Ad Hoc Networks 9 (2011) 1083–1103

Contents lists available at ScienceDirect

Ad Hoc Networks

journal homepage: www.elsevier.com/locate/adhoc

# Clustered Zigbee networks with data fusion: Characterization and performance analysis

P. Medagliani *, M. Martalò, G. Ferrari

*University of Parma, Department of Information Engineering, Viale G.P. Usberti 181/A, I-43100 Parma, Italy*

ARTICLE INFO

*Article history:*
Received 18 September 2009
Received in revised form 8 February 2010
Accepted 7 October 2010
Available online 13 January 2011

*Keywords:*
Zigbee
Clustering
Network lifetime
Data fusion
Design guidelines

ABSTRACT

In this paper, we analyze the performance of *clustered* Zigbee wireless sensor networks (WSNs) with data fusion. Performance indicators at both physical (probability of decision error) and network (network transmission rate, throughput, aggregate throughput, delay, and network lifetime) layers are considered. Data fusion is carried out at the access point (AP) and, in clustered configurations, also at the clusterheads, which act as intermediate fusion centers (FCs). The goal of this paper is to shed light on the joint impact of topology and data fusion on the network performance. The presented results, mainly obtained through Opnet-based simulations, show clearly that the operational point of a Zigbee WSN with data fusion lies over a characteristic multi-dimensional surface, whose shape remains the same regardless of the number of nodes in the network. The existence of this peculiar surface highlights fundamental performance trade-offs in Zigbee networks.



## 1. Introduction

Wireless sensor networks (WSNs) have recently become an interesting research topic, both in military [1–3] and civilian scenarios [4,5]. In particular, remote/environmental monitoring, surveillance of reserved areas, etc., are important fields of application of wireless sensor networking techniques. These applications often require very low power consumption and low-cost hardware [6], and *clustering* has been proposed as a possible strategy for saving energy. In [7], the authors present a system-level design methodology, based on a semi-random communication protocol, for clustered WSNs. In [8], after the derivation of an energy consumption model of WSNs, the transmission range is optimized. In [9], the authors investigate how the energy efficiency of a clustered WSN is affected by the transmit power distribution, the numbers of sensors in the clusters, the required end-to-end packet error rate, and the relative lengths of intra-cluster and inter-cluster distances.

Another important aspect of wireless sensor networking is the presence of an embedded data decision strategy, often involving *data fusion*. In [10], the authors present theoretical results on optimal decision rules and their application to data fusion. In [11], the authors follow a Bayesian approach for the minimization of the probability of decision error at the access point (AP). The data fusion mechanism has also a strong impact on practical applications. In [12], the authors analyze several methods of multi-sensor data fusion, such as Bayesian estimation, Kalman filtering, and Dempster–Shafer evidence theoretical methods, in order to design a move-in-mud robot. In [13], the impact of source–destination placement and communication network density on the energy costs and delay associated with data aggregation are evaluated.

In complex systems, such as WSNs, basic design approaches may no longer be sufficient to effectively improve the performance. Therefore, it is preferable to resort to joint optimization strategies which involve more than one layer of the communication/networking protocol stack, in order to significantly improve the performance.

* Corresponding author. Tel.: +39 0521 905759; fax: +39 0521 905758.
*E-mail addresses:* paolo.medagliani@unipr.it (P. Medagliani), martalo@tlc.unipr.it (M. Martalò), gianluigi.ferrari@unipr.it (G. Ferrari).

1570-8705/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.adhoc.2010.10.009

In [14], the authors present a cross-layer design framework, based on the use of an optimization agent to exchange information between different protocol layers, to improve the performance in WSNs. In [15], however, the authors show that unintended cross-layer interactions can have undesirable consequences on the overall system performance. Therefore, care has to be taken in developing cross-layer frameworks for WSNs.

In this paper, we shed light, through a simulation-based study, on the impact of *clustering* and *data fusion* on the performance of Zigbee networks. We first characterize the behavior of the network transmission rate as a function of the network tolerable death level, denoted as $\chi_{net}$ and representative of the network lifetime. We then evaluate the network transmission rate and the delay as functions of the packet generation rate. In addition, we provide a complete simulation-based characterization of Zigbee WSNs through the evaluation of network and physical layer performance indicators, such as network transmission rate, throughput, delay, network lifetime, and probability of decision error. The goal of this paper is to analyze the performance of clustered Zigbee networks with or without a simple data fusion mechanism, clearly understanding the interplay between network configuration and data fusion and, thus, deriving useful network design guidelines.

The structure of this paper is the following. In Section 2, preliminaries are given: Section 2.1 contains an overview of the Zigbee standard; in Section 2.2, the network tolerable death level is clearly defined; in Section 2.3, the clustering configurations of interest are introduced; in Section 2.4, we summarize the considered data fusion mechanisms from an analytical point of view; finally, in Section 2.5 the implementation characteristics of the used Opnet simulation model are accurately described. In Section 3, simulation-based performance results are shown and discussed: in Section 3.1, network transmission rate and delay are studied as functions of the network tolerable death level; in Section 3.2, the behavior of the same performance indicators is analyzed, for fixed values of network tolerable death level, as a function of the packet generation rate; in Section 3.3, we analyze the impact of data fusion on the throughput, the delay, and the probability of decision error. In Section 4, a few guidelines for the design of Zigbee WSNs, on the basis of the performance analysis in Section 3, are given: in Section 4.1, we show that the operational point of a Zigbee WSN lies over a characteristic multi-dimensional surface, whose shape is independent of the number of nodes and clearly shows the existence of fundamental performance trade-offs between $\chi_{net}$, network transmission rate, and delay; in Section 4.2, we derive simple analytical approximation of the obtained multi-dimensional Zigbee performance surfaces. Section 5 concludes the paper.

## 2. Preliminaries

### 2.1. Zigbee standard overview

The Zigbee standard is suited for the family of Low-Rate Wireless Personal Area Networks (LR-WPANs), allowing network creation, management, and data transmission over a wireless channel with high energy savings. Three different types of nodes are foreseen by the Zigbee standard: (i) coordinator, (ii) router, and (iii) end device. In the absence of a direct communication link, the router is employed to relay the packets towards the correct destination. The coordinator, in addition to relaying the packets, can also create the network, exchange the parameters used by the other nodes to communicate (e.g., a network IDentifier, ID, a synchronization frame, etc.), and send network management commands. The router and coordinator are referred to as Full Function Devices (FFDs), i.e., they can implement all the functions required by the Zigbee standard in order to set up and maintain communications. The end devices, referred to as Reduced Function Devices (RFDs), can only collect data values from sensors, insert these values into proper packets, and send them to their destinations. In the remainder of this paper, coordinator and AP will be used interchangeably, and the same will hold for RFDs and sensors. In Fig. 1, we show illustrative examples of the topologies allowed by the Zigbee standard, namely star, cluster-tree, and mesh.

The Zigbee standard is based, at the first two layers of the ISO/OSI stack, on the IEEE 802.15.4 standard [16], which employs a non-persistent Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) Medium Access Control (MAC) protocol and operates in the 2.4 GHz band (similarly to the IEEE 802.11 standard [17]). In addition, the IEEE 802.15.4 standard provides an *optional* ACK message to confirm correct delivery of a data packet. In a scenario with ACK messages, the access mechanism of the non-persistent CSMA/CA MAC protocol is slightly modified. Whereas a generic data packet is sent according to the CSMA/CA protocol, an ACK message is sent back to the source immediately after the message is received by the destination. As soon as the ACK message is received, both the source and the destination nodes wait for a time interval, referred to as Long InterFrame Spacing (LIFS), which allows them to perform internal stack operations and process the data. This interval is used also in the absence of ACK messages' transmission. In this case, once a node has sent its packet, it waits for a LIFS interval, whereas the receiving node waits for a shorter interval, referred to as Turn Around Time (TAT), used to take into

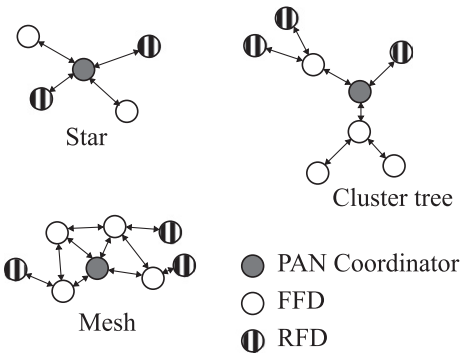


**Fig. 1.** Illustrative examples of the topologies foreseen in the Zigbee standard.

account radiofrequency interface recalibration. During the TAT, the receiving node cannot accept new incoming packets. In the remainder of this paper, we will assume that no ACK messages is used during network optimization.

We remark that the non-persistent CSMA/CA MAC protocol provides a medium access mechanism which tries to avoid, through a technique referred to as Bynary Exponential Back-off (BEB), packet collisions. A node, before transmitting a new packet, waits for an interval randomly chosen within a range defined during the network start-up phase. This time interval is referred to as Contention Window (CW). After this time has elapsed, the node, before trying to send its packet, senses the status of the channel for a given time interval. This operation is referred to as Clear Channel Assessment (CCA). If the channel is free, the node transmits its packet; otherwise if an incoming transmission is detected, the node doubles the previously chosen CW and waits. This procedure is repeated for three times, after which the waiting interval is kept fixed to its maximum value. After five unsuccessful retransmission attempts, the packet is dropped. The use of the BEB technique makes it likely that a node will eventually manage to transmit its packet.

### 2.2. Network tolerable death level

A critical issue in wireless sensor networking is the network lifetime, since nodes are typically equipped with a limited-energy battery and may be subject to failures. First, one has to define when the network has to be considered "alive," and several definitions have been proposed in the literature. In general, the network can be considered alive until a proper Quality of Service (QoS) condition is satisfied. Obviously, the more stringent this QoS, the shorter the network lifetime. In this paper, we consider, as network lifetime QoS, the percentage of RFDs' deaths at which the overall network is assumed to be dead. This percentage is defined as *network tolerable death level* and we denote it with $\chi_{\text{net}}$. This choice is motivated by the fact that $\chi_{\text{net}}$ quantifies the intuitive idea that a minimum number of observations (or a minimum spatial density of observations) may be required for proper network operations. In other words, if the lifetime QoS condition is stringent, the network is considered dead just after few RFDs' deaths. According to results in the field of reliability theory [18], we model the lifetime of a single RFD as an exponentially distributed random variable with mean value $\mu$ (dimension: [d]). We point out that our approach can be extended to account for any RFD lifetime distribution.

### 2.3. Possible network configurations

In order to allow the deployment of large-scale WSNs, the sensors may be grouped into *clusters*, i.e., they transmit their data to intermediate nodes (denoted as clusterheads), which may properly modify these data and relay them to the AP [19].

An illustrative representation of some of the network configurations of interest with $N = 16$ RFDs is given in Fig. 2. In particular, the presented schemes can be grouped into two main classes: (i) networks with *uniform* clusters (all clusters have the same number of sensors, as in Fig. 2a) and (ii) networks with *non-uniform* clusters (the dimensions of the clusters may vary, as in the cases in Fig. 2b–d). In the scenarios with uniform clustering, the considered network configurations include: (i) one 16-node cluster and no relay (i.e., direct transmission from each RFD to the AP); (ii) one 16-node cluster with one relay (i.e., transmission from each RFD to the AP through a relay); (iii) two 8-node clusters with two relays (one relay per cluster); (iv) four 4-node clusters with four relays (one relay per cluster); and (v) eight 2-node clusters with eight relays (one relay per cluster). In the simulations with non-uniform clustering, instead, the adopted network configurations are: (i) 8–2–2–2–2 with five relays (i.e., the network is divided into five clusters, one formed by eight RFDs and each of the other four formed by two RFDs, and each cluster is connected to the AP through a relay); (ii) 10–2–2–2 with four relays (i.e., the network is divided into four clusters, one formed by 10 RFDs and other formed by two RFDs each, connected to the AP through a relay), as shown in Fig. 2b; (iii) 14–1–1 with three relays (i.e., one cluster is composed by 14 RFDs and two clusters are composed by one node each), as shown in Fig. 2c; and (iv) 14–1–1 with one relay (i.e., only one cluster composed by 14 nodes and connected to the AP through a relay, while the other two RFDs communicate directly to the AP), as shown in Fig. 2d. The network configurations adopted in scenarios with a different number of RFDs will be described similarly.

The average delay is computed as the arithmetic average of all delays experienced by each RFD, taking into account the presence of a relay with a supplementary delay. For example, in the 14–1–1 scenario with only one relay node (shown in Fig. 2d), the average delay is computed as follows:

$$D = \frac{1}{16}\sum_{i=1}^{16}\overline{D}_i = \frac{1}{16}\left[\overline{D}_{\text{AP}} + \overline{D}_{\text{AP}} + 14\cdot\left(\overline{D}_{\text{AP}} + \overline{D}_{\text{relay}}\right)\right] = \overline{D}_{\text{AP}} + \frac{14}{16}\cdot\overline{D}_{\text{relay}} \quad (1)$$

where $\overline{D}_{\text{AP}}$ is the average delay of a direct transmission to the AP (either from an RFD or the relay) and $\overline{D}_{\text{relay}}$ is the average delay introduced by the relay.

