Research Article

International Journal of
Distributed
Sensor Networks

International Journal of Distributed
Sensor Networks
2017, Vol. 13(1)

DOI: 10.1177/1550147716689682
journals.sagepub.com/home/ijdsn
SAGE

Increasing network lifetime using data compression in wireless sensor networks with energy harvesting

Sunyong Kim[1], Chiwoo Cho[1], Kyung-Joon Park[2] and Hyuk Lim[1]

Abstract

In wireless sensor networks powered by battery-limited energy harvesting, sensor nodes that have relatively more energy can help other sensor nodes reduce their energy consumption by compressing the sensing data packets in order to consequently extend the network lifetime. In this article, we consider a data compression technique that can shorten the data packet itself to reduce the energies consumed for packet transmission and reception and to eventually increase the entire network lifetime. First, we present an energy consumption model, in which the energy consumption at each sensor node is derived. We then propose a data compression algorithm that determines the compression level at each sensor node to decrease the total energy consumption depending on the average energy level of neighboring sensor nodes while maximizing the lifetime of multihop wireless sensor networks with energy harvesting. Numerical simulations show that the proposed algorithm achieves a reduced average energy consumption while extending the entire network lifetime.

Keywords

Energy efficiency, wireless sensor networks, data compression, network lifetime, energy consumption model

Date received: 28 July 2016; accepted: 25 December 2016

Academic Editor: Jemal Abawajy

Introduction

Wireless sensor networks (WSNs) typically consist of numerous sensor nodes, which accomplish their task over a specific field. Each sensor node is capable of sensing, processing, and communication using its own sensor, processor, and wireless transceiver. It usually monitors some surrounding environmental phenomena, collects the sensing data, and forwards the data toward a designated sink node, which is responsible for data storage and analysis. These WSNs have been widely deployed in real environments[1] in order to provide information gathering to users for better understanding of the environment.

In many applications of WSNs, a large number of sensor nodes with a limited battery capacity should be deployed at remote places that are difficult to access. Due to the limited battery capacity, sensor nodes should be energy-efficient to extend their lifetime. Many studies have focused on achieving the maximum utilization with limited energy. In El Gamal and colleagues,[2,3] energy-efficient scheduling algorithms have been proposed and studied, in which the energy for wireless data transmission is minimized by varying the packet transmission time. Pal et al.[4] have proposed a balanced cluster-size clustering algorithm in order to

[1]School of Information and Communications, Gwangju Institute of Science and Technology (GIST), Gwangju, Republic of Korea
[2]Department of Information & Communication Engineering, Daegu Gyeongbuk Institute of Science and Technology (DGIST), Daegu, Republic of Korea

Corresponding author:
Hyuk Lim, School of Information and Communications, Gwangju Institute of Science and Technology (GIST), Gwangju 500-712, Republic of Korea.
Email: hlim@gist.ac.kr

extend the lifetime of WSNs. The proposed algorithm has improved the cluster quality while providing a lower death rate of the sensor nodes. In Sankarasubramaniam et al.,[5] a fixed frame-size optimization for energy efficiency has been proposed. A study of the feasibility of forward error correction in WSNs has also been carried out. In He,[6] many multipath routing algorithms have been introduced to guarantee the network latency and energy efficiency. Oh and Chae[7] have suggested grid-based solutions, however, which are based on unrealistic assumptions. As a promising technology for overcoming the limited battery capacity, energy harvesting in WSNs has also been studied.[8–10] By gathering energy from the environment to extend the lifetime of the sensor nodes, the network connectivity can be sustained longer. Gunduz et al.[9] have provided mathematical tools and analytical models for designing reliable communication systems with energy harvesting. In particular, energy harvesting from a radio-frequency (RF) signal in a relay network is considered in Nasir et al.[10] Given the structure of a sensor-relay-destination node, a relay node is capable of harvesting energy from the RF signal transmitted by the sensor node, and the harvested energy is used to forward data to the destination node.

A sensor node usually performs three operations, which are sensing, processing, and data communication. It was reported in Anastasi et al.[11] and Medeiros et al.[12] that among the three operations, the data communication is the most energy-consuming operation in most cases. For this reason, many researchers have studied to reduce the energy consumption for data communication in WSNs mainly by two approaches: duty-cycling and in-network processing. In the schemes based on duty-cycling approach, the energy is saved by coordinating the schedules of wakeup/sleep time at sensor nodes. Meanwhile, in-network processing-based schemes, such as data aggregation and data compression, mainly attempt to reduce the amount of data to be transmitted.

