Maximizing Area Throughput in Clustered Wireless Sensor Networks

Roberto Verdone, *Member, IEEE,* Flavio Fabbri, *Student Member, IEEE,* and Chiara Buratti, *Member, IEEE*

***Abstract*—In this paper we present a mathematical model to study a multi-sink Wireless Sensor Network (WSN). Both sensors and sinks are assumed to be Poisson distributed in a given finite domain. Sinks send periodic queries, and each sensor transmits its sample to a sink, selected among those that are audible, thus creating a clustered network. Our aim is to describe how the *Area Throughput*, defined as the amount of samples per unit of time successfully transmitted to the sinks from the given area, depends on the density of sensors and the query interval. We jointly account for radio channel, Physical (PHY), Medium Access Control (MAC) and Network (NET) aspects (i.e., different network topologies, packet collisions, power losses and radio channel behaviour), and we compare the performance of two different simple data aggregation strategies. Performance is evaluated by varying the traffic offered to the network (i.e., the density of sensors deployed), the packet size, and, by considering IEEE 802.15.4 as a reference case, the number of Guaranteed Time Slots allocated, and the Superframe Order. The mathematical model shows how the *Area Throughput* can be optimized.**

***Index Terms*—Wireless Sensor Networks, Multiple Sinks, MAC, Connectivity, Data Aggregation, IEEE 802.15.4.**

I. Introduction

MANY applications of Wireless Sensor Networks (WSNs) deal with the estimation of spatial/temporal random processes [1], [2], [3]: sensors are deployed in the target area, which is observed through query/response mechanisms; queries are periodically generated by the application, and sensor nodes react by sampling and sending data to a fusion center. By collecting samples taken from different locations and observing their temporal variations, estimates of the target random process realization can be generated. Good estimates require sufficient data taken from the area. The greater the amount of samples, the better the estimation accuracy [4], [5], [6]. Sometimes, data must be sampled from a specific area, even though sensor nodes are distributed over a larger portion of space. The aim of the query/response mechanism is then to acquire the largest possible number of samples from the area of interest upon each query.

We denote as *Area Throughput* the amount of samples per unit of time originated at the target area and successfully transmitted to a fusion center. According to the characteristics of the observed process, and the area size, the amount of data to be forwarded to the fusion center can be very large. Energy, power, cost and complexity constraints can pose severe limitations to the network design, especially in case of large-scale networks. Hence, simple yet efficient, techniques must be implemented on the network nodes, to maximize the *Area Throughput* given the network cost, related to the number of sensor nodes deployed.

The samples are transmitted through the WSN to sinks, which forward the data to the fusion center by means of a proper network infrastructure. When the number of sensors or the target area are large, the network is usually partitioned for scalability reasons: each sink coordinates transmissions inside its sub-network, forwarding queries to the sensors, and collecting responses. The network is said to be *clustered*, with the sub-networks denoted as *clusters* [7]–[9].

Sinks are sometimes specifically deployed in optimized and planned locations with respect to sensors. However, opportunistic exploitation of the presence of sinks, connected to the infrastructure through any mobile radio interfaces, is an interesting option in some cases [10]. Under these circumstances, many sinks can be present in the monitored space, but their positions are unknown and unplanned; therefore, achievement of a sufficient level of samples is not guaranteed, because the sensor nodes might not reach any sinks (and thus be isolated) owing to the limited transmission range. In such an uncoordinated environment, network connectivity (i.e., the property of making every node able to reach at least one sink) is a relevant issue, and it is basically dominated by the transmission techniques implemented at physical layer (PHY), the wireless medium behavior and the density of sinks: in any case, one would expect that the *Area Throughput* is larger if the density of sensor nodes is larger.

On the other hand, if simple contention-based MAC (Medium Access Control) techniques are implemented at the Data Link (DL) layer of the protocol stack, as is usual in WSNs [2], [3], the density of nodes significantly affects the packet collision probability (i.e., the event of simultaneous transmissions by separate nodes): if the number of sensor nodes per cluster is very large, collisions and backoff procedures can make data transmission impossible under time-constrained conditions, and the samples taken from sensors do not reach the sinks.

The maximization of the *Area Throughput* then, requires proper dimensioning of sensor density, in a framework model where both MAC and connectivity issues are considered.

Furthermore, the network topology implemented at network layer (NET) inside each cluster can affect the amount of data received by the sink, since different topologies (e.g., a tree instead of a star) can provide different levels of reliability [11].

One option to reduce the need for transmission of large amounts of packets is data aggregation, which consists in

Manuscript received 1 May 2009; revised 16 February 2010.
The authors are with WiLab at University of Bologna, Italy.
Digital Object Identifier 10.1109/JSAC.2010.100924.

accumulating samples at a given node, and transmitting them to the sink through one single packet, possibly of larger size. In this context, we might consider two simple options:

i) concatenation (i.e., aggregation at sensor level): sensor nodes take and record one sample after each query for a given interval of time, and send all of them in one single batch at the end of the interval;

ii) aggregation at router level: if multi-hop links are followed by packets to reach the sink, intermediate nodes can process the samples received by separate sensors and aggregate them, to reduce packet payload and duration.

With the former technique, delays are introduced, as some samples are queued for several query intervals before being forwarded to the sink; although simple, this strategy can be applied only as long as the temporal variation of the observed process is not too considerable. On the other hand, the latter requires exploitation of potential spatial correlations between samples, which requires proper signal processing capabilities.

This paper jointly accounts for PHY, MAC, NET and data aggregation issues of clustered WSNs, with the aim of mathematically deriving the conditions for maximization the *Area Throughput*. A general analytical framework is introduced, covering two separate cases: small networks, where the transmission range of sensors is in the same order of the area side, and large networks where border effects can be neglected thus reducing mathematical complexity. The latter case brings to some interesting discussions.

The rest of the paper is organized as follows. Next Section describes the scenario, the objectives of the paper, and discusses related publications. Section III introduces the model assumptions at Application, PHY, DL and NET layers, for the radio propagation channel, and on the two simple data aggregation policies considered. In Section IV the *Area Throughput* is evaluated, by computing the success probability for the transmission of a packet accounting for connectivity, NET and MAC issues. Finally, in Sections V and VI numerical results and conclusions are presented, respectively. Some details about the IEEE 802.15.4 case are reported in the Appendix.

II. Scenario, Objectives and Related Works

We consider an infinite area where sensors and sinks are both uniformly distributed at random. Sinks forward all collected samples to the fusion center without losses. Then, we define a specific portion of space, of finite size and given shape (without loss of generality, we consider a square), as the target area; both the sensors and sinks are then distributed in such area according to a homogeneous Poisson Point Process (PPP) (see Figure 1).

We define as *Available Area Throughput* the amount of samples which are available in the area, per unit of time. The basic objective of this paper is thus to determine, through some analytical work and a novel mathematical flexible approach, how the *Area Throughput* depends on the *Available Area Throughput*, as a function of the two different data aggregation policies, the algorithms used at DL (specifically, MAC) and NET layers, and PHY and wireless channel characteristics.

The definitions of *Available Area Throughput* and *Area Throughput* are specifically introduced here for the purpose of WSN design when process estimation applications are considered. However the objective of this work is, more generally, to extend the throughput-offered traffic relations which are well known from seminal works on MAC initiated during the '80s. In particular, we refer the reader to N. Abramson for single sink ALOHA and Slotted ALOHA [12]. Later extension to CSMA and CSMA/CA (Carrier Sensing Multiple Access with Collision Avoidance) can be found in [13]. We also aim at taking account of the multiple sink case, connectivity and simple data aggregation techniques.

It is shown that the maximization of the *Area Throughput* significantly depends on the performance of the MAC protocol. The role of different network topologies within the clusters is also emphasized.

As a reference standard for PHY and MAC, we consider 802.15.4, an IEEE standard for Low Rate Personal Area Networks (PANs) [16]. In our scenario, sinks will act as PAN coordinators. The 802.15.4 standard allows for two types of channel access mechanisms: Beacon- and Non Beacon-Enabled. The Non Beacon-Enabled mode uses a CSMA/CA protocol to access the channel, whereas in the Beacon-Enabled a superframe, composed of an inactive and an active part, where some slots denoted as Guaranteed Time Slots (GTSs) may also be allocated to specific nodes, is defined (see Appendix and Fig. 2).