### 2.4. Sensing and data fusion

In this subsection, we summarize the data fusion mechanism proposed in [20], which will be used in our simulator—the reader is referred to [20] for more details. In the considered scenarios, $N$ RFDs observe (in a noisy manner) a common binary phenomenon $H$, defined as

$$H = \begin{cases} H_0 & \text{with probability } p_0 \\ H_1 & \text{with probability } (1-p_0) \end{cases} \quad (2)$$

where $p_0 \triangleq P(H = H_0)$. Throughout this paper, we will consider equally distributed phenomena, i.e., $p_0 = 1/2$, but the proposed approach is general. As previously introduced, the RFDs may be clustered into $n_c < N$ groups, and each RFD communicates only with its corresponding relay, which, in this case, acts as a fusion center (FC). The FCs

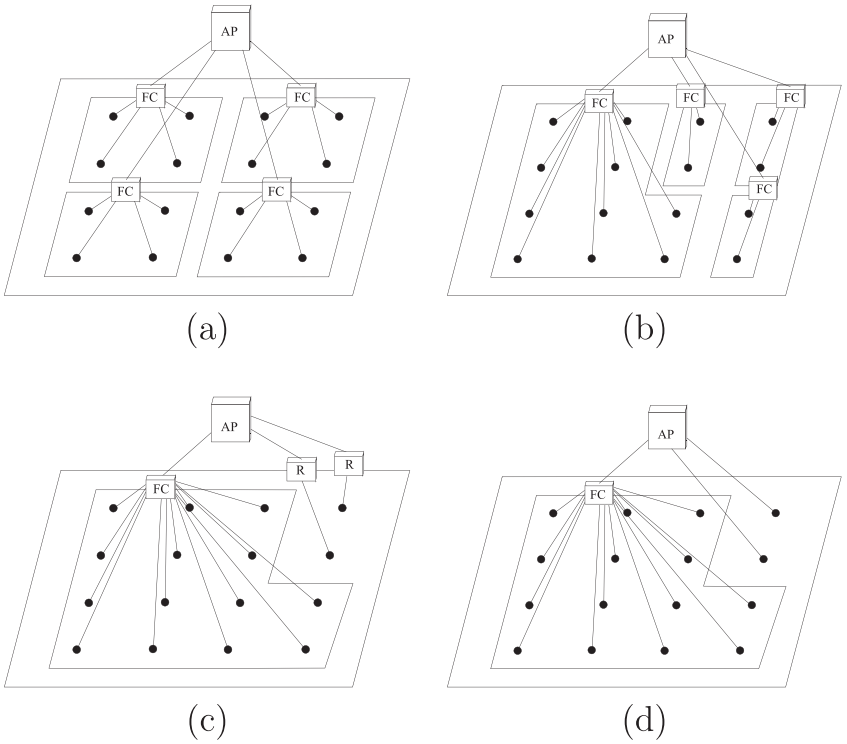


**Fig. 2.** Illustrative clustered topologies: (a) uniform with four clusters and four relays, (b) non-uniform with a 10–2–2–2 configuration with four relays, (c) non-uniform with a 14–1–1 configuration and three relays, and (d) non-uniform with a 14–1–1 configuration and one relay. The nodes denoted with "R" are only dedicated to relay packets from the RFDs to the AP, whereas the nodes denoted with "FC" are also used to fuse data received from the RFDs.

collect data from the RFDs in their clusters and make local decisions on the status of the binary phenomenon. Then, each FC transmits, without noise, its "fused" decisions to the AP, which makes the final decision. Being the observed signal the same across the RFDs and assuming that the observation noises are Gaussian and independent with the same distribution $\mathcal{N}(0,\sigma^2)$, the common signal-to-noise ratio (SNR) at the RFDs can be defined as follows [20]:

$$\mathrm{SNR} = \frac{s^2}{\sigma^2}$$

where $s$ is the intensity of the observed signal. Each sensor makes a decision comparing its observation $r_i$ with a threshold value $\tau_i$ and computes a local decision $u_i = U(r_i - \tau_i)$, where $U(\cdot)$ is the unit step function. Even though, in general, a common value of the decision threshold for all sensors might not be the best choice, in the following we assume a fixed threshold value $\tau$ equal to $s/2$—as shown in [20], this guarantees quasi-optimality.

### 2.4.1. Uniform clustering

In a scenario with *uniform* clustering, the sensors are grouped into identical clusters, i.e., each of the $n_c$ clusters contains $d_c$ sensors, with $n_c \cdot d_c = N$. An illustrative example, in a scenario with $N = 16$ and $n_c = 4$, is given in Fig. 2a.

The fusion rules at the FCs and the AP are majority-like fusion rules with fusion thresholds set, respectively, to $k$ and $k_f$. Using a combinatorial approach (based on the repeated trials formula [21]) and taking into account the majority-like fusion rules, the probability of decision error at the coordinator can be expressed as follows [20]:

$$\begin{aligned} P_e &= P(\widehat{H} = H_1|H_0)P(H_0) + P(\widehat{H} = H_0|H_1)P(H_1) \\ &= p_0 \mathrm{bin}(k_f, n_c, n_c, \mathrm{bin}(k, d_c, d_c, 1-\Phi(\tau))) \\ &\quad + (1-p_0)\mathrm{bin}(0, k_f - 1, n_c, \mathrm{bin}(k, d_c, d_c, 1-\Phi(\tau - s))) \end{aligned} \tag{3}$$

where $\Phi(x) \triangleq \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp(-y^2/2)\mathrm{d}y$ and

$$\mathrm{bin}(a,b,n,z) \triangleq \sum_{i=a}^{b} \binom{n}{i} z^i (1-z)^{(n-i)}$$

In the case of majority fusion, $k = \lfloor d_c/2 \rfloor$ and $k_f = \lfloor n_c/2 \rfloor$.

### 2.4.2. Non-uniform clustering

In this case, the RFDs do not cluster regularly. In particular, the size of the $i$th cluster is denoted as $d_c^{(i)}$ ($i = 1,\ldots,n_c$) and it hods that $\sum_{i=1}^{n_c} d_c^{(i)} = N$. Since the RFDs are not equally distributed among the clusters, the optimized decision threshold at each RFD could depend on the corresponding cluster. We remark that the fusion rules at the FCs and the AP are majority-like fusion rules with proper values of the fusion thresholds, as discussed in Section 2.4.1.

Let us define $p_\ell^{1|1}$ ($p_\ell^{1|0}$, respectively) as the probability that the $\ell$-th FC decides for $H_1$ when $H_1$ ($H_0$, respectively).

After a few manipulations, the probability of decision error can be expressed as follows [20]:

$$P_e = p_0 \sum_{i=k_f}^{n_c} \binom{n_c}{i} \sum_{j=1}^{n_c} \prod_{\ell=1}^{n_c} \left\{ c_{i,j}(\ell) p_\ell^{1|0} + (1 - c_{i,j}(\ell)) \left(1 - p_\ell^{1|0}\right) \right\} + (1 - p_0) \sum_{i=0}^{k_f - 1} \binom{n_c}{i} \sum_{j=1} \prod_{\ell=1}^{n_c} \left\{ c_{i,j}(\ell) p_\ell^{1|1} + \left(1 - c_{i,j}(\ell)\right) \left(1 - p_\ell^{1|1}\right) \right\} \quad (4)$$

where $\boldsymbol{c}_{i,j} = (c_{i,j}(1), \ldots, c_{i,j}(n_c))$ is a vector which designates the $j$th configuration of the decisions from the first-level FCs in a case with $i$ '1's (and, obviously, $n_c - i$ '0's). For example, in the presence of $n_c = 3$ clusters $\boldsymbol{c}_{1,2}$ is the second possible configuration with one '1' (and two '0s'), i.e., 010: the '1' is the decision of the second FC. It can be shown that (4) reduces to (3) in the presence of uniform clustering [20].

### 2.5. Simulation setup

The simulations have been carried out with the Opnet Modeler simulator [22] and a built-in Zigbee network model designed at the National Institute of Standards and Technologies (NIST) [23]. Since this model refers only to the first two layers of the ISO/OSI stack, we have extended it by deriving an Opnet model for a FC, which, in addition to providing relay functionalities, implements the intermediate data fusion mechanisms described in Section 2.4. We assume (ideal) wireless communications between the RFDs and the FCs, and between the FCs and the coordinator. As anticipated in Section 2.1, ACK message are not used, as this may not be realistic in large-scale WSNs. In fact, according to the results in [24], the use of ACK messages in large-scale WSNs drastically decreases the performance in terms of network transmission rate and delay. In small-scale WSNs, our results show that the benefits brought by the use of ACK messages are limited and reduce for increasing values of $N$.

We point out that the considered Zigbee Opnet model is compliant with the Zigbee standard. In particular, this model provides a complete implementation of the physical and MAC layers of the IEEE 802.15.4 standard, taking into account of all parameters characterizing these layers, such as CW, TAT, LIFS, and CCA. We point out that in the NIST model all RFDs interfere with each other, regardless of the clustering configuration. Therefore, all presented results relative to clustering will refer to worst-case communication scenarios. In clustering configurations where only the RFDs belonging to the same cluster interfere, the performance is intuitively expected to be better. Finally, since the application layer has no direct impact on the performance at both networking and MAC layers, we simply model this layer as a transmission queue where the generated/received packets are stored before being transmitted/processed.

#### 2.5.1. Without data fusion

Each simulation result is obtained by averaging over 10 consecutive runs in order to make possible statistical fluctuations negligible. The duration of each simulation has been chosen so that the simulation ends as soon as the percentage of alive RFDs reduces below the network tolerable death level. The packet length is 632 bits (512 bits of data payload and 120 bits introduced by MAC and physical layers).

#### 2.5.2. With data fusion

The RFDs carry out noisy observations of a randomly generated binary phenomenon $H$ and make local decisions on its status. Then, the RFDs embed their decisions into data packets, which are sent either to the coordinator (in the absence of clustering) or to the FCs (in the presence of clustering). The decisions are assumed to be either 0 (no phenomenon) or 1 (presence of the phenomenon).

- In the absence of clustering, the coordinator makes its final decision through a majority rule, directly on the messages received from the RFDs. Obviously, if some packets are lost due to medium access collisions, decisions are made only on the received packets (this leads to a reduced reliability of the final decision). If all the packets related to a set of observations of the same phenomenon are lost, the coordinator decides randomly (i.e., with probability 0.5) for one of the two possible values. Finally, if half of the decisions are in favor of one phenomenon status and the other half are in favor of the other, the coordinator decides for the presence of the phenomenon.
- In the presence of clustering, each FC makes an intermediate decision on the basis of the messages received from the RFDs associated with its cluster and forwards this decision to the coordinator, which makes the final decision on the basis of the received messages (from either an FC or an RFD). The decision rule used by the FCs is the same of that used by the coordinator (i.e., majority decision).

In both scenarios, it is possible to evaluate, by simulation, the probability of decision error at the coordinator by comparing, in each simulation run, the final decision made by the coordinator with the true status of the phenomenon. Together with the probability of decision error, the simulator allows to evaluate: the network transmission rate $S$ (dimension: [b/s]), defined as the ratio between the number of bits correctly received by the coordinator during the simulation time; the throughput $S_{th}$ (adimensional in [0,1]), defined as the ratio between the number of packets correctly delivered at the coordinator and the number of packets sent by the RFDs; and the delay $D$ (dimension: [s]), defined as the time interval between the transmission and the reception instants of a generic packet. The last performance indicator of interest is the aggregate throughput (dimension: [pck/s]), denoted as $S_{agg}$ and defined as $N \cdot g \cdot S_{th}$, where $N$ is the number of transmitting RFDs and $g$ is the packet generation rate at each RFD (dimension: [pck/s]).

In each simulation run, several consecutive independent phenomenon realizations are considered. At each RFD, consecutive decisions (relative to consecutive noisy observations) are then inserted into 96-bit data packets. In order to reliably estimate a probability of decision error at the AP of the order of $10^{-6}$, approximately 115,200 local decisions on the phenomenon status need to be transmitted. Since in the Opnet implementation each decision must be coded into a "char", the memory occupation is therefore equal to eight bits per decision. In addition, in each packet a null char terminator (eight bits) and a 32-bit sensor identifier must be added. Considering a packet length of 96 bits, each packet can thus contain seven local decisions. From this simple analysis, it follows that 16,458 packets need to be transmitted. Since the packet generation rate $g$ is equal to 2 pck/s, the simulation duration is set to 8229 s, in order to guarantee that all required local decisions are sent to the coordinator—this value is chosen as a compromise between simulation duration and achievable probability of decision error at the AP. The effective packet length is equal to 216 bits, since, besides a payload of 96 bits containing the local decisions, a header of 120 bits is introduced by physical and MAC layers. The RFDs send the packets over an ideal wireless channel. At the end of the simulation, the probability of decision error is computed through a direct comparison between the sequence of decisions at the AP and the sequence of true phenomenon realizations.

## 3. Performance analysis

We first evaluate the performance from a networking perspective (in the first two subsections), considering scenarios both with and without clustering. We then consider (in Section 3.3) the presence of data fusion and analyze its impact on the probability of decision error, the throughput, the aggregate throughput, and the delay. In all figures in Sections 3.1 and 3.2, we indicate, for each performance curve, the confidence interval $2\sigma$, where $\sigma$ is the standard deviation, over consecutive simulation runs, with respect to their average value.

We remark that the values of the system parameters are chosen in order to simulate realistic Zigbee networks. For instance, the considered values of $N$ may be suitable for surveillance applications, as considered, in the area of mobile target detection, in [25]. The values of the packet generation rate are typical of scenarios with low data traffic load (e.g., surveillance) or high data traffic load (e.g., target tracking in military applications). As indicated at the end of Section 2.2, the lifetime of a single RFD is exponentially distributed with average value $\mu$, and we will assume, in most of the simulations, that $\mu = 300$ s. The choice of this short average node lifetime (5 min) is dictated by the need of making the simulations manageable. However, the obtained performance results are confirmed considering an average node lifetime equal to 600 s (see Section 3.1.2). Therefore, our conclusions are meaningful also for commercial testbeds where the average lifetime of a single node may be on the order of months.

### 3.1. Impact of tolerable network death level

In the following, we analyze the impact of tolerable network death level on network transmission rate and delay performance indicators, considering either (1) $N = 64$ or (2) $N = 16$ RFDs, which transmit data to the AP either directly or through intermediate clusterheads.

#### 3.1.1. N = 64 RFDs

In this case, we consider scenarios with either no clustering (no relay) or with a number of clusters ranging from 1 to 8 (with a relay per cluster). In Fig. 3, (a) the network transmission rate and (b) the delay are shown as functions of $\chi_{net}$. In Fig. 3c, instead, the curves in the previous two subfigures are combined, obtaining transmission rate–delay curves (parameterized with respect to $\chi_{net}$).

Looking at Fig. 3a, the network transmission rate is first slightly decreasing, for small values of the network tolerable death level, and then suddenly drops in correspondence to $\chi_{net} \simeq 90\%$. In the scenario with many small clusters, the network transmission rate remains low since there is a larger number of transmitting nodes, and this increases the number of collisions.

Considering the results in Fig. 3b, unlike what the intuition may suggest, the delay in the scenario with one large cluster and a single relay (64 configuration) is higher than the delay in the scenario with eight clusters and eight relays (8–8–8–8–8–8–8–8). The delay reduces when the number of clusters increases, since, in this case, the relay of each cluster has to process a smaller number of packets. As expected, the delay is shortest in the absence of clustering, i.e., when all RFDs transmit directly to the AP. Obviously, in all cases the delay decreases for increasing values of percentage of nodes' deaths required to kill the network (i.e., for increasing values of $\chi_{net}$). In fact, the more RFDs die, the more efficiently the surviving RFDs can be served and, thus, the average delay reduces. Comparing the results in Fig. 3a with those in Fig. 3b, it can be noted that the network configuration which guarantees the best performance, in terms of network transmission rate and delay, is the one without relays (lines with circles).

Finally, in Fig. 3c the delay is shown as a function of the network transmission rate (the points of each curve correspond to different values of the network tolerable death level). These curves give a concise (instantaneous) picture of the network operating status: the best operating conditions would correspond, obviously, to the bottom right (low delay and high transmission rate).

We now analyze the network performance in the presence of *non-uniform* clustering. In Fig. 4a, the network transmission rate is shown as a function of $\chi_{net}$, for various clustering configurations.

