of the reliability as obtained by Eq. (11) predicts well the simulation results. The reliability decreases as the probability of bad channel increases due to the greater number of retransmissions. Note that each device retransmits a packet if an ACK is not received within the maximum number of retries. For the same reason, the average packet delay increases as the probability of the bad channel increases.

Now, we analyze the performance of CSMA/CA algorithm in the presence of carrier sensing errors. Two types of carrier sensing errors, i.e., false negative and false positive, are considered, and their impact on the system performance is analyzed using simulation results. A false negative failure is that carrier sensing incorrectly detects that the medium is idle when it is actually busy. A false positive event occurs when a busy state is reported when the medium is idle. We model carrier sensing failures by independent Bernoulli trials with success probability $1 - \gamma$ where $0 < \gamma < 1$. The independence of the trial results is assumed to be over all links and time slots.

Fig. 21 presents the reliability and average packet delay of the CAP without having GTS allocations $\eta_d = 1$ of simulation results as a function of the probabilities of the carrier sensing error $\gamma = 0.1, \ldots, 0.5$, traffic load $\eta_t = \eta_p = 0.2, 0.6, 0.9$, beacon order BO $= 8$, superframe order SO $= 7$ with the packet length $L_p = 7$. Note that "false negative" and "false negative+positive" refer to the consideration of the false negative event and the combined event of the false nega-



(a) Reliability (b) Average packet delay

Fig. 21. Effect of imperfect carrier sensing as a function of false probabilities $\gamma = 0.1, \ldots, 0.5$, traffic load $\eta_t = \eta_p = 0.2, 0.6, 0.9$ given beacon order $\mathrm{BO} = 8$, superframe order $\mathrm{SO} = 7$, number of devices $N = 10$, probability for generating a non time-critical data packet $\eta_d = 1$ and packet length $L_p = 7$.

tive and positive failure for the simulation setup, respectively. The false negative and positive failure decreases the system performance in terms of the reliability and average packet delay. The false negative failure is critical because a device sends a packet even if the medium is actually busy. In [Kim et al. 2011], the authors show that the false positive failure is irrelevant to achievable throughput. However, we remark that this observation does not hold if the number of backoffs is limited and the traffic load is low. The false positive failure significantly decreases the performance in terms of the reliability and average packet delay for the lower traffic load $\eta_t = \eta_p = 0.2, 0.6$ due to the limited number of backoffs. The busy channel probability increases as the false positive failure increases. However, the false positive failure has negligible effects on the reliability of the higher traffic load $\eta_t = \eta_p = 0.9$. The false positive failures tend to defer the transmission and in fact are beneficial to compensate the effect of the false negative failures. Because false positive failures create extra backoffs, they increase the packet delay.

9. DISCUSSIONS

The traffic model of this paper is motivated by typical control applications. In [Park et al. 2010], it is shown how the IEEE 802.15.4 MAC protocol can be used for real-time feedback control systems. We consider multiple plants controlled over a WSN using the IEEE 802.15.4 protocol. N plants contend to transmit sensor measurements to the controller over a wireless network that induces packet losses and varying delays. Independent control applications asynchronously generate packets when their timers expire. When a node sends a packet successfully or discards a packet, it stays in an idle period for h seconds without generating new packets. We remark that controllers can be designed to optimize the sample period h [Araujo et al. 2011; Mazo and Tabuada 2011]. Furthermore, it is possible to assign different probabilities of traffic generation η_t, η_p. For instance, if $\eta_t = 1$, then it becomes saturated traffic, if $\eta_t = 0, \eta_p = 1$, then the traffic pattern is similar to the one of periodic traffic model. We believe that the packet generation probabilities η_t, η_p and the length of idle-queue h allows us to model different traffic patterns for different applications.

One of the interesting problems we highlighted is the high contention at the beginning of the superframe when many devices attempt to transmit packets after a long inactive period. This problem is similar to the batched arrivals described in [Bender et al. 2005]. Since a maximum number of backoffs and retries are strictly limited by the CSMA/CA algorithm, the reliability may significantly decrease at the beginning of the superframe, whereas the network may be underutilized for the remained time except this initial high contention. Hence, each device may delay the activation of the CSMA/CA algorithm at the beginning of the superframe even though it has a packet to trans-

mit. Each device may run an activation timer of the CSMA/CA algorithm to spread the contention equally during the superframe. Therefore, the traffic generation model proposed in this paper can be used for an activation timer of the CSMA/CA algorithm to guarantee a stable network performance.

10. CONCLUSIONS

In this paper, we developed an analytical framework for modelling the behavior of the MAC protocol of the IEEE 802.15.4 standard in the hybrid modality. Since the hybrid MAC provides a random access scheme of the CAP and a deterministic access scheme of the CFP in each superframe, the analysis of the mutual effects of these two schemes provides a fundamental step towards the understanding of the performance of the IEEE 802.15.4 MAC. We first model the contention access scheme of the MAC exponential backoff process through an extended Markov chain, which included channel, retry limits, ACKs, unsaturated traffic, and superframe period. With respect to the previous literature, a substantial novelty of our model is that it explicitly considered the deferred attempt of the backoff state and the extra backoff mechanism for the event of deferred attempts, which are crucial for an accurate analysis of the reliability and delay. Then, we analyzed the behavior of contention free scheme using a new Markov chain. To connect the two access modalities, the probability distribution function of the number of received GTS requests per superframe was used as the input parameter to analyze the performance of the CFP.

We evaluated the network performance in terms of reliability of the CAP, average packet delay of the CAP, queuing delay of the CFP, and throughput of the overall hybrid MAC. The network performance was evaluated through the transient behavior of reliability and packet delay of CAP and number received non-time critical packets, number of received requests, and number of allocated GTSs per superframe of hybrid MAC. We compared the analytical model to an experimental implementation, and thereby presented the first comparison of an analytical model of the IEEE 802.15.4 MAC to real experiments. The protocol was implemented on Contiki, and was tested on the IEEE 802.15.4-based TelosB platform. Experimental results confirmed that the proposed analysis offers a satisfactory accuracy. It was determined under which conditions the superframe structure of IEEE 802.15.4 is stable. Furthermore, we showed that a Poisson distribution approximates well the number of successfully received packets for a given CAP length. The mutual effect between throughput of the random access and the TDMA scheme for a fixed superframe length was shown to be critical to maximize the throughput of the hybrid MAC. We showed that the throughput of the random access mechanism is a dominant factor with respect to the one of GTS allocation mechanism in the high traffic load due to the strict limitation of the GTS allocation in the standard. Furthermore, it is shown that the effect of imperfect channel and carrier sensing failures on system performance heavily depends on the traffic load and limited range of the protocol parameters. The false positive failure affects not only the packet delay but also the reliability in the lower traffic load mainly due to the limited number of backoffs.

APPENDIX

Table II. Main symbols used in the paper. The italic names on the right side of the table are defined in the IEEE 802.15.4 standard, see [IEEE 2006] in the manuscript.

Symbol	Meaning
BO	Beacon order
SO	Superframe order
h	Length of idle-queue state without generating packets
$\mathbb{E}[D]$	Average delay for successfully received packets of CSMA/CA mechanism
$\mathbb{E}[D_q(t)]$	Average queuing delay of the GTS allocation mechanism at superframe t
$\mathbb{E}[\psi_s]$	Average service time for successfully received packet
$\mathbb{E}[\psi_{dc}]$	Average service time for discarded packet due to channel access failure
$\mathbb{E}[\psi_{dr}]$	Average service time for discarded packet due to retry limits
L_s	Successful packet transmission time with ACK
L_c	Packet collision time with ACK mechanism
L_g	Successful packet transmission time without ACK
L_p	Total packet length including overhead and payload
m	Maximum number of backoffs, *macMaxCSMABackoffs*
m_0	Minimum value of backoff exponent, *macMinBE*
m_b	Maximum value of backoff exponent, *macMaxBE*
n	Maximum number of retries, *macMaxFrameRetries*
N	Total number of devices of the network
N_{CAP}	Average number of successfully received non time-critical packets
N_{CFP}	Average number of successfully received GTS requests
N_{DPT}	Number of superframes in which a GTS descriptor exists, *aGTSDescPersistenceTime*
N_{GTS}	Maximum number of GTS descriptors
N_{SS}	Number of slots contained in a superframe
N_{SD}	Average number of successfully received any packets
p	Probability of bad channel
P_c	Packet retransmission probability
P_s	Packet delivery rate, reliability
P_{dc}	Probability that the packet is discarded due to channel access failure
P_{dr}	Probability of a packet being discarded due to retry limits
T_b	Basic time unit, *aUnitBackoffPeriod*
T_{BI}	Length of beacon interval
T_{SD}	Length of superframe duration
T_{CAP}	CAP length
T_{CFP}	CFP length
T_{SP}	Length of inactive period
T_{SS}	Length of superframe slot
T_0	Length of superframe when $\mathrm{SO} = 0$
$T_{\min}$	Minimum CAP length, *aMinCAPLength*
$T^{\star}_{\mathrm{CAP}}$	Fixed point of the CAP length
$T^{\star}_{\mathrm{CFP}}$	Fixed point of the CFP length
α	Busy channel probability of CCA_1
β	Busy channel probability of CCA_2
γ	Probability of carrier sensing failures
Δ_u	Maximum number of GTSs to be allocated to devices
η_t	Packet generation probability after the device sends a packet successfully or discard a packet
η_p	Packet generation probability when the sampling interval is expired
η_d	Probability to generate a non time-critical data packet
Θ	Normalized throughput of the network
$\underline{\theta}$	Minimum number of superframe slots for a singe GTS
λ_i	Probability of i successful requests during the CAP
π^t_k	Probability of state k at time t of Markov chain for the GTS allocation mechanism
ρ_t	Probability of deferred attempts due to the lack of the transmission time in a CAP
ρ_b	Probability of deferred attempts due to the lack of the backoff counter time in a CAP
τ	Channel sensing probability
φ_n	Number of time-critical data packets for each GTS request

A.1. Proof of Proposition 5.1

We follow two steps to compute the stationary probability. First, we derive the state transition probability of Markov chain. Second, a normalization condition is applied to compute the probability $S_{0,0,0,0}$.

