Chapter from the book *Wireless Sensor Networks - Technology and Protocols*
Downloaded from: http://www.intechopen.com/books/wireless-sensor-networks-technology-and-protocols

Interested in publishing with IntechOpen?
Contact us at book.department@intechopen.com
1. Introduction

Wireless sensor networks (WSNs) are attracting increasing research attention, due to their wide spectrum of applications, including military purposes for monitoring, tracking and surveillance of borders, intelligent transportation systems for monitoring traffic density and road conditions, and environmental applications to monitor, for example, atmospheric pollution, water quality, agriculture, etc. [26].

A WSN is composed of a number of sensor nodes (SN) transmitting wirelessly the information they capture. An SN is generally composed of a power unit, processing unit, sensing unit, and communication unit. Power consumption is the main limiting factor of an SN. In fact, SNs are in general required to operate autonomously and independently for a large period of time in areas where power infrastructure may not be available. Thus, battery-powered SNs should be able to operate with very low power consumption. Some SNs have batteries rechargeable by solar power, thus ensuring longer autonomous operation. The processing unit is responsible to collect and process signals captured from sensors before transmitting them to the network. The sensing unit is a device that produces a measurable response to a change in a physical condition like temperature or pressure. The wireless communication unit is responsible for transferring the sensor measurements to the exterior world, e.g., to be stored on a server, where they can be distributed on the internet or accessed by specialized personnel. The wireless communication unit can also ensure a mechanism for ad-hoc communication between SNs forming a WSN [26]. In fact, in some scenarios, it might be more energy efficient to transmit a message via multihop communications over short distances instead of a single hop long distance transmission to the base station (BS).

In this Chapter, a protocol for energy efficient multihop communications in WSNs is presented and analyzed. In the presented approach, SNs form cooperative groups or clusters. Within each cluster, SNs communicate with each other over multihop links, and the SN at the last hop communicates with the BS by relaying the aggregated multihop data. Thus, cooperation between SNs is exploited for the benefit of energy efficiency. Hence, SNs use two wireless
interfaces: one to communicate with the BS over a long-range (LR) wireless technology (e.g., UMTS/HSPA, WiMAX, or LTE), and one to communicate with other SNs over a short-range (SR) wireless technology (e.g., Bluetooth, ZigBee, or WLAN). In addition to freeing bandwidth at the BS and increasing network throughput [19, 20], SR collaboration between SNs leads to a reduced energy consumption [8, 31]. In fact, higher rates can be achieved over SR communications between SNs that are relatively close from each other in a single cooperating cluster. This leads to shorter transmission and reception times and hence less energy consumption from the batteries of the SNs.

In this Chapter, SNs are considered to be distributed throughout the cell area and can form several cooperating clusters. The energy minimization problem during cooperative content distribution in the multiple clusters case is formulated and the solution outline is presented. Multihop communications are studied, and remarkable energy savings are achieved even with the 2-hop scenario, corresponding to a clustering framework where a single SN, the cluster head (CH), is in charge of directly receiving the measurement data from each SN in the cluster on the SR, and for transmitting the aggregated data to the BS on the LR. A general formulation that incorporates both multihop and clustering is presented, and energy efficient suboptimal schemes are proposed.

The paper is organized as follows. Related work is presented and differences with the proposed approach are outlined in Section 2. The system model is presented in Section 3. The problem formulation and solution are discussed in Section 4. Suboptimal schemes leading to significant energy savings at reduced complexity are proposed in Section 5 for the multihop and clustering scenarios. The simulation results are presented in Section 6. Practical implementation aspects are discussed in Section 7. An application example of a WSN for air quality monitoring is presented in Section 8. Potential research directions for future investigation are described in Section 9. Finally, conclusions are drawn in Section 10.

2. Related work

This section presents an overview of related work in energy efficiency in multihop wireless communications. Differences with the approach investigated in this Chapter for energy efficient cooperative multihop data transmission are outlined.

Network topology design in order to achieve different requirements in a service-oriented framework is considered in [32]. Requirements include throughput maximization, delay constraints, security, and reliability. Energy minimization constraints are not considered. Topology control is also considered in [22], where energy constraints are taken into account via transmit power adjustments. Connectivity between nodes is determined based on distance considerations. In [23] and [16], energy efficiency is considered by having a minimum energy path between each pair of nodes in a wireless multihop network. Topology is controlled by varying the transmission power at each node, and the transmission power at the antenna is considered as the criterion for energy efficiency. In this Chapter, the energy drained from the sensors’ batteries, not only the transmit power at the antenna, is used as the criterion for energy efficiency.

Processing capacity is studied in [25] for wireless sensor networks. A cross-layer collaborative in-network processing approach among sensors is adopted, where, in addition to processing information at the application layer, sensors synchronize their communication
activities to exchange partially processed data for parallel processing. Sensor nodes are grouped into clusters, and operations are performed independently inside each cluster. Communications between clusters are performed using channels that are orthogonal to intra-cluster communications. Multihop communications are implemented inside each cluster to perform parallel computing of certain processing tasks. Thus, energy efficiency is considered in the sense of minimizing the processing power during task scheduling and implementation, not in the sense of transmissions and receptions for relaying measurement data of sensors, as is the case in this Chapter.

Small scale networks where sensor nodes are closely located are studied in [7]. TDMA is assumed as an access method. Both transmission and circuit-based energy consumption are considered. Perfect synchronization between nodes is assumed. The joint design of the physical, MAC, and routing layers to minimize network energy consumption is formulated into a convex optimization problem and the solution is provided. The approach presented in this Chapter does not make any assumptions concerning the channel accessing scheme or the scale of the sensor network.