As one effective way to reduce the amount of data to be transmitted, data compression has been actively researched.[12–19] Using an appropriate compression method for sensing data, sensor nodes can transmit and receive smaller-size data packets. As a result, the energy consumption for data transmission at sensor nodes can be significantly reduced. In Kimura and Latifi,[20] Srisooksai et al.,[21] and Razzaque et al.,[22] a detailed comparative survey is provided on various data compression approaches for energy efficiency in WSNs. Kimura and Latifi[20] presented four types of feasible data compression schemes that are used for WSN nodes with limited resources. Srisooksai et al.[21] and Razzaque et al.[22] provided a comprehensive review of data compression algorithms for WSNs, classified the compression algorithms into several categories such as distributed source modeling, transform coding, source coding, and compressive sensing, and compared them with respect to various performance metrics such as the compression performance and amount of power consumption.

As one of the applications for data compression, we consider multihop WSNs, in which packets generated from a source node are relayed by intermediate nodes and travel toward the destination node along a multihop wireless path. The nodes along the multihop path are able to perform data compression in order to reduce their energy consumption for transmission. As the packet is relayed toward the destination node from the source node, the length of the packet becomes smaller by data compression, which leads to a gradual decrease in the energy consumption.

In this article, we consider the energy savings using data compression in order to extend the lifetime of multihop WSNs. A sensor node transmits and receives a smaller-size packet using data compression so that it can reduce its energy consumption. We also consider battery-operated sensor nodes, which are rechargeable by energy harvesting. Depending on the environment, for instance, the amount of sunlight, the energy harvested at each node can be different as time goes by, resulting in an unevenness in the energy levels among the sensor nodes. In order to decrease the average energy consumption of sensor nodes and extend the lifetime of the networks, we propose a data compression algorithm considering the average energy level of the sensor nodes within the next m hops. The compression level for each sensor node is determined such that the difference between the magnitude of its own energy level and the average energy level of the sensor nodes in next m hops is reduced. Numerical results from a simulation of the proposed algorithm verify that the network lifetime is significantly increased while reducing the average energy consumption.

The remainder of this article is organized as follows. An overview of related works is presented in section "Related works." In section "Energy consumption model," we describe the energy consumption model and formulate the amount of energy consumed at each node. We then propose an algorithm to extend the network lifetime using data compression in section "Proposed decision scheme for compression." The simulation results are presented in section "Performance evaluation," and section "Conclusion" concludes this article.

Related works

There has been a significant amount of work on the data compression for energy conservation in WSNs. We briefly summarize the related work into two categories, standalone compression at a single node and cooperative compression among multiple nodes.

Standalone compression at a single node

The data compression schemes in this category focus on the compression process carried out at a single sensor node without exchanging energy information with other nodes.[12–16] The main purpose of these schemes is to achieve the increase in the compression ratio of data packets, that is, how much the data packet size can be reduced. Mercelloni and Vecchio[13,14] have proposed a simple lossless entropy compression (LEC) algorithm for temperature and relative humidity sensing data. LEC algorithm exploits the characteristics of high correlation between consecutive data measured by a sensor node. To compress the measured data, LEC algorithm first computes the differences of consecutive measured data and divides them into a small number of groups. These groups are then entropy encoded using a small, fixed dictionary table based on Huffman coding, resulting in the compression ratio of around 67%. Since LEC algorithm can be implemented using a small, fixed dictionary, it requires low memory, but the data resolution is very limited. Marcelloni and Vecchio[12] have proposed a more improved lossless compression method based on Huffman coding, called lightweight data compression. Compared to LEC that always uses the same dictionary, this method utilizes a reference dataset to generate a dictionary under measurement. With the modest computational and memory requirements, it achieves a higher compression ratio compared to LEC, which varies between 46% and 82%. Some lossy data compression methods have also been researched.[15,16] In contrast to lossless data compression, lossy data compression tolerates a certain level of inaccuracy induced by the data compression, but the significantly higher compression ratio can be provided. Capo-Chichi et al.[15] have proposed a lossy data compression algorithm called K-RLE,where K is a parameter of precision and RLE means run-length encoding. The main idea of K-RLE algorithm is that the consecutive sequence of the same bits b or bits in the range $[K - b, K + b]$ are compressed as a combination of the bit b and its count n. Simulation results show that K-RLE provides higher compression ratio than RLE, but it consumes more energy to compress and decompress. In Alsalaet and Ali,[16] a modified discrete cosine transform (MDCT)-based data compression method for vibration signals has been proposed. Since the vibration signals from machinery parts such as bearings and gears are almost stationary and deterministic, the authors have addressed that MDCT is the most suitable for compressing the vibration signals. To increase the compression ratio, the authors have suggested to encode MDCT coefficients using embedded harmonic coding (EHC).

