
Messaoud Doudou$^{ab}$, Djamel Djenouri$^a$, Jose M. Barcelo-Ordinas$^c$, Nadjib Badache$^{ab}$

$^a$: CERIST Research Center, Algiers, Algeria
$^b$: University of Sciences and Technology Houari-Boumediene (USTHB), Algiers, Algeria
$^c$: DAC Department, Universitat Politecnica de Catalunya-BarcelonaTECH (UPC), Barcelona, Spain
E-mails: {doudou,badache}@cerist.dz, ddjenouri@acm.org, joseb@ac.upc.edu

Abstract—The real-world node deployment aspect is investigated, while considering cost minimization for resolving the energy hole around the sink, which represents a serious problem in typical sensor networks with uniform distribution. A novel strategy is proposed that is based on the use of two sinks and a few extra relay nodes close to the sinks' areas. The traffic is then alternatively sent to the sinks in every other cycle. As a second contribution, an efficient data collection mechanism has been developed to determine the optimal data rate that meets delay requirements of individual sensor reports and improves the network lifetime. The comparison of the proposed node deployment strategy with uniform, non-uniform geometric and linear increase node distributions demonstrates that the cost of the proposed solution is very close to that of the uniform distribution and much lower than all the others, while achieving a load balancing at the same order of the state-of-the-art solutions.

Index Terms—Wireless Sensor Networks, Energy-Efficiency; Node Deployment, Data Collection.

I. INTRODUCTION

Energy management and sensor deployment are two important and strongly related research topics in wireless sensor networks (WSN). The objective of energy management is to increase network operational lifetime through energy efficient protocols (Routing, MAC, etc.), while the main goal in sensor deployment on the other hand is to determine the location of the sensor nodes that minimizes the cost, provides high coverage and resilience to failures, and notably prevent energy hole. Due to the multi-hop sink-centric traffic in typical WSNs, the network often experiences unbalanced traffic distribution where sensor nodes act as data originators and relay nodes. Since the entire network traffic flows toward the sink, the nodes nearer to the sink will carry heavier traffic and will deplete their energy faster. This tends to create energy holes around the sink and partitioning the whole network, while the energy of the large majority of nodes remains unused.

Uniformly exploiting the energy of all nodes is challenging in a multi-hop WSN, but also essential for maximizing the network service. Numerous works have been conducted for this purpose in the last few years. However, most of them rely on increasing the number of node around the sink to balance traffic which inevitably increases the cost of deployment. Moreover, the application requirements in terms of end-to-end delay and optimum reporting rate at individual sensors have not been considered. These issues are considered in this paper, where a novel cost-effective node distribution strategy is introduced. The proposed solution takes the application requirement into account and balances energy depletion of the whole network.

The remainder of the paper is organized as follows. Section II presents related work. Section III describes different notation and models being used in this paper. Section IV details earlier works on node distribution and our novel strategy. Next, Section V presents the evaluation and comparison results of our novel strategy with uniform, non-uniform geometric and linear increase node distributions. Finally, the paper is concluded in Section VI.

II. RELATED WORK

Data gathering with load balancing in terms of residual energy at each node may be an efficient approach to address the energy hole problem. Routing with data fusion/compression has been considered for energy balancing in [4]. However, computing the most balanced data gathering routes is an NP-hard problem [5]. For the same purpose, Clustering with data aggregation have also been considered. For example, LEACH [1] and EECS [11] use clusters to distribute the energy depletion among all nodes. Energy hole prevention by load balancing using power control has been studied in [13], [16], and [17]. Power control may play a significant role, but it is far from being sufficient to eliminate the energy hole problem. Aside from power control for balancing energy, node distribution is an appealing method to avoid the energy hole. With uniform node distribution and continuous traffic, authors of [4] have shown that the energy hole problem is unavoidable in WSNs. A non-uniform node distribution has been considered in [2], [3], [6], [7], [15], where the node density in regions closer to the sink is increased progressively to cope with the increasing traffic construction in those regions. The deployment of extra relay nodes around the sink can also be helpful in solving energy imbalanced problem [9], [10]. Hou et al. [19] propose a solution to discover the optimal locations
of relay nodes. The drawback of these solutions is that they considerably increase the number of nodes around the sink, which consequently increases the deployment cost.