In [13], energy efficiency is studied in wireless sensor networks. Sensors having data to transmit should relay this data to a single source using multihop. Nodes that do not have data to transmit or that are not relaying the data of other nodes can be put to sleep. Energy efficiency is achieved by reducing the number of active nodes. An energy efficient routing technique in multihop wireless sensor networks is presented in [28]. For each node, the energies consumed during reception, transmission, and sensing are considered in the analysis. In the model of [28], frame nodes relay the content of the source to the destination. If the communication fails between the source and a frame node, or between two frame nodes, assistant nodes come into play and relay the data to the next frame node. Hence the use of opportunistic transmissions depending on the fading conditions of the channel. The optimal number of nodes that should be included in a path is determined. The purpose is to reduce the energy consumption by reducing the number of nodes relaying the data from source to destination. In the scenario investigated in this Chapter, all nodes are assumed to have data to transmit, and hence cannot be put to sleep to achieve energy savings. This scenario corresponds, for example, to WSNs deployed for the purpose of air quality monitoring in a given area, where each sensor will periodically send measurement data to a central processing system.

In [3], multipath routing based on spatial relationships among nodes is considered. Stochastic geometric and queueing models are used for the evaluation of different types of scenarios. Energy aware routing with the possibility of energy replenishment of nodes in multihop wireless sensor networks is presented in [17]. An algorithm that only requires short term energy replenishment information is also presented. However, channel conditions are not taken into consideration in the approach of [17], conversely to the work in this Chapter where channel state information (CSI) is exploited in order to build the energy efficient routes from SNs to the BS.

Several papers in the literature consider implementation scenarios related to a particular standard. For short range multihop communications, IEEE 802.11s is receiving significant attention. In [6], a tutorial is presented for multihop communications and mesh capabilities in IEEE 802.11. Task group 802.11s is handling this issue. In the draft 802.11s proposal, the mesh network is implemented at the link layer and relies on MAC addresses instead
of IP addresses, which provides layer-2 multihop communication. A survey of the unicast admission control schemes designed for IEEE 802.11-based multi-hop mobile ad-hoc networks (MANETs) is presented in [10], where different admission control protocols are discussed and analyzed. In [27], cooperative rate adaptation in multihop IEEE 802.11 is considered. The problem is formulated as an optimization problem and shown to be NP-hard. Thus, a suboptimal method is presented. Energy efficiency is considered in terms of reducing the transmission power at the SNs’ antennas. Enhancements of the performance of IEEE 802.11-based multihop ad hoc wireless networks from the perspective of spatial reuse were surveyed in [2]. Techniques adopting transmit power control, tuning the carrier sensing threshold, performing data rate adaptation, and using directional antennas were discussed. In this Chapter, the presented approach is general and not confined to a particular standard, it does not only consider transmit energy at the antenna, but also the energy drained from the battery during transmission and reception. Compared to mesh networks, not every SN needs to communicate with all other SNs. Instead, each SN needs to transmit the measured data using an optimum energy minimizing path to the BS. This path remains the same as long as the channel conditions remain constant.

In addition to multihop, energy efficient clustering methods are also investigated in the literature. An algorithm is presented in [14] as an improvement on the methods in [12] and [15]. In [12, 14, 15], each node volunteers to be a cluster head in a probabilistic manner, and non-cluster nodes associate themselves with cluster heads based on the announcements received from these cluster heads. The actual energy drained from the battery of the device is considered. However, the problem is not formulated and solved as an optimization problem (as in this Chapter), but rather an efficient clustering algorithm that ensures fairness in energy consumption between nodes, due to the probabilistic selection, is presented. In [15], the use of a proxy node was added to the approach of [12], whereas in [14] the additional use of a main cluster head was implemented, with the main cluster head relaying the data from cluster heads to the BS. The work of [12] was extended in [4] to include multihop communications in addition to clustering. In addition, an approach to determine the optimal number of cluster heads is proposed. Clustering is performed on distance based criteria and a probabilistic random approach is adopted for the election of cluster heads. A cluster head selection based on proximity was adopted in [30], where the residual energy of the node is also considered in the selection process. A multihop time reservation using adaptive control for energy efficiency (MH-TRACE) is presented in [24]. Cluster formation is probabilistic and it is not based on connectivity information. In MH-TRACE, the interference level in the different time-frames is monitored continuously in order to minimize the interference between clusters. MH-TRACE clusters use the same spreading code or frequency and time division is adopted. In this Chapter, cluster head selection is not probabilistic or simply proximity based. Fading is considered in the selection approach since CSI affects the achievable rates and is thus incorporated in the optimization problem.

3. System model

The energy minimization problem in a WSN is considered. The data is to be delivered to the BS from \( K \) SNs distributed throughout the cell area of the BS. The SNs can communicate with the BS using a long range communication technology (e.g., UMTS/HSPA, WiMAX, or LTE), or with neighboring SNs using a short range technology (e.g., Bluetooth or WLAN). SNs form cooperating clusters for the purpose of energy minimization during cooperative
data transmission. Within each cooperating cluster, the data is delivered from the SNs in that cluster to the BS using multihop communications. Fig. 1 shows the scenario considered. The maximum number of hops allowed $H$ can be specified as a parameter. With two-hop communications (case $H = 2$), the problem becomes a clustering problem that consists of finding the best grouping of SNs into cooperating clusters, as shown in Fig. 2.

Each SN transmits its measured data to a single destination, which could be either the BS or another SN. We consider the energy minimization problem with multihop/clustering. The BS and SNs are denoted as “nodes”, with node $k = 0$ corresponding to the BS and nodes $k = 1, ..., K$ corresponding the SNs. As shown in Fig. 1, these nodes appear to form a direct
acyclic graph (DAG) starting from the node \( k = 0 \). If node \( j \) receives the data of node \( k \) on hop \( h \), a parameter \( \alpha_{kj}^h \) is set to one, marking the existence of an edge in the graph between \( k \) and \( j \). Otherwise, \( \alpha_{kj}^h \) is set to zero.