Very few papers jointly consider MAC and connectivity issues under a mathematical approach. Some analysis is performed through simulations: as examples, [17] related to ad hoc networks, and [18], to WSNs.

Many papers devoted their attention to connectivity issues of wireless ad-hoc and sensor networks in the past (e.g., [19]). Single-sink scenarios have attracted more attention so far. An example of multi-sink scenario can be found in [20]. All the previously cited works do not account for MAC issues.

Concerning the analytical study of CSMA-based MAC protocols, in [21] the throughput for a finite population when a persistent CSMA protocol is used, is evaluated. An analytical model of the 802.11 CSMA-based MAC protocol, is presented by Bianchi in [22]. In these works no physical layer or channel model characteristics are accounted for. Capture effects with CSMA in Rayleigh channels, are considered in [23] whereas [24] addresses CSMA/CA. However, no connectivity issues are considered in these papers. In [25] the per-node saturated throughput of an 802.11b multi-hop ad hoc network with a uniform transmission range, is evaluated.

The model proposed in this paper is based on the following previous works. In [26] the authors presented a mathematical model for the evaluation of the degree of connectivity of a multi-sink WSN in unbounded and bounded domains. [28], [29] provide a mathematical model to derive the success probability for the transmission of a packet in an 802.15.4 single-sink scenario when the Non Beacon-Enabled mode is used; in [11], [30] the mathematical model for the 802.15.4 Beacon-Enabled mode, is provided.

III. System Model

We denote sensors and sinks densities as $\rho_s[m^{-2}]$ and $\rho_0[m^{-2}]$, respectively, and with $\mathcal{A}$ the target domain of area

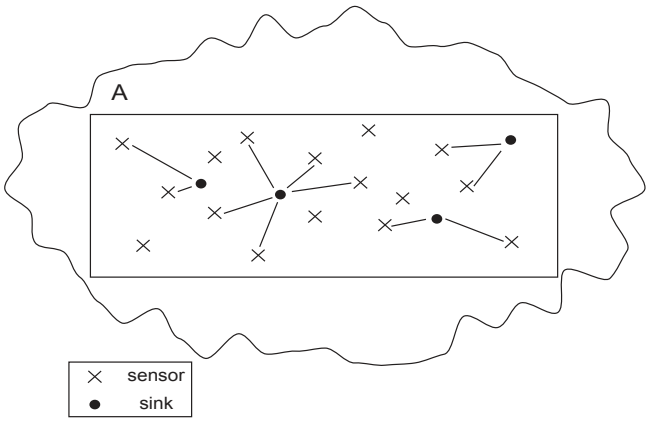

Fig. 1. The reference scenario considered.

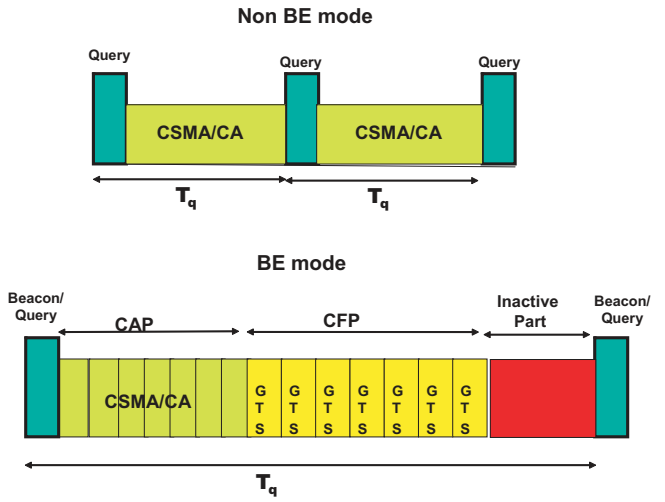

Fig. 2. Above part: The IEEE 802.15.4 Non Beacon-Enabled mode. Below part: The IEEE 802.15.4 Beacon-Enabled mode.

A. Denoting by K the number of sensor nodes in $\mathcal{A}$, K is Poisson distributed with mean $\bar{K} = \rho_s \cdot A$ and probability mass function (p.m.f.).

$$\Pr(K = k) = \frac{\bar{K}^k e^{-\bar{K}}}{k!}. \quad (1)$$

We also denote as $I = \rho_0 \cdot A$ the average number of sinks in $\mathcal{A}$.

A. The Application Layer

Sinks send queries to sensors periodically, every $T_q = 1/f_q$ seconds. In case a sensor node receives queries from multiple sinks, it selects the one providing the largest received power. Upon reception of the query, each node will take one sample from the environment and will attempt a channel access according to the MAC protocol, to transmit the data to the sink. Depending on the network topology, transmission will occur either directly or through multiple hops. We assume that each sensor node generates a packet having size $D{\cdot}10$ Bytes, with $D = H + xP$. H represents the header, P is the payload and x depends on the data aggregation technique implemented. When an 802.15.4 air interface is considered, the time needed to transmit a packet is equal to $D \cdot T$, where $T = 320\ \mu sec$ (a bit rate of 250 kbit/sec is used). We set the size of the query equal to 60 Bytes.

B. The Wireless Channel

The wireless channel model that we consider accounts for both a distance-dependent path loss and the random channel fluctuations caused by possible obstructions.

The power loss in decibel scale at distance d is expressed in the form

$$L = k_0 + k_1 \ln d + s, \quad (2)$$

where $k_0 = 20\ log_{10}\frac{4\pi}{\lambda}$, being λ the wavelength and $k_1 = \beta\frac{10}{ln(10)}$, being β the propagation coefficient, s is a Gaussian random variable (r.v.) with zero mean, variance σ^2, which represents channel fluctuations. This channel model was also adopted in [32], and is shown to fit some WSN scenarios [3]. The random component in (2) statistically accounts for the possible obstructions encountered by the radio wave when traveling across the wireless medium.

C. The PHY Layer

Two nodes (and the correspondent link) are usually said to be connected if the Packet Error Probability (PEP), denoted as P_p, at the receiver is below a given threshold (this is known as threshold-based model). The PEP depends on the signal-to-noise ratio SNR (in the absence of interference), the receiver characteristics, the packet structure and the transmission techniques implemented at the PHY layer. As a reference, one might think of an uncoded BPSK system with matched filter: the Bit Error Probability (BEP), denoted as P_b, in AWGN (Additive White Gaussian Noise), is given by $P_b = 0.5 \cdot \mathrm{erfc}(\sqrt{\gamma})$, where γ represents the SNR. If the packet is not protected by any error correction coding techniques, then $P_p = 1 - (1 - P_b)^{80 \cdot D}$, being $80 \cdot D$ the packet size in bits. When error correction coding is used, the PEP is reduced for large values of γ, making the PEP versus γ curve steeper. Generally speaking, the more complex are the transmission techniques used at PHY, the more the $P_p(\gamma)$ curve takes the form of a step, with PEP $\approx$ 1 or 0, if γ is smaller or larger than a threshold γ_0, respectively; when the latter conditions hold, two nodes are connected if γ is larger than γ_0 (threshold assumption). The SNR γ can be computed in dB as $\gamma = P_T - L - P_N$ where P_T and P_N are the transmit and noise powers (in dBm), respectively. Therefore, under the threshold assumption, two nodes can be said to be connected if $L < L_{\text{th}}$, where $L_{\text{th}} = P_T - P_N - \gamma_0$ represents the maximum loss tolerable by the link.

The event $L < L_{\text{th}}$ happens with probability, P_{CON}, that depends on the statistics of s and d introduced in (2). By considering an average transmission range as in [32], an average (over such statistics) connectivity area of the sensor node can be defined as

$$A_\sigma = \pi e^{\frac{2(L_{\text{th}} - k_0)}{k1}} e^{\frac{2\sigma^2}{k_1^2}}. \quad (3)$$

This threshold-based model is used in many papers, also dealing with WSNs. However, when simple techniques are used at the PHY layer (e.g., no error correction coding), the threshold assumption is not accurate, as the $P_p(\gamma)$ curve is

not stepwise, and a link might suffer from packet losses (with small probability) even with large values of γ.