Sink roaming is introduced in [12] [14]. The mobility is a good approach to avoid the energy hole, but cannot be easily adopted in real sensor applications. Finally, multiple sinks can be deployed [8] to collect data over a certain sub-region of the monitored area where the optimal assignment of traffic to multiple sinks is found using a method similar to electrostatic theory. An effective evaluation framework with a generalized network lifetime and cost models was presented in [18]. Authors state that the effectiveness of the solution must trade-off the efficiency of energy depletion strategy with the deployment cost. The application requirements are not considered thus far. In this paper, we introduce a novel node deployment strategy that takes into consideration these above mentioned factors and resolves the energy hole problem as well.

III. NOTATION AND ASSUMPTIONS

In the following, we assume that: (1) all nodes are homogeneous and have an ID number, (2) nodes are static with uniform initial energy and non-renewable energy budget; (3) Each node has a maximum communication radius denoted by, \( R_c \), and a sensing radius denoted by, \( R_s \); (4) Nodes are required to sense and send their data whereas relay nodes only forward traffic being received; (5) Every node sends its traffic at a certain rate, say \( r_i \) (6). in the following, the different models used in the paper are defined.

To ensure full sensing coverage of a given region of interest, every point in the latter must be covered by at least one sensor. Given the required number of nodes, \( N_i \), to cover an Area, \( A_i \), the density, \( \rho_c \), is given by \( \rho_c = \frac{N_i}{A_i} \) [nodes/m²]. As depicted in Fig. 1, uniform random, a Tri-Hexagon Tiling (THT), a square grid, and triangle lattice based coverage are example of converge models in WSN. We consider square model in our proposed strategy.

It is supposed in the rest of the paper that a set of sensors are be deployed in a grid region of length, \( L \), and width, \( W \) (Fig. 2.a), or in a circular region (Fig. 2.b) of network radius, \( R \). The grid area is also partitioned into \( D \) belts with the same size, \( W \times R_c \). The disk region is divided into \( D \) adjacent rings with the same width of \( R_c \), where \( R_i \) denotes the \( i^{th} \) ring. Obviously, the ring \( R_i \) is composed of nodes whose distances to the sink are between \((i-1)\) and \( i \). For example, in Fig. 2.b, \( D = 3 \) and the area is partitioned into three rings \( R_1 \), \( R_2 \), and \( R_3 \). The network topology is represented by a graph \( G = (V, E) \), in which \( V = V_b \cup V_c \cup V_r \) is the set of nodes; It includes a set of sinks, \( V_b \), of sensor nodes, \( V_c = \{1, ..., N_s\} \), and relay nodes, \( V_r = \{1, ..., N_r\} \), \((i, j) \in E \) if nodes \( i \) and \( j \) are in the transmission range of each other.

The initial energy of each sensor is denoted, \( \varepsilon > 0 \). But the sink is supposed to have no energy limitation. The energy consumed in communication is supposed to follow the First Order Radio Model [1]. In this model, the amount of energy to transmit a bit can be represented as: \( E_{\text{tx}} = E_{\text{elec}} + \epsilon_{\text{amp}}d^\alpha \), and the amount of energy to receive a bit can be represented as, \( E_{\text{rx}} = E_{\text{elec}} \), where \( E_{\text{elec}} \) represents the electronics energy, \( \epsilon_{\text{amp}} \) is determined by the transmitter amplifier’s efficiency and the channel conditions, \( d \) represents the distance over which data is being communicated, and where \( \alpha \) represents the path loss exponent, \( 2 < \alpha < 6 \).

IV. NODE DEPLOYMENT STRATEGIES

A. Uniform Node Deployment

When considering a disk region with uniform node distribution, the density of nodes in each ring/belt is given by, \( \rho_i = \frac{N_i}{A_i} = \frac{N_i}{R_i^2} \rho_j \) \( \forall \) \( i, j \). The number of nodes in each ring, \( R_i \), is then given by

\[
N_i = \frac{2i - 1}{2i - 3} 
\]

(1)

Where \( A_i = \pi(2i - 1)R_c^2 \), \( 1 \leq i \leq D \). Given a number of nodes, \( N_D \) in ring \( R_D \), by iterating Eq.(1) we obtain \( N_i = N_D \frac{2i - 1}{2D - 1} \), and the total number of deployed nodes is given by

\[
N = \sum_{i=1}^{D} N_i = N_D \frac{D^2}{2D - 1} 
\]

(2)

Obviously, when uniformly deploying sensors in each ring, inner nodes (close to the sink) will consume more energy than outer nodes. This is because inner nodes have to relay a large amount of traffic than outer nodes. As a result, the inner nodes will die faster (deplete their batteries), which creates the energy hole around the sink. The non-uniform node
distribution strategies have been proposed to cope with this problem.