First, the state transition probabilities associated to the Markov chain of Fig. 3 are

$$P(i,j,k,0|i,j+1,k,0) = 1-\rho_b, \text{ for } k \geq 0\,, \tag{32}$$
$$P(i,j,k,1|i,j+1,k,1) = 1, \text{ for } k \geq 0\,, \tag{33}$$
$$P(i,j,k,0|i-1,0,k,0) = \frac{(1-\rho_t)(\alpha+(1-\alpha)\beta)}{W_i}, \text{ for } i \leq m\,, \tag{34}$$
$$P(0,j,k,0|i,0,k-1,0) = \frac{P_c(1-\rho_t)(1-\alpha)(1-\beta)}{W_0}, \text{ for } k \leq n\,, \tag{35}$$
$$P(0,j,k,1|i,0,k-1,1) = \frac{P_c(1-\alpha)(1-\beta)}{W_0}, \text{ for } k \leq n\,, \tag{36}$$
$$P(Q_0|m,0,k,0) = (1-\eta_t)(1-\rho_t)(\alpha+(1-\alpha)\beta), \text{ for } k < n\,, \tag{37}$$
$$P(Q_0|m,0,k,1) = (1-\eta_t)(\alpha+(1-\alpha)\beta), \text{ for } k < n\,, \tag{38}$$
$$P(Q_0|i,0,n,0) = (1-\eta_t)(1-\rho_t)(1-\alpha)(1-\beta), \text{ for } i < m\,, \tag{39}$$
$$P(Q_0|i,0,n,1) = (1-\eta_t)(1-\alpha)(1-\beta), \text{ for } i < m\,, \tag{40}$$
$$P(0,j,0,0|Q_0) = \frac{\eta_p}{W_0}, \text{ for } j \leq W_0 - 1\,. \tag{41}$$

Eq. (32) corresponds to the decrement of backoff counter of non-deferred transmission, which happens with probability $1-\rho_b$. Eq. (33) models the decrement of backoff counter after the event of deferred attempts, which happens with probability 1. Recall that we assume that the maximum number of deferred attempts of a single packet transmission is 1 in Section 4. Eq. (34) represents the probability of selecting uniformly a state in the next backoff stage of non-deferred transmission after finding busy channel in CCA_1 or CCA_2. Eqs. (35) and (36) give the probability of picking uniformly a state in the next retransmission stage after the packet collision or bad channel of non-deferred and deferred transmission, respectively. Eqs. (37) and (38) represent the probability of going back to the idle-queue stage due to the channel access failure of non-deferred and deferred transmission, respectively. Eqs. (39) and (40) give the probability of going back to the idle-queue stage due to the retry limits of non-deferred and deferred transmission, respectively. Finally, Eq. (41) models the probability of going back to the first backoff stage from the idle-queue stage.

Next, we derive the closed form expression for the chain of Fig. 3. Let us first consider the stationary probability of non-deferred transmission $l = 0$. Owing to the chain regularities and Eqs. (32) and (34), we have $S_{i,0,k,0} = y_i\, S_{i-1,0,k,0}$ where

$$y_i = \frac{(1-\rho_t)(\alpha+(1-\alpha)\beta)}{W_i} \sum_{j=0}^{W_i-1} (1-\rho_b)^j\,, \qquad i \geq 1\,,$$

and $y_0 = 1$. By using the product of y_i, the stationary probability $S_{i,0,k,0}$ is rewritten as follows

$$S_{i,0,k,0} = \xi_i\, S_{0,0,k,0}\,, \tag{42}$$

where $\xi_i = \prod_{r=1}^{i} y_r$ and $\xi_0 = 1$. From Eqs. (32) and (34), we have

$$S_{i,j,k,0} = \frac{1}{W_i}\left(1 + \sum_{r=j+1}^{W_i-1} (1-\rho_b)^{W_i-r}\right)(1-\rho_t)x\, S_{i-1,0,k,0}\,,$$

where $x = \alpha + (1-\alpha)\beta$. Then, we obtain

$$S_{i,j,k,0} = \varpi_{i,j} S_{i,0,k,0}\,, \tag{43}$$

where $\varpi_{i,j} = \frac{1-(1-\rho_b)^{W_i-j}}{1-(1-\rho_b)^{W_i}}$. From Eqs. (35) and (42), we have

$$S_{i,0,k,0} = P_c(1-\rho_t)(1-\alpha)(1-\beta)\sum_{i=0}^{m}\sum_{j=0}^{W_0-1}\frac{(1-\rho_b)^j}{W_0} S_{i,0,k-1,0} = g^k S_{0,0,0,0}\,, \tag{44}$$

where $g = P_c(1-\rho_t)(1-\alpha)(1-\beta)\sum_{i=0}^{m}\xi_i\sum_{j=0}^{W_0-1}\frac{(1-\rho_b)^j}{W_0}$. Now, we characterize the stationary probability of the deferred transmission $l = 1$. From Eqs. (36) and (42)–(44), we derive

$$\begin{aligned}
S_{i,0,k,1} =&(\alpha + (1-\alpha)\beta) S_{i-1,0,k,1} u(i-1) + \rho_t S_{i,0,k,0} + \rho_b \sum_{j=1}^{W_i-1} S_{i,j,k,0} \\
&+ P_c(1-\alpha)(1-\beta)\sum_{i=0}^{m} S_{i,0,k-1,1}\delta(i)u(k-1) \\
=&x S_{i-1,0,k,1} u(i-1) + \left(\rho_t + \rho_b \sum_{j=1}^{W_i-1}\varpi_{i,j}\right)\xi_i\, g^k\, S_{0,0,0,0} \\
&+ P_c(1-x)\sum_{i=0}^{m} S_{i,0,k-1,1}\delta(i)u(k-1)\,,
\end{aligned}$$

$u(i)$ is the unit step function and $\delta(i)$ is the unit delta function. For $k = 0$, we obtain

$$S_{i,0,0,1} = \zeta_i S_{0,0,0,0}\,,$$

where $\zeta_i = x^i c + x^i \sum_{r=1}^{i}\frac{z_r}{x^r}$, $z_i = \rho_t + \rho_b \sum_{j=1}^{W_i-1}\varpi_{i,j}$, $c = \zeta_0 = \xi_0 = \rho_t + \rho_b \sum_{j=1}^{W_0-1}\varpi_{0,j}$. Analogously, we derive the following recursive formula:

$$S_{i,0,k,1} = v_{i,k} S_{0,0,0,0} \tag{45}$$

where $v_{i,k} = \zeta_i g^k + x^i a_k$, $v_{0,0} = \rho_t + \rho_b\,\zeta_0 \sum_{j=1}^{W_0-1}\varpi_{0,j}$,

$$a_k = P_c(1-x)\sum_{i=0}^{m} v_{i,k-1}, \qquad k \geq 1\,, \tag{46}$$

and $a_0 = 0$. By putting together Eqs. (36) and (42)–(44), we obtain

$$\begin{aligned}
S_{i,j,k,1} =&\rho_b S_{i,j,k,0} + \frac{\rho_t}{W_i} S_{i,0,k,0} + \sum_{r=j+1}^{W_i-1} S_{i,r,k,1} + \frac{\alpha + (1-\alpha)\beta}{W_i} S_{i-1,0,k,1} u(i-1) \\
&+ \frac{P_c(1-\alpha)(1-\beta)}{W_0}\sum_{i=0}^{m} S_{i,0,k-1,1}\delta(i)u(k-1) \\
=&\left[\sum_{r=j}^{W_i-1}\rho_b\varpi_{i,r}\xi_i g^k + \rho_t\frac{W_i-j}{W_i}\xi_i g^k + x\frac{W_i-j}{W_i} v_{i-1,k} u(i-1)\right. \\
&\left. + P_c(1-x)\frac{W_0-j}{W_0}\sum_{i=0}^{m} v_{i,k-1}\delta(i)u(k-1)\right] S_{0,0,0,0}\,.
\end{aligned} \tag{47}$$