A more accurate model to account for the PHY layer is then the following. Two nodes are said to be (short-term) connected if a packet is correctly received. This happens with a probability, P_{CON-ST}, which can be defined as

$$P_{CON-ST} = \mathbb{E}\{1 - P_p(\gamma)\}; \quad (4)$$

the expectation is taken over the statistics of s and d in (2). Therefore, the entire $P_p(\gamma)$ curve impacts on link connectivity. Clearly, in the case of a stepwise curve with threshold γ_0, then P_{CON-ST} formally becomes equal to the probability that γ is larger than γ_0, that is, P_{CON}.

D. The DL Layer

At DL layer, MAC issues play a fundamental role in WSNs, as contention-based protocols are often used. In this paper we focus on CSMA/CA, taking 802.15.4 as a reference case. With CSMA/CA, packet collisions due to simultaneous transmissions are possible even with the implementation of backoff and retransmission policies; moreover, under time-constrained conditions, interference may cause some packet losses even if the SNR is very large. As a result, a packet might be lost because of MAC failures such as maximum number of retransmissions reached, time allowed for transmission expired, etc.

Then, two nodes are connected if both the following two events hold: neither is the packet affected by collisions or any MAC failure, nor is short-term connectivity lost. Assuming the two events are disjoint, we can define the probability of a link to be active as $P_{LINK} = P_{CON-ST} \cdot P_{MAC}$, where P_{MAC} is the probability that MAC procedures succeed in transmitting the packet.

In general, $P_{MAC} = P_{MAC}(n, D)$ is a function of the number of nodes, n, competing for the receiver, and depends on the packet size, that is function of D. By increasing n and D, in fact, the probability that two or more packets collide, that is they are partially or completely overlapped, gets larger. To demonstrate this, we show in Fig. 3 the behavior of P_{MAC} by varying n and D for an IEEE 802.15.4 network working in Beacon-Enabled mode, as an example. Lines in the figure were obtained through the mathematical model used in this paper and described in [11], [30]; points are obtained through simulation analysis by using a simulator written in C language. A perfect agreement can be observed. Details about IEEE 802.15.4 are given in the Appendix.

E. The NET Layer

According to our assumptions, multiple sinks are present in the target area. Each of them transmits queries to all sensor nodes in the area, which then select the PAN to associate to, based on the strongest received signal. Therefore, assuming the links are symmetric, the nodes forming a PAN could establish direct links with the PAN coordinator (i.e., the sink), and create a star topology. However, if the number of nodes, n, associated to a sink is very large, P_{MAC} might be very small. The sink might then randomly elect as routers n_1 out of the n nodes, with the remaining $n_2 = n - n_1$ nodes acting as leaves of a two-level tree rooted at the sink. Although the optimization of the number of nodes per level is beyond the scope of this work, a possible optimization strategy is presented below.

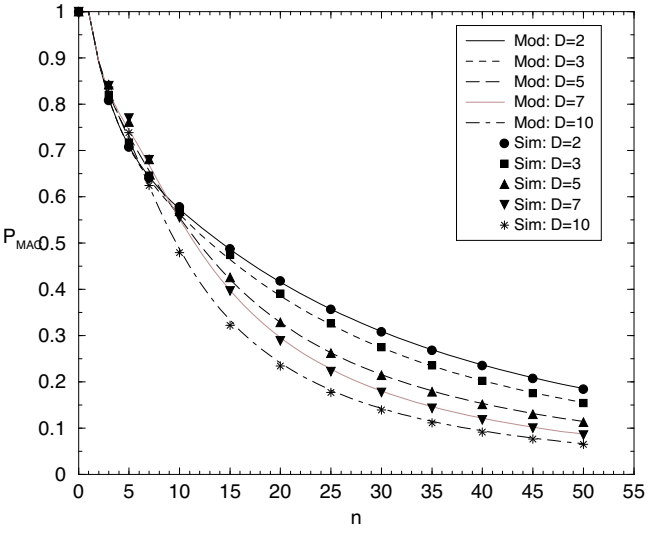

Fig. 3. P_{MAC} as a function of n for different values of D, in 802.15.4 Beacon-Enabled networks.

While two hops are needed for leaves to reach the sink, the advantage of this solution stands in the lower number of nodes competing for the same receiver, either at the router or sink level, provided that the transmission from leaves to routers, and from routers to sink, are not conflicting. This can be obtained, through allocation of different resources (i.e., time or frequency) to the different sub-clusters of nodes in the tree, formed by a router (or the sink) and its children. As an example, in IEEE 802.15.4/Zigbee tree-based topologies nodes bound to different parents in the tree will compete for the channel in different portions of the superframe. This means that a sort of time division between sub-clusters is implemented in order to limit competition to nodes bound to the same parent (see Appendix). Therefore, the tree leads to a larger P_{MAC} at each hop, but some samples need two hops to reach the sink. Note that, being competition only between children of the same parent and also that in general the time allowed to the different sub-clusters is the same, the best strategy for setting n_1 and n_2 is the one that leads to the same number of children per parent. This could be achieved by setting $n_1 \approx \lfloor \sqrt{n} \rfloor$ [11].

Let us denote as P_{NET} the probability that the network topology can forward the sample taken at a given sensor node to the sink it is associated to, assuming no connectivity problems exist (accounted for by P_{CON-ST}).

By assuming no conflicts between leaves attached to different routers (see Appendix) and that routers aggregate the received data generating a packet of size $D \cdot 10$ bytes (see the following subsection), the following holds. For a leaf in the tree the network will forward the packet to the sink with probability $P_{NET} = P_{MAC}(n_{2,1}, D) \cdot P_{MAC}(n_1, D)$ where $n_{2,1}$ is the number of leaf nodes competing for a router; if the sink creates a balanced tree, then $n_{2,1}$ is equal to n_2/n_1, assuming it is an integer value. For a sensor node elected as router by its PAN coordinator, then $P_{NET} = P_{MAC}(n_1, D)$.

Therefore, on average, if a tree topology is used, the probability that a sensor node can reach the sink it is associated to, is

$$P_{NET} = \frac{n-n_1}{n} P_{MAC}(n_{2,1}, D) P_{MAC}(n_1, D) + \frac{n_1}{n} P_{MAC}(n_1, D). \quad (5)$$

On the other hand, if a star topology is implemented, then

$$P_{NET} = P_{MAC}(n, D). \quad (6)$$

Under such formalism, the probability of packet success, P_s, taking network topology and MAC into account, is given by

$$P_s = P_{CON-ST} \cdot P_{NET}, \quad (7)$$

with P_{NET} given by (5) or (6). The probability P_s is a function of the number, n, of sensor nodes associated to the sink in the cluster.

F. Data Aggregation

In the previous subsection, it is implicitly assumed that all packets have equal duration, $D \cdot T$. However, this depends on the network topology and the data aggregation policy.

If concatenation is used, then sensor nodes do not reply to all queries. We assume in our model the simplest form of concatenation: nodes take one sample per query, then reply only every x queries, sending the x samples recorded. The packet has a longer duration in this case, since we have $D = H + xP$. The advantage stands in the smaller average number of nodes competing for access, which brings to a larger value for P_{MAC}. On the other hand, as D increases, $P_{MAC}(n, D)$ gets smaller because of the larger probability of packet collisions (see Fig. 3).

When data aggregation is performed at router level (this policy can be implemented only if a tree is used at NET layer), the samples received by the router can be processed in order to reduce the amount of data to be forwarded to the sink. In this paper, we assume the processing is such that, upon reception of n samples, the router transmits data equivalent to one single sample. If the two aggregation policies are integrated in the tree case, then the routers, upon reception of $x \cdot n$ samples, transmit x samples. It is out of the scope of this paper to determine how the processing can be performed.

Therefore, in (5) and (6), P_{MAC} should be computed as a function of the proper values of D, depending on the data aggregation policy used.

IV. Evaluation of The Area Throughput

A. Probability of Successful Transmission

Let us consider an arbitrary sensor node located in $\mathcal{A}$, and denote its position as (x, y) with respect to a reference system with origin centered in $\mathcal{A}$. We aim at computing the probability that the node can connect to one of the sinks in the area and successfully transmit its data sample. To this aim, we define $P_{s|K}(x, y)$ as the probability of successful transmission conditioned on the overall number, K, of sensors in the area. This probability can be computed by averaging $P_s(x, y)$ of (7) over the number of nodes, n, associated to a sink. The dependence on (x, y), previously not emphasized to avoid complex notations, is due to the well-known border effects in connectivity [19].