As in the case with uniform clustering, the network transmission rate is first slightly decreasing (for small values of $\chi_{net}$) and then rapidly decreasing (for values of $\chi_{net}$ higher than 90%). In fact, for small values of $\chi_{net}$ the RFDs' deaths are balanced by the reduced number of collisions in the channel, so that the network transmission rate remains high. When the number of RFDs' deaths becomes large, instead, the overall traffic generated by the RFDs reduces

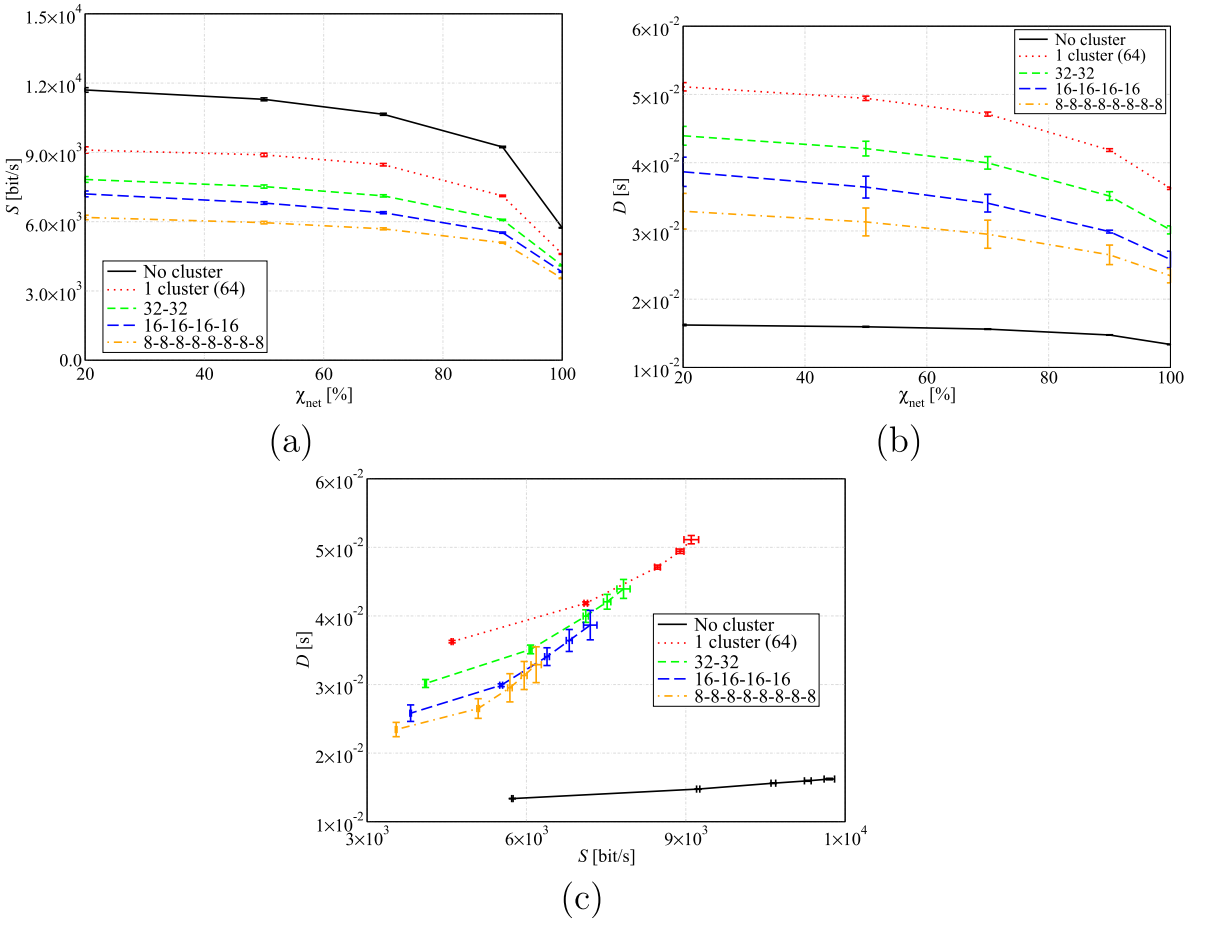


**Fig. 3.** Performance evaluation in a scenario with $N = 64$ RFDs and various uniform clustering configurations: (a) network transmission rate and (b) delay as functions of $\chi_{net}$; (c) delay as a function (parameterized with $\chi_{net}$) of the network transmission rate. The packet generation rate is equal to $g = 2$ pck/s.

and, consequently, the network transmission rate necessarily reduces. The best performance is obtained in the 56–4–4 clustering configuration with one relay (associated with the large cluster). In this case, in fact, the presence of the eight directly-connected RFDs has a beneficial impact on the network transmission rate, since the packets can be transmitted without the need of being relayed. The worst performance, instead, is obtained in the scenario with the 32–8–8–8–8 clustering configuration and five relays: in this case, the number of transmitted packets is larger because of the larger number of relays and, thus, the higher probability of finding the channel busy.

In Fig. 4b, the delay is shown as a function of $\chi_{net}$. The performance, in the presence of non-uniform clustering, is similar in all cases where the RFDs are connected to the AP through a relay. In these cases, in fact, when the number of RFDs' deaths increases, it is more likely that a node finds the channel free and, therefore, can transmit its data. The network performance in the scenario with the 56–4–4 clustering configuration with one relay, instead, shows a different behavior. In this case, both throughput and delay (the latter is obtained according to equation (1)) are affected by the eight RFDs directly connected to the coordinator and, therefore, able to transmit more frequently, since they are not affected by the presence of a relay acting as a bottleneck. In Fig. 4c, the delay is shown as a function of the network transmission rate.

### 3.1.2. N = 16 RFDs

We first consider a scenario with *uniform* clustering. The performance results, in terms of network transmission rate and delay, are shown in Fig. 5, where various network configurations are compared. Looking at the results in Fig. 5a, where the network transmission rate is shown as a function of the network tolerable death level, one can observe that the best performance is obtained, as in the case with 64 RFDs, in the absence of clustering. Similarly to the case with 64 RFDs, the transmission rate is a decreasing function of $\chi_{net}$ and the shape of the curves is the same regardless of the number of clusters. From the results in Fig. 5b, where the delay is shown as a function of $\chi_{net}$, one can notice that there is a substantial agreement with the delay performance in the case with 64 RFDs (compare, for instance, Fig. 5b with Fig. 3b). As in the case with 64 RFDs, the lowest possible delay is obtained without clustering. In the presence of clustering, instead, the best performance is obtained in the case with eight 2-sensor clusters. This can intuitively be explained considering that

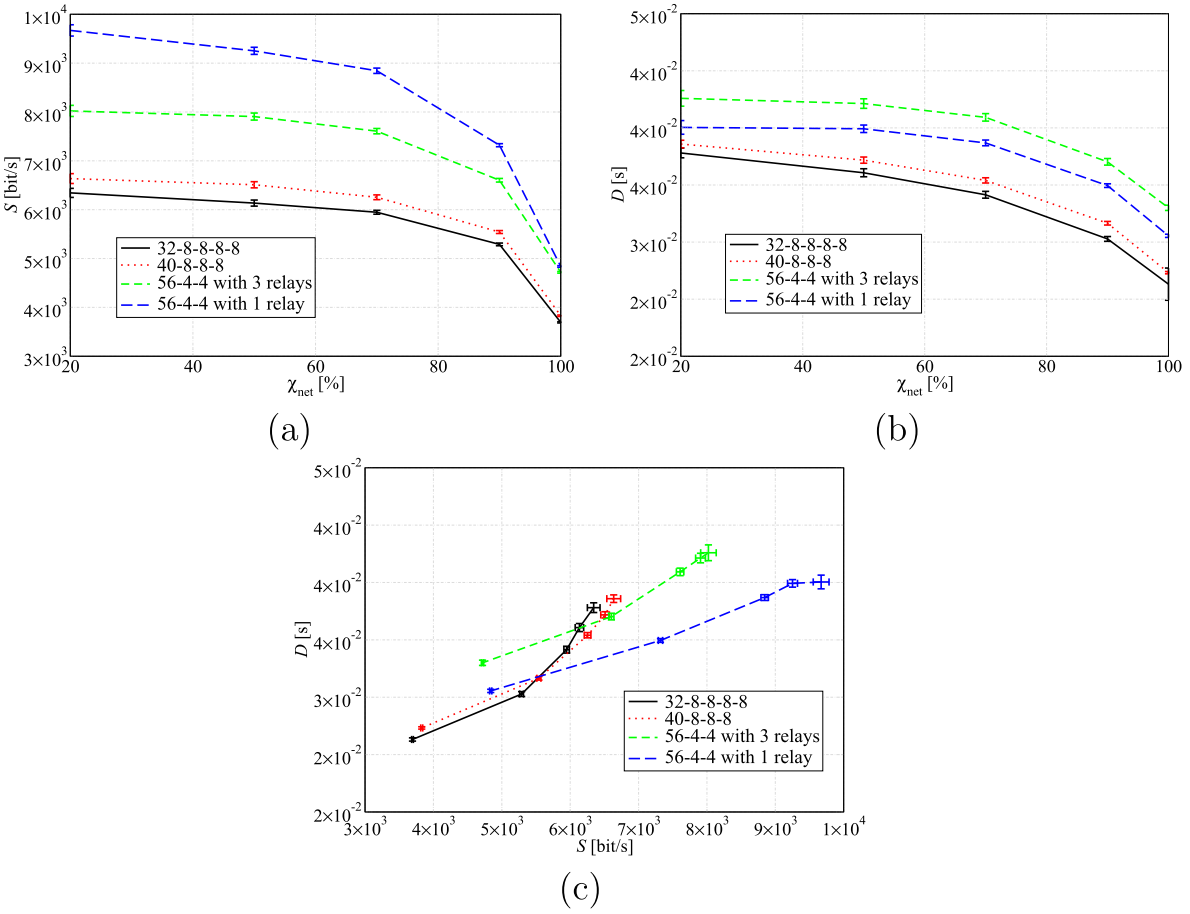


**Fig. 4.** Performance evaluation in a scenario with $N = 64$ RFDs and various non-uniform clustering configurations: (a) network transmission rate and (b) delay as functions of $\chi_{net}$; (c) delay as a function (parameterized with $\chi_{net}$) of the network transmission rate.

when a relay is connected to a large number of RFDs (i.e., a large cluster), it is likely that the RFDs find the channel "busy" and, therefore, have to wait longer in order to transmit their packets. In the case with only a few RFDs connected to a relay, instead, the latter has to manage a limited number of packets.

Let us now turn our attention to a scenario with *non-uniform* clustering. The performance results, in terms of network transmission rate and delay, are shown in Fig. 6. Considering the performance results in Fig. 6a, where the network transmission rate is shown as a function of $\chi_{net}$, one can observe that the highest network transmission rate is obtained, regardless of the value of $\chi_{net}$, with the configuration 14–1–1 with one relay, formed by one big cluster (connected to the AP through a relay) and two single RFDs connected directly to the AP. This can be explained observing that, unlike the scenarios where an RFD is connected to the AP through a relay, in this case the two RFDs connected to the AP can send their data packets directly, avoiding a "bottleneck" relay. For all the other non-uniform clustered configurations, one can observe that there is no significant difference. In particular, this difference becomes negligible at large values of $\chi_{net}$. Observing the delay performance in Fig. 6b, the lowest delay is guaranteed almost everywhere by the 8–2–2–2–2 configuration. However, the delay associated with 14–1–1 configuration with one relay, i.e., the one which guarantees highest network transmission rate, has a peculiar behavior. More precisely, the delay curve of this configuration is a steep function of $\chi_{net}$: for small values of $\chi_{net}$, the delay is approximately equal to that associated with the 10–2–2–2 clustering configuration; when $\chi_{net}$ becomes 100%, however, the corresponding delay becomes lowest. Finally, in Fig. 6c the delay is shown as a function of the network transmission rate, for various values of the network tolerable death level $\chi_{net}$. As one can see, the curves are closer to each other than in a scenario with uniform clustering. In this case as well, the choice of the network operating point depends on the specific user needs.

In Figs. 7 and 8, the performance of the networking schemes considered in Figs. 5 and 6 is analyzed by increasing the packet generation rate $g$ from 2 pck/s to 10 pck/s, considering both uniform and non-uniform clustering. The network transmission rate, in the case with $g = 10$ pck/s, has the same shape than in the case with $g = 2$ pck/s, the only difference being the fact that it is five times higher. In fact, even if the traffic load is five times higher, this load is not critical, so that the number of

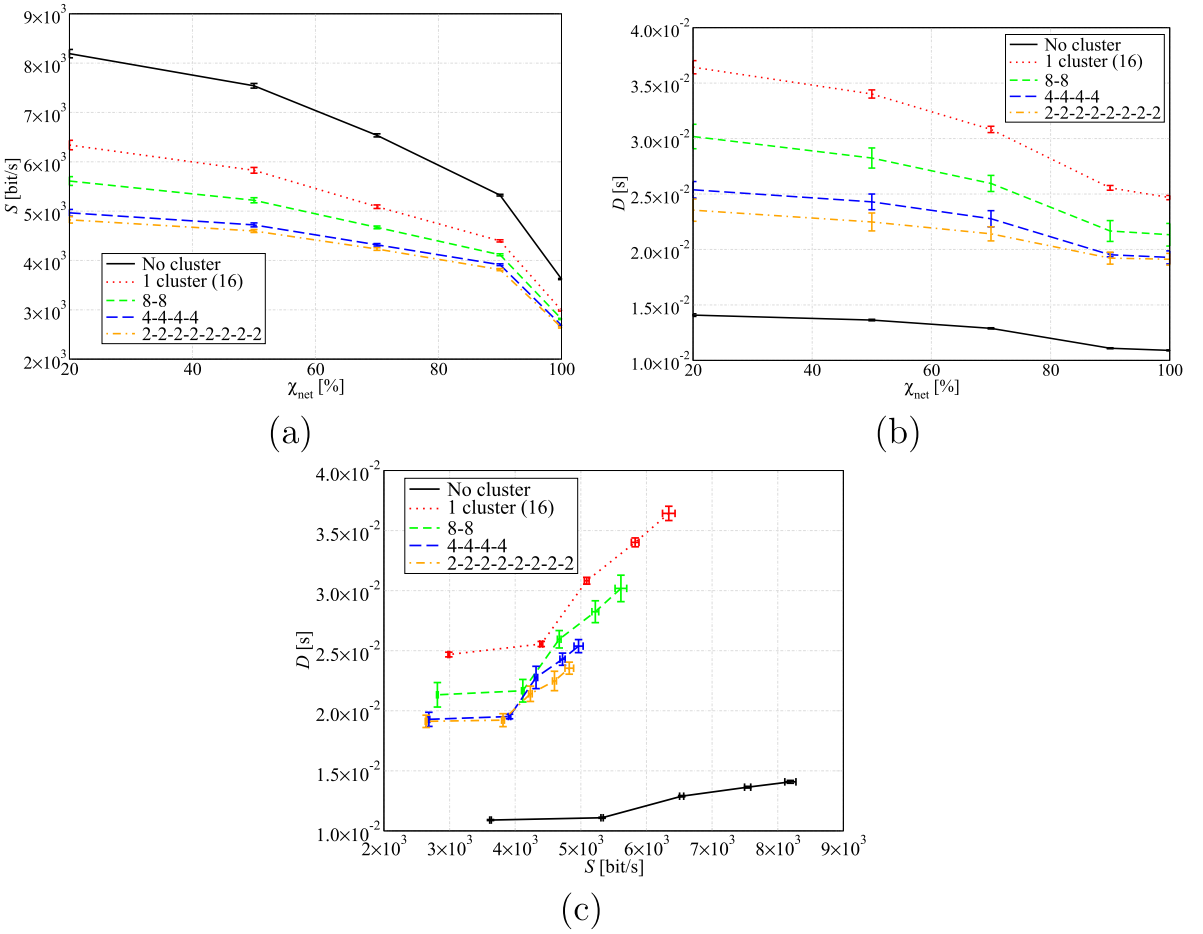


**Fig. 5.** Performance evaluation in a scenario with $N = 16$ RFDs and various *uniform* clustering configurations: (a) network transmission rate and (b) delay as functions of $\chi_{net}$; (c) delay as a function (parameterized with $\chi_{net}$) of the network transmission rate. The packet generation rate is equal to $g = 2$ pck/s.

collisions is limited and each generated packet can be immediately transmitted.