We define \( C_j \) as the set of children of \( j \), i.e., the set of nodes sending their data directly to \( j \):

\[
C_j = \left\{ k, \sum_{h=1}^{H} \alpha_{kj}^h = 1 \right\} 
\] (1)
The set $D_j$ is defined as the sub-DAG starting from $j$, i.e., having $j$ as its root. It includes $j$, its children, the children of its children, etc. Thus, it can be expressed as:

$$D_j = \{j\} \cup \bigcup_{k \in C_j} D_k$$  \hspace{1cm} (2)

### 3.1. Data rates

Given for each node: the transmit power $P_{t,kj}$ that node $k$ is using in order to transmit to node $j$, the channel gain $H_{kj}$ of the channel between $k$ and $j$, and the thermal noise power $\sigma^2$, the received signal-to-noise ratio (SNR) $\gamma_{kj}$ on the link between $k$ and $j$ can be calculated following

$$\gamma_{kj} = \frac{P_{t,kj}H_{kj}}{\sigma^2}$$. Given the target bit error rate $P_e$ and the SNR, the bit rates on the link between any two nodes $k$ and $j$ can be calculated as follows:

$$R_{kj} = W_{kj} \cdot \log_2(1 + \beta \gamma_{kj})$$  \hspace{1cm} (3)

In (3), $W_{kj}$ is the passband bandwidth of the channel between $k$ and $j$, and $\beta$ is called the SNR gap. It indicates the difference between the SNR needed to achieve a certain data transmission rate for a practical M-QAM system and the theoretical Shannon limit [9, 21]. It is given by:

$$\beta = \frac{-1.5}{\ln(5P_e)}$$. The channel gain is expressed as:

$$H_{kj,\text{dB}} = (-\kappa - v \log_{10} d_{kj}) - \xi_{kj} + 10 \log_{10} F_{kj}$$  \hspace{1cm} (4)

In (4), the first factor captures propagation loss, with $d_{kj}$ the distance between nodes $k$ and $j$, and $v$ the path loss exponent. The second factor, $\xi_{kj}$, captures log-normal shadowing with a standard deviation $\sigma_{\xi}$, whereas the last factor, $F_{kj}$, corresponds to Rayleigh fading (generally considered with a Rayleigh parameter $a$ such that $E[a^2] = 1$).

### 4. Multihop problem formulations

With each SN transmitting the data in blocks of size $S_T$ bits, the time needed to transmit this content on a link between nodes $k$ and $j$ having an achievable rate $R_{kj}$ bps is given by $S_T / R_{kj}$. Denoting the power drained from the battery of node $j$ to receive the data from node $k$ by $P_{R,kj}$, then the energy consumed by $j$ to receive the data from $k$ is given by $S_T \cdot P_{R,kj} / R_{kj}$. Similarly, denoting by $P_{T,kj}$ the power drained by the battery of node $k$ to transmit the data to node $j$, then the energy consumed by $k$ to transmit the content to $j$ is given by $S_T \cdot P_{T,kj} / R_{kj}$. It should be noted that $P_{T,kj}$ can be expressed as:

$$P_{T,kj} = P_{T,\text{ref},kj} + P_{t,kj}$$  \hspace{1cm} (5)

where $P_{T,\text{ref},kj}$ corresponds to the power consumed by the circuitry of node $k$ during transmission on the communication interface with node $j$, and $P_{t,kj}$ corresponds to the power transmitted over the air on the link from node $k$ to node $j$.

In this section, a flexible formulation is presented that accommodates power adaptive or rate adaptive transmission. In the case of adaptive rate control, the node transmit power is
constant, i.e., $P_{t,kj} = P_t$ and $P_{Tx,kj} = P_{Tx}$. Consequently, the rate $R_{kj}$ on the link between nodes $k$ and $j$ is the rate achievable with the transmit power $P_t$. It is varied adaptively depending on the channel conditions between nodes $k$ and $j$. High data rates result in low energy per bit consumption, thus leading to a gain in total energy consumption. For example, the WLAN technologies apply rate control [11].

In the case of adaptive power control, the nodes communicate at a constant rate $R_{0j} = R_L$ on the LR or $R_{kj} = R_S$ (with $k > 0$) on the SR. The transmit power $P_{t,kj}$ is varied adaptively depending on the channel conditions between nodes $k$ and $j$ in order to achieve the target data rate $R_L$ or $R_S$. Thus, nodes that are in proximity of each other will communicate with lower power than nodes that are further apart. This will result in a reduction of consumed energy. Some technologies such as Bluetooth apply power control [5].

Hence, the energy consumed during cooperative multihop content distribution can be expressed as follows:

$$ E_{coop} = S_T \cdot \sum_{k=1}^{K} \sum_{j=0,j\neq k}^{K} \sum_{h=1}^{H} \frac{\alpha_{kj}^h \cdot |D_k| \cdot P_{Tx,kj}}{R_{kj}} $$

$$ + S_T \cdot \sum_{k=1}^{K} \sum_{j=0,j\neq k}^{K} \sum_{h=1}^{H-1} \frac{\alpha_{kj}^h \cdot |D_k| \cdot P_{Rx,kj}}{R_{kj}} $$

$$ = S_T \cdot \sum_{k=1}^{K} \sum_{j=0,j\neq k}^{K} \sum_{h=1}^{H} \frac{\alpha_{kj}^h \cdot |D_k| \cdot (P_{Tx,kj} + P_{Rx,kj})}{R_{kj}} $$

where the first term corresponds to the energy consumed by the nodes for transmission and the second term corresponds to the energy consumed by the nodes for reception. Hop $h = H$ corresponds to transmission on the LR and node $k = 0$ corresponds to the BS. The multiplication by $|D_k|$, with $\cdot$ denoting set cardinality, is used to indicate that an SN aggregates the data of its sub-DAG before transmitting it on the next hop. To be able to write the last equality in (6), it is assumed that $P_{Rx,k0} = 0$ for all $k$. This corresponds to excluding the energy consumed at the BS to receive the data at hop $H$. In fact, power consumption of the BS is not considered in the energy minimization process since the interest is in the battery life of the SNs. This is justified by the fact that most BSs rely on power line cables and not on batteries and thus do not have as stringent power limitations as the SNs.