Therefore,

$$\begin{aligned} P_{s|K}(x,y) &= \mathbb{E}_n\{P_s(x,y)\} \\ &= \mathbb{E}_n\{P_{NET}(n) \cdot P_{CON-ST}(x,y)\} \\ &= \mathbb{E}_n\{P_{NET}(n)\} \cdot P_{CON-ST}(x,y). \end{aligned} \quad (8)$$

In [33], Orriss *et al.* showed that the number of sensors uniformly distributed on an infinite plane that hear one particular sink as the one with the strongest signal power (i.e., the number of sensors competing for access to such sink) is Poisson distributed with mean

$$\bar{n} = \mu_s \frac{1 - e^{-\mu_{sink}}}{\mu_{sink}}, \quad (9)$$

where $\mu_{sink} = \rho_0 A_\sigma = I A_\sigma / A$ is the mean number of audible sinks on an infinite plane from any position [32]; $\mu_s = \rho_s A_\sigma$ is the mean number of sensors that are audible by a given sink. Such a result is relevant toward our goal even though it was derived on the infinite plane. In fact, n can still be considered Poisson distributed even inside a finite area. The only two things that change are:

- n is upper bounded by K (i.e., the pdf is truncated);
- the density ρ_s is to be computed as the ratio K/A $[m^{-2}]$, thus yielding $\mu_s = K \frac{A_\sigma}{A}$.

Therefore, we assume $n \sim \text{Poisson}(\bar{n})$, with

$$\bar{n} = \bar{n}(K) = K \frac{A_\sigma}{A} \frac{1 - e^{-\mu_{sink}}}{\mu_{sink}} = K \frac{1 - e^{-I A_\sigma / A}}{I}. \quad (10)$$

By making the average in (8) explicit, we get

$$P_{s|K}(x,y) = P_{CON-ST}(x,y) \cdot \frac{1}{M} \sum_{j=1}^{K} P_{NET}(j) \frac{\bar{n}^j e^{-\bar{n}}}{j!}, \quad (11)$$

where

$$M = \sum_{j=1}^{K} \frac{\bar{n}^j e^{-\bar{n}}}{j!} \quad (12)$$

is a normalizing factor.

1) Small Networks: $P_{CON-ST}(x, y)$ represents the probability that the sensor is not isolated (i.e., it receives a sufficiently strong signal from at least one sink), which is computed in [26] for a scenario similar to the one considered here (e.g., squared and rectangular areas), under the threshold-based assumption. We let the reader refer to [26], [27] for details and validation of the model, for the sake of conciseness. Since the position of the sensor is in general unknown, $P_{s|K}(x, y)$ of (8) can be deconditioned as

$$\begin{aligned} P_{s|K} &= E_{x,y}[P_{s|K}(x,y)] \\ &= E_{x,y}[P_{CON-ST}(x,y)] \cdot \frac{1}{M} \sum_{j=1}^{K} P_{NET}(j) \frac{\bar{n}^j e^{-\bar{n}}}{j!}. \end{aligned} \quad (13)$$

The average $E_{x,y}[P_{CON-ST}(x, y)]$ does not result in closed form expressions (see [26] for details).

2) Large Networks: It is shown in [26] that under the threshold-based assumption, border effects are negligible when $A_\sigma < 0.1A$. In this case the following holds [33]:

$$P_{CON-ST}(x,y) = P_{CON} = 1 - e^{-\mu_{sink}}. \quad (14)$$

Under such assumption, we get

$$P_{s|K} = (1 - e^{-IA_\sigma/A}) \cdot \frac{1}{M} \sum_{j=1}^{K} P_{NET}(j) \frac{\bar{n}^j e^{-\bar{n}}}{j!}. \quad (15)$$

B. Area Throughput

The average number of samples per query that can be generated by the network is given by the mean number of sensors in $\mathcal{A}$, $\bar{K}$.

Now denote by G the average number of samples that can be generated per unit of time, given by

$$G = \bar{K} \cdot f_q = \rho_s \cdot A \cdot \frac{1}{T_q} \text{ [samples/sec]}. \quad (16)$$

From (16) we have $\bar{K} = G \cdot T_q$. We denote G as *Available Area Throughput*, since it represents the maximum amount of samples achievable, if there are no connectivity losses and MAC/NET failures.

We can now compute the *Area Throughput*, that is, the average amount of data received by the infrastructure per unit of time, denoted as S, as

$$S = \sum_{k=1}^{+\infty} S(k) \cdot \Pr(K = k) \text{ [samples/sec]}, \quad (17)$$

where $S(K) = \frac{K}{T_q} P_{s|K}$, $\Pr(K = k)$ as in (1) and $P_{s|K}$ as in (13) or (15).

The most interesting case is when border effects can be negligible (large scale networks). By means of (15), (12) and (16), equation (17) may be rewritten as

$$S = \frac{1 - e^{-IA_\sigma/A}}{T_q} \cdot \sum_{j=1}^{+\infty} \frac{\sum_{j=1}^{k} P_{NET}(j) \frac{\bar{n}^j e^{-\bar{n}}}{j!}}{\sum_{j=1}^{k} \frac{\bar{n}^j e^{-\bar{n}}}{j!}} \cdot \frac{(GT_q)^k e^{-GT_q}}{(k-1)!}. \quad (18)$$

This expression will be used in the numerical results to determine the values of G providing maximum *Area Throughput*.

However, such expression can be simplified by means of proper approximations, under the assumption that I and $G \cdot T_q$ are both much larger than one. In this case, we get (19) which becomes (assuming k large) approximately equal to

$$\frac{1 - e^{-IA_\sigma/A}}{T_q} \cdot \sum_{k=1}^{+\infty} \frac{\sum_{j=1}^{+\infty} P_{NET}(j) \frac{\bar{n}^j e^{-\bar{n}}}{j!}}{\sum_{j=1}^{+\infty} \frac{\bar{n}^j e^{-\bar{n}}}{j!}} \cdot k \cdot \frac{(GT_q)^k e^{-GT_q}}{k!}.$$

Moreover, by means of an approximation, we have $\mathbb{E}_K\{\bar{n}(K)\} \approx \bar{n}(\mathbb{E}\{K\}) = \bar{n}(\bar{K}) = \bar{N}$, with $\bar{N} = \bar{K} \cdot \frac{1-\exp(-IA_\sigma/A)}{I}$. We then obtain (20). Now the order of the two sums (over k and j) can be reversed, obtaining

$$S \approx \frac{1 - e^{-IA_\sigma/A}}{T_q} \cdot \sum_{j=1}^{+\infty} \frac{P_{NET}(j) \frac{\bar{N}^j e^{-\bar{N}}}{j!}}{1 - e^{-\bar{N}}} \cdot \sum_{k=1}^{+\infty} k \cdot \frac{(GT_q)^k e^{-GT_q}}{k!}.$$

$$= \frac{1 - e^{-IA_\sigma/A}}{T_q} \cdot \sum_{j=1}^{+\infty} \frac{P_{NET}(j) \frac{\bar{N}^j e^{-\bar{N}}}{j!}}{1 - e^{-\bar{N}}} \cdot GT_q.$$

This result can be approximated to

$$S \approx G(1 - e^{-IA_\sigma/A}) \cdot \sum_{j=1}^{+\infty} P_{NET}(j) \frac{\bar{N}^j e^{-\bar{N}}}{j!}. \quad (21)$$

The numerical results will show that expression (21) approximates very well (18), even for small values of G, while being much simpler to compute. Moreover, (21) allows further elaborations.

C. Optimum Available Area Throughput

By using (21) the value of G which yields the maximum value of S, denoted as G_m hereafter, can be evaluated by computing the derivative of S with respect to G. The value of G_m is relevant because it determines the number of sensors to deploy in $\mathcal{A}$ to maximize the *Area Throughput*.

By computing such derivative, one gets easily to the following equation:

$$\sum_{j=1}^{+\infty} P_{NET}(j) \cdot (1+j) \cdot \frac{\bar{N}^j e^{-\bar{N}}}{j!} = \sum_{j=1}^{+\infty} P_{NET}(j) \cdot (\bar{N}) \cdot \frac{\bar{N}^j e^{-\bar{N}}}{j!}, \quad (22)$$

where in this case $\bar{N} = G_m T_q (1 - e^{-IA_\sigma/A})/I$.