Cooperative compression among multiple nodes

The data compression methods presented in this category focus on how sensor nodes cooperatively compress the data packets transmitted on the network by sharing the network information such as the position of each node or the spatial correlation between sensing data at different sensor nodes.[17–19] In Tavli et al.,[17] an optimal data compression scheme was formulated as an optimization problem that maximizes the minimum lifetime of the sensor nodes in order to increase the network lifetime. It is assumed that each node can compress and decompress raw data with multiple compression levels. The energy consumption was modeled by introducing two types of virtual nodes that perform data compression and decompressions. Because of computation complexity of solving the optimization, the authors proposed a heuristic approach, which enables the nodes farther away from the base station to always compress their data and those closer to the base station not to compress data. Incebacak et al.[18] have considered the data compression in a stealth mode of WSNs, where each sensor node has a different limited transmission power depending on its position in order not to be detected from adversaries. They proposed an optimization framework that jointly considers the network privacy preservation and the multi-level compression presented in Tavli et al.[17] in order to increase the network lifetime while retaining the network privacy. Five different compression strategies based on the multi-level compression have been compared to investigate the impacts of these strategies on the network lifetime in the stealth mode of WSNs. The data compression for delay-tolerant applications in WSNs has also been considered in Ali et al.[19] The authors have focused on both spatial and temporal correlations between the data collected by different sensor nodes over a long period of time. Accordingly, they proposed an adaptive hybrid compression (AHC) scheme that fuses both spatial and temporal compression in order to increase the compression ratio with a guaranteed data recovery accuracy.

Our proposed data compression algorithm falls into the second category because it utilizes the information from the other sensor nodes on the network to reduce the energy consumption and increase the network lifetime in WSNs. The proposed approach is distinct from existing methods in that it focuses on an energy-harvesting scenario of WSNs, in which the sensor nodes have significantly different amount of energy consumed and harvested at each node. In order to deal with this unevenness of energy, each sensor node gathers the usable energy levels of other sensor nodes and compare them with its own energy level when deciding how much it needs to compress data packets to support the other sensor nodes that are relatively lack of energy. Under the proposed method, all sensor nodes cooperate with other nodes to fairly consume their energy so that they are almost simultaneously exhausted and consequently the network lifetime is increased.

Energy consumption model

We consider multihop data transmission in a WSN, in which data packets are relayed by intermediate nodes, as shown in Figure 1. In our network topology, N nodes are randomly distributed in the network. Each node is battery powered with an energy-harvesting capability. Therefore, the energy level in the battery could be different among the sensor nodes in the same network because they are geographically distributed in an area and are exposed to different surrounding environments.

We derive an energy consumption model for compressed data transmission in the multihop WSN. Let L be the length of a data packet. Each data packet consists of b blocks, and the block size is $L_b = L/b$. For energy conservation, each data packet can be compressed and then transmitted in a shorter form. Note that the sensing data may include repeated and redundant information and can be compressed.

Each node can compress a certain number of blocks among all blocks before it transmits the packet. For example, if a source node compresses x blocks among the total number of b blocks, a relay node receives a packet with $(b - x)$ uncompressed blocks. Then, the relay node can further compress a certain number of blocks among the uncompressed blocks and transmits the packet to the next relay node. Therefore, the packet size becomes smaller as it is relayed toward the destination node. Depending on the energy levels of the nodes, the number of blocks compressed at each node could be different in order to extend the network lifetime.