By considering the state transition probability of the Markov chain, we apply the normalization condition to compute the state probability $S_{0,0,0,0}$. The normalization condition is

$$\eta_d S_\sigma(\rho_{t,d}, L_{p,d}) + (1-\eta_d) S_\sigma(\rho_{t,r}, L_{p,r}) = 1\,, \tag{48}$$

where

$$\begin{aligned} S_\sigma(\rho_t, L_p) =& \sum_{i=0}^{m}\sum_{j=0}^{W_i-1}\sum_{k=0}^{n}\sum_{l=0}^{1} S_{i,j,k,l} + \sum_{i=0}^{m}\sum_{k=0}^{n}\sum_{l=0}^{1} S_{i,-1,k,l} \\ &+ \sum_{k=0}^{n}\left(\sum_{j=0}^{L_s-1}\sum_{l=0}^{1} S_{+*,j,k,l} + \sum_{j=0}^{L_c-1}\sum_{l=0}^{1} S_{-*,j,k,l}\right) + \sum_{i=0}^{h-1} Q_i\,, \end{aligned}$$

$*$ is 1 for a non time-critical data packet and 2 for a GTS request of a time-critical data packet and $\rho_{t,d}$ and $\rho_{t,r}$ are the probabilities of the event of deferred transmission for a non time-critical packet with packet length $L_{p,d}$ and a GTS request for time-critical packet with request length $L_{p,r}$, respectively. In Eq. (48), the first and second term are related to a non time-critical data packet and a GTS request of time-critical data packet, respectively. We next derive the expressions of each term in Eq. (48). From Eqs. (42)–(44), we have

$$\sum_{i=0}^{m}\sum_{j=0}^{W_i-1}\sum_{k=0}^{n} S_{i,j,k,0} = \sum_{i=0}^{m}\sum_{j=0}^{W_i-1}\sum_{k=0}^{n} \varpi_{i,j}\xi_i g^k = \frac{1-g^{n+1}}{1-g}\sum_{i=0}^{m}\xi_i \sum_{j=0}^{W_i-1}\varpi_{i,j} S_{0,0,0,0}\,. \tag{49}$$

Similarly, from Eqs. (42)–(47), we have

$$\begin{aligned} \sum_{i=0}^{m}\sum_{j=0}^{W_i-1}\sum_{k=0}^{n} S_{i,j,k,1} = \Bigg[& \rho_b \sum_{k=0}^{n} g^k \sum_{i=0}^{m}\xi_i \sum_{j=0}^{W_i-1}\sum_{r=j}^{W_i-1} \varpi_{i,r} + \rho_t \sum_{k=0}^{n} g^k \sum_{i=0}^{m}\xi_i \sum_{j=0}^{W_i-1}\frac{W_i-j}{W_i} \\ & + x\sum_{i=0}^{m}\sum_{j=0}^{W_i-1}\sum_{k=0}^{n}\frac{W_i-j}{W_i} v_{i-1,k} u(i-1) \\ & + P_c(1-x)\sum_{j=0}^{W_0-1}\frac{W_0-j}{W_0}\sum_{i=0}^{m}\sum_{k=0}^{n} v_{i,k-1}\delta(i)u(k-1)\Bigg] S_{0,0,0,0}\,. \end{aligned} \tag{50}$$

From Eqs. (42) and (44), it follows

$$\sum_{i=0}^{m}\sum_{k=0}^{n} S_{i,-1,k,0} = \sum_{i=0}^{m}\sum_{k=0}^{n}(1-\rho_t)(1-\alpha) S_{i,0,k,0} = (1-\rho_t)(1-\alpha)\frac{1-g^{n+1}}{1-g}\sum_{i=0}^{m}\xi_i S_{0,0,0,0}\,. \tag{51}$$

Similarly, from Eq. (45), we obtain

$$\sum_{i=0}^{m}\sum_{k=0}^{n} S_{i,-1,k,1} = \sum_{i=0}^{m}\sum_{k=0}^{n}(1-\alpha) S_{i,0,k,1} = (1-\alpha)\sum_{i=0}^{m}\sum_{k=0}^{n} v_{i,k} S_{0,0,0,0}\,. \tag{52}$$

Analogously, the packet transmission state follows:

$$\begin{aligned} & \sum_{k=0}^{n}\left(\sum_{j=0}^{L_s-1}\sum_{l=0}^{1} S_{+*,j,k,l} + \sum_{j=0}^{L_c-1}\sum_{l=0}^{1} S_{-*,j,k,l}\right) \\ & = \left(L_s(1-P_c)(1-\beta) + L_c P_c(1-\beta)\right)\sum_{i=0}^{m}\sum_{k=0}^{n}\sum_{l=0}^{1} S_{i,-1,k,l}\,, \end{aligned} \tag{53}$$

where the last triple sum is the sum of Eqs. (51) and (52). To derive the idle-queue state probability, we first compute the state probability of the failure events due to the limited number of backoff

stages m and the retry limit n and the successful transmission. From Eqs. (37) and (38), the state probability of the failure events due to the limited number of backoff stages m is

$$\sum_{k=0}^{n}(1-\rho_t)xS_{m,0,k,0}+\sum_{k=0}^{n}xS_{m,0,k,1}=\left((1-\rho_t)x\xi_m\frac{1-g^{n+1}}{1-g}+x\sum_{k=0}^{n}v_{m,k}\right)S_{0,0,0,0}\,. \quad (54)$$

Similarly, from Eqs. (39), (40), (42) and (44), the state probability of the failure events due to the limited number of the retry limit n is

$$\begin{aligned}&\sum_{i=0}^{m}P_c(1-\beta)S_{i,-1,n,0}+P_c(1-\beta)S_{i,-1,n,1}\\&\quad=\sum_{i=0}^{m}P_c(1-\rho_t)(1-\alpha)(1-\beta)S_{i,0,n,0}+P_c(1-\alpha)(1-\beta)S_{i,0,n,1}\\&\quad=\left(P_c(1-\rho_t)(1-\alpha)(1-\beta)\sum_{i=0}^{m}\xi_i g^n+P_c(1-\alpha)(1-\beta)\sum_{i=0}^{m}v_{i,n}\right)S_{0,0,0,0}\,. \quad (55)\end{aligned}$$

The state probability of the successful transmission follows by summing Eqs. (51) and (52):

$$\sum_{i=0}^{m}\sum_{k=0}^{W_i-1}\sum_{l=0}^{1}(1-P_c)(1-\beta)S_{i,-1,k,l}\,. \quad (56)$$

By putting together Eqs. (54), (55) and (56), the idle-queue state probability is

$$\begin{aligned}Q_0=&\frac{1-\eta_t}{\eta_p}\left((1-\rho_t)x\xi_m\frac{1-g^{n+1}}{1-g}+x\sum_{k=0}^{n}v_{m,k}+P_c(1-\rho_t)(1-\alpha)(1-\beta)\sum_{i=0}^{m}\xi_i g^n\right.\\&+P_c(1-\alpha)(1-\beta)\sum_{i=0}^{m}v_{i,n}+(1-P_c)(1-\beta)\left((1-\rho_t)(1-\alpha)\frac{1-g^{n+1}}{1-g}\sum_{i=0}^{m}\xi_i\right.\\&\left.\left.+(1-\alpha)\sum_{i=0}^{m}\sum_{k=0}^{n}v_{i,k}\right)\right)S_{0,0,0,0}\,. \quad (57)\end{aligned}$$

Hence, $\sum_{i=0}^{h-1}Q_i = hQ_0$ where h is the length of idle-queue state without generating packets as described in Section 4. Note that Eqs. (49)–(53) and (57) give the state values $S_{i,j,k,l}$ as a function of $S_{0,0,0,0}$. By replacing Eqs. (49)–(53) and (57) in the normalization condition given by Eq. (48), we obtain the expression for $S_{0,0,0,0}$.

A.2. Proof of Proposition 5.3

To compute the average delay, we need some intermediate technical steps. In particular, we characterize (a) the expected value of the backoff delay due to busy channel and (b) the expression of the delay due to collision or bad channel after two successful CCAs. We first address the average backoff delay due to busy channel in the following.

Let d_i be the random time associated to the successful CCAs of a packet at the i-th backoff stage. Denote by $\mathscr{A}_i$ the event of two successful CCAs at time $i+1$ after i-th events of unsuccessful CCAs. Let $\mathscr{A}$ be the event of successful CCAs within the total attempts m. Then, the backoff delay for two successful CCAs after the i-th unsuccessful attempt is

$$d=\sum_{i=0}^{m}\mathbb{1}_{\mathscr{A}_i|\mathscr{A}}\,d_i\,,$$

where $\mathbb{1}_{\mathscr{A}_i|\mathscr{A}}$ is 1 if $\mathscr{A}_i|\mathscr{A}$ holds, and 0 otherwise.

By considering the deferred attempt, we divide the events of two successful CCAs at time $i+1$, given i previous unsuccessful CCAs, as follows: (a) the event of successful CCAs without any deferred attempts during $i+1$ attempts (b) the event of successful CCAs at the deferred attempt (c)

the event of successful CCAs after the deferred attempts during i previous unsuccessful attempts before time $i+1$. We assume that the maximum number of deferred attempts for packet transmission is one, i.e., every device needs to transmit a packet within $2T_{\text{BI}}$. In the following, we derive the average backoff delay of these three events.

First, we determine the event of successful CCAs without any deferred attempts during $i+1$ attempts. There are two events which depend on the previous deferred attempt before the current $i+1$ attempts as follows: the event of successful CCAs without any deferred attempts and with a deferred attempt before current $i+1$ attempts.