B. Non-Uniform Node Deployment

A circular area is generally considered in non-uniform node distributions. In [2], authors propose to increase the node density from outer to inner rings in geometric progression to cope with the increasing traffic. The following condition is satisfied:

\[
\frac{N_i}{N_{i+1}} = \begin{cases} 
q, & q > 1, 1 \leq i \leq D - 2, \\
q - 1, & i = D - 1.
\end{cases}
\]  

(3)

The total number of deployed nodes is given by, \(N = N_Dq^{D-1}\). As node density grows in geometric progression towards the sink and each node introduces a packet at each round, large number of redundant packets is generated around the sink, which overloads the network.

In [3], the authors propose to use a linear increase of nodes from outer to inner rings, instead of geometric progression. The main idea is to duplicate the number of nodes of the outer ring and add it to the adjacent inner ring (see Fig. 3).

Considering Eq.(1); if the number of nodes in ring \(R_D\) is \(N_D\), then the number of nodes in ring, \(R_i\), is given by,

\[
N_i = N_D \frac{2i - 1}{2D - 1} + N_{i+1} = N_D \frac{D^2 - i^2 + 2i - 1}{2D - 1} 
\]  

(4)

and the total number of deployed nodes is given by,

\[
N = N_D \times D \frac{4D^2 + 3D - 1}{6(2D - 1)} 
\]  

(5)

In [2] a q-switch routing algorithm was proposed in which any node in the network, except in ring \(R_D\), has \(q\) forwarder nodes from the next inner ring. The choice of relay nodes is based on the maximum residual energy to balance energy depletion among nodes. Each intermediate node sends one packet to each relay and drop the other packets.

C. Novel Node Deployment Strategy

To cope with the increasing number of nodes, a novel node deployment strategy is proposed as follows. A rectangular grid monitoring area is considered as shown in Fig. 2.a. The grid area with length, \(L\), and width, \(W\), is divided into \(D\) belts, with the same width, \(R_c\), and length, \(W\), such as their areas satisfy, \(A_1 = A_2 = \ldots = A_D\). Square lattice coverage model is assumed. It ensures 1-coverage, with \(R_c \geq 2 + R_s\), which provides a connected network. As depicted in Fig. 4, the solid circle is the communication range and the narrowed shadowed circle is the sensing range of each sensor. Sensors are uniformly deployed in each belt such that the node density in each belt satisfies \(\rho_i = \frac{N_i}{A_i} = \frac{N}{A} = \rho_j, \forall i,j\).

The set of sinks is denoted by \(V_p\), the set of relays as by \(V_r\), and the set of sensor nodes by \(V_s\). The same number of sensors is deployed in each belt, ie. \(N_i = N_j \forall i, j \in \{1, \ldots, D\}\). Hence, the total number of sensors is given by, \(N_s = N_D \times D\).

According to the coverage model, the network is connected and the sensors generate data packets and forward them to the sink. We refer to the two smallest edges of the grid as the top side and the down side (Fig. 4). To efficiently balance the energy among sensor nodes, two data sinks are deployed to collect data from sensors. \(Sink_1\) is centered at the borderline of the top side of the grid, and \(Sink_2\) at the bottom one (see Fig. 4). Every sensor chooses only one forwarder node from the top adjacent belt and one from the down adjacent belt (Fig. 4). Sensors in belts, 1, and, \(D\), located one hop away from \(Sink_1\) and \(Sink_2\), send traffic directly to these sinks, while other sensors use dedicated relays. Instead of using uniform sensor nodes at these belts used for sensing coverage, we propose to add some relays around \(Sink_1\) and \(Sink_2\) whose role is to forward traffic from sensors at these belts that cannot reach the sink in one hop. The relay nodes do not generate traffic but just forward data packets from their connected sensors as shown in Fig. 4. By doing so, the energy
depletion among sensors in belt 1 and $D$ is balanced. The relay nodes are deployed to ensure a uniform energy depletion among sensors in belt 1 and belt $D$. However, the energy between different belts is still unbalanced since belts close to the sink carry more traffic.