Consequently, the optimization problem can be formulated as follows:

$$ \min_{\alpha} E_{coop} = S_T \cdot \sum_{k=1}^{K} \sum_{j=0,j\neq k}^{K} \sum_{h=1}^{H} \frac{\alpha_{kj}^h \cdot |D_k| \cdot (P_{Tx,kj} + P_{Rx,kj})}{R_{kj}} $$

subject to

$$ \alpha_{k0}^h = 0 \text{ for } h < H \text{ and } k = 1, ..., K $$

$$ \sum_{j=0}^{K} \sum_{h=1}^{H} \alpha_{kj}^h = 1 \text{ for } k = 1, ..., K $$

$$ \alpha_{kj}^h \in \{0, 1\} \forall k, j, h $$
The first constraint (8) indicates that transmissions to the BS take place at the last hop \( h = H \) only. The second constraint (9) indicates that each SN should transmit its collected data exactly once to a single destination on one of the \( H \) hops (hop \( H \) on the LR and \( H - 1 \) hops on the SR). Finally, constraint (10) specifies that the optimization variable \( \alpha_{h}^{kj} \) is a binary variable.

In the problem formulated in (7), the maximum number of hops can be specified as a parameter. Setting \( H = K \) allows full multihop communications, although the actual hops might be less than \( K \), and in this case the parameters \( \alpha_{h}^{kj} \) corresponding to the unnecessary hops will be set to zero in the optimal solution. Setting \( H = 2 \) corresponds to reducing the problem into a clustering problem where SNs are grouped into clusters. In each cluster, an SN selected as cluster head (CH) in the optimal solution sends the data on the LR to the BS after aggregating the data it receives on the SR from the SNs in its cluster. Furthermore, setting \( H = 1 \) corresponds to the non-cooperative approach where all SNs send the data on the LR to the BS. In this case, the energy is denoted by \( E_{\text{No-coop}} \). The normalized energy consumption \( \eta \) can be calculated as follows:

\[
\eta = \frac{E_{\text{coop}}}{E_{\text{No-coop}}}
\]

(11)

The value of \( \eta \) indicates whether the cooperation is beneficial in terms of energy consumption or not; if \( \eta < 1 \), then the cooperation results in a gain of energy consumption while \( \eta > 1 \) reflects a non-beneficial cooperation.

The formulation in (7) is applicable to any number of hops, allows communication using different wireless interfaces (different values of \( P_{\text{Rx},kj} \) and \( P_{\text{Tx},kj} \) can be set for each wireless link between any two nodes \( k \) and \( j \)), and permits any combination of power adaptive/rate adaptive transmissions. For example, a node may be transmitting to its parent in the DAG using rate adaptive transmission while another can be using power adaptive transmission. The values of the parameters in the different implementation scenarios are detailed in Table 1. Using, for each node in the network, the appropriate parameters from Table 1 according to its communication scheme adopted, then the formulation (7) can be customized to a huge variety of node combinations and hybrid wireless interfaces.

The problem formulated in (7) appears as a binary integer program that can be solved using known software solvers. However, this is not the case due to the dependence of \(|D_k|\) on the parameters \( \alpha_{h}^{kj} \), which makes the problem intractable. In addition, even when the problem can be considered as a binary integer program, the complexity of finding the optimal solution of the problem (7) using software solvers increases tremendously when the number of nodes increases and is not suitable for real time implementation. In fact, binary integer programming is known to be NP-hard. In the next section, low complexity suboptimal schemes are presented that are able to achieve efficient multihop routing of sensor data with significant energy savings compared to the non-cooperative approach.
5. Suboptimal energy-efficient WSN data routing methods

In this section, we present algorithms that perform energy efficient routing of sensor data. Section 5.1 presents a multihop approach whereas Section 5.2 presents a clustering-based approach. Section 5.3 presents a complexity analysis that applies to both methods.

5.1. Suboptimal multihop approach

In this section, we present an algorithm that performs energy efficient multihop routing of sensor data. Starting with the SNs having worst channel conditions on the LR (and hence worst achievable rates and highest energy consumption), we find for each SN \( k \) the parent \( p_k \) to which it can send the data with the minimum energy consumption. When the turn comes to SN \( p_k \), a parent \( p_{p_k} \) is found to which \( p_k \) can send the data with the minimum energy consumption, thus leading to an additional hop if \( p_{p_k} \neq 0 \). The details of the proposed approach are presented below:

- **Step 1**: Sort the SNs in decreasing order of energy consumption without cooperation. After this step, SN \( k = 1 \) would be the one having the worst channel conditions on the LR and SN \( j = K \) would be the one having the best channel conditions on the LR.

- **Step 2**: Start from SN \( k = 1 \).

- **Step 3**: For SN \( k \), find the parent node (could be another SN or the BS) \( p_k \) to which \( k \) can forward the data with the least energy consumption. The search is done over the nodes \( j \) having better LR channel conditions than \( k \), i.e., such that \( j > k \). Energy consumption to distribute the content includes the energy of \( k \) to transmit and the energy of \( p_k \) to receive. i.e.:

\[
p_k = \arg \min_{j; j > k} |D_k|S_T \cdot \frac{(P_{Tx,kj} + P_{Rx,kj})}{R_{kj}} \tag{12}
\]

- **Step 4**: break the connection of \( k \) with the BS and set \( p_k \) as the direct parent of \( k \) if \( \frac{(P_{Tx,pp_k} + P_{Rx,pp_k})}{R_{pp_k}} < \frac{P_{Tx,0} + P_{Rx,0}}{R_{0}} \), i.e., if it is more energy efficient for \( k \) to send the data to \( p_k \) rather than sending it directly to the BS. Then update \( D_{p_k} \) as: \( D_{p_k} = D_{p_k} \cup D_k \).

- **Step 5**: increment \( k \) and repeat Steps 3-5 on the SNs whose order is \( > k \) in the sorted list.