The previous results have been obtained considering an average node lifetime $\mu$ equal to 300 s. However, the conclusions on the impact of the network configuration on the system performance hold regardless of value $\mu$ of the average lifetime. In order to verify this, we have compared the performance results in Fig. 5a and b with those obtained considering $\mu = 600$ s. A direct comparison, in terms of network transmission rate and delay, is considered in Fig. 9a and b, respectively. As one can observe, the discrepancy between the performance with 300 s and with 600 s is negligible, thus making the conclusions of this work applicable to Zigbee WSNs with longer lifetime.

### 3.2. Impact of the packet generation rate

For the sake of conciseness, we investigate the impact of the packet generation rate only in scenarios with $N = 16$ RFDs—similar results, trend-wise, hold also for other values of $N$. In particular, we consider the same network topologies described in Section 2.3, i.e., with both uniform and non-uniform clustering. The network tolerable death level is set to 50%, i.e., the simulation stops when half of the nodes die.

We first analyze the network performance with the following uniform clustering configurations: (i) no clustering (i.e., direct transmission from the RFDs to the AP), (ii) one 16-node cluster with one relay, (iii) two 8-node clusters with two relays, (iv) four 4-node clusters with four relays, and (v) eight 2-node clusters with two relays. In Fig. 10a and b, the network transmission rate and delay are analyzed as functions of $g$, respectively.

From the results in Fig. 10a, one can conclude that the network transmission rate shows a similar behavior in all considered network scenarios, except for variations of the confidence interval due to the limited duration of simulations.[1] All curves shown in Fig. 10a first increase till a maximum value, after which they slowly decrease. In this case,

[1] The Opnet simulator is based on a pseudo-random value generator, which can be initialized through a user-defined seed. In our simulations, we have experienced that there exist some values (not predictable) of this seed, which lead to a worse performance. If a larger number of simulations were considered, i.e., a larger number of seeds (e.g., 100 different seeds instead of 10) were used, the confidence interval would reduce even further.

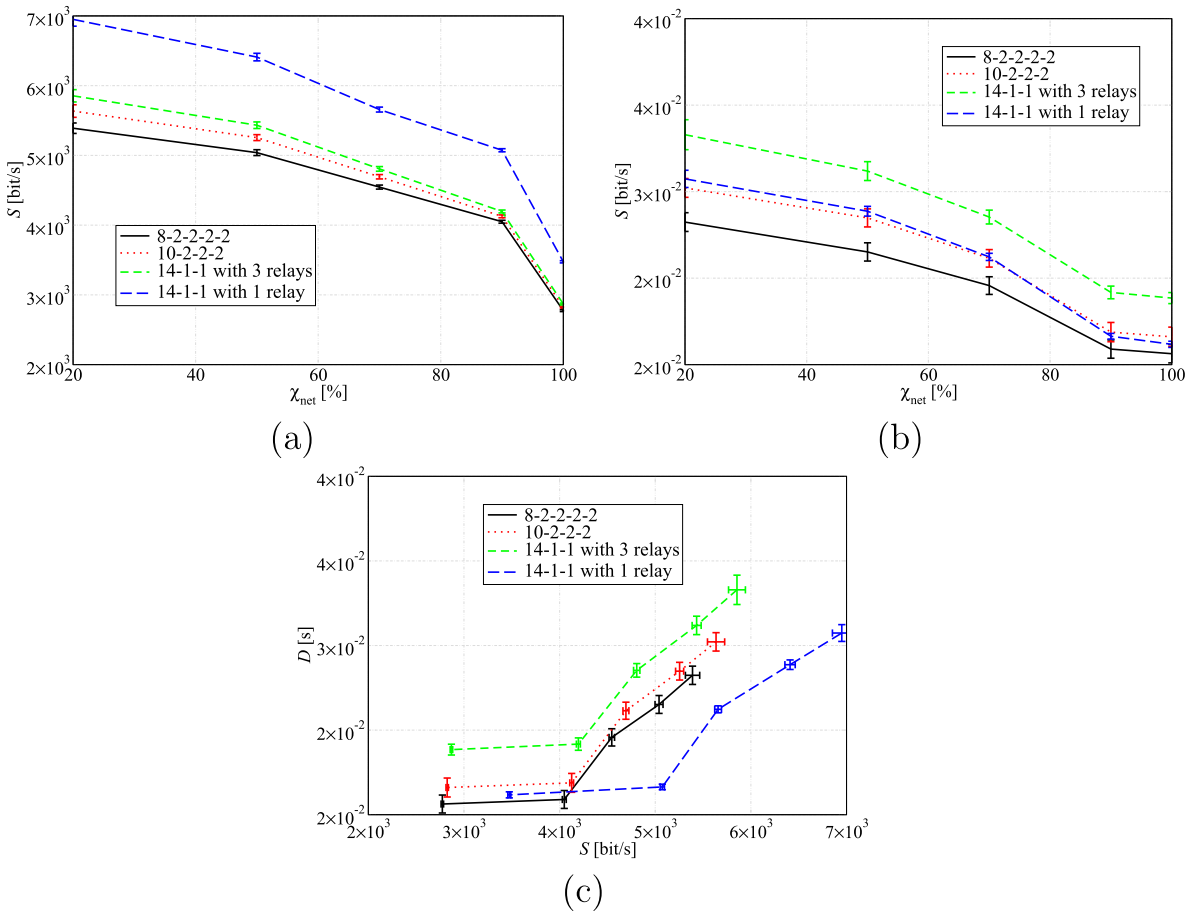


**Fig. 6.** Performance evaluation in a scenario with $N = 16$ RFDs and various *non-uniform* clustering configurations: (a) network transmission rate and (b) delay as functions of $\chi_{net}$; (c) delay as a function (parameterized with $\chi_{net}$) of the network transmission rate. The packet generation rate is equal to $g = 2$ pck/s.

as in the other previously described cases, the best performance is obtained without clustering. In the presence of clustering, instead, the network transmission rate curves change their behavior, approximately in correspondence to a packet generation rate equal to 20 pck/s, because of the increasing traffic load. In a scenario with low traffic load, the best clustered configuration is that with one 16-node cluster and one relay, since the latter node can efficiently manage all packets sent by the RFDs. When the traffic load increases, instead, the best network configuration is that with eight 2-node clusters with eight relays. In fact, in the presence of higher traffic load, it is more likely to find the relays busy, therefore the probability of finding a relay idle increases when number of relays used in the network is larger.

In Fig. 10b, the delay is investigated as a function of the packet generation rate. All curves remain constant and low for small values of the packet generation rate, then the delay quickly increases and reaches a maximum value which depends on the specific clustering configuration. Considering clustered scenarios, the best configuration is that with eight 2-node clusters with eight relays, whereas the worst performance is obtained in the scenario with one 16-node cluster with one relay. Intuitively, in the presence of a few small clusters, it is likely that the relays are ready to receive new incoming packets. On the opposite, when there is only one 16-node cluster with one relay, it is likely that a node, in need of sending a data packet, finds the relay node occupied by another transmitting node.

In Fig. 11a and b, instead, we carry out the same performance analysis (in terms of network transmission rate and delay) obtained in the presence of non-uniform clustering. The considered network configurations are the following: (i) 8–2–2–2–2 with five relays (i.e., the network is divided into five clusters, one formed by eight RFDs and the other four formed by two RFDs each, and each cluster is connected to the AP through a relay); (ii) 10–2–2–2 with four relays (i.e., the network is divided into four clusters, one formed by 10 RFDs and other formed by two RFDs each, connected to the AP through a relay), as shown in Fig. 2b; (iii) 14–1–1 with three relays (i.e., one cluster is composed by 14 RFDs and two clusters are composed by one node each), as shown in Fig. 2c; and (iv) 14–1–1 with one relay (i.e., only one cluster composed by 14 nodes and connected to the AP through a relay, while the other two RFDs communicate directly to the AP), as shown in Fig. 2d.

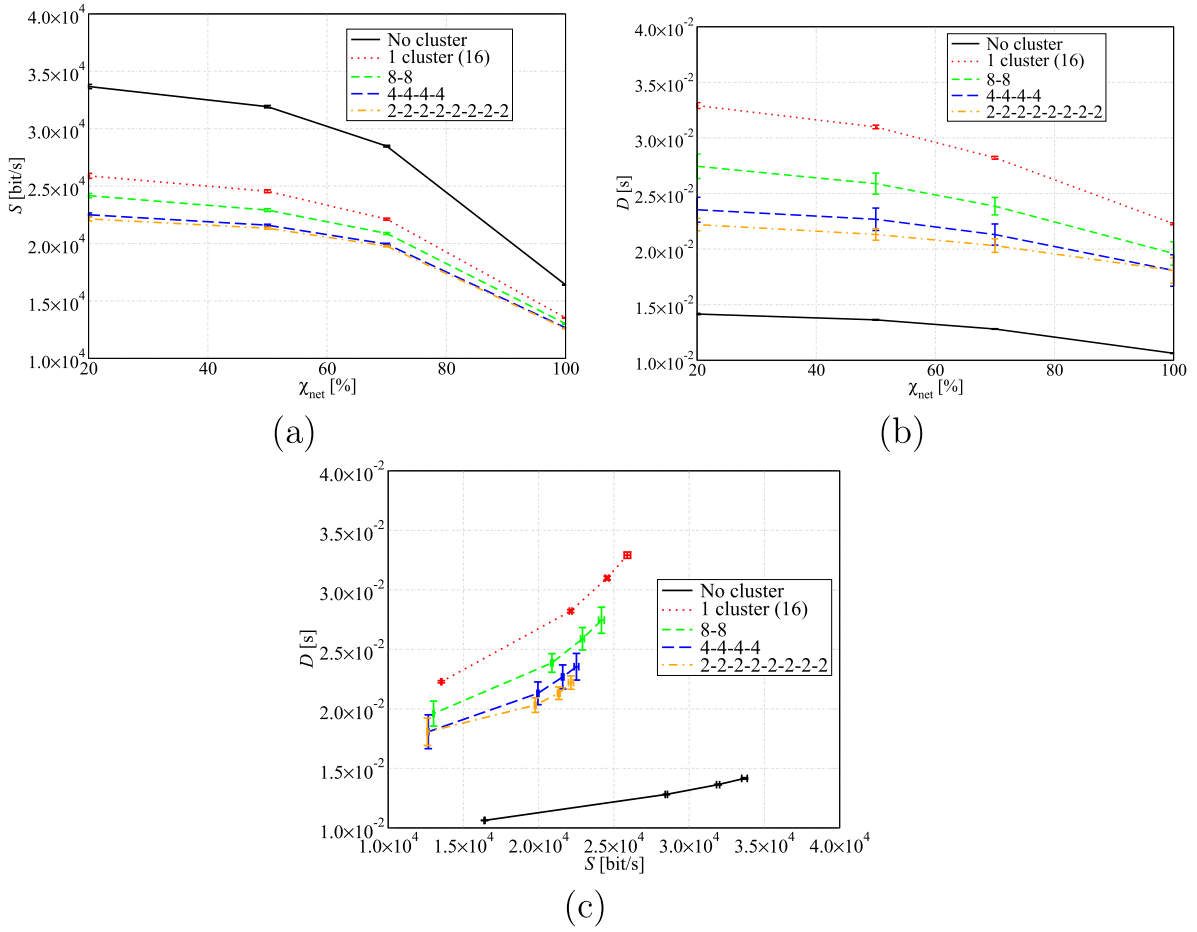


**Fig. 7.** Performance evaluation in a scenario with $N = 16$ RFDs and various *uniform* clustering configurations: (a) network transmission rate and (b) delay as functions of $\chi_{net}$; (c) delay as a function (parameterized with $\chi_{net}$) of the network transmission rate. The packet generation rate is equal to $g = 10$ pck/s.

In Fig. 11a, the network transmission rate is shown as a function of the packet generation rate. As one can see, all curves present a very similar behavior: for low values of the packet generation rate, the network transmission rate quickly increases, then reaches a maximum, and finally decreases to a saturation value which depends on the network configuration. When the packet generation rate is high, the best performance is obtained with the 14–1–1 clustering configuration with one relay. On the opposite, the worst performance is obtained with the 14–1–1 clustering configuration with three relays. The presence of two RFDs connected directly to the AP, in fact, has a beneficial effect on the network transmission rate, since it is likely that these two directly connected RFDs transmit without waiting. If we consider scenarios where all RFDs are connected to the AP through a relay, according to the results in Fig. 10a the best performance is obtained in the scenario with many small clusters.

In Fig. 11b, the delay is shown as a function of the packet generation rate. Unlike the case with uniform clustering (see Fig. 10b), in this case the curves are close to each other. For small values of the packet generation rate the delay is low, but it rapidly increases for values of the packet generation rate between 20 pck/s and 30 pck/s. Finally, when the packet generation rate is high, the delay saturates to a maximum value, which is approximately the same for all network configurations. In this case as well, the best performance, in terms of delay, is obtained in the scenario with many small clusters. In the scenario with 14–1–1 clustering configuration and one relay, unlike the scenarios where all RFDs are connected to the AP through one relay, the overall delay is affected by the presence of the two directly connected RFDs, which are likely to reserve communications with the AP.