Even though the numerical computation of the value G_m requires the specific knowledge of $P_{NET}(n)$, some interesting considerations can be done.

Expression (22) depends on G_m only through $\bar{N}$. Therefore, the equality holds for a given value of $\bar{N}$; let us denote such (unknown) value as Y. Consequently, $G_m T_q (1 - e^{-IA_\sigma/A})/I = Y$, and therefore $G_m = Y \cdot I \; T_q (1 - e^{-IA_\sigma/A})$. In words, the optimum value of G is proportional to the product of the average number of sinks in the area and P_{CON}.

Furthermore, by substituting $G_m = Y \cdot I \; T_q (1 - e^{-IA_\sigma/A})$ in (21), the maximum value of the *Area Throughput*, denoted as S_m, is shown not to depend on P_{CON}.

Finally, the most interesting result that can be achieved through this approximated analysis stands in the behavior of the S versus G curves for G tending to infinity. In fact, by assuming $P_{NET}(n) = 1/n$ and letting G tend to infinity, expression (21) brings to $S = I/T_q$. It can be shown that if the tail of the $P_{NET}(n)$ function is heavier than $1/n$ (e.g., it goes like $1/n^2$), for G tending to infinity, expression (21) brings to $S = 0$. Therefore, if the MAC and network topology are chosen such that the tail of the $P_{NET}(n)$ function is lighter than $1/n$ or, for large n, it follows a $1/n$ law, then the *Area Throughput* has an horizontal asymptote when G is large; increasing the density of sensors, S does not reach zero and the network behavior is more stable. On the other hand, if the tail of the $P_{NET}(n)$ function is heavier than $1/n$, the network becomes instable for large G.

All these considerations, though obtained through the approximate expression (21), are shown in the numerical results to be correct.

$$S = \frac{1 - e^{-IA_\sigma/A}}{T_q} \cdot \sum_{k=1}^{+\infty} \frac{\sum_{j=1}^{k} P_{NET}(j) \frac{\bar{n}^j e^{-\bar{n}}}{j!}}{\sum_{j=1}^{k} \frac{\bar{n}^j e^{-\bar{n}}}{j!}} \cdot k \cdot \frac{(GT_q)^k e^{-GT_q}}{k!} \tag{19}$$

$$S \approx \frac{1 - e^{-IA_\sigma/A}}{T_q} \cdot \sum_{k=1}^{+\infty} \frac{\sum_{j=1}^{+\infty} P_{NET}(j) \frac{\bar{N}^j e^{-\bar{N}}}{j!}}{1 - e^{-\bar{N}}} \cdot k \cdot \frac{(GT_q)^k e^{-GT_q}}{k!} \tag{20}$$

V. Numerical Results

In this Section we show the behavior of the *Area Throughput* as a function of G (see (18)) for different network and MAC settings. In particular, we consider both small and large network cases and we emphasize the role played by different choices at DL layer and the effects of data aggregation performed at both sensor and router level. We also test the simplified expression derived in (21) against the exact model of Eq. (17) (directly employing Eq. (13)). In the following plots we consider a square domain and $k_0 = 40$ [dB] (being $\lambda = 0.125$, since the 2.4 GHz frequency is used), $\beta = 3$ (i.e., $k_1 = 13.03$), $\sigma = 4$ [dB], $L_{\text{th}} = 106$ [dB] (assuming that nodes transmit power is $P_T = 4$ [dBm] and the receiver sensitivity is -102 [dBm]).

In Figure 4 we take a domain of area $A = 500$ [m^2] and S as a function of G is shown when the 802.15.4 MAC in Non Beacon-Enabled mode is employed. In this small network case we can evaluate the impact of border effects on connectivity: not properly accounting for them does result in an overestimation of the *Area Throughput*. Other considerations regarding the behavior of S when varying G are left to the following plots.

For the following figures we consider the case where border effects are negligible by setting $A = 1000$ [m^2]. In Figure 5, S as a function of G is shown when the 802.15.4 MAC in Beacon- and Non Beacon-Enabled mode is employed. We vary SO, the number of GTSs allocated, denoted as N_{GTS} (see the Appendix for details about the GTSs), and T_q, having fixed $D = 10$ and $P_{CON} = 1$. It is visible that, once we fix SO (Beacon-Enabled case), by increasing N_{GTS}, S increases, since P_{MAC} increases. Moreover, once we fix N_{GTS} there exists a value of SO maximizing S. When GTSs are not allocated, an increase of SO results in a decrease of S. In fact, even though P_{MAC} gets larger (nodes have more time to access the channel), the query interval increases too and the number of samples per second received by the sink decreases. On the other hand, when GTSs are allocated, the time allowed to nodes to access the channel through CSMA/CA, in the case $SO = 0$, is too small (see Appendix) and many packets are lost: in this case $SO = 1$ provides the best performance. The same holds for the Non Beacon-Enabled mode: the value of SO maximizing S is the one that gives an optimum trade-off between time allowed to nodes to access the channel and number of packets generated per unit of time. Even if it is not shown in Fig. 5, note also that the optimum value of SO depends on D: if SO is small, nodes receive queries at a higher frequency but large packets will undergo a greater number of losses. On the contrary, a larger SO ensures correct reception of packets but penalizes the query frequency.

The comparison in Figure 6 highlights the accuracy of the asymptotic model derived in (21) against the exact one. Once again we consider the IEEE 802.15.4 MAC protocol in Beacon-Enabled mode, with $D = 2, 10$ and a different number of GTSs allocated, namely 0 and 6. As expected, when $I = 20$ the throughput is higher because a greater number of clusters are formed and hence a smaller number of nodes are competing for channel access in each of them. Overall, the asymptotic approximation appears very tight, especially in the proximity of optimal operating points: this allows the use of the simplified model for optimizing the values of G.

In Figure 7, S as a function of G is shown when different laws of success probability are employed. In particular the two bottom curves are obtained with $P_{MAC}(n) = 1/n^\alpha$ while the two upper curves with $P_{MAC}(n) = 1, \quad n \leq n^*$, $P_{MAC}(n) = n^*/n, \quad n > n^*$. It is worth noting how throughput is affected by the tail of the P_{MAC} law used. In particular when it is of the kind of $1/n^\alpha$ (see the two curves below), with $\alpha = 1$ it is asymptotically finite. On the contrary, it vanishes for $\alpha = 2$. We can then state that $1 < \alpha < 2$ for realistic cases (e.g., 802.15.4). The two upper curves highlight the importance of $P_{MAC}(n)$ for small values of n: while in both cases the asymptotic behavior is like $1/n^\alpha$, with $\alpha = 1$, by doubling n^*, throughput is almost doubled as well.

In the next two figures we consider 802.15.4 MAC protocol in Beacon-Enabled mode and use the simplified model. In Figure 8 we set $D = 2$ and plot the value of G, G_m, which maximizes S as a function of the number of guaranteed time slots assigned, for different values of SO. As expected, the value of G needed to saturate the network is larger when a larger number of guaranteed time slots are given. Moreover, SO is related to T_q (see Appendix) and G is also a function of T_q (see (16)) in such a way that smaller values of SO result in a larger G. As a consequence, when $SO = 0$, the maximum *Area Throughput* is reached for larger values of G.

Finally, the two different aggregation strategies are analyzed in Figure 9. S is reported here as function of G when employing a star topology with aggregation performed at sensor level (i.e., concatenation) and a 2-level tree topology with aggregation performed at router level. Packet header is set $H = 2$, while payload is $x = 1, 2, 3$. In the case of 2-level tree we fixed $D = 3$ for both the transmission from sensor to router and router to sink. The quantities n_1 and n_2 are chosen such that $n_2 = n_1^2$, according to the optimization strategy described in Section III-E. While the star topology reaches saturation at some point due to the many access attempts, it can be observed that the 2-hop solution does so for larger G, since multi-hopping splits access attempts among a larger number of nodes (although at the expense of greater cost in terms of

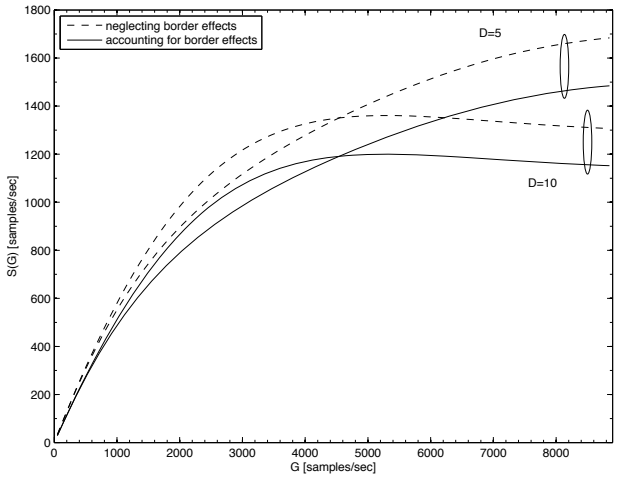

Fig. 4. S as a function of G for the 802.15.4 Non Beacon-enabled case, with $D = 5, 10$, $I = 10$, $T_q = 320 \cdot (120 + D)$ μsec.