Let $\mathscr{B}_{0,i}$ be the event of successful CCAs without any deferred attempts before and during $i+1$ attempts with $i \in \{0,\ldots,m\}$. The probability of such an event is

$$\Pr[\mathscr{B}_{0,i}] = \varepsilon_i \sum_{k=0}^{i} (1-\rho_t) \sum_{j=0}^{W_i-1} \frac{(1-\rho_b)^j}{W_i} (1-\alpha)(1-\beta)\,, \tag{58}$$

where

$$\varepsilon_i = \begin{cases} \prod_{r=0}^{i-1} \epsilon_r & \text{if } i \geq 1\,, \\ 1 & \text{otherwise}\,, \end{cases}$$

$$\epsilon_r = (1-\rho_t) \sum_{j=0}^{W_r-1} \frac{(1-\rho_b)^j}{W_r} \max\left(\alpha\,, (1-\alpha)\beta\right)\,. \tag{59}$$

The term ϵ_r gives an approximation for the probability (58) for analytical tractability. For an accurate model, see our previous work [Park et al. 2009] in the manuscript. In Eq. (58), the first term ε_i models i previous unsuccessful attempts without any deferred attempts. Note that the terms $1-\rho_t$ and $1-\rho_b$ give the probability of non-deferred attempts due to the lack of the remaining slot times for packet transmission and of the remaining slot times during backoff time, respectively. $\mathscr{B}_0$ is the event of successful CCAs without any deferred attempts before and during the maximum $m+1$ times:

$$\Pr[\mathscr{B}_0] = \sum_{i=0}^{m} \Pr[\mathscr{B}_{0,i}]\,. \tag{60}$$

Consider the attempt of CCA of the i-th channel sensing. Then, the random CCA delay $d_{\mathscr{B}_0}$ within $m+1$ attempts can be described as

$$d_{\mathscr{B}_0} = \begin{cases} \Phi_0 + 2L_{\text{sc}}, & \text{if } \mathscr{B}_{0,0}|\mathscr{B}_0; \\ \Phi_0 + 2L_{\text{sc}} + \Phi_1 + 2L_{\text{sc}}, & \text{if } \mathscr{B}_{0,1}|\mathscr{B}_0; \\ \quad\vdots & \\ \sum_{k=0}^{m} \Phi_k + (m+1)2L_{\text{sc}}, & \text{if } \mathscr{B}_{0,m}|\mathscr{B}_0; \end{cases}$$

where Φ_k is the random backoff time at $k+1$-th attempts. $2L_{\text{sc}}$ is the successful sensing time and $i2L_{\text{sc}}$ is the unsuccessful sensing time due to busy channel during CCAs. Note that we consider the worst case, i.e., a failure of the second sensing (CCA_2), which implies that $L_{\text{sc}} = T_b$ and that each sensing failure takes $2L_{\text{sc}}$ in Eq. (58). Recall that a device transmits the packet when the backoff counter is 0 and two successful CCAs are detected, see [IEEE 2006] in the manuscript. We can rewrite $d_{\mathscr{B}_0}$ as

$$d_{\mathscr{B}_0} = \sum_{i=0}^{m} \left[\sum_{k=0}^{i} \Phi_k + (i+1)2L_{\text{sc}} \right] \mathbb{1}_{\mathscr{B}_{0,i}|\mathscr{B}_0}\,. \tag{61}$$

The expectation of Φ_k can be computed by recalling the uniform distribution of backoff time:

$$\mu_{\Phi_k} = \sum_{l=0}^{W_k-1} \frac{l}{W_k} T_b \,. \tag{62}$$

Now, it is possible to compute the average value of $d_{\mathscr{B}_0}$ as

$$\mu_{\mathscr{B}_0} = \sum_{i=0}^{m} \frac{\rho_{\mathscr{B}_{0,i}}}{\Pr[\mathscr{B}_0]} \,, \tag{63}$$

where

$$\rho_{\mathscr{B}_{0,i}} = \varepsilon_i \sum_{k=0}^{i} (1-\rho_t) \sum_{j=0}^{W_i-1} \frac{(1-\rho_b)^j}{W_i} (1-\alpha)(1-\beta) \left(\mu_{\Phi_k} + (i+1)2L_{\rm sc}\right) \,,$$

and recall that the first term is given in Eq. (58). Note that the normalization comes by considering all possible events of successful attempts.

Let $\mathscr{B}_{1,i}$ be the event of successful CCAs without any event of deferred attempt during $i+1$ attempts, but with an event of deferred attempt before $i+1$ attempts. The probability of such an event is

$$\Pr[\mathscr{B}_{1,i}] = \sum_{k=0}^{i} \alpha^{i-k} ((1-\alpha)\beta)^k (1-\alpha)(1-\beta) \,. \tag{64}$$

With a similar way of Eq. (61), we can define $d_{\mathscr{B}_1}$ as

$$d_{\mathscr{B}_1} = \sum_{i=0}^{m} \left[\sum_{k=0}^{i} \Phi_k + (i+1)2L_{\rm sc} \right] \mathbb{1}_{\mathscr{B}_{1,i}|\mathscr{B}_1} \,. \tag{65}$$

By using the expectation of Φ_k given in Eq. (62), the average value of $d_{\mathscr{B}_1}$ with the maximum number of times $m+1$ is

$$\mu_{\mathscr{B}_1} = \sum_{i=0}^{m} \frac{\rho_{\mathscr{B}_{1,i}}}{\Pr[\mathscr{B}_1]} \,, \tag{66}$$

where

$$\rho_{\mathscr{B}_{1,i}} = \sum_{k=0}^{i} \alpha^{i-k} ((1-\alpha)\beta)^k (1-\alpha)(1-\beta) \left(\mu_{\Phi_k} + (i+1)2L_{\rm sc}\right) \,,$$

$$\Pr[\mathscr{B}_1] = \sum_{i=0}^{m} \Pr[\mathscr{B}_{1,i}] \,.$$

Here, we have considered the worst case of sensing time given by $2L_{\rm sc}$.

Second, we derive the average backoff delay of the event of successful CCAs at the $i+1$-th attempt with an event of deferred attempt. We remind that an event of deferred attempt is due to two reasons: (i) lack of the remaining time slots for packet transmission, which happens with the probability ρ_t, and (ii) lack of the remaining time slots during the backoff time, which happens with probability ρ_b.

Let $\mathscr{C}_{0,i}$ be the event of successful CCAs after the event of deferred attempt due to the lack of the remaining slot times for packet transmission at the $i+1$-th attempt. $\mathscr{C}_0$ is the successful event of $\mathscr{C}_{0,i}$ during the maximum $m+1$ times. The probability of such an event $\mathscr{C}_{0,i}$ is

$$\Pr[\mathscr{C}_{0,i}] = \varepsilon_i \sum_{k=0}^{i} \rho_t \sum_{j=0}^{W_i-1} \frac{(1-\rho_b)^j}{W_i} (1-\alpha)(1-\beta) \,. \tag{67}$$

The random CCA delay within $m+1$ attempts of the event $\mathscr{C}_0$ is

$$d_{\mathscr{C}_0} = \sum_{i=0}^{m} \left[\sum_{k=0}^{i} \Phi_k + T_{\rm SP} + L_{\rm tx} + \Phi_i + (i+1)2L_{\rm sc} \right] \mathbb{1}_{\mathscr{C}_{0,i}|\mathscr{C}_0} \,. \tag{68}$$

The average value of $d_{\mathscr{C}_0}$ is

$$\mu_{\mathscr{C}_0} = \sum_{i=0}^{m} \frac{\rho_{\mathscr{C}_{0,i}}}{\Pr[\mathscr{C}_0]}\,, \tag{69}$$

where

$$\begin{aligned}\rho_{\mathscr{C}_{0,i}} =&\varepsilon_i \sum_{k=0}^{i} \rho_t \sum_{j=0}^{W_i-1} \frac{(1-\rho_b)^j}{W_i}(1-\alpha)(1-\beta)\left(\mu_{\Phi_k} + T_{\text{SP}} + L_{\text{tx}} + \mu_{\Phi_i} + (i+1)2L_{\text{sc}}\right)\,,\\ \Pr[\mathscr{C}_0] =& \sum_{i=0}^{m} \Pr[\mathscr{C}_{0,i}]\,,\end{aligned}$$

and the expectation of Φ_k in Eq. (62), the total number of time slots that are needed for a single transmission L_{tx} in Eq. (6) and ϵ_r in Eq. (59).

