

Document downloaded from:

<http://hdl.handle.net/10251/152262>

This paper must be cited as:

Leyva-Mayorga, I.; Pla, V.; Martínez Bauset, J.; Rivero-Angeles, ME. (2017). A hybrid method for the QoS analysis and parameter optimization in time-critical random access wireless sensor networks. *Journal of Network and Computer Applications*. 83:190-203. <https://doi.org/10.1016/j.jnca.2017.01.027>



The final publication is available at

<https://doi.org/10.1016/j.jnca.2017.01.027>

Copyright Elsevier

Additional Information

A hybrid method for the QoS analysis and parameter optimization in time-critical random access wireless sensor networks

Israel Leyva-Mayorga^{a,*}, Vicent Pla^a, Jorge Martinez-Bauset^a, Mario E. Rivero-Angeles^b

^a*ITACA, Universitat Politècnica de València, Camí de Vera, s/n, 46022, Valencia, Spain*

^b*Communication Networks Laboratory, CIC-IPN, Mexico City, Mexico*

Abstract

Evolution in electronics has led to the development of complex applications in wireless sensor networks (WSNs), where efficient and swift event reporting is needed. In time-critical applications, achieving an adequate report latency is particularly relevant as it allows a proper reaction from the network to the occurring phenomena. It is evident that mean report latency is insufficient as a QoS indicator for time-critical applications. Instead, high percentiles or the whole distribution are much better suited. In certain applications such as target tracking and positioning, the transmission of a certain number of event packets is required to accurately characterize the occurring phenomena. Building on this, we present a hybrid method for obtaining the probability distribution of report latency in random access (RA) WSN protocols. In this method, the distribution of the number of detecting nodes is obtained by simulation, which then allows us to obtain the desired QoS parameters analytically. In this study, we use our method to obtain and optimize the event report latency and energy consumption in RA WSNs. Results show that modifying the transmission parameters during backoff increases the robustness of RA event reporting and also enhances the performance of the WSN in environments where multiple types of events can be detected.

Keywords: Event reporting, QoS, time-critical applications, wireless sensor networks.

1. Introduction

Wireless sensor networks (WSNs) are a cost-efficient solution to massive monitoring thanks to its capacity of identifying and reporting a wide range of physical conditions inside the area of interest. Hence, WSNs are an essential component of smart urban environments, which aim to improve the quality of life by providing the population with real-time information and services (Rashid and Rehmani, 2015). In such environments, delay-tolerant applications usually coexist with time-critical applications. In the latter, nodes are usually in charge of detecting hazardous conditions. Therefore, a swift response is needed in order to ensure the timely activation of disaster contention mechanisms and reduce the damage caused to the network, the environment or the population.

The overall behavior of a WSN is determined by the selected WSN protocol, so its selection should be based on the monitoring needs of the WSN user. WSN protocols are usually a combination of routing and medium access control (MAC) protocols. In the former, nodes are organized depending on their spatial distribution to optimize data transmission paths. *Cluster-based* protocols are a classic

form of organization widely used nowadays (Zhang et al., 2015; Afsar and Tayarani-N, 2014). In these protocols, nodes are divided in groups named *clusters* during a cluster formation (CF) phase. Each cluster contains a cluster head (CH) node, which is in charge of collecting the data packets from its cluster members (CMs) and its transmission to the sink node.

MAC protocols, on the other hand, establish the communication links and define the manner in which the nodes share the communication resources, which must be in line with the requirements of the target application (Akyildiz et al., 2002; Rault et al., 2014; Asudeh et al., 2016). Note that, when the selected protocol is unable to meet the basic application requirements, the whole network is inoperative.

An important aspect that is typically overlooked by MAC protocols is the number of event reports required at the sink node to characterize the occurring phenomena (Harrison et al., 2016); i.e., in certain applications, receiving a single event report may be sufficient, but in target positioning and tracking applications, which usually use trilateration or triangulation, at least three packets are needed. Besides, when the WSN is in charge of extracting the mobility pattern of targets, the higher the number of transmitted packets, the higher the accuracy (Misra et al., 2015).

Given the basic application requirements are met, the efficiency of WSN protocols is measured in terms of QoS

*Corresponding author

Email addresses: isleyma@upv.es (Israel Leyva-Mayorga), vpla@upv.es (Vicent Pla), jmartinez@upv.es (Jorge Martinez-Bauset), mriveroa@ipn.mx (Mario E. Rivero-Angeles)

parameters. Nodes being battery supplied, energy consumption is the QoS parameter most widely studied in the literature (Rault et al., 2014; Ye et al., 2002; Yahya and Ben-Othman, 2009; AlSkaif et al., 2015) as it directly affects network lifetime (period of time for which the network remains functional). The relevance of other QoS parameters such as report latency and event overlooking probability is application dependent.

In time-critical applications where event reporting is delay-sensitive, these QoS parameters are as important as energy efficiency since they determine whether the network is capable of satisfying the monitoring needs of the end user. Despite its importance, the performance of WSN protocols is oftentimes assessed in terms of mean values of report latency, which is clearly insufficient for time-critical applications. High percentiles or the whole probability distribution of report latency are much better suited and provide the network administrator with more meaningful information regarding the behavior of the system than its mean value, but the research in this area is scarce.

Building on this, we propose a novel method to calculate the probability distribution of report latency and mean energy consumption in cluster-based random access (RA) WSNs. In this method, we first obtain the probability mass function (pmf) of detecting nodes by simulation. Then we use Markov chains to model the process of RA event reporting, from where we analytically calculate the probability distribution of report latency and the mean energy consumption. We use this method to evaluate the performance of RA event reporting in applications that require the transmission of a certain number of packets to fully characterize the event. We also use our method to optimize the QoS of an event reporting WSN by identifying the optimal transmission parameters prior to network deployment for the selected application.

Note that the optimal transmission parameters depend on the number of detecting nodes for each event occurrence. Hence, one of the main challenges to overcome when trying to optimize event reporting is that the number of detecting nodes is not known before the event detection. Specifically, we focus on optimizing event report latency by identifying the ideal transmission parameters of detecting nodes in two different approaches, the *non-adaptive* and *adaptive backoff (ABO)*. In the *non-adaptive backoff*, transmission parameters are selected prior to network deployment and remain constant throughout event reporting. In the *adaptive backoff (ABO)*, transmission parameters are selected prior to network deployment and modified during the backoff of collided CMs. Note that implementing the adaptive backoff consumes a minimal amount of memory and computational power from the nodes, which is highly desirable.

Results show that increasing the time between successive transmissions during backoff can reduce the energy consumption during event reporting when compared to the non-adaptive backoff. Furthermore, the ABO mitigates the negative effects of the inaccurate selection of trans-

mission parameters. This results in a noticeable increase in the stability of QoS parameters when compared to the non-adaptive backoff. These results hold true even when the network operates in a multi-event environment, where it is in charge of monitoring multiple types of events.

The rest of the paper is organized as follows. Section 2 presents research performed in recent years for time-critical applications and QoS analysis of WSN protocols. The network model used for QoS analysis is described in Section 3. Then, we present our novel method for evaluating QoS in RA WSN protocols in Section 4. Here, we use a Markov process to obtain the energy consumption and report latency in cluster-based WSNs. Results are presented in Section 5, where we illustrate the capabilities of our hybrid method, along with the advantages of an adaptive backoff. The article concludes with a summary of results and future work.

2. Related work

As power electronics evolve, the QoS of WSNs in time-critical applications has become an important topic (Monaco et al., 2006). Protocols that were specially designed to improve network lifetime such as (Ye et al., 2002; Heinzelman et al., 2002; Younis and Fahmy, 2004) have served as a base for more complex protocols that aim to reduce report latency and packet loss probability in time-constrained packets without compromising the energy efficiency.

Energy consumption being so important in WSN, techniques such as sleep scheduling strategies and multi-hop delivery have been proposed to further enhance this parameter (Guo et al., 2012). While these techniques reduce energy wastage, they may increase report latency (Ye et al., 2002; Sharif et al., 2010; Chen, 2015) and network congestion (Marco et al., 2011).

Hybrid protocols are a distributed energy-efficient solution for time-critical applications (Younis and Fahmy, 2004; Manjeshwar and Agrawal, 2002). These protocols are capable of adapting its behavior depending on the characteristics of the application. The downside of these protocols is, usually, an increase in complexity (Leyva-Mayorga et al., 2014, 2015b).

In previous studies, we have presented and evaluated the performance of the RA phases of hybrid protocols by means of Markov models (Leyva-Mayorga et al., 2014, 2015b), but our results were limited to mean values of energy consumption and report latency. For instance, in RA WSN protocols, report latency is most frequently assessed in terms of its mean value (Jamieson et al., 2006; Lee and Choi, 2015; Liang et al., 2011), which is clearly insufficient for time-critical applications. In WSN protocols with a scheduled transmission scheme, assessing the report latency by means of its worst-case value may be adequate for time-critical applications. Such is the case of (Zheng et al., 2016), where a MAC protocol for industrial applications is proposed and analyzed by means of a discrete-time Markov chain (DTMC). Nevertheless, the

worst-case latency is not adequate for the QoS analysis of RA protocols. Instead, high percentiles or the whole distribution of report latency are much better suited.

Generic analytic methods for evaluating the performance of RA protocols in wireless networks may not be easily adapted for the analysis of cluster-based WSN due to their particular characteristics. Such is the case of the queuing model for the IEEE 802.11, presented in (Tickoo and Sikdar, 2008). In addition, most of the existing methods for obtaining the probability distribution of report latency in WSNs are protocol specific; e.g., (Souil et al., 2014; Siddiqui and Ghani, 2013). More specifically, a QoS evaluation of the AMPH protocol is conducted in (Souil et al., 2014), but despite the in-depth analysis, the selected network topology and protocol compromise the adaptability of the developed method.

In (Siddiqui and Ghani, 2013), authors obtain the probability distribution of a successful and failed transmission by means of a Markov Chain. Note that in both (Souil et al., 2014; Siddiqui and Ghani, 2013), delay is calculated as the time required for the first successful transmission to occur, whereas the transmission of a minimum number of packets may be necessary for the accurate characterization of the occurring event. In (Wang et al., 2011), authors do consider the need for the transmission of a minimum number of event packets during event reporting and propose a spatio-temporal fluid model, along with a simplified model to obtain the distribution of report delay in multi-hop WSNs.

We identified the need for a method capable of calculating the report latency distribution for a wide range of WSN protocols and environments. Hence we developed a hybrid method for obtaining the distribution of report latency in (Leyva-Mayorga et al., 2015a). We extend our previous study and evaluate its adaptability by analyzing the benefits of adaptive backoff transmission probabilities in RA event reporting WSNs. We now also use this method to identify, prior to network deployment, the transmission probabilities that minimize report latency for the non-adaptive and adaptive backoff and also reduce energy consumption. To the best of our knowledge, no other study has presented a method for calculating the distribution of report latency in WSNs as adaptable as ours. Furthermore, our work differs significantly from most of the research related to multi-object optimization (MOO) in sensor networks, which usually involves the use of evolutionary algorithms and simulators such as NS2 for the QoS analysis (Iqbal et al., 2016).