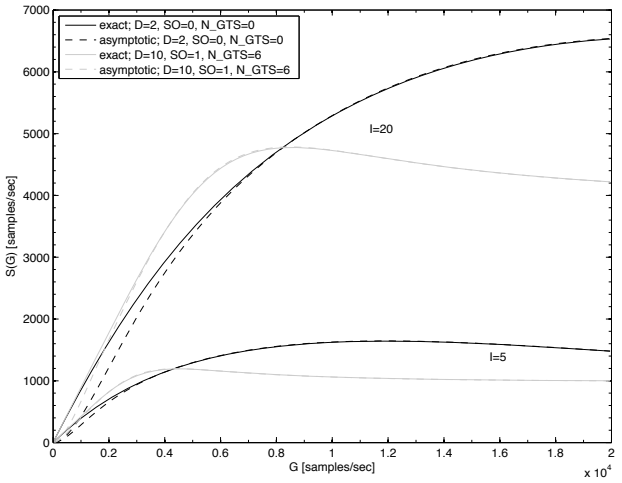

Fig. 6. S as a function of G for the 802.15.4 Beacon-Enabled case, with $D = 2, 10$, $I = 5, 20$, different values of SO and N_{GTS}, computed by means of both the exact and approximated models.

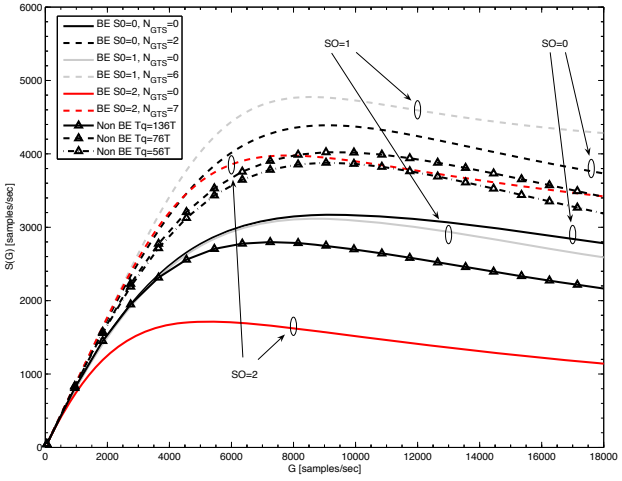

Fig. 5. S as a function of G for the 802.15.4 Beacon- and Non Beacon-Enabled cases, by varying SO, N_{GTS} and T_q, having fixed $D = 10$.

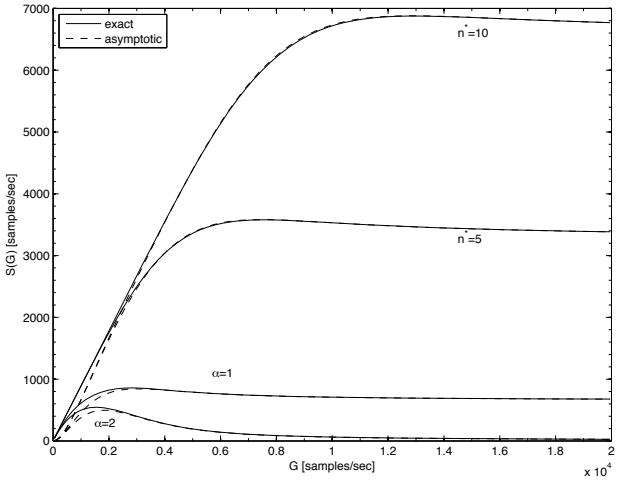

Fig. 7. S as a function of G when different $P_{MAC}(n)$ functions are employed: $P_{MAC}(n) = 1/n^{\alpha}$ for the two bottom curves, $P_{MAC}(n) = 1, \ n \leq n^*$, $P_{MAC}(n) = n^*/n, \ n > n^*$ for the two upper curves.

delay and energy). Hence, on the one hand, the number of nodes competing for channel access is dramatically reduced thanks to sub-clustering. On the other hand, processing at router level significantly reduces packet size in the router to sink transmission. This reveals highly beneficial because packets traveling from router to sink contain data aggregates from all sensor in the sub-tree.

VI. Conclusions

In this paper a mathematical framework to determine the maximum *Area Throughput* S of a WSN designed for temporal/spatial random process estimation, has been determined, accounting for radio channel, PHY, MAC and NET protocol layers, and simple data aggregation techniques. How the *Area Throughput* S depends on the nodes density and the query interval, has been shown by introducing the concept of *Available Area Throughput* G, and determining the relationship between S and G.

The algorithms and techniques considered are simple: this is a plus in WSN environments where network deployment cost should be minimised. On the other hand, the mathematical framework is sufficiently flexible to allow consideration of other schemes.

The use of mathematical approaches permits the derivation of some conclusions which might be very useful to drive the selection of algorithms, such as the impact of MAC behavior on the overall performance, or the role played by separate types of data aggregation techniques (at the sensor or router level).

The determination of the maximum *Area Throughput* allows proper dimensioning of the number of sensors to be deployed in the observed area. Though general, the approach is applied to the specific case of 802.15.4, an IEEE standard widely used in WSN applications.

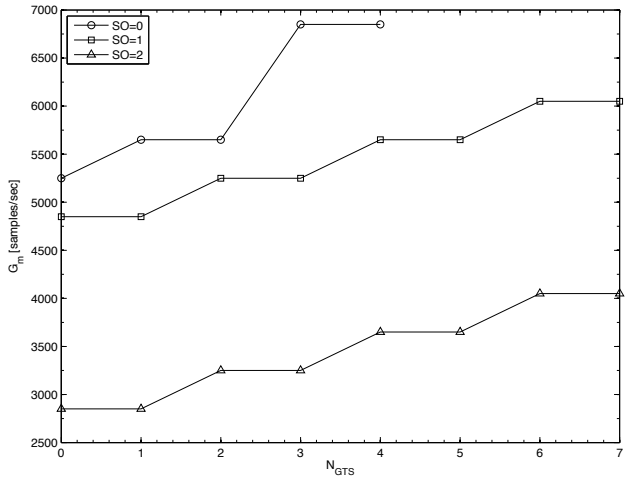

Fig. 8. G_m as a function of the number of guaranteed time slots assigned, with $D = 2$, for different values of SO.

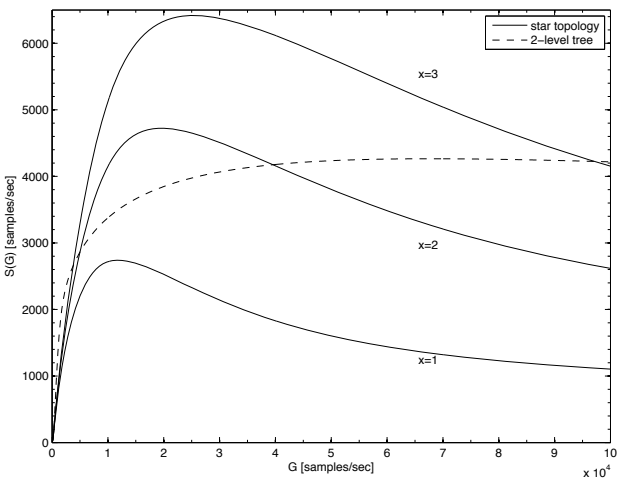

Fig. 9. S as a function of G for the 802.15.4 Beacon-Enabled case, with packet header $H = 2$, different values of payload x, different topologies and aggregation techniques. $SO = BO = 2$ in the star topology case, $SO = 0$ and $BO = 2$ in the 2-level tree case.