### 3.3. Impact of data fusion mechanisms

In *non-clustered* scenarios, various values of the number $N$ of RFDs are considered. In *clustered* scenarios, instead, the number $N$ of RFDs is set to 16, and various clustering configurations are investigated: (i) 8–8 (i.e., two 8-RFD clusters with two FCs); (ii) 4–4–4–4 (i.e., four 4-RFD clusters with four FCs); (iii) 14–1–1 with three FCs (i.e., one cluster is composed by 14 RFDs and two clusters are composed by one RFD each); (iv) 10–2–2–2 with four FCs (i.e., the network is divided into four clusters, one formed by 10

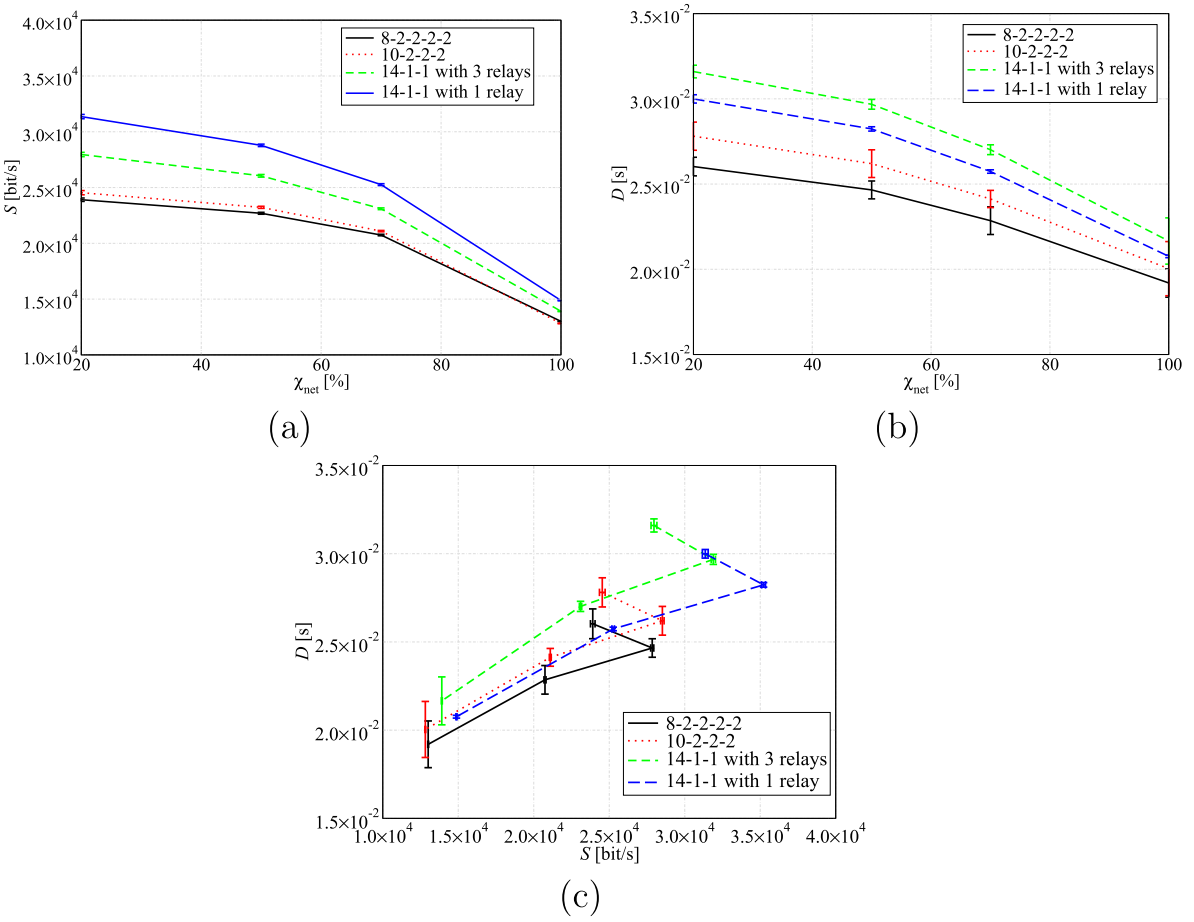


**Fig. 8.** Performance evaluation in a scenario with $N = 16$ RFDs and various *non-uniform* clustering configurations: (a) network transmission rate and (b) delay as functions of $\chi_{net}$; (c) delay as a function (parameterized with $\chi_{net}$) of the network transmission rate. The packet generation rate is equal to $g = 10$ pck/s.

RFDs and others formed by two RFDs each); and (v) 8–2–2–2–2 with five FCs (i.e., the network is divided into five clusters, one formed by eight RFDs and others formed by two RFDs each).

In Fig. 12, the probability of decision error at the AP is shown, as a function of the observation SNR, in scenarios without clustering. In the configuration with $N = 1$ RFD (solid line with circles), i.e., point-to-point communication

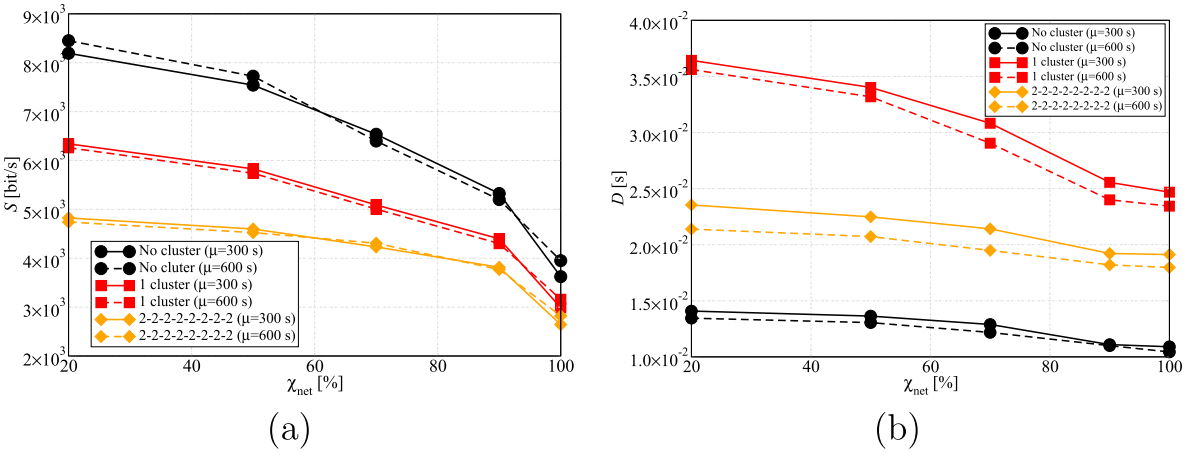


**Fig. 9.** Performance evaluation in a scenario with $N = 16$ RFDs, various uniform clustering configurations and different values of mean node lifetime: (a) network transmission rate and (b) delay as functions of $\chi_{net}$. The packet generation rate is equal to $g = 2$ pck/s.

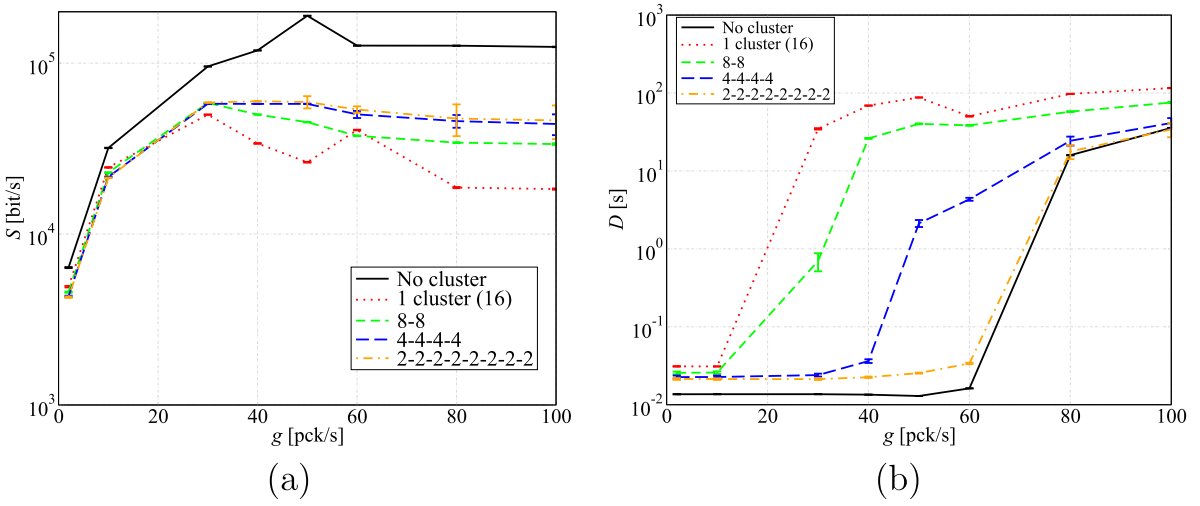


**Fig. 10.** Performance evaluation in a scenario with $N = 16$ RFDs and various *uniform* clustering configurations: (a) network transmission rate and (b) delay as functions of the packet generation rate. The network tolerable death level $\chi_{net}$ is fixed to 50%.

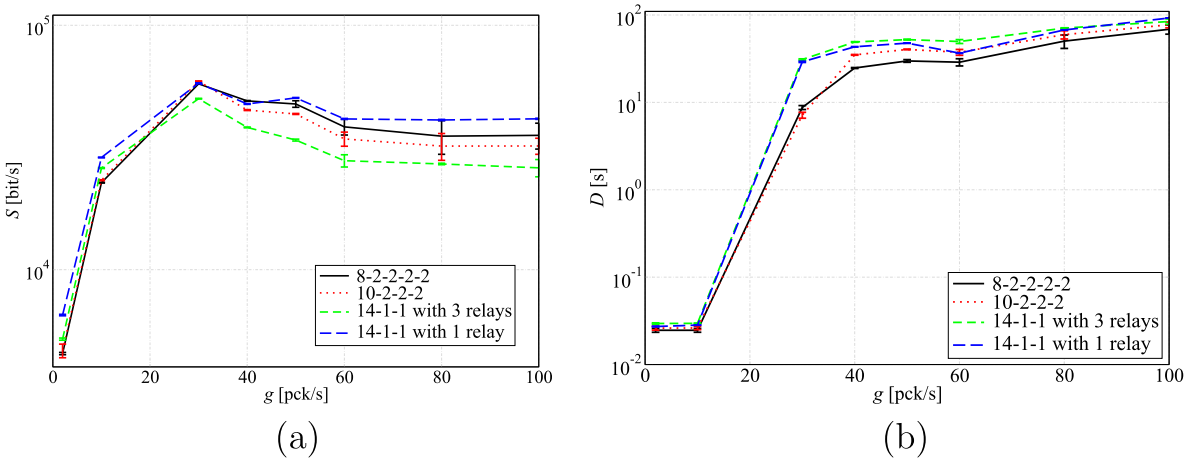


**Fig. 11.** Performance evaluation in a scenario with $N = 16$ RFDs and various *non-uniform* clustering configurations: (a) network transmission rate and (b) delay as functions of the packet generation rate. The network tolerable death level $\chi_{net}$ is fixed to 50%.

between an RFD and the coordinator, the probability of decision error has the typical trend of On Off Keying (OOK)—this is a "sanity check" for our simulator.[2] The coordinator, in fact, may decide only for either 0 or 1 using the (possibly erroneous) decision received from the RFD. As one can see, the probability of decision error reduces for increasing values of $N$. This can be explained recalling the data fusion mechanism described in Section 2.4. In fact, since communication links are modeled as ideal, for a fixed observation SNR, the larger the number of received decisions, the lower the probability of decision error with a majority decision rule. Note that for increasing numbers of RFDs, the relative reduction of the probability of decision error is negligible (e.g., the improvement is higher when the number of RFDs increases from 1 to 10 than when the number of RFDs increases from 20 to 30). In the same figure, we also show theoretical results (dashed lines) obtained with the analytical framework proposed in [20] and summarized in Section 2.4. Since this analytical framework does not take into account the medium access policies, the predicted performance is in agreement with the simulation results only in the scenario with $N = 1$ RFD. In the other scenarios, the probability of decision error predicted by the analytical framework is better than that obtained through simulations. This is due to the fact that our Opnet simulator takes also into account the presence of collisions in the medium access phase. In this case, since some packets may be lost or dropped, the probability

[2] Note that the modulation schemes foreseen by the Zigbee standard have no effect on the probability of decision error, since the channel is considered ideal.

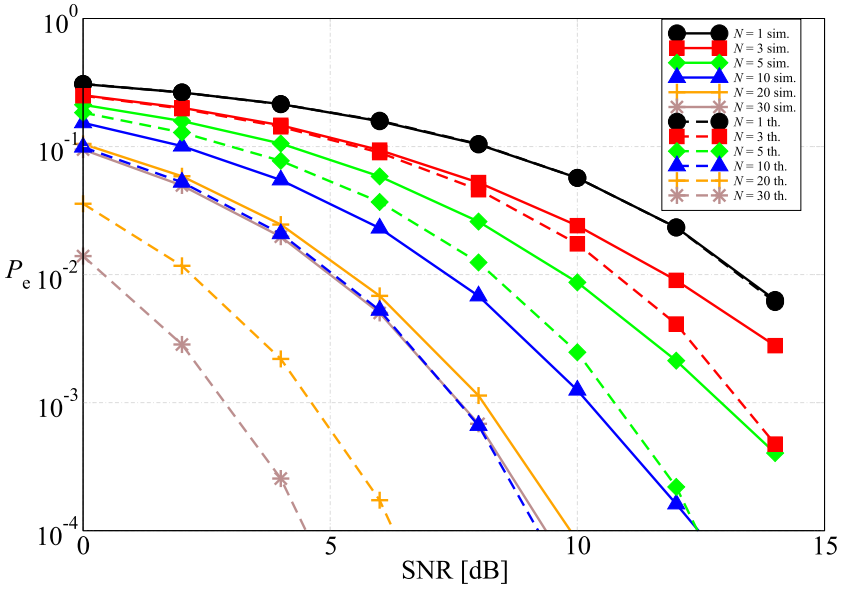


**Fig. 12.** Performance analysis in a scenario without clustering: probability of decision error performance as a function of the SNR at the RFDs. Both simulations (solid lines) and theoretical (dashed lines) results are shown.

of decision error worsens. The negative impact of the collisions exacerbates when the number of transmitting nodes increases.