3. Network model

In this Section, we describe the network model used for the QoS analysis during the present study. It is similar to the one we have used in previous studies (Leyva-Mayorga et al., 2014, 2015b), where a squared $100\text{ m} \times 100\text{ m}$ area, from coordinates $(0,0)$ to $(100,100)$, containing $M = 100$

nodes is considered. The sink node is located outside the supervised area at the coordinates $(200, 0)$.

During event reporting and CF phases, the network operates on a slotted channel. During event reporting, each time slot is the time required for the transmission of a data packet from CM to the CH and its immediate retransmission to the sink node. The size of the data packet, $l = 2\text{ kbits}$, comprises the data payload, the identification field, Id , and a *type* field. The size of the control packet is $l_c = 1\text{ kbits}$, which comprises the same fields but with a shorter payload. Since the transmission bit rate is $R_b = 40\text{ kbps}$ and two data packets are transmitted; i.e., one from the CM to the CH and one from the CH to the sink node, the slot duration is $t_s = 0.1\text{ s}$.

The energy to receive a packet depends on its length and on the energy required per bit by the communication circuits, E_{elec} , as

$$E_{rx}(l) = l \cdot E_{elec}. \quad (1)$$

In this study we adopt a generic energy consumption model that is widely used in the literature, in which $E_{elec} = 50\text{ nJ/bit}$ (Heinzelman et al., 2002; Younis and Fahmy, 2004; Chen, 2015); though any energy consumption model can be selected in our hybrid method.

Our study focuses on cluster-based protocols, where the role of nodes acting as either CHs or CMs shifts constantly throughout the operation of the network in order to avoid fast battery depletion of nodes acting as CHs. Throughout this study we use one of the most important clustering algorithms: LEACH (Heinzelman et al., 2002). Through the years, LEACH has served as a base to develop and to assess the efficiency of other routing protocols (Chen, 2015; Afsar and Tayarani-N, 2014). However, as our hybrid method presents a general structure, any clustering algorithm can be easily incorporated.

During event reporting phases, the communication circuits in every CH must be active whenever CM transmissions can occur so they are able to relay any received data packet containing an alarm message from its CMs to the sink node with minimal report latency. Then, the energy required to transmit a packet depends on both the length of the packet, l , and the selected transmission range, d , as calculated in (Heinzelman et al., 2002). Therefore, the total energy consumed during a packet transmission is given as

$$E_{tx}(l, d) = l \cdot E_{elec} + l \cdot \epsilon_{amp} \cdot d^{P_l}, \quad (2)$$

where the energy required per bit and per square meter by the transmission amplifier is $\epsilon_{amp} = 10\text{ pJ/bit/m}^2$ and P_l is the path loss exponent.

Two power levels are defined for packet transmissions: low and high power. Low power transmissions consume E_{cmtx} J and are used for CM to CH communication. By using this power level, nodes are able to perform transmissions for up to $d_l = 35\text{ m}$. High power transmissions consume E_{chtx} J and are used for CH to sink communication. By choosing this power level, nodes are able to

transmit from the farthest possible coordinates within the network to the sink node; i.e., $d_h = \sqrt{200^2 + 100^2}$ m. This approach eliminates the need of calculating the minimum energy required for transmission as proposed in (Heinzelman et al., 2002). It also reduces packet loss probability due to changes in the wireless environment, i.e., the minimum energy required for packet transmission may vary through time. Furthermore, energy efficiency is not significantly affected by following this approach when compared to the minimum energy method, as the increase in power consumption is minimal.

Events are generated as in (Calafate et al., 2010), where a model for indoor gas propagation is presented. Each event has originating coordinates (w, h) , which are selected by means of a two-dimensional uniform random variable. A node detects an event when the reading of the physical parameter of interest exceeds the threshold in one or more of its sensors; i.e., when the node is located within $R \in \{5, 10, 15, 20, 30\}$ m from the event originating coordinates.

In order to evaluate the performance of RA protocols in extreme conditions, it is assumed that the event is detected simultaneously by all the CMs within a distance, R , from the event originating coordinates; the CMs located at a distance larger than R from the event originating coordinates do not detect the event.

Given the random nature of the event occurring coordinates, the total number of detecting CMs, N_{tot} , the number of clusters with detecting CMs, N_c , and the number of detecting CMs in the i th cluster, N_i ($i \in \{1, 2, \dots, N_c\}$), can be different for each event occurrence.

We assume that a double sliding window scheme (Haining Shu and Qilian Liang, 2006) is implemented in order to limit the number of transmissions generated by each event. As a result, only one event packet is generated for each detecting CM at the very beginning of each event.

As in previous work (Leyva-Mayorga et al., 2014, 2015b), we assume that the buffer on the nodes is limited and can allocate a single event report. The arrival of events while a previous event is being reported in a cluster is not considered because this scenario has an extremely low probability of occurrence.

Upon the occurrence of a new event, detecting nodes attempt transmission with probability τ per time slot and a backoff policy is used for collision handling. The transmission probability of the CMs for the first event report is τ . The transmission probability of collided CMs (CBMs) becomes $\beta = \tau/B$ for the subsequent transmissions, where $B \geq 1$. $B = 1$ and $B > 1$ correspond to the non-adaptive and adaptive backoff respectively. Whenever the CH receives an event packet, it is directly transmitted to the sink node in the same time slot in order to minimize report latency. At the end of event reporting, nodes reset their transmission probabilities to the initial value, τ .

Code Division Multiple Access (CDMA) is used to avoid inter-cluster collisions. For this, a CDMA code is selected per cluster and used for transmissions from CMs to the CH and from the CH to sink (Hu, 1991). Consequently,

during CF and event reporting phases, collisions can only occur between CMs from the same cluster, so each cluster operates independently during RA. As such, we are able to analyze event reporting in terms of independent clusters by means of a discrete-time Markov Chain (DTMC) (Leyva-Mayorga et al., 2015b). A basic diagram of the operation of cluster-based protocols for event detection and reporting is shown in Fig. 1.

The number of messages required to fully characterize the occurring phenomena is denoted by k . Once the sink node receives k event messages, it reacts accordingly. Conversely, in cases where $k > N_{\text{tot}}$, the event is overlooked.

During event reporting, the network can be configured to transmit either N_i or k data packets per cluster. In the former, every detecting CM transmits its data packet, which results in N_i transmissions to the sink node from the i th cluster. In the latter, CMs sense the medium during event reporting until the k th packet is successfully transmitted. Then, $N_i - k$ CMs in the i th cluster discard its data packet, which reduces energy wastage. Fig. 2 illustrates the process of event reporting for the considered set of applications.

It is worth noting that, since CDMA is used for avoiding inter-cluster collisions, CMs are only aware of successful transmissions within its own cluster. As a result, event reports can be restricted within clusters but not within the entire network. As a result, whenever each cluster is set to send k packets and an event is detected in $N_c > 1$ clusters, more than k packets can be received at the sink node.

4. Hybrid method for QoS analysis

In this section we describe our hybrid method for QoS analysis. This method comprises three main phases:

1. *Obtaining the distribution of the number of detecting nodes:* The probability distribution of the number of nodes that detect the event simultaneously is needed to effectively calculate the QoS parameters. Since we focus on cluster-based WSN protocols and each cluster operates independently during event reporting (due to the use of CDMA), we obtain the distribution of the number of detecting CMs per cluster, N , and detecting clusters (number of clusters with, at least, one detecting CM), N_c , by simulation. For this, we have developed a discrete event simulator in C, in which a clustering algorithm is implemented, along with the generation and detection of hazardous events.

In each simulation the nodes are first randomly distributed within the area of interest; their coordinates are selected by means of two uniform random variables. Then the nodes are organized in clusters according to the selected clustering algorithm; 20 CF phases are performed in each simulation. Next, 1000

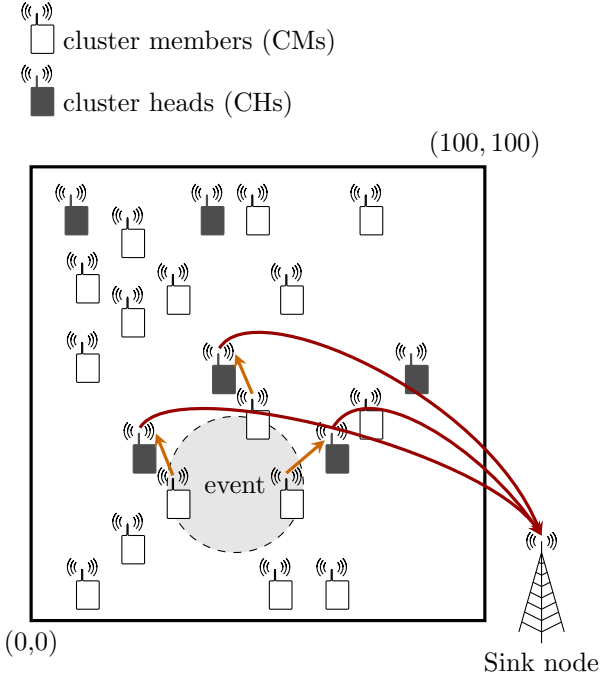


Figure 1: Basic diagram of a cluster-based wireless sensor network (WSN) for event reporting.

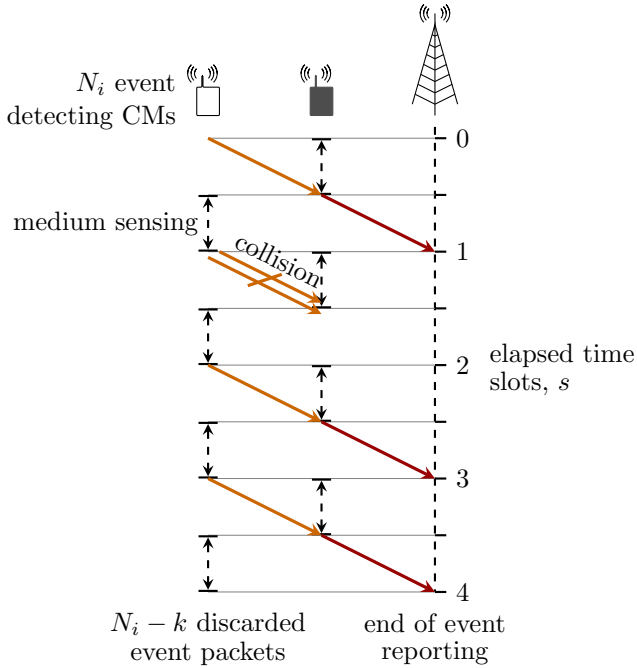


Figure 2: Random access event reporting in time-critical applications with medium sensing, $k = 3$.

events are generated after each CF phase and, finally, the number of detecting nodes is obtained and stored. Simulations were performed until the error between the distribution of detecting CMs obtained until the j and the $(j - 1)$ th simulation is less than $\epsilon = 10^{-5}$.

In this study, we assume that the cluster formation phase is conducted according to that of the LEACH protocol (Heinzelman et al., 2002); still, any clustering algorithm can be selected.