Figure 2 shows the compression process of a data packet while the packet is relayed along a multihop path. The data packet is divided into smaller-sized blocks and can be compressed on a block-by-block basis. Each node performs data compression using a certain compression level and transmits the compressed data to its next node along a multihop routing path. For instance, in Figure 2, the source node compresses the first three blocks of the original data, and the remaining blocks are sequentially compressed by the following relay nodes. Finally, the nth node receives the reduced packet with the size of $l(n-1)$, which is compressed and transmitted by the $(n-1)$th node. As more

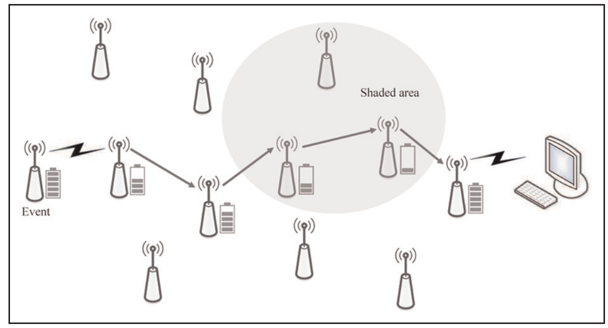

Figure 1. A multihop wireless sensor network.

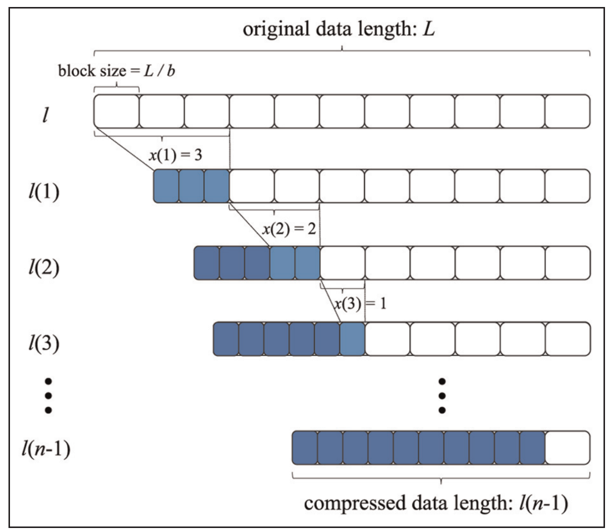

Figure 2. Variation in the data length with compression at each node.

blocks are compressed, the length of the data packet becomes smaller.

As the data packet begins to be compressed at the preceding nodes, the length $l(i)$ of the data packet that would be transmitted at the ith node becomes

$$\begin{aligned} l(i) &= L_b \cdot (1-\alpha) \cdot \sum_{j=1}^{i} x(j) + L_b \cdot (b - \sum_{j=1}^{i} x(j)) \\ &= L_b \cdot ((1-\alpha) \cdot \sum_{j=1}^{i} x(j) + b - \sum_{j=1}^{i} x(j)) \\ &= L_b \cdot (b - \alpha \cdot \sum_{j=1}^{i} x(j)) \end{aligned} \tag{1}$$

where $x(j)$ is the compression level meaning the number of blocks compressed by the jth node, and α is an average compression ratio. For example, if $\alpha = 0.1$, the length of a block is reduced by 10%, and it becomes $0.9L_b$. The first and second terms in equation (1) correspond to the sum of the lengths of compressed blocks and the sum of the lengths of uncompressed blocks, respectively. Depending on $x(j)$, the jth node and its following nodes may transmit a different-size packet. When a block is compressed, it is assumed to have a size of $(1-\alpha)\times L_b$ for simplicity. Note that $x(j)$ cannot be greater than the total number of blocks b, and the sum of $x(j)$ for all nodes should be less than or equal to b, (i.e. $0 \le x(j) \le b$ and $\sum_{j=1}^{n-1} x(j) \le b$).