- **Step 6**: After all the SNs have been assigned to their direct parent based on the most energy efficient path, we check if SN \( K \) can send the data with lower energy than sending it directly on the LR link, since it is still connected to the BS (due to sorting the SNs in decreasing order of LR energy consumption). Hence, if there exists an SN \( x \neq K \) such that \( p_x = 0 \) (i.e. there is another path to the BS that does not go through SN \( K \)), which means that the link between the BS and SN \( K \) can be broken while still being able to send the data from the SNs to the BS), then for all SNs \( j < K \) such that \( p_j \neq K \), the parent of SN \( K \) is selected such that:

\[
p_K = \arg \min_{j; j \neq K} |D_K|S_T \cdot \frac{(P_{Tx,kj} + P_{Rx,kj})}{R_{kj}} \tag{13}
\]

- **Step 7**: We set \( p_K = 0 \), i.e., the best destination for SN \( K \) to send the data to is the BS. If \( p_K \) is set as the parent of \( K \), then update \( D_{p_K} \) as: \( D_{p_K} = D_{p_k} \cup D_k \).
The algorithm presented in this section does not impose a limit on the number of hops. The outcome could be any number $H$ such that $1 \leq H \leq K$, where $H = 1$ indicates that all SNs send their data directly to the BS. This corresponds to a scenario where SNs are scattered to an extent such that collaboration is not energy efficient, and the best for each SN is to send the data directly to the BS. In the next section, we present a similar algorithm that performs node clustering ($H = 2$).

### 5.2. Suboptimal clustering approach

In this section, we present an algorithm that performs energy efficient clustering for sensor data transmission. The algorithm performs a grouping of SNs into cooperating clusters, with each cluster having an SN, the cluster head (CH), receiving the data from the SNs within its cluster and forwarding it to the BS, along with its own measurements. The algorithm could lead to situations where one or more clusters contain a single SN. In this case, that SN is the cluster head and sends its data on the LR without receiving from other SNs on the SR. This corresponds to a situation where other SNs are too far or the links with them are under severe fading, such that collaboration is not energy efficient, and the best solution for that SN is to send the data directly to the BS.

Starting with the SNs having worst channel conditions on the LR (and hence worst achievable rates and highest energy consumption), we find for each SN $k$ the parent $p_k$ to which it can send the data with the minimum energy consumption. If $k$ is a cluster head, all members of $D_k$ are moved to $D_{p_k}$ if the data transmission form $k$ and all the members of $D_{p_k}$ to $p_k$ is more energy efficient than having an independent cluster with $k$ as cluster head. It should be noted that in the special case of clustering, we have $D_k = k \cup C_k$. The details of the proposed approach are presented below:

- **Step 1:** Sort the SNs in decreasing order of energy consumption without cooperation. After this step, SN $k = 1$ would be the one having the worst channel conditions on the LR and SN $k = K$ would be the one having the best channel conditions on the LR.

- **Step 2:** Start from SN $k = 1$.

- **Step 3:** For SN $k$, find the parent node (could be another SN or the BS) $p_k$ to which $k$ and all the members of $D_k$ (if there are any SNs other than $k$) can send their data with the least energy consumption. Energy consumption to distribute the content includes the energy of $p_k$ to receive and the transmission energy of the SNs in $D_k$, i.e.:

$$p_k = \arg \min_{j,j>k} S_j \cdot \sum_{i \in D_k} \frac{P_{Tx,ij} + P_{Rx,ij}}{R_{ij}}$$

- **Step 4:** break the connection of $k$ with the BS, and the connection of all other members of $D_k$ with $k$, and set $p_k$ as the direct parent of $k$ and all other SNs in $D_k$ if

$$\sum_{i \in D_k} \frac{(P_{Tx,ip_k} + P_{Rx,ip_k})}{R_{ip_k}} < \frac{P_{Tx,k0}}{R_{k0}} + \sum_{i \in D_k, i \neq k} \frac{(P_{Tx,ik} + P_{Rx,ik})}{R_{ik}}$$

i.e., move all members of $D_k$ to $C_{p_k}$ if this is more energy efficient than having an independent cluster with $k$ as cluster head sending the data to the BS: $C_{p_k} = C_{p_k} \cup D_k = C_{p_k} \cup k \cup C_k$. 


• **Step 5:** increment \( k \) and repeat Steps 3-5 on the SNs whose order is \( > k \) in the sorted list.

• **Step 6:** After all the SNs have been grouped into clusters based on the most energy efficient method, we check if SN \( K \) can send the data with lower energy than sending on the LR link, since it is still connected to the BS (due to sorting the SNs in decreasing order of LR energy consumption). Hence, if there exists an SN \( x \neq K \) such that \( p_x = 0 \) (i.e. there is another cluster with cluster head other than SN \( K \), which means that the link between the BS and SN \( K \) can be broken while still being able to send the data from the SNs to the BS), then for all SNs \( j < K \) such that \( p_j = 0 \), the parent of SN \( K \) is selected such that:

\[
p_K = \arg \min_{j, p_j = 0} S_T \cdot \sum_{i \in D_K} \frac{(P_{Tx,i,j} + P_{Rx,i,j})}{R_{ij}}
\]

• **Step 7:** We set \( p_K \) as the direct parent of \( K \) if

\[
\sum_{i \in D_K} \frac{(P_{Tx,i,p_K} + P_{Rx,i,p_K})}{R_{ip_K}} < \frac{P_{Tx,K0}}{R_{K0}} + \sum_{i \in D_K, i \neq K} \frac{(P_{Tx,i,K} + P_{Rx,i,K})}{R_{iK}}
\]

Otherwise, we keep \( p_K = 0 \), i.e., SN \( K \) is a cluster head sending the data to the BS. If \( p_K \) is set as the parent of \( K \), then we update \( C_{p_K} \) as: \( C_{p_K} = C_{p_K} \cup D_K = C_{p_K} \cup K \cup C_K \).

### 5.3. Complexity analysis

This section presents a complexity analysis that applies to both methods of Sections 5.1 and 5.2. Step 1 of the algorithms is a sorting step, and hence has a worst-case complexity \( O(K^2) \). In Step 3, the search involves \( K \) nodes when \( j = 1 \), it involves \( (K - 1) \) nodes when \( j = 2 \), etc., and 2 nodes when \( j = (K - 1) \). Hence, the complexity of Steps 2 to 5 is: \( K + (K - 1) + \cdots + 2 = \frac{k(k+1)}{2} - 1 \). In Steps 6-7, the search involves at most \( K \) nodes. Consequently, the worst-case complexity of the algorithms is: \( K^2 + \frac{k(k+1)}{2} - 1 + K = \frac{3K^2}{2} + \frac{3K}{2} - 1 \). This is a quadratic complexity of order \( O(K^2) \). Hence, the proposed suboptimal methods are significantly easier to implement than the optimal solution of the NP-hard problem of Section 4.