The distributed clustering algorithm of the LEACH protocol is performed as follows. The operation of the network is divided into rounds and a CF phase is performed in each round. At the beginning of each round, each node has a certain probability of being elected as a CH; this probability increases with the number of rounds. Once a node has been elected as a CH, the probability of (once more) being elected as a CH in the next few rounds becomes 0. Afterwards, the CHs inform its status to the other nodes by broadcasting an advertisement message (using CDMA). The remaining nodes join a cluster based on the received signal strength of the CH transmission; this is, typically, the one whose CH is located at the shortest distance from the node. Please refer to (Heinzelman et al., 2002) for more details on the LEACH protocol.

2. *Defining the Markov reward process:* A discrete-time Markov chain (DTMC) that describes the Markov reward process that occurs during RA event reporting is defined. For this, rewards are determined by the energy consumption at each system transition. This is the phase in which the specific characteristics of the WSN protocol are included, which makes the method highly adaptable.
3. *Obtaining the QoS parameters:* Once the Markov reward process has been defined, the probability distribution of report latency and mean energy consumption during event reporting for a specific scenario can be calculated. Specifically, the probability distribution of report latency is calculated as the s step probability for the transmission of the first k packets. The mean energy consumption during event reporting is obtained by solving a set of Bellman equations.

These phases are now explained in detail.

4.1. Obtaining the distribution of the number of detecting nodes

In our method, the probability mass function (pmf) of the number of detecting CMs and clusters is needed for accurately calculating the energy consumption and report latency. We have developed a discrete event simulator to obtain $P(N = n|N_c = i)$, the pmf of the number of detecting CMs per cluster, N , conditioned on the number of detecting clusters, N_c . For the sake of illustration, Fig. 3 shows the cumulative distribution function (CDF) of the number of detecting CMs, $P(N \leq n|N_c = i)$, for a collection of event radii, R .

We also used our simulator to obtain the probability of an event affecting N_c clusters for the studied event detection radii, $P(N_c = i)$; the CDF, $P(N_c \leq i)$ is shown in Fig. 4.

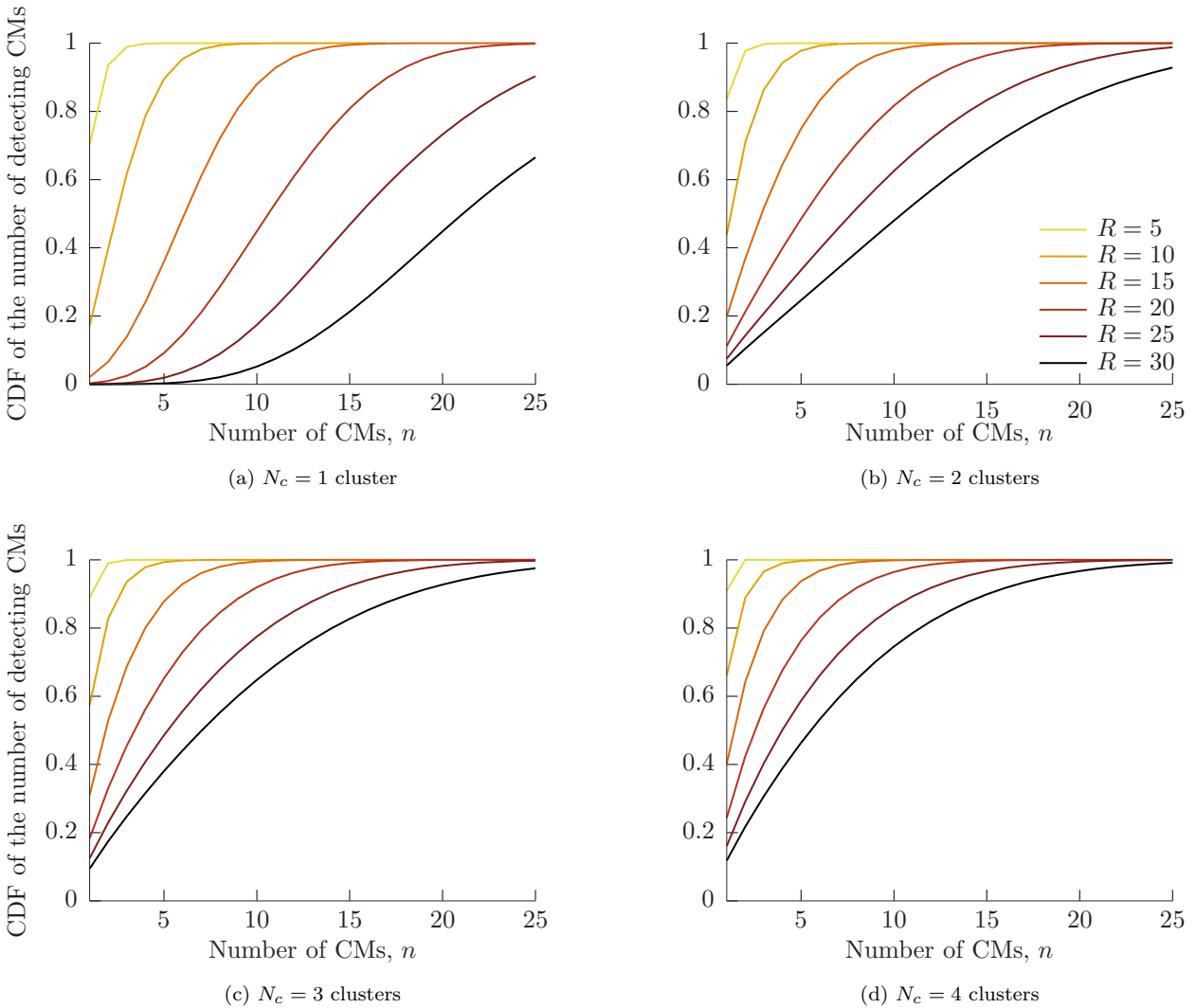


Figure 3: CDF of the number of detecting CMs per cluster, N , given the event is detected in N_c clusters, $P(N \leq n | N_c = i)$, for $N_c \in \{1, 2, 3, 4\}$ and $R \in \{5, 10, 15, 20, 25, 30\}$ m.

Once we have obtained the distribution of N , N_{tot} and N_c for the selected detection radii, we proceed to define the Markov reward process that describes the event reporting WSN protocol.

4.2. Defining the Markov Reward Process

Since we are able to analyze the system in terms of independent clusters due to the use of CDMA, we build a discrete-time Markov Chain (DTMC) that describes the Markov reward process of event reporting within each cluster. Fig. 5 shows a basic model that describes event reporting with transmission probability τ and a non-adaptive backoff. This model is similar to the one we have used in our previous studies (Leyva-Mayorga et al., 2014, 2015a,b), where the transmission probability remains unaffected during backoff ($B = 1$).

The model starts at state (N), where N is the number of CMs that have detected the event. Note that N can be

different for each event. The states represent the number of CMs with pending event transmissions, n . At each time slot, the DTMC can either transition towards the absorbing state ($N - k$), $1 \leq k \leq N$, or remain in the same state. The transition probability from state (n) to ($n - 1$), p_n , is the probability of a successful transmission, which occurs whenever a single CM sends a data packet. Conversely, the probability of remaining in the same state is the probability of an unsuccessful event report, $q_n = 1 - p_n$, i.e., none of the CMs attempts transmission or a collision occurs. Since each of the n CMs attempts transmission with probability τ , transition probabilities depend on τ and n as

$$p_n = P(S_n = 1) = n\tau(1 - \tau)^{n-1}, \quad (3)$$

where S_n is the number of event packet transmissions at a specific time slot.

The absorbing DTMC that describes the process of event reporting; i.e., the transmission of the first $\ell =$

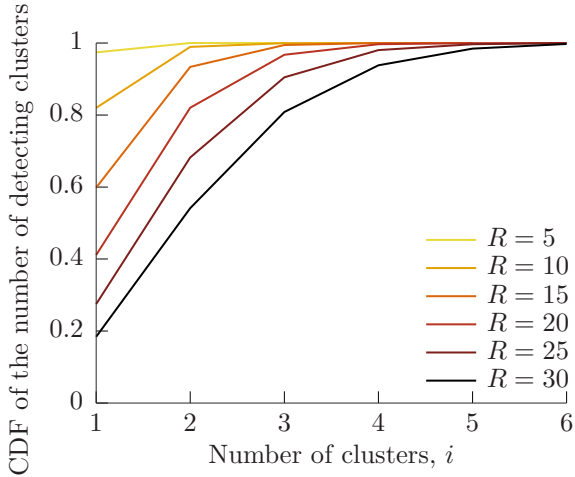


Figure 4: Probability distribution of the number of detecting clusters, N_c , for $R \in \{5, 10, 15, 20, 25, 30\}$ m.

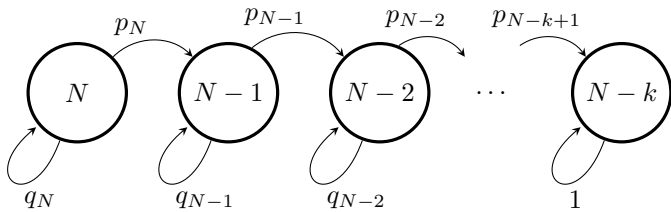


Figure 5: Markov model for random access event reporting over a slotted channel.

$\min\{k, N\}$ messages within a cluster is depicted in Fig. 5. Then, the substochastic matrix that represents the transitions within transient states is given as

$$\mathbf{P} = \begin{bmatrix} q_N & p_N & 0 & \cdots & 0 \\ 0 & q_{N-1} & p_{N-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & p_{N-\ell+2} \\ 0 & 0 & 0 & \cdots & q_{N-\ell+1} \end{bmatrix}$$

Note that we use the variable ℓ instead of k . The reason for this is that the number of detections within a cluster, N , can be less than k ; still, these cases must be taken into account in order to obtain the pmf of report latency (see Section 4.3).

Also, from Fig. 5 we derive a Markov reward process that will allow us to calculate the mean energy consumption during event reporting, \bar{E} . For this, rewards are given as the mean energy consumption at each system transition. In basic RA (no medium sensing) successfully transmitting an event packet requires a CM to CH (E_{cmtx}) and a CH to sink (E_{chtx}) transmission. Thus, the transition from state (n) to $(n-1)$ has a reward

$$\hat{r}(p_n) = E_{\text{cmtx}} + E_{\text{chtx}}. \quad (4)$$

In case no transmission is attempted by the CMs or a collision occurs, the system remains in the same state. As

a result, the reward for remaining in state n ,

$$\hat{r}(q_n) = E_{\text{cmtx}} \cdot \mathbb{E}[S_n | S_n \neq 1] = \frac{n\tau - p_n}{1 - p_n} E_{\text{cmtx}}, \quad (5)$$

is given as the energy consumed by a CM transmission multiplied by the mean number of transmissions, conditioned to the fact that no successful transmission occurred.