We derive the energy consumption at each node in a multihop network. Let ϵ_t, ϵ_r and ϵ_c denote the energy consumed for one-bit transmission, reception, and compression, respectively, and let $\Delta e(i)$ represent the energy consumed at the ith node for delivering a data packet. By equation (1), $\Delta e(1)$ at the first node, which is the source node, is given by

$$\begin{aligned}\Delta e(1) &= L_b \cdot x(1) \cdot \epsilon_c + l(1) \cdot \epsilon_t \\ &= L_b \cdot x(1) \cdot \epsilon_c + L_b \cdot (b - \alpha \cdot x(1)) \cdot \epsilon_t\end{aligned} \tag{2}$$

The first and second terms in equation (2) are the energies consumed for compression and transmission at the first node, respectively. Note that since we focus on how the energy consumption for packet transmission, compression, and reception affects the network lifetime, the energy consumed for other functionalities such as sensing and routing is not included in equation (2). Intermediate nodes receive a packet, compress it, and transmit the reduced-size packet. The energy consumption at the ith relay node for $i \in \{2, \ldots, (n-1)\}$ can be represented by

$$\begin{aligned}\Delta e(i) &= l(i-1) \cdot \epsilon_r + L_b \cdot x(i) \cdot \epsilon_c + l(i) \cdot \epsilon_t \\ &= L_b \cdot (b - \alpha \sum_{j=1}^{i-1} x(j)) \cdot (\epsilon_t + \epsilon_r) + L_b \cdot x(i) \cdot (\epsilon_c - \alpha\epsilon_t)\end{aligned} \tag{3}$$

where each term represents the energies consumed for reception, compression, and transmission, respectively. The destination node receives the compressed packet and retrieves the original packet by decompressing it. The energy consumption at the nth node (destination node) is given by

$$\begin{aligned}\Delta e(n) &= l(n-1) \cdot \epsilon_r + L_b \cdot \sum_{j=1}^{n-1} x(j) \cdot \epsilon_d \\ &= L_b \cdot (b - \alpha \cdot \sum_{j=1}^{n-1} x(j)) \cdot \epsilon_r + L_b \cdot \sum_{j=1}^{n-1} x(j) \cdot \epsilon_d\end{aligned} \tag{4}$$

where ϵ_d is the energy consumed for decompression per bit of the decompressed packet

$$\begin{aligned}\begin{pmatrix}\Delta e(1)\\ \Delta e(2)\\ \Delta e(3)\\ \vdots\\ \Delta e(n-1)\\ \Delta e(n)\end{pmatrix} &= L\begin{pmatrix}\epsilon_t\\ \epsilon_r+\epsilon_t\\ \epsilon_r+\epsilon_t\\ \vdots\\ \epsilon_r+\epsilon_t\\ \epsilon_r\end{pmatrix} \\ &- L_b\left[\begin{pmatrix}\epsilon_t\alpha & 0 & \ldots & 0\\ (\epsilon_r+\epsilon_t)\alpha & \epsilon_t\alpha & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots\\ (\epsilon_r+\epsilon_t)\alpha & (\epsilon_r+\epsilon_t)\alpha & \ldots & \epsilon_t\alpha\\ \epsilon_r\alpha & \epsilon_r\alpha & \ldots & \epsilon_r\alpha\end{pmatrix}\right. \\ &\left. - \begin{pmatrix}\epsilon_c & 0 & \ldots & 0\\ 0 & \epsilon_c & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \ldots & \epsilon_c\\ \epsilon_d & \epsilon_d & \ldots & \epsilon_d\end{pmatrix}\right] \cdot \begin{pmatrix}x(1)\\ x(2)\\ \vdots\\ x(n-1)\end{pmatrix}\end{aligned} \tag{5}$$

Equations (2)–(4) can be represented in the matrix form given in equation (5). Then, the energy consumption for each compressed data transmission in a multihop network is simply written as follows

$$\mathbf{\Delta e} = L \cdot \Upsilon_d - L_b \cdot (\Upsilon_r - \Upsilon_c) \cdot \mathbf{x} \tag{6}$$

where $\mathbf{\Delta e} = [\Delta e(1), \ldots, \Delta e(n)]^{\mathrm{T}}$,$\Upsilon_d$ is the energy consumption for data transmission and reception, Υ_r is the energy reduction due to the reduced packet length, Υ_c is the energy consumption for data compression and decompression, and $\mathbf{x} = [x(1), \ldots, x(n-1)]^{\mathrm{T}}$. Note that Υ_r depends on α, whereas Υ_c is a constant matrix for the given values of ϵ_c and ϵ_d.