Instead of increasing the number of nodes in all belt, a different approach is proposed. The traffic forwarding is alternatively dispatched toward $Sink_1$ and $Sink_2$ (see Fig. 5) in each other round (cycle), so that all nodes will carry the same amount of traffic and thus exhaust their energy progressively and at the same time. To determine the number of relay nodes required to ensure the energy depletion around the two sinks, the distance separating each sensor in belt 1, and belt $D$, from $Sink_1$ and $Sink_2$ is considered. Let this distance be denoted by $d_i$. For each sensor, $i$, in these belts, the number of relays required to forward traffic from this sensor to the closest sink is determined by the ratio $d_i/R_e$. Hence, the number of deployed relays is given by,

$$N_r = \sum_{i \in N_1 \cup N_D} \lfloor d_i/R_e \rfloor$$

The total number of deployed nodes is given by,

$$N = N_D \times D + \sum_{i \in N_1 \cup N_D} \lfloor d_i/R_e \rfloor$$

### D. Distributed Data Collection Protocol

Every sensor forwards traffic to exactly one forwarder node from the top adjacent belt, and one from the down adjacent belt. This can be achieved by running a simple distributed algorithm at each node to discover its neighbors and choose its forwarders. Alternatively in each round, sensors send their traffic to one of the two sinks. That is, for a given round $i$, the traffic is directed to $Sink_1$, and then belts closer to $Sink_1$ will carry more traffic than belts closer to $Sink_2$. While in the next round $i + 1$, the traffic will be directed to $Sink_2$, and consequently belts closer to $Sink_2$ will carry more traffic than belts closer to $Sink_1$. This way, after two rounds, and given the assumption of homogeneous traffic generation, all nodes will carry the same amount of traffic and their energy will be balanced. For instance, consider the example of Fig. 5 with 5 sensors and 2 relays in the forwarding path. Each sensor generates 1 packet and forwards its own and the received traffic towards $Sink_1$ in a current round, and towards $Sink_2$ in the next round. By doing this, each sensor will forward exactly 6 packets and each relay forwards 5 packets after 2 rounds. In general, the average load carried by each node $i$ per round is given by,

$$Load(i) = \begin{cases} \frac{D+1}{2}, & i \in N_s \\ \frac{D}{2}, & i \in N_r \end{cases}$$

### E. Delay Award Rate Control

Although the load and the energy is fairly balanced among nodes, the amount of load carried by each node is relatively high, which limits the service lifetime of the network. The aim is to maximize this lifetime while satisfying the quality of service required by the application. The traffic load is determined by the traffic generation rate of each node, and this is in fact related to the maximum end-to-end delay required by the application, denote $L_{max}$. We propose to control the traffic rate $r_i$ of each node and to find the optimal rate $r_i^*$ required to satisfy a given application requirement, in terms of maximum latency $L_{max}$. Let us assume that all traffic flows in one direction and that there are $D$ sensors in each path towards the sink ($D$ is the number of belts). The duration of a round, $T$, is equal to the maximum time for a data packet from any node to reach the sink. We have,

$$T = \sum_{l \in P} (T_{cw} + T_{data} \times P_l)$$

where $T_{data} = T_{hdr} + P/R$, and $P_l$ is the probability of successful transmission over link $l$, belonging to the path, $P$. Symbols used in this equation and others are described in table I with their typical values. Making each node forwarding at most 1 packet per round $T$ is a targeted objective. In the worst case, if the data rate of a node $i$, is $r_i = 1/T$, then the average load per round is given by Eq.(8). Now, if the maximum end-to-end delay required by the application is $L_{max} = D \times T$, then the optimal data rate will be $r_i^* = \frac{1}{D \times T}$. In this case each sensor generates 1 packet every $D$ rounds, and each node forwards at most 1 packet per round. Thus, when considering the second direction of traffic flow, all nodes will forward the same load each $2D$ rounds as shown by the example in Fig. 6 with 5 sensor nodes. The average load in this case will be,

$$Load(i) = \begin{cases} \frac{D+1}{2D}, & i \in N_s \\ \frac{1}{2}, & i \in N_r \end{cases}$$

In general, lets $k = \lceil \frac{L_{max}}{D \times T} \rceil$, then, the optimal data rate of each node is expressed by,

$$r_i^* = \frac{1}{k \times D \times T}$$
A. Setup and Parameters

The proposed solution is evaluated analytically through simulations and compared against others node deployment solutions namely: uniform, non-uniform geometric progression, and linear increase strategy. All the solutions are implemented in TinyOS under the network topology and conditions defined by each strategy. The traffic rate of each node is set to $1 \text{pkt/s}$ and is set to $r^*$ for our solution. The default TinyOS LPL MAC protocol and CC2420 radio transceiver are used by all the solutions. Table I sketches the simulation setup and the parameter values used in this evaluation. Comparison is carried out from the number of deployed nodes, the deployment cost, and the energy depletion perspective as described in the following:

B. Number of Deployed Nodes

The total number of deployed nodes is first investigated, as well as its impact on the system performance and cost. It’s obvious that the lower the number of deployed nodes is, the lower the system cost will be. Fig. 7.a) shows the comparison of total number of deployed nodes following uniform, non-uniform geometric progression, linear increase, and the proposed strategy. It can be observed that in case of geometric and linear increase distributions, the total number of nodes increases rapidly with the network size.