In order to identify the successful and failed transmissions within the cluster, the CMs must be set to perform medium sensing during event reporting. This enables the CMs to discard any remaining event packets once k messages have been successfully transmitted. Here, the reward for the transition from state (n) to $(n-1)$ becomes

$$\begin{aligned} r(p_n) &= r(p_n) + (n-1)E_{\text{elec}} \\ &= E_{\text{cmtx}} + E_{\text{chtx}} + (n-1)E_{\text{elec}}, \end{aligned} \quad (6)$$

and remaining in the same state, n , has a reward given as

$$\begin{aligned} r(q_n) &= r(q_n) + (n - \mathbb{E}[S_n | S_n \neq 1]) \cdot E_{\text{elec}} \\ &= \frac{(E_{\text{cmtx}} - E_{\text{elec}})(n\tau - p_n)}{1 - p_n} + n \cdot E_{\text{elec}}. \end{aligned} \quad (7)$$

We use these rewards to construct the reward matrix,

$$\mathbf{R} = \begin{bmatrix} r(q_N) & r(p_N) & 0 & \cdots & 0 \\ 0 & r(q_{N-1}) & r(p_{N-1}) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & r(p_{N-\ell+2}) \\ 0 & 0 & 0 & \cdots & r(q_{N-\ell+1}) \end{bmatrix}$$

4.2.1. Adaptive backoff (ABO)

Given the adaptive backoff (ABO) is implemented, N nodes initiate each event reporting process with a transmission probability, τ . Then, whenever a collision occurs, the implicated CMs modify their transmission probability during backoff by $\beta = \tau/B$, where $B \geq 1$ ¹ as collisions indicate that several CMs are indeed competing for medium access and higher transmission probabilities increase collision probability. $B = 1$ corresponds to the non-adaptive backoff approach. The number of CMs that have caused a collision and perform backoff (BCMs) is ν . Once event reporting is concluded, the transmission probability of CMs is restored to its original value, τ . This is a more general case than the one presented in the previous section, where we developed a DTMC for the case where $B = 1$. Hence, we now construct a two-dimensional DTMC for the case where $B \geq 1$.

In this model, the state of the system, (n, ν) , is determined by the number of detecting CMs that have not yet attempted transmission, n , and the number of BCMs, ν . The transmission probabilities of the n and ν CMs are τ and β respectively. The DTMC begins at state $(N, 0)$; i.e.,

¹The case where $B < 1$ will rarely enhance the performance of RA protocols

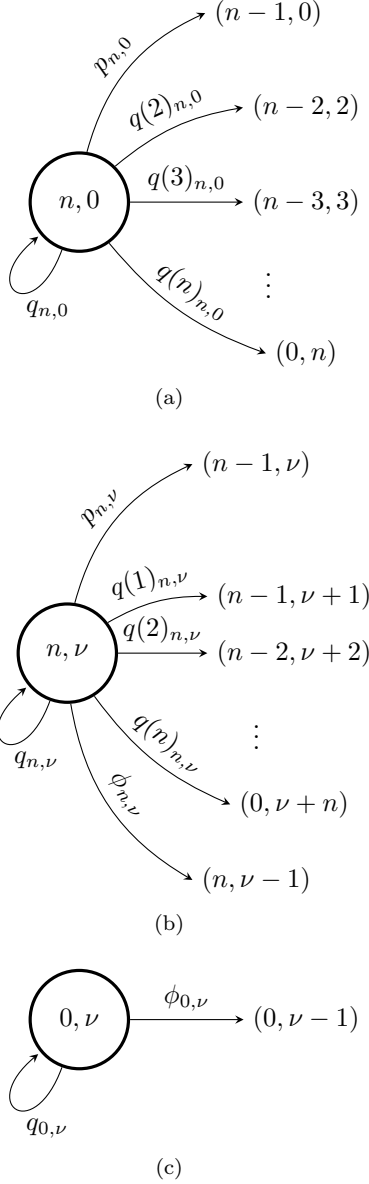


Figure 6: Characteristic states and transitions of the two-dimensional DTMC: (a) $(n, 0)$, (b) (n, ν) and (c) $(0, \nu)$.

N CMs detecting the event and 0 CMs performing backoff, BCMs, and evolves towards state $(0, 0)$.

For the sake of simplicity, we define every possible state as either $(n, 0)$, (n, ν) or $(0, \nu)$. Consequently, the possible transitions from states $(n, 0)$, (n, ν) and $(0, \nu)$ (shown in Fig. 6) define every possible transition of the two-dimensional DTMC. Note that transitions depend on the number of CMs and BCMs that perform a transmission at a given time slot, defined by S_n and S_ν respectively.

Transitions from state (n, ν) to $(n-1, \nu)$ occur with probability

$$\begin{aligned} p_{n,\nu} &= P(S_n = 1) \cdot P(S_\nu = 0) \\ &= n\tau(1-\tau)^{n-1}(1-\beta)^\nu, \end{aligned} \quad (8)$$

which represents a successful transmission from one out of

the n CMs. Transitions from state (n, ν) to $n, \nu - 1$ occur with probability

$$\begin{aligned} \phi_{n,\nu} &= P(S_n = 0) \cdot P(S_\nu = 1) \\ &= \nu\beta(1-\beta)^{\nu-1}(1-\tau)^n, \end{aligned} \quad (9)$$

which represents a successful transmission from a BCM. Therefore, the probability of a successful transmission is

$$p_{n,\nu} + \phi_{n,\nu}. \quad (10)$$

The probability of remaining in the same state is

$$\begin{aligned} q_{n,\nu} &= P(S_n = 0) \cdot P(S_\nu \neq 1) \\ &= (1-\tau)^n [1 - \nu\beta(1-\beta)^{\nu-1}], \end{aligned} \quad (11)$$

which occurs whenever none of the n CMs perform transmission and no successful transmission occurs.

Transitions from (n, ν) to $(n-i, \nu+i)$, $q^{(i)}_{n,\nu}$, are divided in two cases, namely $i = 1$ and $i \geq 2$. In the former, one out of the n CMs and at least one of the ν BCMs attempt transmission, which occurs with probability

$$q^{(1)}_{n,\nu} = P(S_n = 1) \cdot P(S_\nu \neq 0). \quad (12)$$

In the latter, $i \geq 2$ of the n CMs cause a collision, which occurs with probability

$$q^{(i)}_{n,\nu} = P(S_n = i | i \geq 2). \quad (13)$$

Note that in case $i \geq 2$, the number of transmitting BCMs, ν , is irrelevant.

Now that every transition probability has been defined, we build the substochastic matrix that represents the transitions between transient states as

$$\mathbf{P}_A = \begin{bmatrix} q_{N,0} & 0 & q^{(2)}_{N,0} & \cdots & 0 \\ 0 & q_{N-1,1} & q^{(1)}_{N-1,1} & \cdots & 0 \\ 0 & 0 & q_{N-2,2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & q^{(1)}_{1,N-\ell} \\ 0 & 0 & 0 & \cdots & q_{0,N-\ell+1} \end{bmatrix}$$

The dimension of \mathbf{P}_A , δ , depends on the initial number of detecting nodes, N , and k as

$$\delta = \sum_{j=0}^{\ell-1} N + 1 - j = \ell(N+1) - \sum_{j=0}^{\ell-1} j. \quad (14)$$

That is, each state $(n, 0)$ has $n+1$ possible transitions to states in which the total number of contending CMs remains unaffected; only one transition, $p_{n,0}$, reduces the total number of contending CMs by 1.

From there, we build the reward matrix, \mathbf{R}_A , for obtaining the energy consumption in the same manner as for \mathbf{R} . A successful transmission has a reward given by

$$\begin{aligned} r_A(p_{n,\nu}) &= r_A(\phi_{n,\nu}) \\ &= E_{\text{cmtx}} + E_{\text{chtx}} + (n + \nu - 1)E_{\text{elec}}. \end{aligned} \quad (15)$$

The reward for remaining in the same state, i.e., when none of the n CMs and either none or more than two of the ν BCMs perform a transmission, $S_\nu \neq 1$, is given as

$$r_A(q_{n,\nu}) = \frac{(E_{\text{cmtx}} - E_{\text{elec}})(\nu\beta - \phi_{n,\nu})}{1 - \phi_{n,\nu}} + (n + \nu)E_{\text{elec}} \quad (16)$$

and the reward for transition from the state (n, ν) to $(n - i, \nu + i)$ is

$$r_A(q(i)_{n,\nu}) = r_A(q_{n,\nu}) - i \cdot E_{\text{elec}} + i \cdot E_{\text{cmtx}}. \quad (17)$$

Building on this, the resulting reward matrix is

$$\mathbf{R}_A = \begin{bmatrix} r_A(q_{N,0}) & 0 & r_A(q(2)_{N,0}) & \cdots & 0 \\ 0 & r_A(q_{N-1,1}) & r_A(q(1)_{N-1,1}) & \cdots & 0 \\ 0 & 0 & r_A(q_{N-2,2}) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & r_A(q(1)_{1,N-\ell}) \\ 0 & 0 & 0 & \cdots & r_A(q_{0,N-\ell+1}) \end{bmatrix}$$

As in the transition matrix, we generate a different reward matrix for each possible value of N . By building these matrices, we are now able to conduct the performance analysis of event reporting for any given value of $B \in [1, \infty)$.

4.3. Obtaining the QoS parameters

Once we have constructed the transition and reward matrices, \mathbf{P} , and \mathbf{R} respectively, we proceed to calculate the energy consumption during event reporting. First we obtain the energy consumption in each state when for the RA with the non-adaptive backoff, $B = 1$, by solving the Bellman equations (Bellman, 1957),

$$\overline{E}_n = q_n [r(q_n) + \overline{E}_n] + p_n [r(p_n) + \overline{E}_{n-1}]. \quad (18)$$

That is, (18) can be solved either as a set of linear equations or recursively for all n , given the initial condition $\overline{E}_{N-\ell} = 0$; i.e., the energy consumption at the absorbing state is 0. If $B \neq 1$, the energy consumption in each state is calculated similarly for each possible state, (n, ν) , is given as

$$\begin{aligned} \overline{E}_{n,\nu} = & q_{n,\nu} [r(q_{n,\nu}) + \overline{E}_{n,\nu}] \\ & + p_{n,\nu} [r(p_{n,\nu}) + \overline{E}_{n-1,\nu}] \\ & + \phi_{n,\nu} [r(\phi_{n,\nu}) + \overline{E}_{n,\nu-1}] \\ & + \sum_{j=1}^n q(j)_{n,\nu} [r(q(j)_{n,\nu}) + \overline{E}_{n-j,\nu+j}]. \end{aligned} \quad (19)$$

Recall that we have obtained the probability distribution of the number of clusters with detecting CMs, $P(N_c = i)$, and the probability distribution of the number of detecting CMs given N_c detecting clusters, $P(N = n|N_c = i)$,

by simulation. With this information, the mean energy consumption during event reporting in a given scenario, \overline{E} , is obtained by multiplying the energy consumed when a given number of CMs and clusters detect the event by its probability of occurrence as

$$\overline{E} = \sum_{i=1}^{N_{\text{cmax}}} i \cdot P(N_c = i) \times \sum_{n=1}^{N_{\text{max}}} P(N = n|N_c = i) \overline{E}_n; \quad (20)$$

where N_{max} and N_{cmax} are the maximum values of N and N_c . Note that the mean energy consumption is calculated similarly for the case where $B \neq 1$ and is given as

$$\overline{E} = \sum_{i=1}^{N_{\text{cmax}}} i \cdot P(N_c = i) \times \sum_{n=1}^{N_{\text{max}}} P(N = n|N_c = i) \overline{E}_{n,0}, \quad (21)$$

since $\nu = 0$ at the beginning of event reporting.