From equation (6), we derive a dynamic model for the energy consumption of multihop transmission. Let $\mathcal{E}_k$ denote the energy level of the nodes along the path after the kth multihop transmission. Then, the model for $\mathcal{E}_k$ is given by

$$\begin{aligned}\mathcal{E}_{k+1} &= \mathcal{E}_k - \mathbf{\Delta e} + \bar{\mathbf{e}}_k \\ &= \mathcal{E}_k + A_c \cdot \mathbf{x} + B_c\end{aligned} \tag{7}$$

where $A_c = L_b \cdot (\Upsilon_r - \Upsilon_c)$, $B_c = -L \cdot \Upsilon_d + \bar{\mathbf{e}}_k$, and $\bar{\mathbf{e}}_k$ is the average amount of energy harvested during the time elapsed between two transmissions.

Proposed decision scheme for compression

In this section, we consider the energy consumption using data compression for multihop data delivery. As a data packet is forwarded along a multihop path in

WSNs, the packet length gradually becomes smaller by data compression, and then, the energies consumed for receiving and sending the packet at each node decrease. Here, we propose a scheme that determines the compression level of each node in order to minimize the energy consumption and to maximize the network lifetime.

Energy consumption minimization

First, we focus on the minimization of the sum of the energies consumed by the nodes along a multihop path. This problem is formulated as follows

$$\begin{aligned} &\underset{\mathbf{x}}{\text{minimize}} \;\; \Delta\mathbf{e}^{\mathrm{T}} \cdot l_n \\ &\text{subject to} \quad x(i) \geq 0 \text{ for } \forall i \in \{1, \ldots, (n-1)\} \\ &\sum_{i=1}^{n-1} x(i) \leq b \end{aligned} \tag{8}$$

where $l_n \in \mathbb{R}^n$ is an *all-ones column vector*.

Proposition 1. The optimal solution for minimizing the sum of the energies consumed is to compress all blocks at the source node.

Proof. The sum of the energies consumed can be written as

$$\Delta\mathbf{e}^{\mathrm{T}} \cdot l_n = L(n-1)\epsilon_{\mathrm{tr}} - \underbrace{L_b \sum_{i=1}^{n-1} \{\alpha(n-i)\epsilon_{\mathrm{tr}} - \epsilon_{\mathrm{cd}}\} \cdot x(i)}_{= e_{\mathrm{saved}}} \tag{9}$$

where $\epsilon_{\mathrm{tr}} = (\epsilon_t + \epsilon_r)$ and $\epsilon_{\mathrm{cd}} = (\epsilon_c + \epsilon_d)$. Note that e_{saved} is the energy reduction due to data compression, whereas the first term corresponds to the energy consumption for multihop transmission without compression. Since the first term in equation (9) is constant with respect to $\mathbf{x}$, the minimization problem in equation (8) is now the same as the maximization of e_{saved}. Under the same constraints of equation (8), we can easily get the following result:

$$\frac{\partial e_{\mathrm{saved}}}{\partial x(i)} > \frac{\partial e_{\mathrm{saved}}}{\partial x(j)} \text{ for any } i,j \in \{1, \ldots, (n-1)\} \text{ and } i<j.$$

Hence, the optimal solution is obtained as $\mathbf{x} = [b, 0, \ldots, 0]^{\mathrm{T}}$ if $n \geq \left\lceil \frac{\epsilon_{cd}}{\alpha \cdot \epsilon_{tr}} \right\rceil$. Otherwise, $\mathbf{x} = 0$.

The above proposition implies that if ($n \geq \left\lceil \frac{\epsilon_{cd}}{\alpha \cdot \epsilon_{tr}} \right\rceil$) is satisfied, compressing all of the blocks at the source node is optimal for energy saving. If the inequality is not satisfied, data compression is wasteful since data communication with compression consumes more energy than that without compression.

Network lifetime maximization

We now further consider the problem of maximizing the network lifetime. To this end, we need to consider the energy level of each sensor node for each packet transmission. The network lifetime can be defined differently depending on the type of applications. For example, it can be the instant when the first node exhausts all its energy, a certain portion of nodes die, the network is partitioned, or the loss of sensing coverage occurs.[23] Here, we adopt the first one, that is, if one of the sensor nodes runs out of energy, it is considered that the network lifetime ends.