In Fig. 13, the performance is analyzed in scenarios without clustering: more precisely, (a) the probability of decision error is shown as a function of the number of transmitting RFDs and (b) the throughput and the delay are shown as functions of the number of nodes. In Fig. 13a, the probability of decision error is a monotonically decreasing function of $N$ for all considered values of the observation SNR at the RFDs. For each network configuration, the probability of decision error reaches a floor which is strictly related to the collisions and, therefore, to the maximum achievable throughput in each scenario—in the cases with high SNR (12 dB and 14 dB), instead, the floor is not visible since it appears for very low (out of scale) values of the probability of decision error.

In Fig. 13b, the throughput and the delay are shown as functions of the number of transmitting RFDs. These curves are obtained by considering a fixed SNR (equal to 0 dB) at the RFDs. Our results, however, show that the throughput and the delay are not affected by the value of the observation SNR at the RFDs.[3] We consider, in fact, ideal communication channels, so that the noise affects only the observations at the RFDs and not the transmissions from the RFDs or the FCs. Consequently, throughput and delay do not depend on the observation SNR. The throughput curve (solid line with circles) is a decreasing function of $N$. In particular, for small values of $N$, it remains close to 1 (the maximum possible). When $N$ increases, instead, the number of collisions in the channel increases as well and the throughput reduces. Comparing the results in Fig. 13a with those in Fig. 13b, the negative impact that a larger number of RFDs has on the throughput is compensated, in terms of probability of decision error, by the data fusion mechanism. In Fig. 13b, we also show the delay (dotted line with diamonds). As the intuition may suggest, the delay is small for small values of $N$. When the traffic increases, instead, due to a larger number of collisions, the delay is higher, since the channel is busy for a longer period of time and the probability of finding the channel idle reduces. Finally, for large values of $N$, the delay seems to start saturating to a maximum value. In this case, in fact, due to the increased offered traffic, it is likely that there is at least an RFD ready to send its packet as soon as the channel is idle.

In Fig. 14, we show the aggregate throughput, computed on the basis of the previous results in scenarios without clustering, as a function of the number of nodes. When the number of transmitting nodes is small, the aggregate throughput is high (close to the maximum possible for each network configuration). When the number of transmitting RFDs increases, instead, the aggregate throughput tends to saturate. Once the saturation is reached, the number of collisions is so large than an increase of the traffic load has no longer effect on the aggregate throughput.

In Fig. 15, we analyze the impact of non-uniform clustering on the probability of decision error—as a performance benchmark, the probability of decision error in the case with uniform clustering is also shown. We consider scenarios with $N = 16$ RFDs and the following clustering configurations: (i) no clustering; (ii) two 8-RFD clusters with two FCs; (iii) four 4-RFD clusters with four FCs; (iv) 14–1–1 with three FCs; (v) 10–2–2–2 with four FCs; and (vi) 8–2–2–2–2 with five FCs. According to the analytical results presented in [20] and the previously shown simulation results, the best network performance is obtained in the absence of clustering. The worst performance, instead,

[3] This is expected, since the quality of the observations influences the decision at the AP but *not* the communications within the network.

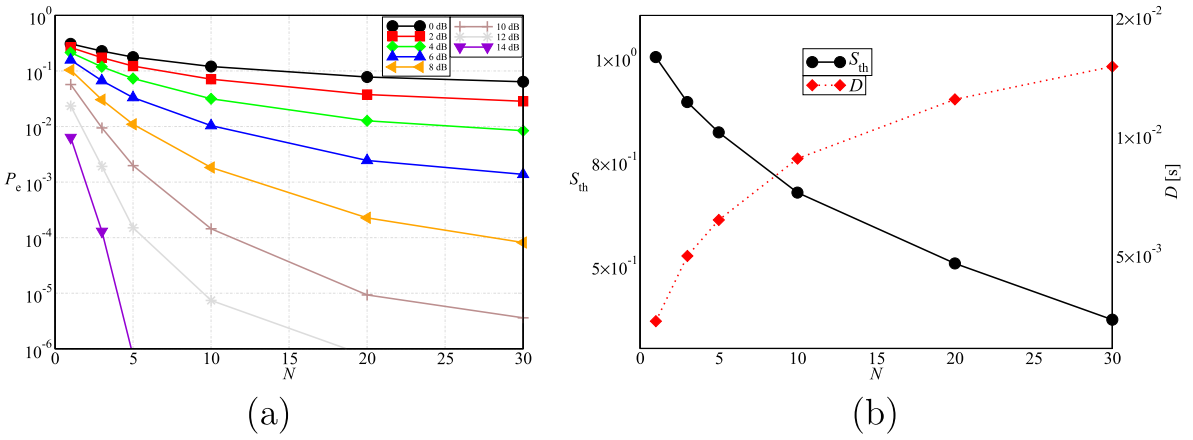


**Fig. 13.** Performance analysis in a scenario without clustering: (a) probability of decision error at the AP as a function of the number $N$ of transmitting RFDs, considering various SNR at the nodes, and (b) throughput and delay as functions of the number $N$ of transmitting RFDs.

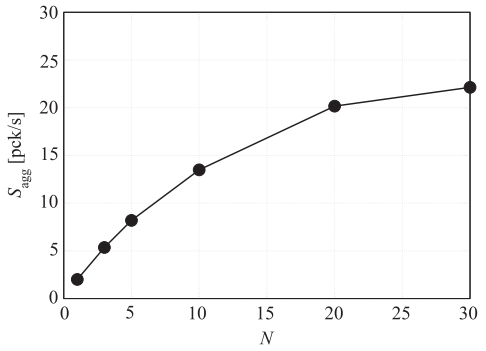


**Fig. 14.** Aggregate throughput, in a scenario without clustering, as a function of the number $N$ of transmitting RFDs.

is obtained in the 14–1–1 scenario, i.e., with three FCs and non-uniform clustering. In this case, in fact, the information collected by the RFDs associated with the largest cluster is very reliable. On the other hand, the information collected by the other two clusters is more likely to be noisy, and the final decision is thus likely to be wrong. Observing the results in Fig. 15, one can conclude that, in the presence of non-uniform clustering, the best performance is obtained with slightly unbalanced clusters. In this case, in fact, decisions made by intermediate FCs are more reliable, so that it is more likely that the final decision made by the coordinator is correct. In the case of uniform clustering, instead, the probability of decision error is *not* affected by the number of clusters in the network, as long as the number of RFDs remains the same. Observing Fig. 15, one can note that the curves corresponding to the scenarios with four 4-RFD clusters and two 8-RFD clusters are almost overlapped. This is due to the fact that a smaller number of clusters is compensated by a higher quality of the intermediate decisions. This result is in agreement with the theoretical analysis presented in [20].

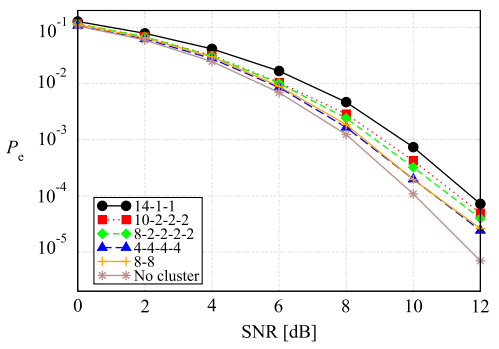


**Fig. 15.** Probability of decision error at the AP in scenarios with $N = 16$ RFDs both in uniform and non-uniform clustering configurations. The considered topologies are the following: (i) no clustering, (ii) two 8-RFD clusters and two FCs, (iii) four 4-RFD clusters and four FCs, (iv) 14–1–1 clustering configuration with three FCs, (v) 10–2–2–2 clustering configuration with four FCs, and (vi) 8–2–2–2–2 clustering configuration with five FCs.

## 4. Design guidelines

On the basis of the extensive performance analysis carried out in Section 3, we first characterize a Zigbee network with a three-dimensional performance surface which clearly shows the existing trade-offs, imposed by the clustering configuration, between network transmission rate, delay, and tolerable network death level. Then, we propose a simple, yet insightful, analytical characterization of the Zigbee performance surface. This allows to derive useful guidelines for the design of Zigbee WSNs with a desired performance level.

### 4.1. The Zigbee performance surface

In Fig. 16, we present combined results of network transmission rate, delay, and network tolerable death level in different (both uniform and non-uniform clustering) clustering configurations, explicitly indicated, with $N = 16$ RFDs.

In Fig. 16a, the delay is shown as a function of both network tolerable death level and network transmission rate. On the $\chi_{net}$–$S$ plane, the contour curves of the delay surface

are shown. The shape of the three-dimensional surface is similar to a skewed half-tube: in fact, there is a peak (in terms of delay), with maximum value in correspondence to $\chi_{net} = 20\ \%$ and $S \simeq 6$ kb/s. This means that the one cluster with one relay configuration leads to the highest delay. In an application where a small delay is required, the acceptable configurations are either the ones with no relay or the ones with many relays (i.e., high clustering level). In the former case, the network transmission rate is high, as shown in Fig. 5a and b; in the latter case, instead, the network transmission rate is limited by the presence of the relays.

In Fig. 16b, the network transmission rate is shown as a function of $\chi_{net}$ and $D$. From the results in this figure it is evident that the network transmission rate reduces as the number of relays in the network increases. In particular, for each configuration and for specific values of $\chi_{net}$ and $D$, there exists a maximum achievable network transmission rate and a suitable network configuration which guarantees the best performance. For a specific value of the network transmission rate, there exists at least a pair of values of $D$ and $\chi_{net}$—and, therefore, a specific pair of network configurations—in correspondence to which the required network performance is met. In Fig. 16c, $\chi_{net}$ is shown as a function of the network transmission rate and the delay. As in the previous figures as well, in this figure, it can be seen that, for a fixed value of $\chi_{net}$, there exists a configuration that guarantees the best trade-off between network transmission rate and delay.

Summarizing, one can conclude that:

- the best network configuration is always without any intermediate relay;
- in clustering configurations where there is at least one RFD directly connected to the AP, the network transmission rate is higher than in fully clustered configurations—this conclusion does not hold in terms of delay, even if acceptable values of delay are obtained;
- both the network transmission rate and the delay are decreasing functions of the number of relays in the network: therefore, in the presence of a high clustering level both the network transmission rate and the delay are low, whereas in the presence of a low clustering level the opposite holds.

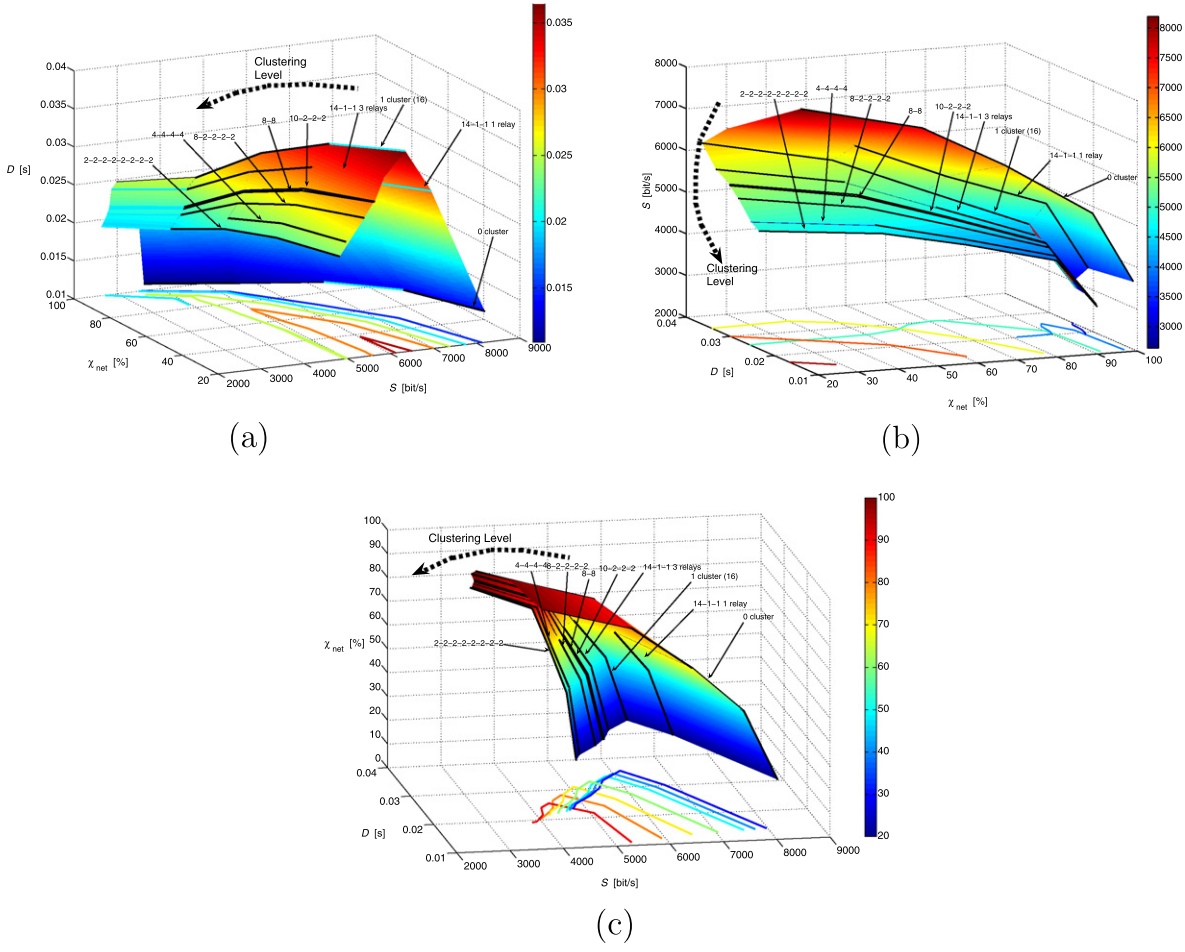


**Fig. 16.** Performance evaluation in a scenario with $N = 16$ RFDs and various uniform and non-uniform clustering configurations; (a) delay as a function of network transmission rate and $\chi_{net}$, (b) network transmission rate as a function of delay and $\chi_{net}$, and (c) $\chi_{net}$ as a function of network transmission rate and delay. The considered packet generation rate is $g = 2$ pck/s.

In Fig. 17, the same results of Fig. 16 are presented in scenarios with $N = 64$ RFDs. The considerations, carried out for scenarios with $N = 16$ RFDs, are still valid, the only difference between the two network configurations being the fact that the network transmission rate and the delay in the scenario with $N = 64$ RFDs are higher than in the scenario with $N = 16$ RFDs. As before, the most important observation is that the network performance trend is strictly related to the number of relays in the network.

### 4.2. A simple analytical model for the Zigbee performance surface

We now provide a simple approach for deriving a closed-form approximation of the three-dimensional Zigbee performance surface introduced in Section 4.1. The derivation of an analytical expression for this surface is very useful, since simulation results may be difficult to obtain, especially in the case of large-scale Zigbee WSNs.