To evaluate the report latency, let T_ℓ be the random variable that defines the number of time slots elapsed between the occurrence of an event detected by N CMs and the end of the event reporting process in a single cluster. Recall that event reporting in a given cluster is completed when $\ell = \min\{k, N\}$ out of a total N packets are transmitted successfully from the detecting CMs to the sink node. As such, T_ℓ is the time to absorption in an absorbing DTMC; absorption occurs when the ℓ th packet is successfully transmitted.

T_ℓ follows a phase-type distribution with representation $(\boldsymbol{\alpha}, \mathbf{P})$, where vector $\boldsymbol{\alpha}$ defines the probabilities that the system starts from each of the transient states and \mathbf{P} is the substochastic matrix of transient states. Let $\mathbf{h} = \mathbf{1} - \mathbf{P}\mathbf{1}$ denote the absorption probabilities; $\mathbf{1}$ is a column vector of 1s of the same dimension as \mathbf{P} .

For the system depicted in Fig. 5 (non-adaptive backoff), $\boldsymbol{\alpha} = [1 \ 0 \ \dots \ 0]$ and $\mathbf{h} = [0 \ \dots \ 0 \ p_{N-\ell+1}]$; the dimension of both vectors is ℓ .

For the system depicted in Fig. 6 (ABO), $\boldsymbol{\alpha} = [1 \ 0 \ \dots \ 0]$ and $\mathbf{h} = [0 \ \dots \ 0 \ p_{N-\ell+1,0} \ \dots \ p_{0,N-\ell+1}]$; the dimension of both vectors is δ (see eq. (14)).

It is known that $P(T_\ell = s) = \boldsymbol{\alpha} \mathbf{P}^{s-1} \mathbf{h}$, and the expected value of T_ℓ is given by (Alfa, 2010),

$$\overline{T}_\ell = \boldsymbol{\alpha} (\mathbf{I} - \mathbf{P})^{-1} \mathbf{1}, \quad (22)$$

where \mathbf{I} is the identity matrix of the same dimension as \mathbf{P} . Then, the mean report latency for a given scenario, described by $P(N_c = i)$ and $P(N = n|N_c = i)$, can be calculated as

$$\overline{T} = \sum_{i=1}^{N_{\text{cmax}}} P(N_c = i) \times \sum_{n=1}^{N_{\text{max}}} P(N = n|N_c = i) \overline{T}_\ell. \quad (23)$$

However, as stated above, we are especially interested in obtaining the probability distribution of the report latency, T . For this, let $f_c(j, s|i)$ be the probability that $j \in \{0, 1, \dots, \ell\}$ packets are successfully transmitted within c clusters in s time slots conditioned on the number of detecting clusters, i . Also let $f(j, s)$ be the probability that

j packets are successfully transmitted in one cluster in s time slots or less after an event is detected by N CMs. Then, $f_c(j, s|i)$ can be determined by the following recursion,

$$f_c(j, s|i) = \sum_{u=0}^j f_{c-1}(u, s|i) f_1(j-u, s|i), \quad \forall c \in \{1, 2, \dots, i\} \quad (24)$$

where

$$f_1(j, s|i) = \sum_{n=1}^{N_{\max}} P(N=n|N_c=i) f(j, s) \quad (25)$$

is the probability that j packets are successfully transmitted within one cluster in s time slots or less after an event is detected by i clusters.

For the system depicted in Fig. 5 (non-adaptive back-off), it is given by,

$$f(j, s) = \alpha \mathbf{P}^s \mathbf{e}_j, \quad j \in \{1, 2, \dots, \ell-1\}, \quad (26)$$

where \mathbf{e}_j is a column vector of zeros of dimension ℓ , except it has a 1 at position $j+1$. For the system depicted in Fig. 6 (ABO), $f(j, s)$ is obtained in a similar manner.

For $j = \ell$, we know that,

$$P(T_\ell > s) = \alpha \mathbf{P}^s \mathbf{1}. \quad (27)$$

Note that in (27) if $s \leq \ell-1$, then $P(T_\ell > s) = 1$. Clearly,

$$f(\ell, s) = 1 - P(T_\ell > s). \quad (28)$$

Finally, when the number of affected clusters is N_c and the number of event packets required at the sink is k , the cumulative distribution function (CDF) of T can be written as

$$\begin{aligned} P(T \leq s; N_c = i, k) &= \sum_{j \geq k} f_i(j, s|i) \\ &= 1 - \sum_{j < k} f_i(j, s|i). \end{aligned} \quad (29)$$

Now, we are able to obtain the report latency distribution for specific environments as

$$P(T \leq s; k) = \sum_{i=1}^{N_{\max}} P(T \leq s; N_c = i, k) \cdot P(N_c = i), \quad (30)$$

given the distribution of the number of detecting nodes and clusters is known. As such, in Section 5 we study the impact that transmission probabilities and event detection radii have on performance.

Table 1: Network parameters.

Parameter	Value
Area Size	100 m × 100 m
Sink Node location	(200, 0)
Deployed nodes	$M = 100$
Data packet length	$l = 2$ kbits
Control packet length	$l_c = 1$ kbits
Data bit rate	$R_b = 40$ kbps
Time slot duration	$t_s = 0.1$ s
Energy per bit required by the communication circuits	$E_{\text{elec}} = 50$ nJ/bit
Transmission range	$d_l = 35$ m
	$d_h = \sqrt{200^2 + 100^2}$ m
Path loss exponent	$P_l = 2$
Amplifier energy	$\epsilon_{\text{amp}} = 10$ pJ/bit/m

5. QoS analysis

Results presented in this section are used to highlight the capabilities of the proposed method for QoS analysis and also to showcase the robustness of an adaptive backoff (ABO) to the inadequate selection of parameters. The relevant network parameters we use for QoS analysis are listed in Table 1.

Herein, we assume that, at least, k must be received at the sink node to accurately characterize the event. In cases where $N_{\text{tot}} < k$, event reporting is unsuccessful as the sink is unable to obtain the necessary information for characterizing the event. Since we have already obtained the probability distribution of the number of detecting nodes by simulation, we can easily calculate the event overlooking probability, $P(N_{\text{tot}} < k)$, for the given values of R and k .

From Fig. 7 we observe that the event overlooking probability, $P(N_{\text{tot}} < k)$, sharply reduces as R increases. Hence, the network administrator must configure the thresholds in the sensors to achieve an adequate detection radius, which consequently reduces event overlooking probability. Obviously, the lower limit for the event detection radius, R , highly depends on the density of deployed nodes within the network. For the considered node density, $D = 0.01$ nodes/m², and $k = 3$, we assume that $R \geq 15$ m results in an adequate $P(N_{\text{tot}} < 3) < 0.1$. As such, for the rest of the paper we assume that event reporting is terminated at the time in which the sink node receives $k = 3$ event packets. Building on this, we now focus on assessing the energy consumption and report latency for large event detection radii.

To study the energy consumption in RA event reporting, we compare two approaches for event reporting. In the first, each of the N detecting nodes is set to transmit a packet to the CH. In the second, CMs sense the medium in order to identify the k th successfully sent packet. Then, the remaining packets are discarded in order to reduce en-

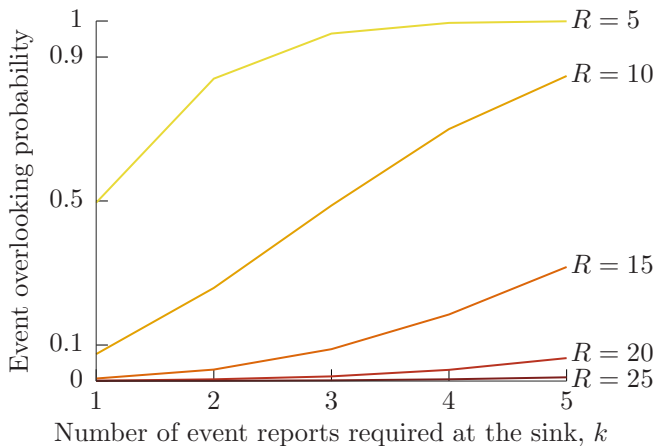


Figure 7: Event overlooking probability, $P(N_{\text{tot}} < k)$, for the considered node density, $D = 0.01$ nodes/m²; $k \in \{1, 2, \dots, 5\}$ and $R \in \{5, 10, 15, 20, 25\}$.

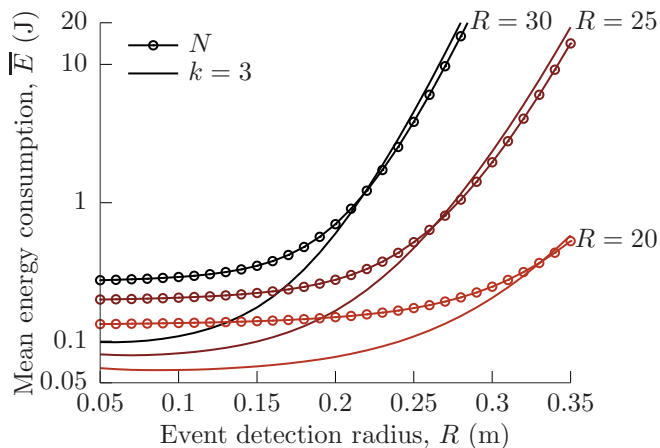


Figure 8: Mean energy consumption, \bar{E} , for the transmission of N and $k = 3$ event packets for several detection radii, $R \in \{20, 25, 30\}$ m, and transmission probabilities, $\tau \in \{0.05, 0.06, \dots, 0.35\}$.

energy wastage due to the transmission of redundant packets. Fig. 8 shows the mean energy consumption during event reporting, \bar{E} , for both approaches. It is clear that energy consumption is greatly affected when high values of τ are selected. Large detection radii, R , also affect the energy consumption but to a lower extent. In other words, by selecting lower values of τ and R , a higher energy efficiency is obtained. Also, note that restricting the number of transmitted packets by means of medium sensing reduces energy consumption for low values of τ in every studied environment. In these cases, the network is not highly congested. Conversely, restricting the number of transmitted packets increases the energy consumption when high values of τ are selected.