Similarly, we denote by $\mathscr{C}_{1,i}$, the event of successful CCAs at the event of deferred attempt due to the lack of the remaining slot times for backoff time counter at the $i+1$-th attempt. $\mathscr{C}_1$ is the successful event of $\mathscr{C}_{1,i}$ during the maximum $m+1$ times. The probability of such an event $\mathscr{C}_{1,i}$ is

$$\Pr[\mathscr{C}_{1,i}] =\varepsilon_i \sum_{k=0}^{i} \left(1 - \sum_{j=0}^{W_i-1} \frac{(1-\rho_b)^j}{W_i}\right)(1-\alpha)(1-\beta)\,. \tag{70}$$

The random CCA delay within $m+1$ attempts of the event $\mathscr{C}_1$ is

$$d_{\mathscr{C}_1} = \sum_{i=0}^{m} \left[\sum_{k=0}^{i} \Phi_k + T_{\text{SP}} + \Phi_i + (i+1)2L_{\text{sc}}\right] \mathbb{1}_{\mathscr{C}_{1,i}|\mathscr{C}_1}\,. \tag{71}$$

The average value of $d_{\mathscr{C}_1}$ is

$$\mu_{C_1} = \sum_{i=0}^{m} \frac{\rho_{\mathscr{C}_{1,i}}}{\Pr[\mathscr{C}_1]}\,, \tag{72}$$

where

$$\begin{aligned}\rho_{\mathscr{C}_{1,i}} =&\varepsilon_i \sum_{k=0}^{i} \left(1 - \sum_{j=0}^{W_i-1} \frac{(1-\rho_b)^j}{W_i}\right)(1-\alpha)(1-\beta)\left(\mu_{\Phi_k} + T_{\text{SP}} + \mu_{\Phi_i} + (i+1)2L_{\text{sc}}\right)\,,\\ \Pr[\mathscr{C}_1] =& \sum_{i=0}^{m} \Pr[\mathscr{C}_{1,i}]\,.\end{aligned}$$

Third, we derive the average backoff delay of the event of successful CCAs at the time $i+1$ given i previous unsuccessful CCAs but one of i unsuccessful attempts with an event of deferred attempt. With a similar way of previous events $\mathscr{C}_0$ and $\mathscr{C}_1$, we consider two different reasons of events of deferred attempt.

Let $\mathscr{F}_{0,i}$ be the event of successful CCAs at the time $i+1$ after an event of deferred attempt during i previous unsuccessful attempts due to the lack of the remaining slot times for packet transmission. $\mathscr{F}_0$ is the successful event of $\mathscr{F}_{0,i}$ during the maximum $m+1$ times. The probability of such an event $\mathscr{F}_{0,i}$ is

$$\Pr[\mathscr{F}_{0,i}] =\sum_{f=0}^{i-2} \sum_{h=0}^{i-2-(f-1)} \varepsilon_f \sum_{k=0}^{f} (\widetilde{\alpha}_f + \widetilde{\beta}_f)\alpha^{i-2-(f-1)-h}((1-\alpha)\beta)^h(1-\alpha)(1-\beta)u(i-2)\,, \tag{73}$$

where $\widetilde{\alpha}_r = \rho_t \sum_{j=0}^{W_r-1} \frac{(1-\rho_b)^j}{W_r}\alpha$, $\widetilde{\beta}_r = \rho_t \sum_{j=0}^{W_r-1} \frac{(1-\rho_b)^j}{W_r}(1-\alpha)\beta$. The random CCA delay within $m+1$ attempts of the event $\mathscr{F}_0$ is

$$d_{\mathscr{F}_0} = \sum_{i=0}^{m} \left[\sum_{f=0}^{i-2} \sum_{k=0}^{f} \Phi_k + T_{\text{SP}} + L_{\text{tx}} + \Phi_f + (i+1)2L_{\text{sc}} \right] \mathbb{1}_{\mathscr{F}_{0,i}|\mathscr{F}_0} . \tag{74}$$

The average value of $d_{\mathscr{F}_0}$ is

$$\mu_{\mathscr{F}_0} = \sum_{i=0}^{m} \frac{\rho_{\mathscr{F}_{0,i}}}{\Pr[\mathscr{F}_0]} , \tag{75}$$

where

$$\begin{aligned}
\rho_{\mathscr{F}_{0,i}} =& \sum_{f=0}^{i-2} \sum_{h=0}^{i-2-(f-1)} \varepsilon_f \sum_{k=0}^{f} (\widetilde{\alpha}_f + \widetilde{\beta}_f)\alpha^{i-2-(f-1)-h}((1-\alpha)\beta)^h(1-\alpha)(1-\beta) \\
&\times \left(\mu_{\Phi_k} + T_{\text{SP}} + L_{\text{tx}} + \mu_{\Phi_f} + (i+1)2L_{\text{sc}}\right) u(i-2) , \\
\Pr[\mathscr{F}_0] =& \sum_{i=0}^{m} \Pr[\mathscr{F}_{0,i}] .
\end{aligned}$$

Similarly, we denote by $\mathscr{F}_{1,i}$, the event of successful CCAs at the time $i+1$ after a deferred attempt during i previous unsuccessful attempts due to the lack of the remaining slot times for backoff time counter. $\mathscr{F}_1$ is the successful event of $\mathscr{F}_{1,i}$ during the maximum $m+1$ times. The probability of such an event $\mathscr{F}_{1,i}$ is

$$\Pr[\mathscr{F}_{1,i}] = \sum_{f=0}^{i-2} \sum_{h=0}^{i-2-(f-1)} \varepsilon_f \sum_{k=0}^{f} (\widehat{\alpha}_f + \widehat{\beta}_f)\alpha^{i-2-(f-1)-h}((1-\alpha)\beta)^h(1-\alpha)(1-\beta)u(i-2) , \tag{76}$$

where $\widehat{\alpha}_r = \left(1 - \sum_{j=0}^{W_r-1} \frac{(1-\rho_b)^j}{W_r}\right)\alpha$, $\widehat{\beta}_r = \left(1 - \sum_{j=0}^{W_r-1} \frac{(1-\rho_b)^j}{W_r}\right)(1-\alpha)\beta$. The random CCA delay within $m+1$ attempts of the event $\mathscr{F}_1$ is

$$d_{\mathscr{F}_1} = \sum_{i=0}^{m} \left[\sum_{f=0}^{i-2} \sum_{k=0}^{f} \Phi_k + T_{\text{SP}} + \Phi_f + (i+1)2L_{\text{sc}} \right] \mathbb{1}_{\mathscr{F}_{1,i}|\mathscr{F}_1} . \tag{77}$$

The average value of $d_{\mathscr{F}_1}$ is

$$\mu_{\mathscr{F}_1} = \sum_{i=0}^{m} \frac{\rho_{\mathscr{F}_{1,i}}}{\Pr[\mathscr{F}_1]} , \tag{78}$$

where

$$\begin{aligned}
\rho_{\mathscr{F}_{1,i}} =& \sum_{f=0}^{i-2} \sum_{h=0}^{i-2-(f-1)} \varepsilon_f \sum_{k=0}^{f} (\widehat{\alpha}_f + \widehat{\beta}_f)\alpha^{i-2-(f-1)-h}((1-\alpha)\beta)^h(1-\alpha)(1-\beta) \\
&\times \left(\mu_{\Phi_k} + T_{\text{SP}} + \mu_{\Phi_f} + (i+1)2L_{\text{sc}}\right) u(i-2) , \\
\Pr[\mathscr{F}_1] =& \sum_{i=0}^{m} \Pr[\mathscr{F}_{1,i}] .
\end{aligned}$$

In the following, we consider the packet loss due to collision or bad channel. We remind that each device transmits a packet when the channel sensing is successful within the maximum number of m backoff stages. Hence, we derive the average packet delay based on the average backoff delay for the event of successful CCAs. By considering the event of deferred attempt, we categorize the

events of successful transmission at the time $k+1$ given k previous packet collisions or bad channel conditions as follows: (a) the event of successful packet transmission without any deferred attempts (b) the event of successful packet transmission at the event of deferred attempt (c) the event of successful packet transmission after an event of deferred attempt during k previous unsuccessful attempts.

First, the average delay of the successful packet transmission at the time $k+1$ after k events of packet collisions or bad channel conditions without any events of deferred attempts is

$$\Gamma_k = (1-P_c)P_c^k \Pr[\mathscr{B}_0]^{k+1} \left((k+1)\mu_{\mathscr{B}_0} + L_s + kL_c\right) , \tag{79}$$

$$P_k = (1-P_c)P_c^k \Pr[\mathscr{B}_0]^{k+1} , \tag{80}$$

where $\Pr[\mathscr{B}_0]$ in Eq. (60), $\mu_{\mathscr{B}_0}$ in Eq. (63), and L_s and L_c are the packet successful transmission time and the packet collision time given in Eqs. (3) and (4), respectively. A packet transmission is successful with probability $1-P_c$, or collide with probability P_c given by Eq. (80).

Second, the average delay of the successful transmission at the deferred attempt of the time $k+1$ after k packet collisions or bad channel conditions is

$$\hat{\Gamma}_{*,k} = (1-P_c)P_c^k \Pr[\mathscr{B}_0]^k \Pr[*] \left(k\mu_{\mathscr{B}_0} + \mu_* + L_s + kL_c\right) , \tag{81}$$

$$\hat{P}_{*,k} = (1-P_c)P_c^k \Pr[\mathscr{B}_0]^k \Pr[*] , \tag{82}$$

where $*$ is one of the events $\{\mathscr{C}_0, \mathscr{C}_1, \mathscr{F}_0, \mathscr{F}_1\}$ given in Eqs. (66), (72), (75), and (78), respectively.