Without loss of generality, we focus on the Zigbee performance surface, which returns the delay as a function of the network transmission rate and the network tolerable death level. Our goal is to fit the simulation-based delay surface with a simple function $\widehat{D} = \widehat{D}(S, \chi_{\text{net}})$. In particular, we consider a fitting function with the following form:

$$\widehat{D} = \alpha_1 f_1(S, \chi_{\text{net}}) + \alpha_2 \tag{5}$$

where the values of $\alpha_1$ and $\alpha_2$ are obtained through the minimization of the mean square error (MSE) between simulation and analytical data and an expression for $f_1$ is derived in Appendix A.

In Fig. 18, the "true" Zigbee performance surface (the one with lines) and its analytical approximation (the continuous one) are compared in scenarios with various values of $N$: (a) 16, (b) 32, (c) 64, and (d) 128, respectively. As one can observe, for all considered scenarios the heuristically derived surfaces approximate well the simulation-based surfaces. However, some differences between the true surfaces and their approximations can be observed, and this is due to the approximations considered in Appendix A. In Table 1, the optimized values of $\alpha_1$ and

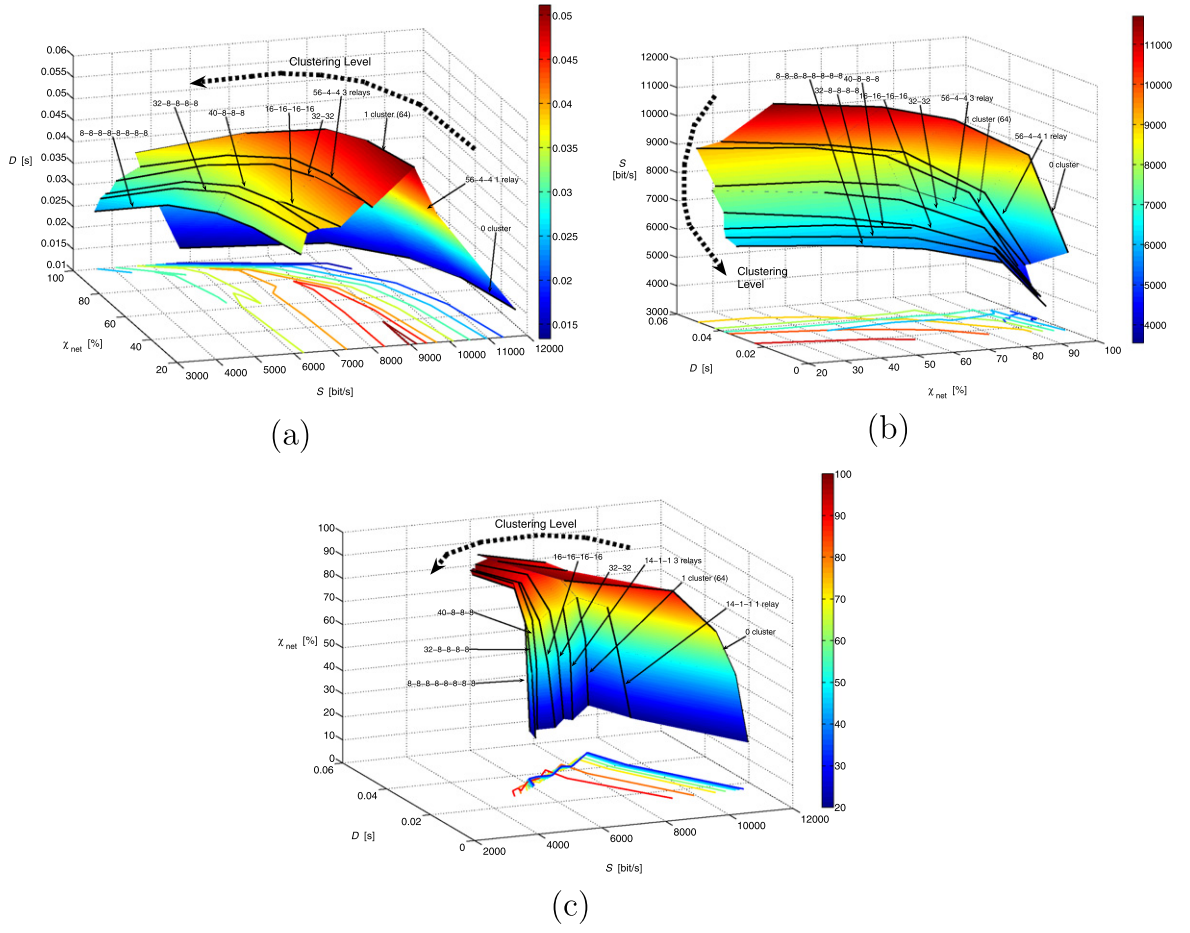


**Fig. 17.** Performance evaluation in a scenario with $N = 64$ RFDs and various uniform and non-uniform clustering configurations; (a) delay as a function of network transmission rate and $\chi_{\text{net}}$, (b) network transmission rate as a function of delay and $\chi_{\text{net}}$, and (c) $\chi_{\text{net}}$ as a function of network transmission rate and delay. The considered packet generation rate is $g = 2$ pck/s.

$\alpha_2$, and the corresponding delay minimum mean square errors (MMSEs) are shown, in scenarios with 16, 32, 64, and 128 RFDs, respectively. More precisely, the normalized MMSE is computed as

$$\text{MMSE} = \frac{|D - \widehat{D}|^2}{|D|^2}$$

where the notation $|\boldsymbol{x}|$ stands for the norm of $\boldsymbol{x}$. From the results in Table 1, the reader should observe that the normalized MMSE is relatively small (lower than 9%) in all scenarios. Therefore, the proposed approximation for the Zigbee network surface can be applied also to large-scale WSNs.

In order to provide an intuitive comparison of the approximating surfaces, in Fig. 19 the four surfaces of Fig. 18, relative to different values of $N$, are shown together. It can be clearly seen that the surfaces have the same shape, with the position of the peak on the $D$–$\chi_{\text{net}}$ and $S$–$\chi_{\text{net}}$ simply translated according to the considered value of $N$. The result of this analysis is that, given the number of nodes $N$ in the network, it is always possible to characterize the performance of a Zigbee WSN through the approximation expression given by (5). Further details can be found in Appendix A.

**Table 1**
Coefficients of the linear combination in (5) and corresponding normalized MMSEs for various values of $N$.

| | $N = 16$ | $N = 32$ | $N = 64$ | $N = 128$ |
|---|---|---|---|---|
| $\alpha_1$ | 0.154 | 0.4716 | 0.128 | 0.3952 |
| $\alpha_2$ | 0.015 | 0.0196 | 0.0028 | 0.0261 |
| MMSE (%) | 1.53 | 2.29 | 3.9 | 8.95 |

The derivation of an accurate approximation for the Zigbee performance surface allows to simplify the analysis and design of WSNs. In particular, we are currently working on the derivation of a general analytical model for the coefficients in (5), in order to have an accurate approximation of the performance surface for a generic value of $N$. Our simple analytical model could also help in analyzing the network behavior in dynamical scenarios, where the network architecture (e.g., the clustering configuration) may change as a consequence of RFDs' failures or adaptive reclustering. Finally, taking into account the performance trade-offs implied by the Zigbee performance surface, one could design the WSN in order to maintain a given performance level, and this goal could be achieved by properly "moving" the network operating point on the surface.

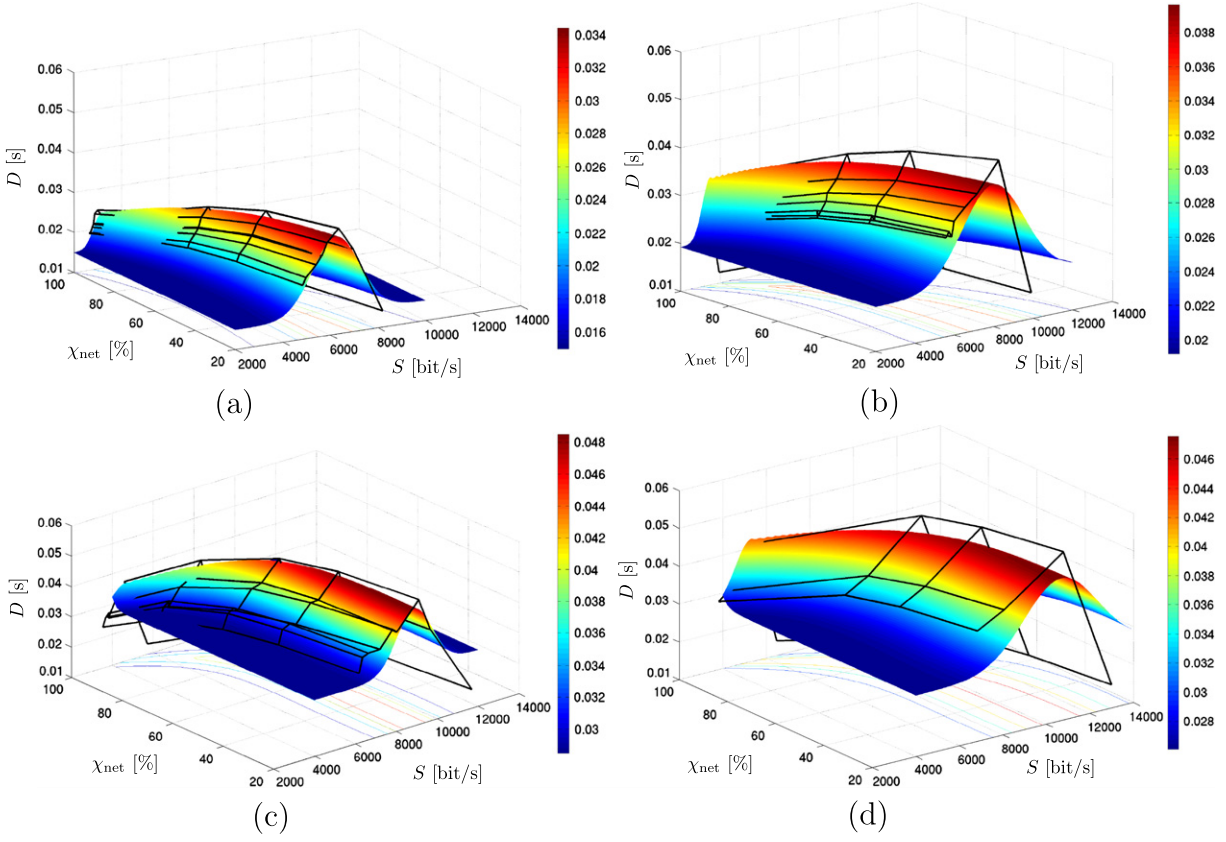


**Fig. 18.** Performance evaluation in a scenario with the following values of $N$: (a) 16, (b) 32, (c) 64, and (d) 128. Both the realistic surface (lines) and its approximation (continuous) are shown.

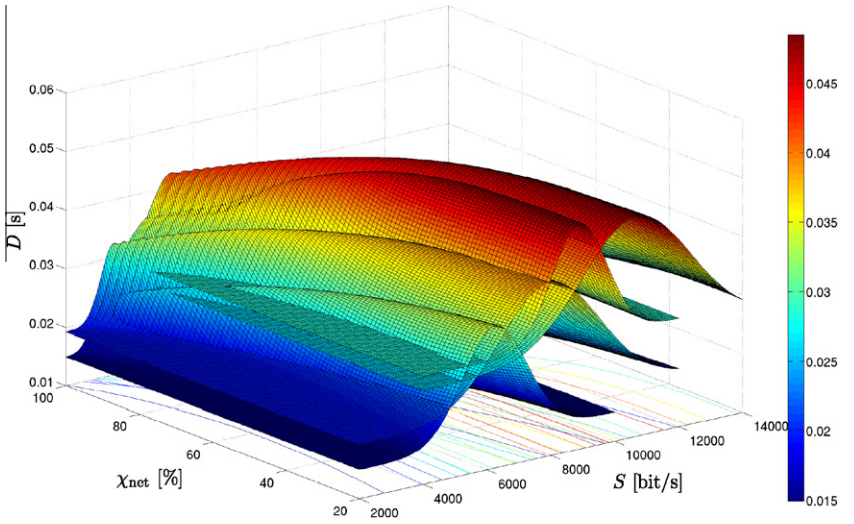


**Fig. 19.** Overlapped approximating Zigbee performance surfaces in scenarios with 16, 32, 64, and 128 RFDs.

## 5. Concluding remarks

In this paper, we have analyzed, through Opnet-based simulations, the performance of Zigbee sensor networks, using physical (probability of decision error at the AP) and network layer (network transmission rate, throughput, aggregate throughput, delay, and network lifetime) performance indicators. In non-clustered scenarios, the presence of a large number of transmitting RFDs has a positive effect on the probability of decision error, at the price of throughput and delay performance degradation. We have also extensively analyzed the performance of Zigbee networks in clustered configurations with and without data fusion. In the presence of data fusion, our results confirm the theoretical results in [20]. In addition, through a simulation-based analysis of the impact of the network tolerable death level, it appears that the best network configuration is always the one without any relay. Finally, we have drawn a few simple guidelines and an approximating model for the design of clustered Zigbee WSNs. A three-dimensional characterization of the network performance, in terms of $D$, $S$, and $\chi_{net}$, shows that the network operating point lies over a characteristic surface, denoted as Zigbee performance surface. Given the number of nodes and the required performance level, one can identify over this surface the network configuration which guarantees the best possible trade-off.

## Acknowledgements

We would like to thank Andrea Muzzini (A.E.B., Cavriago, Italy) for his help in part of the derivation of the data fusion mechanism implemented in the Opnet simulator.

## Appendix A. Analytical approximation of the performance surfaces

In order to derive an analytical model of the delay Zigbee performance surface, we first extract an equation which describes the "peak" (in terms of delay) of the half-tube. Since the maximum of each surface corresponds to the configuration with one cluster with one relay, we first obtain an approximating expression for the one-cluster projection curve on the $S$–$\chi_{net}$ plane (i.e., considering the surface in Fig. 17a). This curve can be accurately approximated as follows[4]:

$$g(\chi_{net}) \triangleq a_1 \cdot \chi_{net}^{b_1} + c_1 \tag{6}$$

where $a_1$, $b_1$, and $c_1$ are proper constants whose values are obtained by minimizing the MSE with respect to the simulation-based points. In Fig. 20, we show simulation results and the proposed analytical approximation for the above curve relative to the Zigbee performance surface with $N = 64$ RFDs. The values of $a_1$, $b_1$, and $c_1$, obtained using the Curve Fitting Toolbox of Matlab [26] with a 95% confidence, are shown in Table 2.