We begin our analysis of report latency by showing its mean value in Fig. 9. Here we observe that for small de-

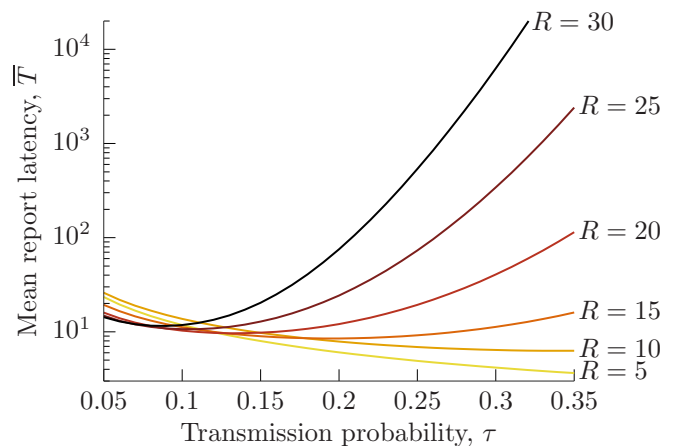


Figure 9: Mean report latency, \bar{T} , assuming $k = 3$ event packets must be received at the sink node for several event detection radii $R \in \{5, 10, 15, 20, 25, 30\}$ m, and transmission probabilities, $\tau \in \{0.05, 0.06, \dots, 0.35\}$.

tection radii, high values of τ reduce report latency. On the other hand, high values of τ increase report latency for large detection radii. Hence, for large detection radii, report latency is correlated to energy consumption. This is an important fact because selecting a large detection radius reduces event overlooking probability. Therefore, the selection of a large R and low τ is the most efficient solution for time-critical applications, as it sharply reduces report latency and event overlooking probability, while maintaining an adequate energy efficiency.

As stated earlier, obtaining the probability distribution of report latency provides with much more valuable information regarding the behavior of the system than mean report latency, which is especially important in time-critical applications. For the sake of simplicity, we obtained the 90th percentile of report latency,

$$T_{90} = \min\{s : P(T \leq s; N_c, k) \geq 0.9\}, \quad (31)$$

to confirm the results presented in Fig. 9. Since the mean report latency and event overlooking probability are enhanced for higher values of R , we obtained the T_{90} for $R = 30$ m. Specifically, Fig. 10 shows the 90th percentile of report latency, T_{90} , for $N_c \in \{1, 2, 3, 4, 5\}$ and the overall $T_{90} = \min\{s : P(T \leq s; k) \geq 0.9\}$ for the given scenario. Here, two tendencies are clearly observed: report latency steeply increases with τ and, the more clusters are involved in event reporting, the lower the time needed for receiving k packets at the sink. Both of these tendencies confirm that the combination of low transmission probabilities with a large detection radius enhances event reporting. Note that the report latency for $N_c = 1$ highly contributes to the overall report latency, whereas the contribution to this parameter fades considerably as N_c increases.

The minimum 90th percentile of report latency, $T_{90} = \min\{s : P(T \leq s; k) \geq 0.9\} = 10$ time slots, is achieved by selecting $\tau = 0.06$, hence we show its CDF in Fig. 11.

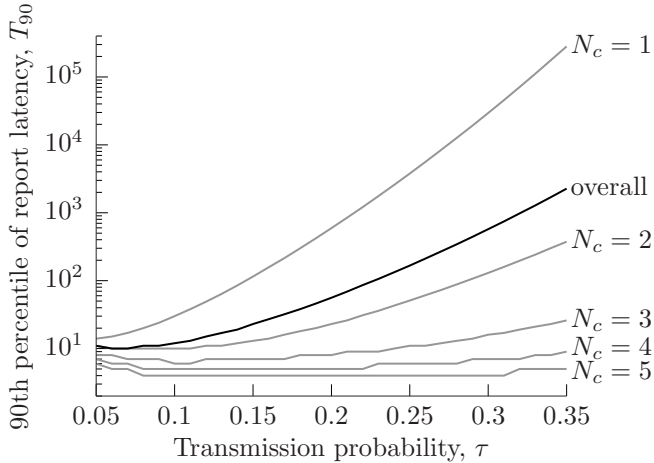


Figure 10: 90th percentile of the report latency, $T_{90} = \min\{s : P(T \leq s; N_c, k) \geq 0.9\}$, for $R = 30$ m.

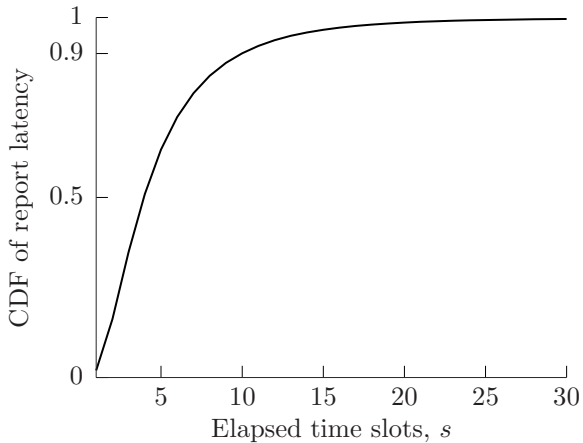


Figure 11: CDF of report latency, $P(T \leq s; k)$, for $R = 30$ m and $\tau = 0.06$.

Hereafter, we use the energy consumption and report latency obtained by selecting $\tau = 0.06$ to assess the benefits of implementing an adaptive backoff (ABO).

5.1. QoS analysis of ABO

We use the hybrid method presented in Section 4 to analyze the performance of RA event reporting when an adaptive backoff (ABO) is implemented. As described previously, we first use our simulator to obtain the probability distribution of the number of detecting nodes, which allows us to analytically calculate the energy consumption and report latency during event reporting.

We evaluate the energy efficiency of the ABO by obtaining the mean energy consumption during event reporting given $R = 30$ m, $\tau \in (0, 1)$ and $B \in \{1, 2, 3, 5, 10\}$. Herein, we denote $B = 1$ as the case where no adaptive backoff is performed and $\tau_e^*(B)$ as the value of τ that optimizes energy consumption for a given B . The mean en-

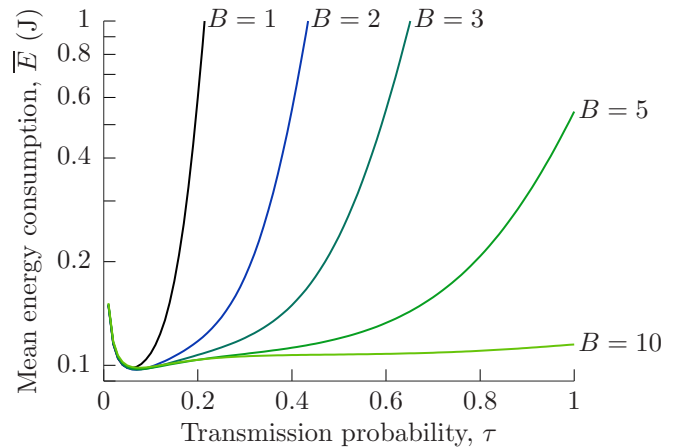


Figure 12: Mean energy consumption during event reporting, \bar{E} , for $R = 30$ m.

Table 2: Transmission probabilities, $\tau_e^*(B)$, that minimize the mean energy consumption, \bar{E} , for the given B .

B	$\tau_e^*(B)$	\bar{E} (J)
1	0.06	0.09812
2	0.07	0.09707
3	0.07	0.09714
5	0.07	0.09749
10	0.08	0.09838

ergy consumption is shown in Fig. 12 and Table 2 shows the achieved values of \bar{E} for the given $\tau_e^*(B)$.

Note that $\tau_e^*(2) = 0.07$ leads to the overall minimum energy consumption in the network. However, extremely similar values are obtained when selecting other low values of B for the ABO. Therefore, the use of an ABO is capable of reducing the mean energy consumption when compared to the optimal selection of τ with a non-adaptive backoff ($B = 1$). Fig. 12 also shows that the mean energy consumption, \bar{E} , increases rapidly if the non-adaptive backoff is implemented and $\tau_e^*(1)$ is not selected. Note that the ABO increases the range of values of τ that lead to an almost-optimal \bar{E} when compared to the non-adaptive backoff. Thus, the implementation of an ABO leads to the overall minimum \bar{E} and also reduces the negative impact on performance of the inadequate selection of transmission probabilities, τ . Hence, the network performance is much more robust.

We now proceed to investigate the behavior of report latency for each possible value of $\tau \in (0, 1)$ and $B \in \{1, 2, 3, 5, 10\}$ given $R = 30$ m. We show the overall T_{90} obtained in this scenario in Fig. 13. Herein, we denote $\tau^*(B)$ as the value of τ that optimizes the T_{90} for a given B . The values of $\tau^*(B)$ and the resulting T_{90} are listed in Table 3.

Here we observe that the same minimum T_{90} is achieved

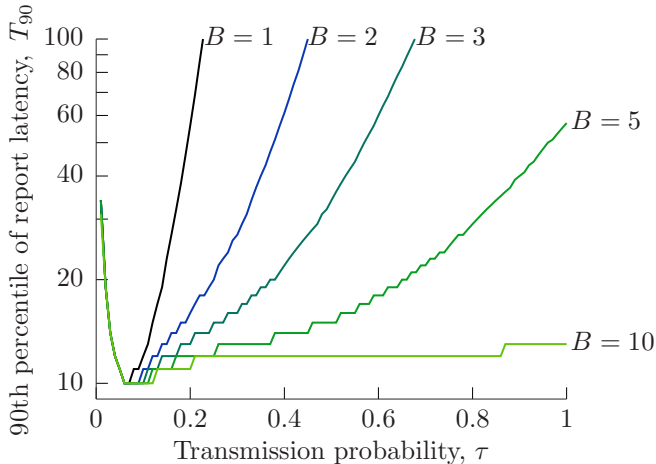


Figure 13: 90th percentile of the report latency, $T_{90} = \min\{s : P(T \leq s; k) \geq 0.9\}$, for $R = 30$ m.

Table 3: Transmission probabilities, $\tau^*(B)$, that minimize the 90th percentile of report latency, T_{90} , for the given B .

B	$\tau^*(B)$	T_{90}
1	0.06	10
2	0.07	10
3	0.07	10
5	0.07	10
10	0.08	10

for each $\tau^*(B)$ and also that $\tau^*(B) = \tau_e^*(B)$; hence the selection of $\tau^*(B)$ leads to the optimal report latency and energy consumption for each B . Note that while $\tau^*(B)$ is rather constant, setting the exact $\tau^*(B)$ becomes less critical as B increases; i.e., the performance degradation if the exact $\tau^*(B)$ is not selected is less important for high values of B . This is the major advantage of the ABO with respect to the non-adaptive backoff and will be further showcased in Fig. 15 and Fig. 16. As such, the implementation of an ABO produces a similar effect in both \bar{E} and T_{90} . This is highly desirable as the overall robustness of the network performance to the selection of transmission parameters increases.

For an in-depth look at the behavior of report latency, we have obtained its CDF, which is shown in Fig. 14. It is observed that selecting $\tau^*(B)$ leads to an almost identical probability distribution of report latency for $B \in \{1, 2, 10\}$. This confirms that implementing an ABO is as time-efficient as the non-adaptive backoff, but its performance is much more robust; i.e., an extremely precise selection of transmission probabilities is not required to achieve an adequate performance.