While energy consumption is, for many applications of WSNs, a relevant issue, it was not considered in this paper for the sake of brevity. However, in [36] such issue has been considered and might be extended to the analysis of this paper.

ACKNOWLEDGMENT

This work is inspired by the activity performed within the FP7 Network of Excellence in Wireless COMmunications NEWCOM++ (contract no. 216715).

APPENDIX

THE IEEE 802.15.4 MAC PROTOCOL

The Non Beacon-Enabled mode uses a CSMA/CA protocol to access the channel, whereas in the Beacon-Enabled case both contention-based and contention-free protocols, are implemented. In the latter case a superframe is defined, starting with a packet denoted as Beacon (which coincides with the query in our scenario), and divided into two parts: an active and an inactive part. The active part is composed of the Contention Access Period (CAP), where a CSMA/CA protocol is used, and the Contention Free Period (CFP), where a maximum number of seven Guaranteed Time Slots (GTSs) could be allocated to specific nodes (see Figure 2, below part). The use of GTSs is optional.

The duration of the whole superframe and of its active part depends on the value of two integer parameters ranging from 0 to 14, called superframe order, SO, and Beacon order, BO, with $BO \geq SO$. In particular, the interval of time between two successive Beacons, that is the query interval T_q in our scenario, is given by: $T_q = 16 \cdot 60 \cdot 2^{BO} \cdot T_s$, where $60 \cdot 2^{SO} T_s$ is the slot size, and $T_s = 16\ \mu sec$ is the symbol time. The duration of the active part, denoted as T_A, is given by: $T_A = 16 \cdot 60 \cdot 2^{SO} \cdot T_s$.

Each GTS must contain the packet to be transmitted and an inter-frame space, equal to 40 T_s. This is, in fact, the minimum interval of time that must be guaranteed between the reception of two subsequent packets. The PAN coordinator may allocate up to seven GTSs; however, a sufficient portion of the CAP must remain for contention-based access. The minimum CAP size is equal to 440 T_s. We denote as N_{GTS} the number of guaranteed time slots allocated. By changing the packet size (i.e., D) and the slot size (i.e., SO), the number of slots occupied by each GTS and the maximum number of GTSs will vary. As an example, if $D = 2$ and $SO = 0$, two slots are needed for a GTS to contain the packet and the inter-frame space and a maximum number of 4 GTSs could be allocated. In case $SO = 2$, instead, each GTS will occupy one slot and 7 GTSs could be allocated.

We assume that in case a node does not succeed in accessing the channel by the end of the superframe (in the Beacon-Enabled case) or before the reception of the subsequent query (in the Non Beacon-Enabled case), the packet will be lost. Note that in the Beacon-Enabled case, T_q may assume only a finite set of values (depending on the values of BO); whereas in the Non Beacon-Enabled case T_q may assume any value. Being $(120 + D) \cdot T$ the maximum delay with which a packet can be received by the sink [29] and having set the query size equal to 60 bytes in order to ensure that all nodes have completed the CSMA/CA algorithm, the sink should set $T_q \geq (126 + D) \cdot T$. In case lower values of T_q are set, a node may receive a new query when it is still trying to access the channel: the old packet will be lost.

In this paper we use the function $P_{MAC}(n, D)$ derived in [28], [29] and [11], [30] for the Non Beacon- and Beacon-Enabled modes, respectively. In these papers, in fact, mathematical models of the two modalities are provided. For details on the protocols and the models we refer the reader to the standard and to the aforementioned papers as well.

A. The Tree-Based Topology

The tree-based topology defined by Zigbee Alliance [34] is considered here.

According to the Zigbee specifications, the tree formation procedure is started by the PAN coordinator, which broadcasts

Beacons to nodes. A candidate node receiving the Beacon may request to join the network at the coordinator. If the coordinator allows the node to join, it will start transmitting Beacons to allow other candidate nodes to join the network as well.

Nodes work in Beacon-Enabled mode: each child node tracks the Beacon of its parent and transmits its own Beacon at a predefined offset with respect to the beginning of its parent Beacon. The offset must always be larger than the active part of the superframe and smaller than T_q (see Figure 10). This implies that the Beacon and the active part of child superframe reside in the inactive period of the parent superframe: no overlap between the active portions of the superframes of child and parent is present. This concept can be expanded to cover more than two nodes: the selected offset must not result in Beacon collisions with neighboring nodes. Each child will transmit its packet to the parent in the active part (CAP or CFP) of the parent superframe.

We assume that all the active parts of the superframes generated by the routers (level 1 nodes) and the sink have the same duration (i.e., we set a unique value of SO). In these conditions, once we set the value of BO, the number of routers (including the sink) that will have a portion of superframe available for receiving data from their children, will be equal to 2^{BO-SO} (see Figure 10). If more than 2^{BO-SO} routers are present, some of those will not have a portion of superframe available. In this case the packets generated by their children will be lost, since the children cannot access the channel.

We assume that the active part of the superframe defined by the sink is used by level 1 nodes to transmit toward the coordinator itself. The remaining $2^{BO-SO}-1$ superframe portions are randomly allocated to level 1 nodes for receiving data from their children. Under such assumption, there exists a certain probability, denoted as p_{frame}, that a level 1 node has not a portion of superframe available. This probability is given by

$$p_{frame} = \frac{2^{BO-SO}-1}{n_1}, \tag{23}$$

being n_1 the number of level 1 nodes.

According to this, the probability that a node can reach its sink, P_{NET}, given by eq. (5) in the case of Zigbee-compliant trees, is given by:

$$P_{NET} = \frac{n-n_1}{n} p_{frame} P_{MAC}(n_{2,1}, D) P_{MAC}(n_1, D) + \frac{n_1}{n} P_{MAC}(n_1, D). \tag{24}$$

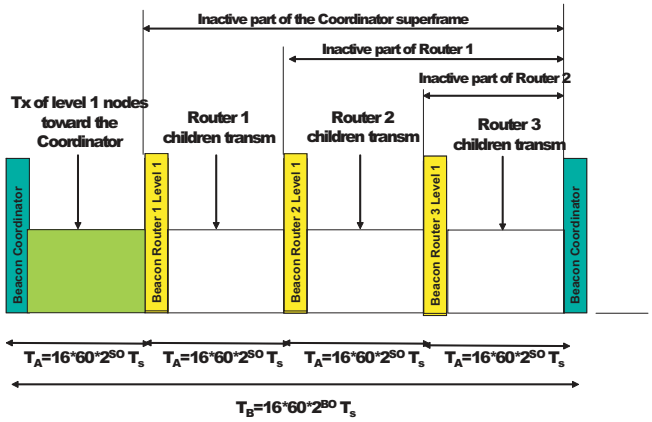

Fig. 10. The superframe structure used in the tree-based topology.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, "A Survey on Sensor Networks," *IEEE Commun. Mag.*, Aug. 2002, 102-114.

[2] H. Karl, A. Willig, "Protocols and Arechitectures for Wireless Sensor Networks," John Wiley and Sons, 2005.

[3] R. Verdone, D. Dardari, G. Mazzini, A. Conti, *Wireless sensor and actuator networks,* Elsevier, first ed., 2008.

[4] J. Matamoros, C. Antn-Haro, "Bandwidth Constraints in Wireless Sensor-based Decentralized Estimation Schemes for Gaussian Channels," in Proc. IEEE Global Conference on Communications (GLOBECOM 2008), 30 November-4 December 2008, Lousiana, New Orleans (USA).

[5] G. Ferrari, M. Martalo' and M. Sarti, "Reduced-complexitydecentralized detection of spacially non-constant phenomena," in Proc. Int. Workshop on Distributed Cooperative Laboratories (Ingrid), Santa Margherita Ligure, Italy, April 2007.

[6] D. Dardari, A.Conti, C. Buratti and R. Verdone, "Mathematical evaluation of environmental monitoring estimation error through energy-efficient Wireless Sensor Networks," *IEEE Trans. Mobile Comput.*, vol. 6, n. 7, July 2007, pag. 790-803.