Third, the average delay of the successful packet transmission at the time $k+1$ given k previous events of packet collisions or bad channel conditions but one of k events of packet collisions or bad channel conditions with an event of deferred attempts is

$$\begin{aligned}\check{\Gamma}_{*,k} =& (1-P_c)P_c^k \sum_{r=0}^{k-1} \Pr[\mathscr{B}_0]^{k-1-r} \Pr[*] \Pr[\mathscr{B}_1]^{r+1} \\ & \times \left((k-1-r)\mu_{\mathscr{B}_0} + \mu_* + (r+1)\mu_{\mathscr{B}_1} + L_s + kL_c\right) u(k-1) ,\end{aligned} \tag{83}$$

$$\check{P}_{*,k} = (1-P_c)P_c^k \sum_{r=0}^{k-1} \Pr[\mathscr{B}_0]^{k-1-r} \Pr[*] \Pr[\mathscr{B}_1]^{r+1} u(k-1) , \tag{84}$$

where $*$ is one of the events $\{\mathscr{C}_0, \mathscr{C}_1, \mathscr{F}_0, \mathscr{F}_1\}$ given in Eq. (66), (72), (75), and (78), respectively.

Now, we are in the position to derive the average delay for successfully received packets. By normalizing Eqs. (79), (82), and (84), the expected value of delay is

$$\begin{aligned}\mathbb{E}[D] =& \sum_{k=0}^{n} \Big(\Gamma_k + \hat{\Gamma}_{\mathscr{C}_0,k} + \hat{\Gamma}_{\mathscr{C}_1,k} + \hat{\Gamma}_{\mathscr{F}_0,k} + \hat{\Gamma}_{\mathscr{F}_1,k} \\ & + \left(\check{\Gamma}_{\mathscr{C}_0,k} + \check{\Gamma}_{\mathscr{C}_1,k} + \check{\Gamma}_{\mathscr{F}_0,k} + \check{\Gamma}_{\mathscr{F}_1,k}\right) u(k-1) \Big) P_{\text{tot}}^{-1} ,\end{aligned} \tag{85}$$

where

$$P_{\text{tot}} = \sum_{k=0}^{n} P_k + \hat{P}_{\mathscr{C}_0,k} + \hat{P}_{\mathscr{C}_1,k} + \hat{P}_{\mathscr{F}_0,k} + \hat{P}_{\mathscr{F}_1,k} + \left(\check{P}_{\mathscr{C}_0,k} + \check{P}_{\mathscr{C}_1,k} + \check{P}_{\mathscr{F}_0,k} + \check{P}_{\mathscr{F}_1,k}\right) u(k-1) .$$

A.3. Proof of Proposition 6.1

We follow two steps to compute the average queuing delay of the GTS allocation mechanism. We first derive the expected delay $D_{i,j,t}$ of j new requests that observe a queue size of i waiting requests at superframe t. Then, the expected delay experienced by j new requests arriving at superframe t is computed by considering $D_{i,j,t}$.

Now, we compute the expected delay $D_{i,j,t}$. If the new requests are $j < \omega_j$ (i.e., $\upsilon_j = 0$), then

$$D_{i,j,t} = \frac{T_{\text{CAP}}^t}{2} + \sum_{k=t}^{t+\upsilon_i} T_{\text{BI}}^{k,k+1} + \underline{\theta} T_{\text{SS}}^{t+\upsilon_i+1} \left(\omega_i + \frac{j+1}{2}\right) , \tag{86}$$

where υ_i, υ_j is the quotient and ω_i, ω_j is the remainder of old and new requests, respectively. The average arrival delay is half of T_{CAP}. The offset $\sum_{k=t}^{t+\upsilon_i} T_{\text{BI}}^{k,k+1}$ takes into account the delay of i old requests. In addition, the average service delay is half of the service time of $j+1$ new requests with a single GTS duration $\underline{\theta} T_{\text{SS}}^{t+\upsilon_i+1}$ for each request. Note that ω_i remainders are considered as an offset before the GTS allocation of j new requests.

If j new requests are more than ω_j (i.e., $\upsilon_j > 0$), then

$$
\begin{aligned}
D_{i,j,t} =& \frac{\Delta_u - \omega_i}{j} \left[T_{\text{CAP}}^t \left(1 - \frac{\Delta_u - \omega_i}{2j} \right) + \sum_{k=t}^{t+\upsilon_i} T_{\text{BI}}^{k,k+1} + \underline{\theta} T_{\text{SS}}^{t+\upsilon_i+1} \left(\omega_i + \frac{\Delta_u - \omega_i + 1}{2} \right) \right] \\
&+ \frac{\Delta_u}{j} \sum_{l=t+\upsilon_i+1}^{t+\upsilon_i+\upsilon_j-1} \left[T_{\text{CAP}}^t \left(1 - \frac{\Delta_u - \omega_i}{j} - \frac{(l-(t+\upsilon_i+1))\Delta_u}{j} - \frac{\Delta_u}{2j} \right) \right. \\
&\left. + \sum_{k=t}^{t+\upsilon_i} T_{\text{BI}}^{k,k+1} + \sum_{k=t+\upsilon_i+1}^{l} T_{\text{BI}}^{k,k+1} + \frac{\underline{\theta} T_{\text{SS}}^{l+1} (\Delta_u + 1)}{2} \right] u(\upsilon_j - 2) \\
&+ \frac{\omega_j}{j} \left[T_{\text{CAP}}^t \left(1 - \frac{\Delta_u - \omega_i}{j} - \frac{(\upsilon_j - 1)\Delta_u}{j} - \frac{\omega_j}{2j} \right) \right. \\
&\left. + \sum_{k=t}^{t+\upsilon_i+\upsilon_j} T_{\text{BI}}^{k,k+1} + \frac{\underline{\theta} T_{\text{SS}}^{t+\upsilon_i+\upsilon_j+1} (\omega_j + 1)}{2} \right] .
\end{aligned} \tag{87}
$$

It is possible to categorize the delay into three groups (first, middle, last group). The different groups of delay are also combined with the arrival, offset, and service delay components. The delay of the first group is the delay of first allocated GTSs $\Delta_u - \omega_i$ out of j new requests at superframe $t+\upsilon_i+1$, see Fig. 6. The average arrival delay of the first group considers the ratio $1-(\Delta_u-\omega_i)(2j)^{-1}$ since the arrival time of j new requests are uniformly distributed in the CAP. The offset $\sum_{k=t}^{t+\upsilon_i} T_{\text{BI}}^{k,k+1}$ takes into account the delay of i old requests. With a similar approach of $j < \omega_j$, the average of the service delay considers ω_i remainders and $\Delta_u - \omega_i$ new requests.

The delay of the middle group is the delay between the first and last allocated GTSs $(\upsilon_j - 1)\Delta_u$ out of j new requests from the superframe $t+\upsilon_i+2$ to $t+\upsilon_i+\upsilon_j$, see Fig. 6. Hence, the number of allocated GTSs is the maximum number of GTSs Δ_u out of j new requests. The average arrival delay of the middle group considers the term $1-(\Delta_u-\omega_i)j^{-1}-(l-(t+\upsilon_i+1))\Delta_u j^{-1}-\Delta_u(2j)^{-1}$ since other requests of the number $\Delta_u - \omega_i$ of the first group and the number $(l-(t+\upsilon_i+1))\Delta_u$ of the middle group obtain GTSs before Δ_u new requests out of j. Note that the number of allocated GTS is dependent on the time progress l in a superframe unit. With a similar approach used in the case $j < \omega_j$, we consider Δ_u new requests to derive the aveage service delay.

The delay of the last group is the delay of last allocated GTSs ω_j out of j new requests at superframe $t+\upsilon_i+\upsilon_j+1$. The average arrival delay of the last group considers the term $1-(\Delta_u-\omega_i)j^{-1}-((\upsilon_j-1)\Delta_u)j^{-1}-\omega_j(2j)^{-1}$ since other requests of the number $\Delta_u-\omega_i$ of the first group and the number $(\upsilon_j-1)\Delta_u$ of the middle group obtain GTSs before ω_j remainders out of j new requests. By a similar approach to the case $j < \omega_j$, we consider ω_j remainders in order to compute the average service delay.

By considering the Eqs. (86) and (87), the expected delay experienced by j new requests arriving at the superframe t is

$$
\mathbb{E}[D_q(t)] = \sum_{i=0}^{\Delta_u - 1} \sum_{j=1}^{\bar{\lambda}} \left[\pi_i^t \frac{\lambda_j}{\sum_{k=1}^{\bar{\lambda}} \lambda_k} \left(D_{i,j,t}\, u(q-j) + D_{i,q,t}\, u(j-(q+1)) \right) \right]
$$

$$+\sum_{i=\Delta_u}^{q}\sum_{j=1}^{\bar{\lambda}}\left[\pi_i^t\frac{\lambda_j}{\sum_{k=1}^{\bar{\lambda}}\lambda_k}\left(D_{i,j,t}\,u(q-i+\Delta_u-j)\right.\right.$$

$$\left.\left.+D_{i,q-i+\Delta_u,t}\,u(j-(q-i+\Delta_u+1))\right)\right]$$

$$+\pi_{\bar{q}}^t\sum_{j=1}^{\bar{\lambda}}\left[\frac{\lambda_j}{\sum_{k=1}^{\bar{\lambda}}\lambda_k}\left(D_{q,j,t}\,u(\Delta_u-j)+D_{q,\Delta_u,t}\,u(j-(\Delta_u+1))\right)\right], \tag{88}$$

where $\bar{\lambda}$ is the maximum number of requests. The first term of Eq. (88) gives the average delay of new requests when the old requests $i \leq \Delta_u - 1$. By considering the Markov chain, if there are more than q new requests, then the extra requests out of q arrivals will be dropped. The unit function takes into account the queue size q. The second term considers the average delay when the old requests $\Delta_u \leq i \leq q$. Note that the Markov chain represents the feasible number of new requests without the dropped requests. By a similar approach to first term, it is possible to consider the maximum number of new requests, $q-i+\Delta_u$, without dropped requests when i old requests wait in the queue. The third term computes the average delay at the dropped state $\bar{q}$. At the dropped state $\bar{q}$, Δ_u new requests are considered since the queue size of old requests is q.