To further illustrate the benefits provided by the ABO, we show the relative increase in the 90th percentile of report latency, T_{90} , and in the mean energy consumption, \bar{E} , due to slight deviations from $\tau^*(B)$ for $B \in \{1, 2, 10\}$

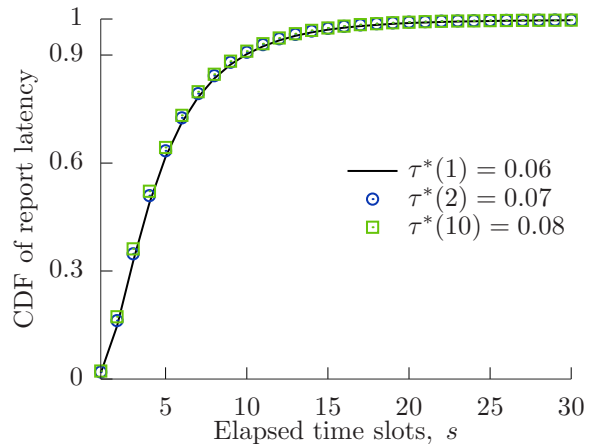


Figure 14: CDF of report latency, $P(T \leq s; k)$, $R = 30$ m.

in Fig. 15. The relative increase is calculated as the ratio between the achieved QoS parameter and its overall achieved minimum. From Tables 3 and 2 it can be seen that $\min\{T_{90}\} = 10$ time slots and $\min\{\bar{E}\} = 0.09707$ J. It is observed that failing to select $\tau^*(1)$ highly affects performance. This is clear even for errors as low as $\tau = \tau^*(1) \pm 0.03$. On the other hand, the energy consumption and report latency that the ABO provides are much more robust to the selection of parameters.

Based on our results, implementing the ABO in a single-event scenario may lead to the exact same minimum report latency and to a lower energy consumption when compared to the implementation of a non-adaptive backoff. In addition, the network performance is not significantly affected when $\tau^*(B)$ is not selected. Hence the robustness in the selection of transmission parameters that the ABO provides is highly valuable.

5.2. Multi-event environments

In multi-event environments several types of events occur and each of them presents different characteristics. This is a typical scenario in complex WSN applications as nodes may be in charge of monitoring a wide range of environmental parameters.

To evaluate the performance of RA event reporting, we have selected an environment where two types of events occur. Specifically, we assume an environment in which the detection radius of 75% of the events is $R_1 = 30$ m and the detection radius of the remaining 25% of the events is $R_2 = 15$ m. We use our hybrid method to obtain $\tau^*(B)$ and the combination of $\tau^*(B)$ and B that optimizes report latency.

Fig. 16 shows the relative increase in the T_{90} and \bar{E} when compared to $\min\{T_{90}\} = 13$ time slots and $\min\{\bar{E}\} = 0.081878$ J. The values of $\tau^*(B)$, along with the achieved T_{90} and \bar{E} are enlisted in Table 4. Here we observe that the implementation of an ABO in multi-event environments leads to a similar behavior as in single-event environments,

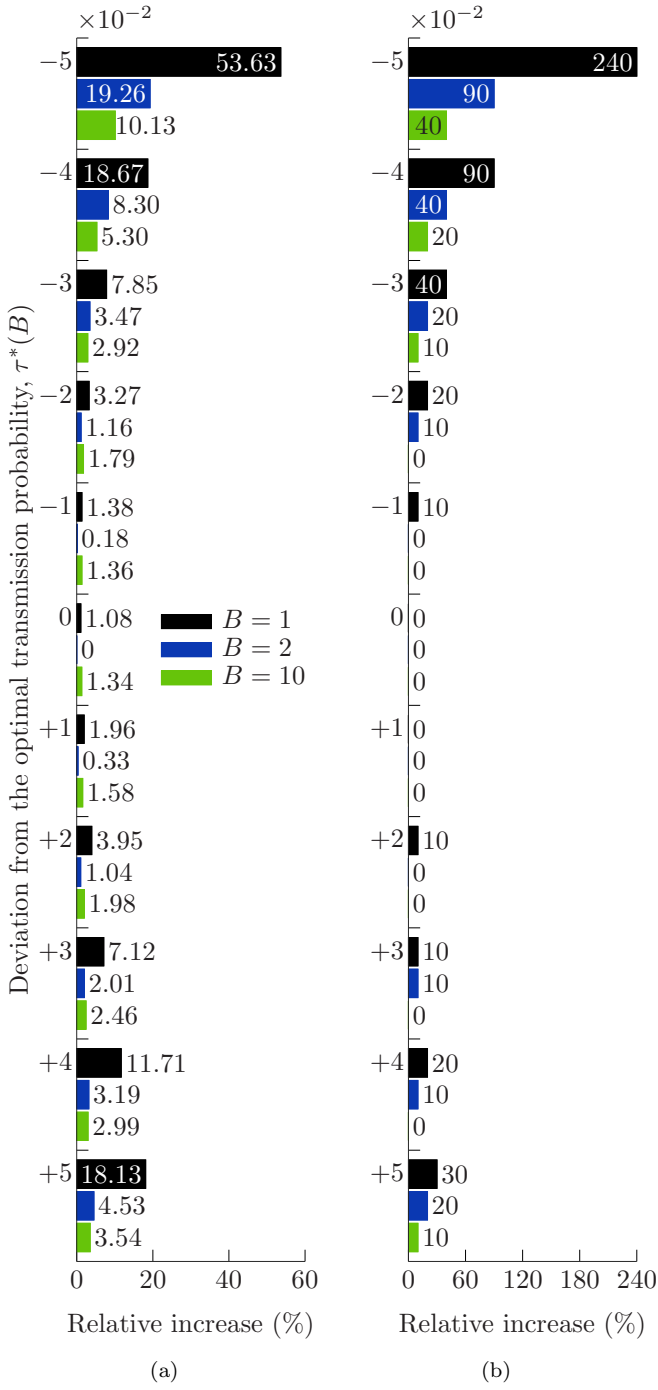


Figure 15: Relative increase in the (a) mean energy consumption, \bar{E} , and (b) 90th percentile of report latency, T_{90} , due to slight deviations from $\tau^*(B)$ with $R = 30$.

i.e., the performance provided by an ABO in much more robust when compared to the non-adaptive backoff.

Table 4 reveals that selecting $\tau^*(B)$ for each B results in the same $T_{90} = 13$ time slots and also that the overall minimum energy consumption is achieved when selecting $\tau^*(2)$ and increases with B .

To further observe the advantages of an ABO in multi-

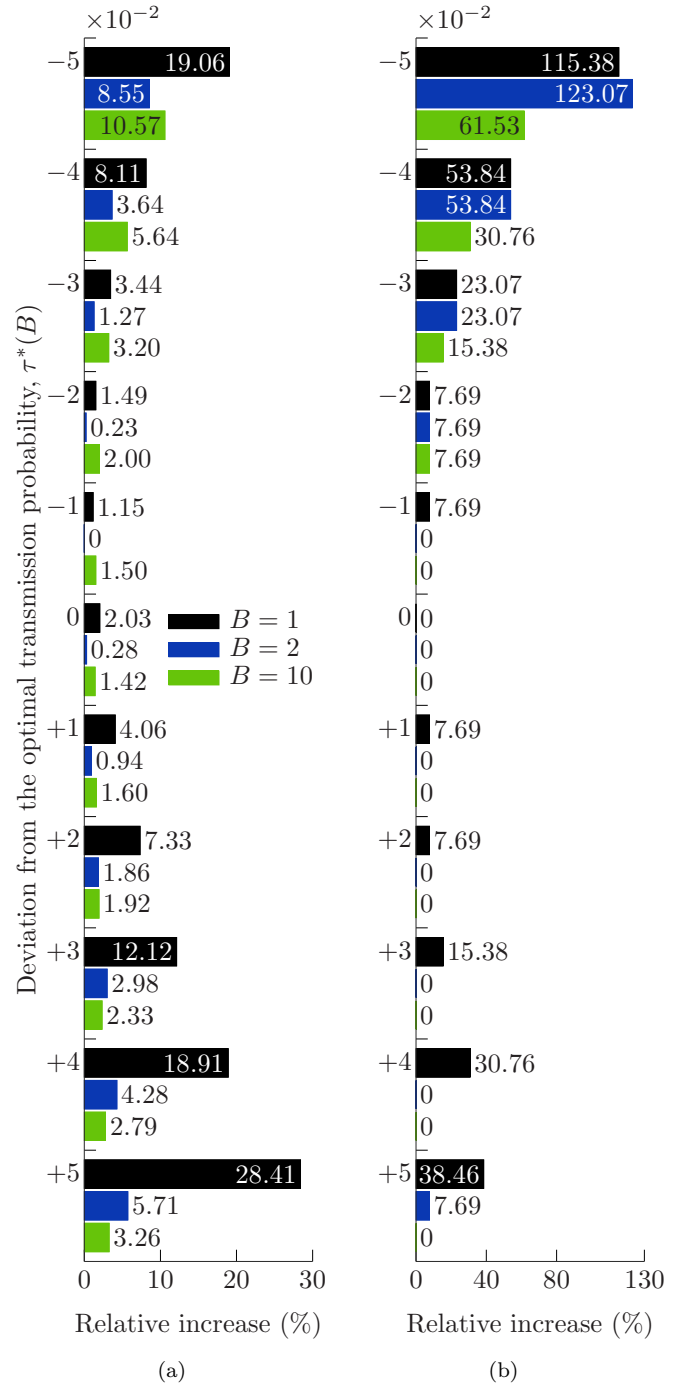


Figure 16: Relative increase in the (a) mean energy consumption, \bar{E} , and (b) 90th percentile of report latency, T_{90} , due to slight deviations from $\tau^*(B)$, $P(R = 30) = 0.75$ and $P(R = 15) = 0.25$.

event environments, the CDF of report latency within each cluster for $B \in \{1, 2, 10\}$, given the selection of $\tau^*(B)$, is presented in Fig. 17. Here we observe slight differences between the selected $\tau^*(B)$. For instance, the CDF of report latency increases more rapidly for $\tau^*(10)$.

As such, an ABO is capable of optimizing report latency and energy consumption in multi-event environments.

Table 4: Transmission probabilities, $\tau^*(B)$, that minimize the 90th percentile of report latency, T_{90} , for the given B in the considered multi-event environment.

B	$\tau^*(B)$	T_{90}	\bar{E} (J)
1	0.07	13	0.08456
2	0.07	13	0.08287
3	0.07	13	0.08293
5	0.08	13	0.08321
10	0.08	13	0.08406

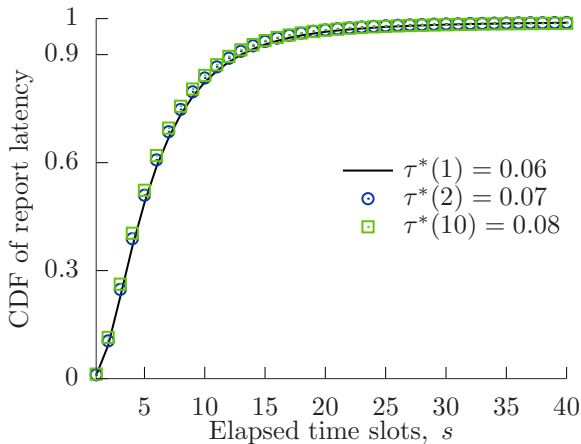


Figure 17: CDF of report latency, $P(T \leq s; k)$, $P(R = 15 \text{ m}) = 0.25$, $P(R = 30 \text{ m}) = 0.75$.