In these two strategies, rings closer to the sink have increased node density. Consequently, large number of redundant data packets are produced and forwarded to the sink. Intuitively, the proposed strategy will perform better with much less number of nodes and less traffic, which are tuned to just be sufficient to cover the monitored area.

V. EVALUATION AND COMPARISON

TABLE I
DIFFERENT SYMBOLS USED IN THIS PAPER WITH TYPICAL VALUES [20]

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P^*$</td>
<td>Payload [byte]</td>
<td>28</td>
</tr>
<tr>
<td>$R$</td>
<td>CC2420 Radio Rate [kbyte/s]</td>
<td>31.25</td>
</tr>
<tr>
<td>$P_f$</td>
<td>Probability of successful transmission over link $l$</td>
<td>$[0, 1]$</td>
</tr>
<tr>
<td>$T_{cw}$</td>
<td>Time [ms] to turn the radio on and probe the channel (carrier sense) + contention window duration</td>
<td>$2.60 + 15*0.62$</td>
</tr>
<tr>
<td>$L_{pmb}$</td>
<td>CC2420 Minimal preamble length [byte]</td>
<td>4</td>
</tr>
<tr>
<td>$T_{hdr}$</td>
<td>$L_{pmb} / R$ [ms]</td>
<td>$9/31.25 = 0.288$</td>
</tr>
</tbody>
</table>

The cost is defined as a function of a given deployment, $C = C_n + C_b$, where $C_n$ is the cost of deploying a set of sensor and relay nodes, and $C_b$ is the cost of deploying a set of sinks. The circular area has a network radius, $R = 200m$, and the grid has a length, $L = 200m$, and a width, $W = 100m$, while a communication radius is, $R_c = 20m$, and $N_D = 15$. Following the cost model in [18], the cost of a micro-sensor is assumed to be $75\$, and that of the sink is $500\$. Fig. 7.b) shows the deployment cost of different strategies when increasing the network size. Although our strategy introduces 2 data sinks and additional relays, the engendered cost reasonably exceeds that of uniform deployment, but it is much lower than that of geometric and linear increase. This advantage is achieved to the small number of total deployed nodes, which makes the proposed solution cost effective.

D. Energy Depletion

To evaluate the energy consumption the network is divided into 5 rings/belts, and the transmission radius of each node is set to $20m$. Every node starts with the same initial energy of $\varepsilon = 0.1$ Joules. Packets are of length $l = 200\text{bits}$, $E_{elec} = 50\text{nJ/bit}$, and $E_{amp} = 0.0013\text{pJ/bit/m}^4$. All the readings are taken as the average of 50 independent runs. Fig. 8.a) shows the average per-node energy consumption in each ring or belt when $N_D = 27$, and the coefficient $q = 2$. It has been observed that the variation of the energy consumed by nodes in inner rings for geometric and linear increase distribution
residual energy (%)

10 20 30 40 50 60 70 80

0

Novel node distribution strategy that balances energy depletion tackled the unbalanced energy problem herein by proposing a costly and ineffective for real sensor network deployments. We unbalanced energy are presented. This makes these solutions nodes around the sink to cope with the increased traffic and the distribution, load balancing cannot be achieved in multi-hop energy balanced network. It is known that with uniform node distribution, and in terms of cost-

VI. CONCLUSION

This paper considers the cost-effective node deployment for energy balanced network. It is known that with uniform node distribution, load balancing cannot be achieved in multi-hop many-to-one sensor networks. The existing non-uniform node deployment strategies that rely in increasing the number of nodes around the sink to cope with the increased traffic and the unbalanced energy are presented. This makes these solutions costly and ineffective for real sensor network deployments. We tackled the unbalanced energy problem herein by proposing a novel node distribution strategy that balances energy depletion of the whole network and takes the application requirement into account. This is by determining the optimal data rate that meets delay need of individual sensor reports. The comparison of the proposed node deployment strategy with uniform, non-uniform geometric, and linear increase node distributions clearly demonstrates that our solution avoids the energy hole, and guarantee an energy balanced network required to achieve a specific service lifetime at the same order to non-uniform distribution, while considerably reducing the cost in terms of additional nodes.

REFERENCES