In the next section, we compare the methods of Sections 5.1 and 5.2 to each other and to the non-cooperative approach.

### 6. Results and discussion

In this section, simulation results are presented and analyzed. The simulation parameters are presented in Table 2. Channel parameters are obtained from [1], whereas energy consumption parameters are taken as in [18], where measurements are made with 3G communications on the LR, and 802.11 b on the SR using the rate adaptive approach.

In Sections 6.1 to 6.3, we investigate a scenario corresponding to multihop data transmission in a WSN. We consider that each sensor sends its measurement data in a file of size \( S_T = 1 \) Mbits, to be routed to the BS in an energy efficient manner. Two main SN deployment scenarios are investigated:
Table 2. Simulation Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa$</td>
<td>-128.1 dB</td>
</tr>
<tr>
<td>$\upsilon$</td>
<td>3.76</td>
</tr>
<tr>
<td>$\sigma_{\xi}^2$ (dB)</td>
<td>8 dB</td>
</tr>
<tr>
<td>$P_{TX}$</td>
<td>1.425 Joules/s</td>
</tr>
<tr>
<td>$P_{B,Rx}$</td>
<td>0.925 Joules/s</td>
</tr>
<tr>
<td>$P_{L,Rx}$</td>
<td>1.8 Joules/s</td>
</tr>
</tbody>
</table>

- In the first deployment scenario, SNs are assumed to be uniformly distributed in a rectangular area of size 200m x 200m, whose origin is at a distance $d_{LR}$ m from the BS. Different values of $d_{LR}$ are investigated in the simulations. This scenario corresponds, for example, to a WSN monitoring air pollution in a particular area of interest, e.g., near a power plant, or an area where a high density of lung disease was detected.

- In the second scenario, the SNs are assumed to be uniformly distributed throughout the whole cell. We consider a single BS placed at the center of a 1 x 1km cell. This scenario corresponds to a case where the whole cell needs to be monitored by the WSN, not a particular or specific area. This scenario will be referred to by “BS at center of 1 x 1km cell” in the figures.

Results are averaged over 50 iterations. In each iteration, new random SN locations are determined and 50 fading realizations are considered (thus results are averaged over $50 \times 50 = 2500$ fading realizations). We compare the methods of Sections 5.1 (denoted as “multihop” in the results) and 5.2 (denoted as “clustering” in the results) to the non-cooperative approach.

6.1. Example on the gap between the optimal and suboptimal methods

<table>
<thead>
<tr>
<th>$d_{LR}$ (m)</th>
<th>300</th>
<th>500</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal</td>
<td>0.6761</td>
<td>1.2015</td>
<td>5.0010</td>
</tr>
<tr>
<td>Proposed Multihop</td>
<td>0.7342</td>
<td>1.2974</td>
<td>5.1023</td>
</tr>
<tr>
<td>Proposed Clustering</td>
<td>0.7423</td>
<td>1.3255</td>
<td>5.1455</td>
</tr>
<tr>
<td>No Cooperation</td>
<td>1.6185</td>
<td>5.3847</td>
<td>45.0133</td>
</tr>
</tbody>
</table>

Table 3. Energy (in Joules) Results for $K = 3$

In this section, the proposed methods of Section 5 are compared to the optimal multihop solution of Section 4 (with $H = K$) for a low number of SNs (in order for the optimal solution to be tractable). Selecting $K = 3$, all the possible cases are shown in Fig. 3. Hence, the optimal solution will be one of the 16 cases presented in Fig. 3, depending on the fading conditions. The results obtained after implementing the optimal solution and the proposed methods are listed in Table 3. It can be clearly seen that the gap between the suboptimal multihop and clustering results from the optimal solution is very small. In addition, Table 3 shows that the cooperative techniques lead to huge savings compared to the non-collaborative scenario.

Fig. 4 shows, for each of the 16 cases, the percentage of times that this case occurs as the optimal solution. When the distance to the BS is small, Case 1 (no collaboration) seems to be optimal for a significant percentage of the time. However, this percentage decreases as the
distance to the BS increases. Cases 2-7 form a group of similar cases where the only variation is a permutation of the SNs involved in the connections. As expected, these cases have almost equal probability of being the optimal case for a given value of $d_{LR}$. The same reasoning applies for Cases 8-13 and Cases 14-16. Interestingly, Cases 8-13 were never optimal in the obtained results.

In fact, with Cases 2-7, and considering Case 2 as an example, SN A transmits $S_T$ bits on the LR, SN C transmits $S_T$ bits on the SR, and SN B transmits $2S_T$ bits (its own data in addition to the data of SN C) on the LR. With Cases 14-16, and considering Case 14 as an example, SN B transmits $S_T$ bits on the SR, SN C transmits $S_T$ bits on the SR, and SN A transmits $3S_T$ (its own data in addition to the data of SNs B and C) bits on the LR. In Cases 8-13, and considering Case 8 as an example, SN C transmits $S_T$ bits on the SR, SN B transmits $2S_T$ bits (its own data in addition to the data of SN C) on the SR, and SN A transmits $3S_T$ (its own data in addition to the data of SNs B and C) bits on the LR. Since the SNs are deployed in a confined area of interest, and since SR transmissions in this case can occur at high rates due to the relative proximity of SNs, Cases 14-16 would generally lead to lower energy consumption than Cases 8-13, since both groups have the same LR energy consumption (due to transmitting $3S_T$ on the LR by one SN), but on the SR each of the other two SNs transmits $S_T$ with Cases 14-16. However, with
Figure 4. Percentage of having each of the 16 possible cases as the optimal solution when $K = 3$.

Cases 8-13, SR energy consumption is higher because one SN transmits $S_T$ while the other transmits $2S_T$ on the SR.

Fig. 4 shows that as $d_{LR}$ increases, Cases 14-16 become more favored than Cases 1-7. In fact, a large distance to the BS leads to spending most of the energy during LR transmission, since the achievable rates become significantly lower due to the increased distance. Thus, one LR transmission with an SN having favorable LR channel conditions in Cases 14-16 would be more energy efficient than two LR transmissions with Cases 2-7 (or three LR transmissions with Case 1).