A.4. Proof of Lemma 7.1

Let us denote $T^{\star}_{\text{CAP}}$ and $T^{\star}_{\text{CFP}}$ the fixed point of (27). We derive the stochastic mean convergence condition of these iterations. The geometric series is used to compute the condition of convergence and divergence of the iterations of the CAP and CFP lengths. The iteration of Eq. (27) is rewritten as follows

$$T_{\text{CAP}}^{t+1} = T_{\text{SD}} - \frac{(1-\eta_d)N_{\text{SD}}T_{\text{SS}}}{T_{\text{SD}}}T_{\text{CAP}}^{t}\,.$$

Note that the convergence of an infinite series is not changed by insertion or removal of finite number of terms. We remark that the iteration follows an oscillatory behavior since the slope of iteration is negative. Then, the CAP length is rewritten as the following sum of geometric series

$$T_{\text{CAP}}^{\infty} = T_{\text{SD}}\sum_{k=0}^{\infty}\left(-\frac{(1-\eta_d)N_{\text{SD}}T_{\text{SS}}}{T_{\text{SD}}}\right)^{k}.$$

This geometric series converges to

$$T^{\star}_{\text{CAP}} = \frac{T_{\text{SD}}^2}{T_{\text{SD}}+(1-\eta_d)N_{\text{SD}}T_{\text{SS}}}\,, \tag{89}$$

for $(1-\eta_d)N_{\text{SD}}T_{\text{SS}} < T_{\text{SD}}$ and the CAP length is saturated to

$$T^{\star}_{\text{CAP}} = T_{\text{SD}} - \Delta_u T_{\text{SS}}\,, \tag{90}$$

for $(1-\eta_d)N_{\text{SD}}(T_{\text{SD}}-\Delta_u T_{\text{SS}})/T_{\text{SD}} \geq \Delta_u$. If the number of requests is greater than Δ_u, then the CAP length is $T_{\text{SD}} - \Delta_u T_{\text{SS}}$. By considering the convergence condition (89) and the saturated condition (90), the stability of the CAP and CFP in a superframe are as follows

$$\begin{array}{ll}\text{convergence} & \text{if } N_{\text{SD}} < \frac{N_{\text{SS}}}{1-\eta_d}\,,\\ \text{saturation} & \text{if } N_{\text{SD}} \geq \frac{N_{\text{SS}}\Delta_u}{(1-\eta_d)(N_{\text{SS}}-\Delta_u)}\,,\\ \text{oscillation} & \text{if } \frac{N_{\text{SS}}}{1-\eta_d} \leq N_{\text{SD}} < \frac{N_{\text{SS}}\Delta_u}{(1-\eta_d)(N_{\text{SS}}-\Delta_u)}\,.\end{array}$$

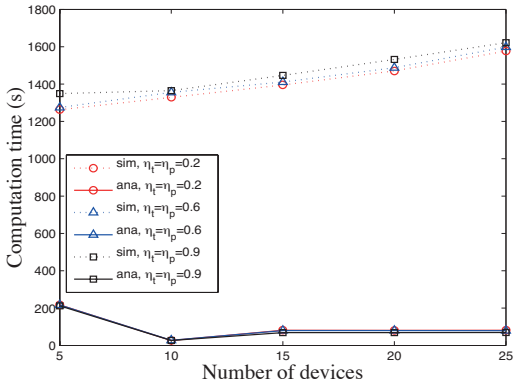
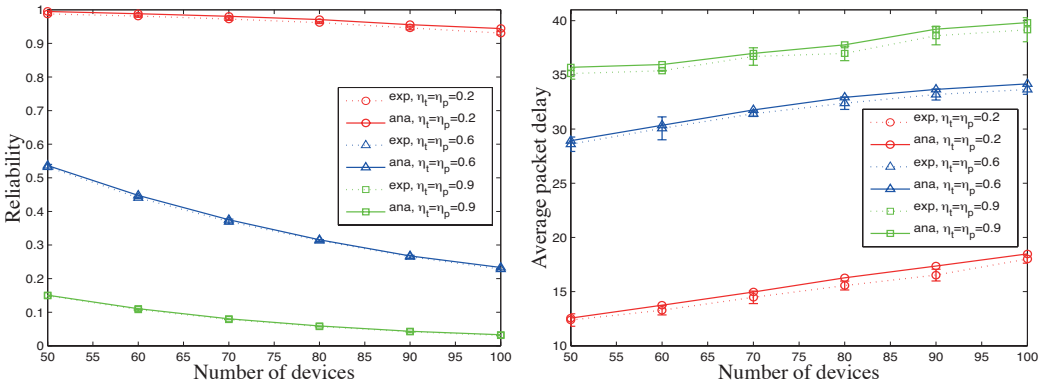


Fig. 22. Computation time of both analytical model and simulation as a function of the number of devices $N = 5, \ldots, 25$, traffic load $\eta_t = \eta_p = 0.2, 0.6, 0.9$ and superframe order $\mathrm{SO} = 5$ with a probability for generating a non time-critical data packet $\eta_d = 1$ and packet length $L_p = 7$.



(a) Reliability (b) Average packet delay

Fig. 23. Reliability and average packet delay as a function of the number of devices $N = 50, \ldots, 100$, traffic load $\eta_t = \eta_p = 0.2, 0.6, 0.9$ given beacon order $\mathrm{BO} = 8$, superframe order $\mathrm{SO} = 7$, probability for generating a non time-critical data packet $\eta_d = 1$ and packet length $L_p = 7$.

A.5. Comparison of computation time

Fig. 22 compares the computation time of both analytical model and simulation as a function of the number of devices $N = 5, \ldots, 25$, traffic load $\eta_t = \eta_p = 0.2, 0.6, 0.9$ and superframe order $\mathrm{SO} = 5$ with a probability for generating a non time-critical data packet $\eta_d = 1$. The computation time for the analytical model is less than the one for the simulation. The computation time for the simulation increases as the number of devices and the traffic load increase. However, the computation time for the analytical model shows weak dependency with the number of devices and the traffic load. We note that the system of nonlinear equations is solved using the common trust-region dogleg algorithm [Conn et al. 2000]. The algorithm is a variant of the Powell dogleg method described in [Powell 1970]. The initial point is critical for the computation time to apply this algorithm. Note that we fix the initial points $\alpha = 0.3, \beta = 0.3, \tau = 0.3$.

A.6. Performance evaluation for a large-scale network

Fig. 23 presents the reliability and average packet delay of the CAP without having GTS allocations $\eta_d = 1$ as a function of the number of devices $N = 50, \ldots, 100$, traffic load $\eta_t = \eta_p = 0.2, 0.6, 0.9$ and superframe order $\mathrm{SO} = 7$ with the packet length $L_p = 7$. As the number of devices increases,

the performance of the reliability and delay decreases due to the high contention to send a packet. The analytical model predicts well the simulation results even though the analytical model is slightly higher than the simulation results. Recall that the analysis using the Markov chain model in this paper assumes the decoupling approximation. The decoupling approximation is to assume that the aggregate attempt process of the other $N - 1$ devices is independent of the backoff process of the given device. However, this hypothesis between consecutive channel slots and statistical homogeneity is not generally true as the number of devices increases.

ACKNOWLEDGMENTS

We acknowledge the support of Yian Qin for implementing the IEEE 802.15.4 MAC using TelosB platform and Contiki OS.

REFERENCES

ARAUJO, J., ANTA, A., JR, M. M., FARIA, J., HERNANDEZ, A., TABUADA, P., AND JOHANSSON, K. H. 2011. Self-triggered control for wireless sensor and actuator networks. In *IEEE DCOSS*.

BENDER, M. A., FARACH-COLTON, M., HE, S., KUSZMAUL, B. C., AND LEISERSON, C. E. 2005. Adversarial contention resolution for simple channels. In *ACM SPAA*.

BIANCHI, G. 2000. Performance analysis of the IEEE 802.11 distributed cordination function. *IEEE Journal on Selected Areas in Communications 18,* 3, 535–547.

BINDER, R. 1975. A dynamic packet switching system for satellite broadcast channels. In *IEEE ICC*.

BOGGIO, G., BURZIO, M., PORTINARO, N., CAI, J., CERUTTI, I., FUMAGALLI, A., TACCA, M., VALCARENGHI, L., CARENA, A., AND GAUDINO, R. 2001. Network designer - artifex - optsim: a suite of integrated software tools for synthesis and analysis of high speed networks. *Optical Networks Magazine* 2, 27–41.