[7] R. Radeke, D. Marandin, F.J. Claudios, P. Todorova, S. Tomic, "On reconfiguration in case of node mobility in clustered wireless sensor networks", *IEEE Wireless Commun.*, Volume 15, Issue 6, December 2008 Page(s):47 - 53

[8] Qingjiang Tian, E.J.Coyle, "Optimal Distributed Detection in Clustered Wireless Senso Networks", *IEEE Trans. Signal Process.*, Volume 55, Issue 7, Part 2, July 2007 Page(s):3892 - 3904

[9] F. Comeau, S. Sivakumar, W.J. Phillips, W. Robertson, "A Clustered Wireless Sensor Network Model Based on LogDistance Path Loss", *Communication Networks and Services Research Conference*, 2008. CNSR 2008. 6th Annual 5-8 May 2008 Page(s):366 - 372

[10] C. Buratti, R. Verdone, "A Hybrid Hierarchical Architecture: From a Wireless Sensor Network to the Fixed Infrastructure," *IEEE EW2008*, 22-25 June 2008, Prague, Czech Republic.

[11] C. Buratti, "Performance Analysis of IEEE 802.15.4 Beacon-Enabled Mode, IEEE Trans. Veh. Technol., 2010, vol. 59, no. 4, pp. 2031-2046, May 2010.

[12] N. Abramson, "The Aloha System - Another Alternative for Computer Communications," Proc. Fall Joint Computer Conference, 1970, pp 281-281.

[13] Y. Yang, T. Yum, "Delay Distribution of Slotted Aloha and CSAMA," IEEE Trans. Commun., vol. 51, n. 11, nov. 2003, pp. 1846-1857.

[14] J. Heidemann and D. Estrin, "An Energy-Efficient MAC Protocol for Wireless Sensor Networks," *Proc. of 12th IEEE International Conference on Computer Networks (INFOCOM 2002)*, New York, USA, Jun 2002.

[15] Y. C. Tay, K. Jamieson, H. Balakrishnan, "Collision-minimizing CSMA and its applications to wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, Issue 6, Aug. 2004, pag. 1048-1057.

[16] IEEE 802.15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs), IEEE, 2003.

[17] P. Stuedi, O. Chinellato, G. Alonso, "Connectivity in the presence of Shadowing in 802.11 Ad Hoc Networks," *IEEE WCNC 2005*.

[18] C. Buratti, R. Verdone, "On the of Cluster Heads minimising the Error Rate for a Wireless Sensor Network using a Hierarchical Topology over IEEE802.15.4," *IEEE PIMRC 06*, Helsinki, FI, 11-14 Sept, 2006.

[19] C. Bettstetter, "On the minimum node degree and connectivity of a wireless multihop network," *Proc. ACM Symp. on Mobile Ad Hoc Networks and Comp (Mobihoc)*, Jun. 2002.

[20] Z. Vincze, R. Vida, and A. Vidacs, "Deploying multiple sinks in multi-hop wireless sensor networks," *IEEE International Conference on Pervasive Secrices*, 15-20 July 2007, pp. 55-63.

[21] H. Takagi, L. Kleinrock, "Throughput Analysis for Persistent CSMA Systems," *IEEE Trans. Commun.*, vol. 33, No. 7, July 1985.

[22] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," *IEEE J. Sel. Areas Commun.*, vol. 18, n. 3, March 2000.

[23] K.J. Zdunek, D.R. Ucci and J.L. Locicero, "Throughput of Nonpersistent Inhibit Sense Multiple Access with Capture," *Electronics Lett.*, vol. 25, no. 1, pp30-31, Jan. 1989.
[24] J. H. Kim, J. K. Lee, "Capture Effects of Wireless CSMA/CA Protocols Rayleigh and Shadow Fading Channels," *IEEE Trans. Veh. Technol.*, vol. 48, No. 4, July 1999.
[25] P. Siripongwutikorn, "Throughput Analysis of an IEEE 802.1lb Multihop Ad Hoc Network," *IEEE TENCON 2006*, Nov. 14-17 2006,pag. 1-4.
[26] F. Fabbri, R. Verdone, "A statistical model for the connectivity of nodes in a multi-sink wireless sensor network over a bounded region," *IEEE EW2008*, 22-25 June 2008, Prague, Czech Republic.
[27] F. Fabbri, C. Buratti and R. Verdone, "A Multi-Sink Multi-Hop Wireless Sensor Network Over a Square Region: Connectivity and Energy Consumption Issues," *GLOBECOM Workshops, 2008 IEEE*, vol., no., pp.1-6, Nov. 30 2008-Dec. 4 2008.
[28] C. Buratti, R. Verdone, " A Mathematical Model for Performance Analysis of IEEE 802.15.4 Non-Beacon Enabled Mode," *IEEE EW2008*, 22-25 June 2008, Prague, Czech Republic.
[29] C. Buratti, R. Verdone, "Performance Analysis of IEEE 802.15.4 Non Beacon Enabled Mode," *IEEE Trans. Veh. Technol. (TVT)*, vol. 58, n. 7, Sept. 2009.
[30] C. Buratti, "A Mathematical Model for Performance of IEEE 802.15.4 Beacon-Enabled Mode," *ACM IWCMC*, Leipzing, Germany, 21-24 June 2009.
[31] R. Verdone, F. Fabbri, C. Buratti, "Area Throughput for CSMA Based Wireless Sensor Network," *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*, vol., no., pp.1-6, 15-18 Sept. 2008.
[32] J. Orriss and S. K. Barton, "Probability distributions for the number of radio transceivers which can communicate with one another," *IEEE Trans. Commun.*, vol. 51, no. 4, pp. 676-681, Apr. 2003.
[33] R. Verdone, J. Orriss, A. Zanella, S. Barton, "Evaluation of the blocking probability in a cellular environment with hard capacity: a statistical approach," *Personal, Indoor and Mobile Radio Communications (PIMRC 02)*, 2002. vol. 2, 15-18 Sept. 2002.
[34] The ZigBee Alliance web site: http://www.zigbee.org/en/index.asp
[35] A. Koubaa, M. Alves, E. Tovar, "Modeling and Worst-Case Dimensioning of Cluster-Tree Wireless Sensor Networks," *27th IEEE International Real-Time Systems Symposium (RTSS'06)*, December 5-8, 2006, Rio de Janeiro, pp. 412-421.
[36] F. Fabbri, C. Buratti, R. Verdone, J. Riihijärvi and P. Mähönen, "Area Throughput and Energy Consumption for Clustered Wireless Sensor Networks," *Proc. IEEE WCNC 2009*, Apr. 2009, Budapest, Hungary.

Roberto Verdone is full Professor at the University of Bologna, Italy, affiliated with WiLab, the Wireless Communication Laboratory, leading Networks@WiLab, the research group dealing with wireless sensor and ad hoc networks. He published about 100 IEEE articles since 1991, when he got his Master degree in Electronic Engineering at the University of Bologna. In 1995 he received his PhD in Telecommunications at the same University, then he joined the National Research Council, till 2001, when he became a full professor. He participated to many National and European projects. He is the general chairman of the European COST Action 2100 on mobile and wireless communications. He is IEEE member.

Flavio Fabbri received the B.S. and M.Sc. degrees in Telecommunication Engineering from the University of Bologna, Italy, in 2004 and 2007, respectively. He has been involved in research activities in wireless communication networks, specifically wireless ad hoc and sensor networks, since 2007. He is currently a PhD candidate with Networks@WiLab, the wireless networks group at University of Bologna, Italy. He is actively involved in the European Network of Excellence on Wireless Communication, NEWCOM++, and has been a visitor of iNETS group at RWTH Aachen University, Germany.

Chiara Buratti received the M.S. degree (summa cum laude) in Telecommunication Engineering from the University of Bologna, Italy, in 2003. On May 2009 she received the Ph.D. degree in Electronics, Computer Science and Systems at the University of Bologna. Her research interest is on Wireless Sensor Networks, with particular attention to routing and connectivity issues for treebased topologies, and MAC protocols, with particular interest in IEEE 802.15.4 standard. Since 2004, she has collaborated to some European Projects, such as NEWCOM, the Network of Excellence in Wireless COMmunications funded by the EC within the FP6, CRUISE, another Network of Excellence, funded by the EC within FP6, related to Wireless Sensor Networks, and NEWCOM++, the follow up of NEWCOM, funded within the FP7. Since 2009 she is involved in the Artemis project eDIANA, being Leader of the Task devoted to the definition of topology and communication protocols. She acted as a TPC member for several international IEEE conferences.