6.2. Energy results

This section presents the energy savings achieved by using the proposed multihop and clustering methods, compared to the non-cooperative scenario. In the non-cooperative approach, each SN sends the data on the LR to the BS without any collaboration with other SNs on the SR. Fig. 5 shows the normalized energy results for the various investigated scenarios. Significant energy savings are achieved compared to the non-cooperative scenario, regardless of the number of hops allowed. In fact, the clustering approach corresponding to $H = 2$ and the multihop approach with $H = K$, thus representing the two extreme cases, have a very comparable performance in terms of normalized energy. Fig. 5 shows that the gains are reduced as the distance to the BS decreases. This is due to a reduction in the energy needed on the LR without cooperation and not to an increase in energy consumption with the proposed approach, since the LR distance was reduced. This leads to an increase in the ratio $\eta$.

In fact, the results of Fig. 6, presenting the energy consumption results without normalization, show that the energy is reduced when the distance to the BS is reduced, as expected. The results of the energy consumption in the non-cooperative scenario are shown in Fig. 6 for reference. Values for $d_{LR} = 1000$ m are not shown, since they are around an order of magnitude larger than the cooperative results, which makes all the plots of the various
Figure 5. Normalized energy consumption vs. the number of SNs for different values of $d_{LR}$.

Figure 6. Energy consumption vs. the number of SNs for different values of $d_{LR}$.
cooperative scenarios appear to overlap. Thus, the combination of Figs. 5 and 6 allows to display both the gains of cooperation compared to the non-cooperative scenario and to understand the variation of the energy gains with the distance to the BS.

6.3. Delay results

![Figure 7. Average delay per SN vs. the number of SNs for different values of $d_{LR}$.](image)

In this section, the impact of multihop-based energy minimization on delay performance is investigated. The transmitter at each hop is considered to wait until it receives all the data from the previous hop before starting transmission. In addition, at each hop, it is considered that transmission is done in parallel using orthogonal channels within the same cluster or within clusters at close proximity. The channels can be reused at clusters located further away. This corresponds, in practice, to the use of OFDMA with different subchannels allocated to each transmitter-receiver link, or to the use of CDMA with different orthogonal codes allocated to each transmitter-receiver link.

Fig. 7 shows the delay results averaged over the SNs. However, in delay sensitive applications, the interest is in the delay incurred by each SN. Therefore, Fig. 8 shows the maximum delay, i.e., the delay incurred by the last SN to send its data to the BS. In other words, this corresponds to the total delay needed to transmit the measurements of all SNs in the network, thus corresponding to the worst case result. Figs. 7 and 8 show that the delay increases with the distance to the BS, since a longer distance leads to lower achievable rates on the LR,
which leads to an increase in data transmission time. In addition, the clustering approach outperforms the multihop approach by leading to shorter delays in all the investigated scenarios. Fig. 7 shows that when SNs are deployed in a confined area at a distance $d_{LR}$ from the BS, the multihop approach leads to average delays comparable to the non-cooperative scenario when $d_{LR} = 300m$, and to better average delay performance when $d_{LR}$ increase to 500m. The trend continues with larger distances. When the BS is placed at the cell center, with the SNs deployed throughout the cell area, the non-cooperative scenario leads to better average delay than the multihop approach, but not than the clustering approach.

Fig. 8 shows that the proposed cooperative methods significantly outperform the non-cooperative case by leading to shorter maximum delay. Particularly, the clustering method leads to considerably shorter maximum delay compared to both the multihop approach and the non-collaborative scenario.

Thus, the suboptimal clustering approach leads to significant energy savings that are comparable to the multihop approach as shown in Figs. 5 and 6, and it leads to much shorter delays in transmitting the measurement data as shown in Figs. 7 and 8, and thus constitutes a suitable approach leading to both energy and delay efficiency in WSNs.

### 6.4. Bandwidth savings

<table>
<thead>
<tr>
<th>$d_{LR}$ (m)</th>
<th>300</th>
<th>500</th>
<th>1000</th>
<th>Centered BS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of clusters</td>
<td>27</td>
<td>16</td>
<td>8</td>
<td>35</td>
</tr>
</tbody>
</table>

**Table 4.** Number of Collaborative Clusters for $K = 50$
In this section, traffic offloading from the BS due to using the proposed approach is investigated. Table 4 shows the number of SNs transmitting the data directly to the BS on the LR, in the case where a network of $K = 50$ SNs is deployed. This corresponds to the number of wireless channels needed on the LR. It should be noted that in the example of Table 4, the multihop and clustering approach lead to the same number of clusters, although the transmission occurs on different routes inside each cluster. From Table 4 it can be seen that a significant portion of the LR bandwidth can be freed due to implementing the proposed approach. In fact, around 46%, 68%, and 84% of the bandwidth can be saved when $d_{LR} = 300$, 500, and 1000 meters, respectively. In addition, when the WSN is deployed throughout the cell area with the BS at the cell center, 30% of the LR bandwidth can be saved. When the proposed approach is implemented network wide, the significantly reduced loads of some BSs might be accommodated by other more loaded BSs. The initial BSs would be switched-off in this case. Hence, the proposed approach would contribute to green communications at the BS level, although its initial purpose was to save battery energy of SNs.

7. Practical implementation aspects

In this section, we discuss some practical limitations of the proposed techniques and propose methods to overcome these limitations.