In addition, the network performance is much more robust to the inadequate selection of transmission probabilities.

6. Conclusion

In this paper, we have presented a hybrid method for the QoS analysis of RA WSN protocols. This method is especially useful to assess the QoS of WSNs in time-critical applications, where event reporting is time-constrained and fault-sensitive. As it considers a basic structure, it can be easily adapted for analyzing a wide range of WSN protocols.

The presented method comprises three main phases. In the first, the probability distribution of detecting nodes is obtained by simulation. In the second phase, a discrete-time Markov reward process is defined. In the third phase, the probability distribution of report latency and the mean energy consumption for one event reporting phase are calculated. We have conducted a thorough analysis of the QoS provided by a non-adaptive and by an adaptive backoff (ABO) scheme in single and multi-event environments to showcase the adaptability of the presented method.

Obtaining the probability distribution of detecting nodes has allowed us to calculate the overlooking probability for the given detection radii, R , and the required number of

messages to characterize the occurring event, k . By comparing the event overlooking probability with the mean energy consumption, we identified a clear trade-off; i.e., large detection radii reduce the event overlooking probability but increase the energy consumption. Hence the network administrator must configure the sensibility (thresholds) of the detecting sensors of the nodes to achieve an adequate balance between these two QoS parameters for the target application.

Regarding the RA event reporting scheme, we observed that restricting the number of transmitted packets leads to a noticeable drop in the energy consumption during event reporting for low values of τ . Conversely, no significant energy savings were achieved by restricting the number of transmitted packets and selecting high values of τ . This implies that in highly congested wireless environments (the number of transmissions per time slot increases with τ), a high percentage of the total energy is consumed during the transmission of the first ℓ messages due to the high probability of collisions and retransmissions. These results were obtained with the selection of an energy consumption model that has been widely used in the literature; if other energy consumption model is employed, different results can be obtained.

Obtaining the CDF of report latency has allowed us to observe that the cases where an event affects a low number of clusters, N_c , contribute the most to the overall T_{90} because these cases occur more frequently and cause a higher latency when compared to cases where a high N_c detect the event. We also observed that this parameter bottomed out at $\tau = 0.06$ when $R = 30 \text{ m}$.

Regarding the adaptive backoff (ABO) in single-event environments, we observed that reducing the transmission probabilities during backoff can indeed result in the same minimum report latency obtained with the non-adaptive backoff. But the main benefits of the ABO is the optimization of energy consumption and the increase in the robustness of the system as the QoS is much less affected by the inadequate selection of transmission probabilities.

Since a wide variety of events may occur in complex applications, we conducted a QoS analysis in a multi-event environment. Here we observed that an ABO may optimize report latency and energy consumption simultaneously. In other words, important time and energy savings are achieved by carefully reducing transmission probabilities during backoff.

In practice, the optimal deployment and configuration of a RA WSN may not be possible. Nevertheless, our results show that a high QoS can be achieved by: a) deploying and configuring the WSN such that the events are detected by several nodes; b) restricting the number of event report packets per cluster and c) implementing a RA protocol with a low transmission probability and an ABO scheme.

Acknowledgment

This work has been supported by the Ministry of Economy and Competitiveness of Spain through the project TIN2013-47272-C2-1-R, by CONACyT under project Basic Science 183370 and by IPN SIP project 20150584. The research of Israel Leyva-Mayorga is partially funded by grant 383936 CONACYT-Gobierno del Estado de México 2014.

References

- Afsar, M. M., Tayarani-N, M.-H., 2014. Clustering in sensor networks: a literature survey. *J. Netw. Comput. Appl.* 46, 198–226.
- Akyildiz, I., Su, W., Sankarasubramanian, Y., Cayirci, E., 2002. Wireless sensor networks: a survey. *Comput. Networks* 38 (4), 393–422.
- Alfa, A. S., 2010. *Queueing Theory for Telecommunications*. Springer US, Boston, MA.
- AlSkaif, T., Guerrero Zapata, M., Bellalta, B., 2015. Game theory for energy efficiency in wireless sensor networks: latest trends. *J. Netw. Comput. Appl.* 54, 33–61.
- Asudeh, A., Zaruba, G. V., Das, S. K., 2016. A general model for MAC protocol selection in wireless sensor networks. *Ad Hoc Networks* 36, 189–202.
- Bellman, R., 1957. *Dynamic Programming*, 1st Edition. Princeton University Press, Princeton, NJ, USA.
- Calafate, C. T., Lino, C., Cano, J.-C., Manzoni, P., 2010. Modeling emergency events to evaluate the performance of time-critical WSNs. In: *Proceedings of the IEEE Symposium on Computers and Communications (ISCC)*. pp. 222–228.
- Chen, D.-R., 2015. An energy-efficient QoS routing for wireless sensor networks using self-stabilizing algorithm. *Ad Hoc Networks* 37, 240–255.
- Guo, P., Jiang, T., Zhang, Q., Zhang, K., 2012. Sleep scheduling for critical event monitoring in wireless sensor networks. *IEEE Trans. Parallel Distrib. Syst.* 23 (2), 345–352.
- Haining Shu, Qilian Liang, 2006. Fundamental performance analysis of event detection in wireless sensor networks. In: *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*. Vol. 4. pp. 2187–2192.
- Harrison, D. C., Seah, W. K. G., Rayudu, R., 2016. Rare event detection and propagation in wireless sensor networks. *ACM Comput. Surv.* 48 (4), 1–22.
- Heinzelman, W., Chandrakasan, A., Balakrishnan, H., 2002. An application-specific protocol architecture for wireless microsensor networks. *IEEE Trans. Wirel. Commun.* 1 (4), 660–670.
- Hu, L., 1991. Distributed code assignments for CDMA packet radio networks. In: *IEEE INFCOM '91. Tenth Annual Joint Conference of the IEEE Computer and Communications Societies*. Vol. 3. pp. 1500–1509.
- Iqbal, M., Naem, M., Anpalagan, A., Qadri, N., Imran, M., 2016. Multi-objective optimization in sensor networks: optimization classification, applications and solution approaches. *Comput. Networks* 99, 134–161.
- Jamieson, K., Balakrishnan, H., Tay, Y. C., 2006. Sift: a MAC protocol for event-driven wireless sensor networks. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 3868 LNCS. pp. 260–275.
- Lee, S. H., Choi, L., 2015. SPEED-MAC: speedy and energy efficient data delivery MAC protocol for real-time sensor network applications. *Wirel. Networks* 21 (3), 883–898.
- Leyva-Mayorga, I., Pla, V., Rivero-Angeles, M. E., 2015a. A hybrid method for obtaining the distribution of report latency in wireless sensor networks. In: *Proceedings of the IFIP Wireless and Mobile Networking Conference (WMNC)*. pp. 9–15.
- Leyva-Mayorga, I., Rivero-Angeles, M. E., Arellano, C. C., 2014. Priority-based multi-event reporting in hybrid wireless sensor networks. In: *Proceedings of the IEEE International Conference on Advanced Information Networking and Applications (AINA)*. pp. 413–420.
- Leyva-Mayorga, I., Rivero-Angeles, M. E., Carreto-Arellano, C., Pla, V., 2015b. QoS analysis for a nonpreemptive continuous monitoring and event-driven WSN protocol in mobile environments. *Int. J. Distrib. Sens. Networks* 2015, 1–16.
- Liang, Z., Feng, S., Zhao, D., Shen, X. S., 2011. Delay performance analysis for supporting real-time traffic in a cognitive radio sensor network. *IEEE Trans. Wirel. Commun.* 10 (1), 325–335.
- Manjeshwar, A., Agrawal, D., 2002. APTEEN: a hybrid protocol for efficient routing and comprehensive information retrieval in wireless. In: *Proceedings of the International Parallel and Distributed Processing Symposium (IPDPS)*. p. 8.
- Marco, A., Casas, R., Sevillano Ramos, J., Coarasa, V., Asensio, A., Obaidat, M. S., 2011. Synchronization of multihop wireless sensor networks at the application layer. *IEEE Wirel. Commun.* 18 (1), 82–88.
- Misra, S., Singh, S., Khatua, M., Obaidat, M. S., 2015. Extracting mobility pattern from target trajectory in wireless sensor networks. *Int. J. Commun. Syst.* 28 (2), 213–230.
- Monaco, U., Cuomo, F., Melodia, T., Ricciato, F., Borghini, M., 2006. Understanding optimal data gathering in the energy and latency domains of a wireless sensor network. *Comput. Networks* 50 (18), 3564–3584.
- Rashid, B., Rehmani, M. H., 2015. Applications of wireless sensor networks for urban areas: a survey. *J. Netw. Comput. Appl.* 60, 192–219.
- Rault, T., Bouabdallah, A., Challal, Y., 2014. Energy efficiency in wireless sensor networks: a top-down survey. *Comput. Networks* 67, 104–122.
- Sharif, A., Potdar, V., Rathnayaka, A., 2010. Prioritizing information for achieving QoS control in WSN. In: *Proceedings of the IEEE International Conference on Advanced Information Networking and Applications (AINA)*. pp. 835–842.
- Siddiqui, S., Ghani, S., 2013. Analytical model for delay distribution of PRMAC. In: *Proceedings of the IEEE International Conference on Frontiers of Information Technology (FIT)*. pp. 1–6.
- Souil, M., Bouabdallah, A., Kamal, A., 2014. Efficient QoS provisioning at the MAC layer in heterogeneous wireless sensor networks. *Comput. Commun.* 43, 16–30.
- Tickoo, O., Sikdar, B., 2008. Modeling queueing and channel access delay in unsaturated IEEE 802.11 random access MAC based wireless networks. *IEEE/ACM Trans. Netw.* 16 (4), 878–891.
- Wang, Y., Vuran, M. C., Goddard, S., 2011. Analysis of event detection delay in wireless sensor networks. In: *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*. pp. 1296–1304.
- Yahya, B., Ben-Othman, J., 2009. Towards a classification of energy aware MAC protocols for wireless sensor networks. *Wirel. Commun. Mob. Comput.* 9 (12), 1572–1607.
- Ye, W., Heidemann, J., Estrin, D., 2002. An energy-efficient MAC protocol for wireless sensor networks. In: *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies*. Vol. 3. pp. 1567–1576.
- Younis, O., Fahmy, S., 2004. HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Trans. Mob. Comput.* 3 (4), 366–379.
- Zhang, R., Pan, J., Liu, J., Xie, D., 2015. A hybrid approach using mobile element and hierarchical clustering for data collection in WSNs. In: *Proceedings of the 2015 IEEE Wireless Communications and Networking Conference (WCNC)*. pp. 1566–1571.
- Zheng, T., Gidlund, M., Akerberg, J., 2016. WirArb: a new MAC protocol for time critical industrial wireless sensor network applications. *IEEE Sens. J.* 16 (7), 2127–2139.