The same procedure can be repeated to approximate the shape of the projection of the "peak" of the surface (for $S = g(\chi_{net})$) on the $D$–$\chi_{net}$ plane. The fitting expression has the following form:

$$\ell(\chi_{net}) \triangleq a_2 \cdot e^{b_2 \cdot \chi_{net}} + c_2 \cdot e^{d_2 \cdot \chi_{net}} \tag{7}$$

where the coefficients $a_2$, $b_2$, $c_2$, and $d_2$ can be obtained with the MMSE-based approach used to determine the values of the coefficient in (6). In Fig. 21, the simulation results are compared with the proposed analytical approximation (7) in the case with $N = 64$ RFDs. The values used for $a_2$, $b_2$, $c_2$, and $d_2$ are shown in Table 2.

Finally, for a given value of $\chi_{net}$, we have approximated the $D$–$S$ two-dimensional curve obtained from the corresponding section of the three-dimensional performance surface. From our analysis, it comes out that the best approximating function, for a fixed value of $\chi_{net}$, is a Gaussian function with (optimized) variance which depends on $\chi_{net}$. Therefore, the final expression for $f_1$ is

[4] Similar considerations can be carried out also for the case with $N = 16$ RFDs.

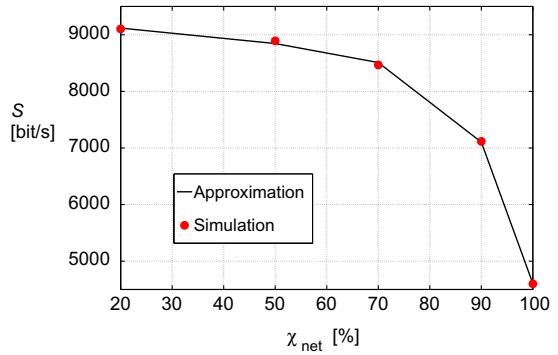


**Fig. 20.** Comparison, between approximation and simulation, of the trajectory of the maximum of the surface in the $S$–$\chi_{net}$ plane. The case with $N = 64$ RFDs is considered.

**Table 2**
Coefficients used for the computation of the approximating function in the case with $N = 16$, $N = 32$, $N = 64$, and $N = 128$, respectively.

| | $N = 16$ | $N = 32$ | $N = 64$ | $N = 128$ |
|---|---|---|---|---|
| $a_1$ | $-0.001302$ | $-2.642 \times 10^{-5}$ | $-2.943 \times 10^{-12}$ | $-7.544 \times 10^{-7}$ |
| $b_1$ | 3.187 | 4.113 | 7.579 | 4.926 |
| $c_1$ | 6256 | 7956 | 8887 | 10,680 |
| $a_2$ | $-0.002132$ | $-2.187 \times 10^{-5}$ | $-0.004571$ | $-0.0001723$ |
| $b_2$ | 0.03209 | 0.06173 | 0.03423 | 0.04315 |
| $c_2$ | 0.1355 | 0.04613 | 0.1585 | 0.05486 |
| $d_2$ | $-0.001969$ | $-0.001568$ | 0.002149 | $-0.0001743$ |
| $p_1$ | $-0.03429$ | $-0.1637$ | $-0.09405$ | $-0.2783$ |
| $p_2$ | $-7.48$ | $-0.797$ | 6.577 | 16.64 |
| $p_3$ | 1576 | 2296 | 1119 | 2177 |

$$f_1(S, \chi_{net}) = \ell(\chi_{net}) \cdot \exp\left\{ -\left[\frac{S - g(\chi_{net})}{c(\chi_{net})}\right]^2 \right\}$$

where $\ell(\chi_{net})$ and $g(\chi_{net})$ are, respectively, given by (6) and (7), and $c(\chi_{net})$ is a proper function which characterizes the standard deviation of the Gaussian approximation and can be expressed as

$$c(\chi_{net}) = p_1 \chi_{net}^2 + p_2 \chi_{net} + p_3 \quad (8)$$

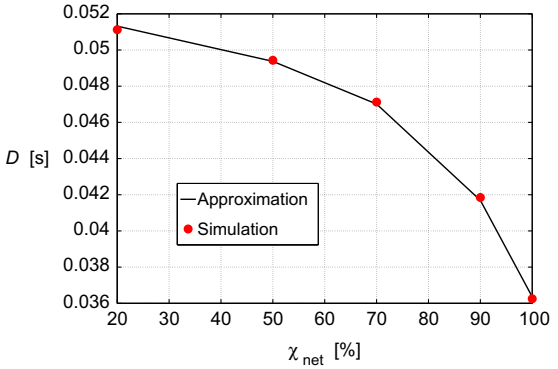


**Fig. 21.** Comparison, between approximation and simulation, of the trajectory of the outline of the surface in the $D$–$\chi_{net}$ plane. The case with $N = 64$ RFDs is presented.

The specific values of the coefficients $p_1$, $p_2$, and $p_3$ can be determined through the MMSE-based approach previously defined and are shown in Table 2.

## References


[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, A survey on sensor networks, IEEE Commun. Mag. 40 (8) (2002) 102–114.
[2] R. Abileah, D. Lewis, Monitoring high-seas fisheries with long-range passive acoustic sensors, in: Proc. Int. Conf. OCEANS'96: Prospects for the 21sr Century, Vol. 1, Fort Lauderdale, FL, USA, 1996, pp. 378–382.
[3] S. Barberis, E. Gaiani, B. Melis, G. Romano, Performance evaluation in a large environment for the awacs system, in: Proc. Int. Conf. on Universal Personal Communications (ICUPC'98), Vol. 1, Florence, Italy, 1998, pp. 721–725.
[4] C.Y. Chong, S.P. Kumar, Sensor networks: evolution, challanges, and opportunities, Proc. IEEE 91 (8) (2003) 1247–1256.
[5] S.N. Simic, S. Sastry, Distributed environmental monitorning using random sensor networks, in: Proc. 2-nd Int. Work. on Inform. Processing in Sensor Networks, Palo Alto, CA, USA, 2003, pp. 582–592.
[6] M. Madou, Fundamentals of Microfabrication, CRC Press, Sound Parkway, NW, USA, 1997.
[7] A. Bonivento, C. Fischione, L. Necchi, F. Pianegiani, A. Sangiovanni-Vincentelli, System level design for clustered wireless sensor networks, IEEE Trans. Ind. Inform. 3 (3) (2007) 202–214.
[8] B. Yin, H. Shi, Y. Shang, Analysis of energy consumption in clustered wireless sensor networks, in: Proceedings on the 2nd International Symposium on Wireless Pervasive Computing (ISWPC'07), San Juan, Porto Rico, 2007, pp. 77–82.
[9] Z. Zhou, S. Zhou, S. Cui, J. Cui, Energy-efficient cooperative communication in clustered wireless sensor networks, in: Proc. IEEE Military Comm. Conf. (MILCOM'06), Washington, DC, USA, 2006, pp. 1–7.
[10] R.R. Tenney, N.R. Sandell, Detection with distributed sensors, IEEE Trans. Aerosp. Electron. Syst. 17 (4) (1981) 501–510.
[11] W. Shi, T.W. Sun, R.D. Wesel, Quasi-convexity and optimal binary fusion for distributed detection with identical sensors in generalized Gaussian noise, IEEE Trans. Inform. Theory 47 (1) (2001) 446–450.
[12] Q. Yang, J. Sun, An underwater autonomous robot based on multi-sensor data fusion, in: Proc. 6th World Congress on Intelligent Control and Automation (WCICA'06), vol. 2, Hong Kong, China, 2006, pp. 9139–9143.
[13] B. Krishnamachari, D. Estrin, S.B. Wicker, The impact of data aggregation in wireless sensor networks, in: Proc. of the 22nd Int. Conf. on Distributed Computing Systems (ICDCS'02), Vienna, Austria, 2002, pp. 575–578.
[14] W. Su, T.L. Lim, Cross-layer design and optimization for wireless sensor networks, in: Proc. of the 7th ACIS Int. Conf. on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD-SAWN'06), Las Vegas, NV, USA, 2006, pp. 278–284.
[15] V. Kawadia, P. Kumar, A cautionary perspective on cross-layer design, IEEE Wireless Commun. 12 (1) (2005) 3–11.
[16] IEEE 802.15.4 Std: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs), IEEE Computer Society Press, 2003, pp. 1–679.
[17] IEEE 802.11 Std: Wireless LAN Medium Access Control (MAC) a Physical Layer (PHY) Specifications, IEEE Computer Society Press, 1997, pp. 1–459.
[18] R.E. Ziemer, Elements of Engineering Probability and Statistics, Prentice-Hall, Upper Saddle River, NJ, USA, 1997.
[19] G. Ferrari, M. Martalò, Extending the lifetime of sensor networks through adaptive reclustering, EURASIP J. Wireless Commun. Network., 2007 (2007) 14. Article ID 81864. doi:10.1155/2007/81864 (special issue on MobileMAN Mobile Multi-hop Ad Hoc Networks: From Theory to Reality).
[20] G. Ferrari, M. Martalò, R. Pagliari, Decentralized detection in clustered sensor networks, IEEE Trans. Aerosp. Electron. Syst., inpress. <http://www.tlc.unipr.it/martalo/privfold/preprint.pdf>.
[21] A. Papoulis, Probability, Random Variables and Stochastic Processes, McGraw-Hill, New York, NY, 1991.
[22] Opnet Website. <http://www.opnet.com>.
[23] National Institute of Standards and Technology (NIST), Opnet Model Website. <http://w3.antd.nist.gov/Health.shtml>.
[24] G. Ferrari, P. Medagliani, M. Martalò, Grid enabled instrumentation and measurement (signals and communication technology), in: F.

Davoli, N. Meyer, R. Pugliese, S. Zappatore (Eds.), Ch. Sensor Networks as Data Acquisition Devices – Performance Analysis of Zigbee Wireless Sensor Networks with Relaying, Springer-Verlag, pp. 55–79.
[25] P. Medagliani, J. Leguay, V. Gay, M. Lopez-Ramos, G. Ferrari, Engineering energy-efficient target detection applications in wireless sensor networks, in: 8th Ann. IEEE Int. Conf. on Pervasive Computing and Communications (PERCOM 2010), Mannheim, Germany, 2010.
[26] Matlab Website. <http://www.matlab.com>.

**Paolo Medagliani** was born in Cremona (CR), on July 1981. He received the "Laurea" degree in Telecommunications Engineering (3-year program) on September 2003 and the "Laurea Specialistica" degree on April 2006, respectively, from the University of Parma, Italy. From January 2007, he is a Ph.D. student in Information Technologies at the Information Engineering Department of the University of Parma, Italy. He is a member, at the Information Engineering Department of the University of Parma, Italy, of the Wireless Ad-hoc and Sensor Networks (WASN) Laboratory. His research interests are in the field of performance analysis and design of wireless sensor networks and ad-hoc networks.

He is reviewer for some international conferences (SECON 2007, WCNC 2007, GLOBECOM 2007 and 2008, WPMC 2007, EWSN 2008, ICC 2008, and PIMRC 2008) and international journals (IEEE Transactions on Wireless Communications). In addition, he is member of the technical committee of International Conference on Advances in Satellite and Space Communications (SPACOMM 2009), Colmar, France, July 19–24 2009 and of International Workshop on Performance Methodologies and Tools for Wireless Sensor NetworksWSNPerf (WSNPerf 2009), Pisa, Italy, October 23 2009.He has also taken part into research projects in collaboration with a few research organizations and private companies.

**Marco Martalò** was born in Galatina (LE), Italy, on June 1981. He received the "Laurea" degree (3-year program) and the "Laurea Specialistica" (3 + 2 year program) degree (summa cum laude) in Telecommunications Engineering on September 2003 and December 2005, respectively, from the University of Parma, Italy. On March 2009, he received the Ph.D. degree in Information Technologies at the University of Parma, Italy. From January 2009, he is a Post-Doc researcher at the Information Engineering Department of the University of Parma, Italy. From October 2007 to March 2008, he has been a "Visiting Scholar" at the School of Computer and Communication Sciences of the Ecole Polytechnique Federale De Lausanne (EPFL), Lausanne, Switzerland, collaborating with the laboratory of Algorithmic Research in Network Information, directed by Prof. Christina Fragouli. He is a member, at the Information Engineering Department of the University of Parma, Italy, of the Wireless Ad-hoc and Sensor Networks (WASN) Laboratory.

He was a co-recipient of a "best student paper award" (with his tutor Dr. Gianluigi Ferrari) at the 2006 International Workshop on Wireless Ad hoc Networks (IWWAN'06). He has been TPC member of the International Workshop on Performance Methodologies and Tools for Wireless Sensor Networks (WSNPERF 2009) and the International Conference on Advances in Satellite and Space Communications (SPACOMM 2009). He also serves as reviewer for many international journals and conferences.

**Gianluigi Ferrari** (http://www.tlc.unipr.it/ferrari) was born in Parma, Italy, in 1974. He received his "Laurea" and PhD degrees from the University of Parma, Italy, in 1998 and 2002, respectively. Since 2002, he has been with the University Parma, where he currently is an Associate Professor of Telecommunications. He was a visiting researcher at USC (Los Angeles, CA, USA, 2000-2001), CMU (Pittsburgh, PA, USA, 2002-2004), KMITL (Bangkok, Thailand, 2007), and ULB (Bruxelles, Belgium, 2010). Since 2006, he has been the Coordinator of the Wireless Ad-hoc and Sensor Networks (WASN) Lab in the Department of Information Engineering of the University of Parma. As of today he has published more than 140 papers in leading international journals and conferences. He is coauthor of a few books, including "Detection Algorithms for Wireless Communications, with Applications to Wired and Storage Systems" (Wiley: 2004), "Ad Hoc Wireless Networks: A Communication-Theoretic Perspective" (Wiley: 2006), "LDPC Coded Modulations" (Springer: 2009), and "Sensor Networks with IEEE 802.15.4 Systems: Distributed Processing, MAC, and Connectivity (Springer: 2011). He edited the book "Sensor Networks: where Theory Meets Practice" (Springer: 2010). His research interests include digital communication systems analysis and design, wireless ad hoc and sensor networking, adaptive digital signal processing, and information theory.

Dr. Ferrari is a co-recipient of a best student paper award at IWWAN'06 and a best paper award at EMERGING'10. He acts as a frequent reviewer for many international journals and conferences. He acts also as a technical program member for many international conferences. He currently serves on the Editorial Boards of "The Open Electrical and Electronic Engineering (TOEEJ) Journal" (Bentham Publishers), the "International Journal of RF Technologies: Research and Applications" (Taylor & Francis), and the "International Journal of Future Generation Communication and Networking" (SERSC: Science & Engineering Research Support Center). He was a Guest Editor of the 2009 EURASIP JWCN Special Issue on Dynamic Spectrum Access: From the Concept to the Implementation."