BURATTI, C. 2010. Performance analysis of IEEE 802.15.4 beacon-enabled mode. *Vehicular Technology, IEEE Transactions on 59,* 4, 2031–2045.

CALI, F., CONTI, M., AND GREGORI, E. 2000. IEEE 802.11 protocol: design and performance evaluation of an adaptive backoff mechanism. *IEEE Journal on Selected Areas in Communications 18,* 9, 1774 –1786.

CHENG, L., ZHANG, X., AND BOURGEOIS, A. G. 2007. GTS allocation scheme revisited. *Electronics Letters 43,* 18, 1005–1006.

CONN, N., GOULD, N., AND TOINT, P. 2000. Trust-region methods. *MPS/SIAM Series on Optimization*.

CROWTHER, W., RETTBERG, R., WALDEN, D., ORNSTEIN, S., AND HEART, F. 1973. A system for broadcast communication: Reservation-aloha. In *IEEE HICSS*.

DUNKELS, A., GRONVALL, B., AND VOIGT, T. 2004. Contiki - a lightweight and flexible operating system for tiny networked sensors. In *IEEE EmNets*.

FARIDI, A., PALATTELLA, M., LOZANO, A., DOHLER, M., BOGGIA, G., GRIECO, L., AND CAMARDA, P. 2010. Comprehensive evaluation of the IEEE 802.15.4 MAC layer performance with retransmissions. *IEEE Transactions on Vehicular Technology 59,* 8, 3917–3932.

FINE, M. AND TOBAGI, F. A. 1984. Demand assignment multiple access schemes in broadcast bus local area networks. *IEEE Transactions on Computers 33,* 12, 1130–1159.

GROSS, D. AND HARRIS, C. M. 1998. *Fundamentals of Queueing theory*. Wesley inter-science.

HESPANHA, J. P., NAGHSHTABRIZI, P., AND XU, Y. 2007. A survey of recent results in networked control systems. *Proceedings of the IEEE 95,* 1, 138–162.

HUANG, Y. K., PANG, A. C., AND HUNG, H. N. 2008. An adaptive GTS allocation scheme for IEEE 802.15.4. *IEEE Transactions on Parallel Distribbuted Systems 19,* 5, 641–651.

IEEE 1999. *IEEE 802.11 standard: Wireless LAN Medium Access Control and Physical Layer Specifications*. IEEE. http://www.ieee802.org/11.

IEEE 2006. *IEEE 802.15.4 standard: Wireless Medium Access Control and Physical Layer Specifications for Low-Rate Wireless Personal Area Networks*. IEEE. http://www.ieee802.org/15/pub/TG4.html.

IEEE 2010. *IEEE 802.15 task group 4e: Wireless Medium Access Control and Physical Layer Specifications for Low-Rate Wireless Personal Area Networks*. IEEE. http://www.ieee802.org/15/pub/TG4e.html.

ISA 2009. *ISA-100.11a-2009 Wireless systems for industrial automation: Process control and related applications*. ISA.

JUNG, C. Y., HWANG, H. Y., SUNG, D. K., AND HWANG, G. U. 2009. 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 58,* 1, 473–478.

KIM, T. H., NI, J., SRIKANT, R., AND VAIDYA, N. H. 2011. On the achievable throughput of CSMA under imperfect carrier sensing. In *IEEE INFOCOM*.

KLEINROCK, L. AND YEMINI, Y. 1978. An optimal adaptive scheme for multiple access broadcast communication. In *IEEE ICC*.

KOUBAA, A., ALVES, M., AND TOVAR, E. 2006. Energy and delay trade-off of the GTS allocation mechanism in IEEE 802.15.4 for wireless sensor networks. *International Journal of Communication Systems 20,* 7, 791 – 808.

MARTALO, M., BUSANELLI, S., AND FERRARI, G. 2009. Markov chain-based performance analysis of multihop IEEE 802.15.4 wireless networks. *Performance Evaluation 66,* 12, 722–741.

MAZO, M. AND TABUADA, P. 2011. Decentralized event-triggered control over wireless sensor/actuator networks. *IEEE Transactions on Automatic Control*. To appear.

MISIC, J. AND MISIC, V. B. 2005. Access delay for nodes with finite buffers in IEEE 802.15.4 beacon enabled PAN with uplink transmissions. *Computer Communications 28,* 10, 1152–1166.

MISIC, J., SHAF, S., AND MISIC, V. B. 2006. Performance of a beacon enabled IEEE 802.15.4 cluster with downlink and uplink traffic. *IEEE Transactions on Parallel and Distributed Systems 17,* 4, 361–376.

NA, C., YANG, Y., AND MISHRA, A. 2008. An optimal GTS scheduling algorithm for time-sensitive transactions in IEEE 802.15.4 networks. *Computer Networks 52,* 13, 2543–2557.

OSTERLIND, F., DUNKELS, A., ERIKSSON, J., FINNE, N., AND VOIGT, T. 2006. Cross-level sensor network simulation with cooja. In *IEEE SenseApp*.

PARK, P. 2011. Modeling, analysis, and design of wireless sensor network protocols. Ph.D. thesis, KTH Royal Institute of Technology, Sweden.

PARK, P., ARAUJO, J., AND JOHANSSON, K. H. 2010. Wireless networked control system co-design. In *IEEE ICNSC*.

PARK, P., FISCHIONE, C., AND JOHANSSON, K. H. 2009. Performance analysis of GTS allocation in beacon enabled IEEE 802.15.4. In *IEEE SECON*.

PARK, P., MARCO, P. D., SOLDATI, P., FISCHIONE, C., AND JOHANSSON, K. H. 2009. A generalized Markov chain model for effective analysis of slotted IEEE 802.15.4. In *IEEE MASS*.

POLASTRE, J., SZEWCZYK, R., AND CULLER, D. 2005. Telos: enabling ultra-low power wireless research. In *ACM/IEEE IPSN*.

POLLIN, S., ERGEN, M., ERGEN, S. C., BOUGARD, B., PERRE, L., MOERMAN, I., BAHAI, A., VARAIYA, P., AND CATTHOOR, F. 2008. Performance analysis of slotted carrier sense IEEE 802.15.4 medium access layer. *IEEE Transactions on Wireless Communication 7,* 9, 3359–3371.

POWELL, M. J. D. 1970. A fortran subroutine for solving systems of nonlinear algebraic equations. In *Numerical Methods for Nonlinear Algebraic Equations*, P. Rabinowitz, Ed.

PYO, C. W. AND HARADA, H. 2009. Throughput analysis and improvement of hybrid multiple access in IEEE 802.15.3c mm-wave WPAN. *IEEE Journal on Selected Areas in Communications 27,* 8, 1414 – 1424.

QIN, Y. AND PARK, P. 2011. Implementation of the IEEE 802.15.4 protocol. Tech. rep., KTH. [Online] Available: http://www.ee.kth.se/∼pgpark/code/hybrid-wpan.zip.

RAMACHANDRAN, I., DAS, A. K., AND ROY, S. 2007. Analysis of the contention access period of IEEE 802.15.4 MAC. *ACM Transactions on Sensor Networks 3,* 1, 641–651.

ROBERTS, L. G. 1973. Dynamic allocation of satellite capacity through packet reservation. In *AFIPS*.

SAHOO, P. K. AND SHEU, J. P. 2008. Modeling IEEE 802.15.4 based wireless sensor network with packet retry limits. In *IEEE PE-WASUN*.

TASAKA, S. AND ISHIBASHI, Y. 1984. A reservation protocol for satellite packet communication - a performance analysis and stability considerations. *IEEE Transactions on Communications 32,* 8, 920– 927.

TICKIOO, O. AND SIKDAR, B. 2004. Queueing analysis and delay mitigation in IEEE 802.11 random access MAC based wireless networks. In *IEEE INFOCOM*.

TOBAGI, F. A. AND KLEINROCK, L. 1976. Packet switching in radio channels: Part III - polling and (dynamic) split-channel reservation multiple-access. *IEEE Transactions on Communications 24,* 8, 832–845.

TSAI, D. AND CHANG, J. 1986. Performance study of an adaptive reservation multiple access technique for data transmissions. *IEEE Transactions on Communications 34,* 7, 725–727.

WHEELER, A. 2007. Commercial applications of wireless sensor networks using ZigBee. *IEEE Communications Magazine 45,* 4, 70 –77.

WILLIG, A. 2008. Recent and emerging topics in wireless industrial communication. *IEEE Transactions on Industrial Informatics 4,* 2, 102–124.

WILLIG, A., MATHEUS, K., AND WOLISZ, A. 2005. Wireless technology in industrial networks. *Proceedings of the IEEE 93,* 6, 1130–1151.

WU, H., PENG, Y., LONG, K., CHENG, S., AND MA, J. 2002. Performance of reliable transport protocol over IEEE 802.11 wireless LAN: Analysis and enhancement. In *IEEE INFOCOM*.

ZHENG, J. AND LEE, M. L. 2004. Will IEEE 802.15.4 make ubiquitous networking a reality?: A discussion on a potential low power, low bit rate standard. *IEEE Communications Magazine 42,* 6, 140–146.