7.1. CSI Exchange for algorithm implementation

In the proposed methods, the BS is assumed to be aware of the channel state information (CSI), and hence of the achievable rates $R_{k0}$ on the LR links in addition to the CSI and rates $R_{kj}$ ($j > 0$) on the SR links. Since the sensors considered are not assumed mobile, this can be achieved by a training phase that precedes the actual data transmission phase. The BS can know the CSI on the LR via feedback from the SNs, which is common in state-of-the-art wireless communication systems. On the SR, SNs can take turns in broadcasting pilot signals. Thus, each SN can estimate its CSI, and hence the rate $R_{kj}$, with every other SN within its transmission range, by measuring the received strength of the pilot signals. The SR pilot broadcasting process can be coordinated by the BS to avoid collisions. When each SN gets a CSI estimate on its SR links with the other SNs, it can feed-back this information to the BS on the LR link. After this training phase, the BS can then coordinate the data transmission process using the proposed methods. The same analysis applies in a limited mobility scenario, without necessarily having the sensors fixed. Hence, in the case of fixed SNs or in a low mobility scenario (portable SNs), the overhead due to the training phase can be considered low since a long time can elapse before the channel conditions change and the need arises to repeat the process.

In addition, it should be noted that SNs form cooperative clusters with other SNs when they can successfully hear their pilot transmission, i.e., when $R_{kj}$ is high enough to allow efficient communication between SNs. When $R_{kj}$ is too low between two SNs $k$ and $j$, these will automatically be in different clusters. Thus, if no CSI feedback is received about the link between SNs $k$ and $j$, then there will not be a possibility for direct communication between these SNs in the multihop approach of Section 5.1. Furthermore, in the clustering approach of Section 5.2, SN $k$ cannot be a cluster head in a cluster of which $j$ is a member and vice versa. This leads to eliminating several candidates in the search conducted in the schemes of Sections 5.1 and 5.2, and hence to a significant reduction in the complexity of the
algorithms. Consequently, the results of Section 5.3 correspond to a worst-case scenario, and the complexity in practical scenarios is generally lower.

7.2. Fairness considerations

The multihop and clustering methods are based on selecting certain nodes that transmit the data of other SNs in addition to their own data. This could lead to an increase in energy consumption for some of these nodes compared to the non-cooperative scenario, although the overall energy consumption in the network is minimized. In [29], it was shown that, within a single cluster, fading variations lead to selecting a different cluster head for each fading realization, and this was shown to lead to fairness in energy consumption in the cluster on the long term. Thus, in the case of WSNs deployed for long term measurement and monitoring of certain parameters, different training phases (as explained in Section 7.1), will occur. Consequently, the techniques presented in this Chapter can be considered to be fair. In fact, different SNs will take turn to relay the SR data when the fading varies, which averages out the energy consumption levels among SNs.

8. Application example - air quality monitoring

The methods presented in this Chapter can be applied to several WSN deployment scenarios. An important application of WSNs is the monitoring of environmental parameters. With the advancements in the production of small, accurate, low power sensors, it is becoming more and more possible to deploy a WSN for continuous monitoring of air quality. The WSN would report the concentration of several pollutants in the atmosphere, and the reported measurements can be made available to the general public via dedicated websites, mobile applications, etc. In addition, the stored measurements can be made available to expert environmental scientists to analyze and assess pollution information in order to submit recommendations to the relevant authorities in order to take appropriate action.

In this section, we present a high level description of the system architecture for air pollution monitoring and describe the role of the SNs where the presented communication protocol will be applied. The system model for air pollution monitoring is displayed in Fig. 9. Each BS covers a cell of certain area, where several SNs are deployed to monitor environmental parameters. The architecture follows a three-tier approach:

1. The sensor nodes (SNs): these include the sensors, measuring pollutants to be monitored, e.g., CO, NOx, Ozone, and Particulate Matter (PM), in addition to other environmental parameters like relative humidity and temperature. An SN usually can accommodate one or more sensors, with each sensor measuring one of the mentioned parameters. The SNs transmit the measured data using the presented communication methods. Thus, the nodes can form cooperative clusters, and relay the data in a multihop fashion ensuring energy efficiency.

2. The database server: the data received at the BS is sent to a database server where it is stored using a common format in order to automate its extraction and analysis. The measured data might contain missing, noisy, or erroneous values. Appropriate data integrity checks should be performed before storing the data for subsequent use. Afterwards, the data becomes ready for analysis and display. Analysis techniques include statistics (for computation of daily, monthly, or yearly averages of a certain air pollutant),
Figure 9. Implementation scenario for air pollution monitoring.

advanced interpolation, neural networks, principal component analysis, and data mining techniques.

3. The Client tier: it consists of client-side applications running on computers or mobile devices, e.g. smart phones. These applications access the network via the server, which forwards the stored data received from the sensors. Examples of applications include periodically updated web sites with data summaries and statistics, data visualization with display of sensor locations on a map (along with each SN’s measurements), and data dissemination applications like SMS alerts relating to pollution levels in certain areas.
9. Future work

After describing the previous contributions in the literature and outlining the differences with the presented approach in Section 2, the problem was defined and formulated in Sections 3 and 4, then the novel proposed method to address the formulated problem was presented in Section 5, and its efficiency was demonstrated in Section 6. Hence, the role of this section is to introduce some interesting future research directions.

In addition to a more thorough and detailed investigation of the topics described in Sections 7 and 8, future work would consist of implementing the proposed methods in a sensor network testbed and of matching the simulation results with actual energy measurements. Another interesting research direction is to consider SNs with variable power sources, and to distinguish between battery powered SNs and SNs having access to renewable energy sources (e.g., solar powered) or mains powered. The problem can be reformulated by imposing a constraint that the latter SNs should be cluster heads since they can transmit large amounts of aggregated data on the LR without suffering from energy shortage.

10. Conclusions

Cooperative data transmission in wireless sensor networks was studied with the objective of energy minimization. The problem was formulated into an optimization problem, and efficient suboptimal methods were presented for the two scenarios: the multihop case where the maximum number of hops is allowed and the clustering case where sensors are grouped into cooperating clusters, each headed by a cluster head in charge of the communication with the base station. The two methods were shown to lead to significant energy savings compared to the non-cooperative scenario, with the clustering approach leading to better delay results than the multihop approach. Practical implementation aspects were also discussed.

Acknowledgment

This work was made possible by an NPRP grant from the Qatar National Research Fund (a member of The Qatar Foundation). The statements made herein are solely the responsibility of the authors.

Author details

Elias Yaacoub and Adnan Abu-Dayya
QU Wireless Innovations Center (QUWIC), Doha, Qatar

11. References