In a given multihop path, there may exist an unevenness in the energy levels among the sensor nodes due to the unbalanced energy consumption for packet transmission or a different amount of recharged energy by energy harvesting at each node. Consequently, for each packet transmission, the problem of maximizing the network lifetime can be formulated as follows:

$$\begin{aligned} &\underset{\mathbf{x}}{\text{maximize}} \quad \min \boldsymbol{\mathcal{E}}_{k+1} \\ &\text{subject to} \quad \boldsymbol{\mathcal{E}}_{k+1} = \boldsymbol{\mathcal{E}}_k + A_c \cdot \mathbf{x} + B_c \\ &x(i) \geq 0 \text{ for } \forall i \in \{1, \cdots, n-1\}, \\ &\sum_{i=1}^{n-1} x(i) \leq b \end{aligned} \tag{10}$$

From equation (10), since $\mathcal{E}_{k+1}(i) = \mathcal{E}_k(i) - \Delta e(i) + \bar{e}_k(i)$ and $\mathcal{E}_k(i) + \bar{e}_k(i)$ is constant with respect ot $\mathbf{x}$, we further have

$$\begin{aligned} \max_{\mathbf{x}} \min_i \{\mathcal{E}_{k+1}(i)\} &= \max_{\mathbf{x}} \min_i \{\mathcal{E}_k(i) - \Delta e(i) + \bar{e}_k(i)\} \\ &\leq \max_{\mathbf{x}} \min_i \{\mathcal{E}_k(i) + \bar{e}_k(i)\} \\ &= \min_i \{\mathcal{E}_k(i) + \bar{e}_k(i)\} \end{aligned} \tag{11}$$

where $\mathcal{E}_k(i)$ denotes the energy level of the ith node for the kth transmission, and $\bar{e}_k(i)$ denotes the amount of energy harvested at the ith node during the time elapsed between the kth and $(k+1)$th transmissions.

Hence, if we denote $i^* = \arg\min\{\mathcal{E}_k(i) + \bar{e}_k(i)\}$, that is, i^* is the index of the node with the smallest energy level, then any $\mathbf{x}$ that satisfies $\mathcal{E}_{k+1}(i) \geq \mathcal{E}_k(i^*) + \bar{e}_k(i^*)$, $\forall i$ can be a solution to equation (10). In other words, the network lifetime is maximized as long as the remaining energy level of every node after compressed transmission is not less than the smallest energy level among the nodes. This result matches with our intuition in that we can exploit the unbalanced energy levels among nodes using the residual energy level of each node above the smallest energy level.

Proposed m-hop averaging compression algorithm

Instead of solving the optimization problems, we propose a heuristic algorithm that determines an appropriate compression level $x(i)$ at each node for maximizing the network lifetime while reducing the total energy consumption of the WSN in practice. Since each node is rechargeable by energy harvesting and may have consumed a different amount of energy in the past, there may exist a certain level of differences in the energy levels among nodes. To reduce these differences and make the energy level of all nodes equal for maximizing the network lifetime, each node determines its compression level by comparing its own energy level with the average energy level of the next m nodes within m hops. Let $\overline{\mathcal{E}_k^m}(i)$ denote the average of the sum of the current energy levels and energy harvested for the m-hop nodes. Then, $\overline{\mathcal{E}_k^m}(i)$ is given by

$$\overline{\mathcal{E}_k^m}(i) = \frac{1}{m}\sum_{j=i+1}^{i+m}\{\mathcal{E}_k(j) + \bar{e}_k(j)\} \tag{12}$$

Here, we assume that each node can predict the amount of energy harvested and obtain the current energy levels of the next m nodes. Since the timeslot in which each packet transmission is carried out is very short (less than several milliseconds), the amount of energy harvested at each node during the time elapsed between two timeslots can be considered steady. Therefore, each node can predict the amount of energy to be harvested for this timeslot using an one-step-ahead linear prediction filter based on the amounts of energy harvested during a few previous timeslots. It is also possible to obtain the energy levels of the next m nodes with very little energy consumption for a small value of m (e.g. $m = 1$ or 2) because sensor nodes can piggyback the energy level on their data packets instead of using any extra control packets for energy information sharing.

Using the calculated average energy level in equation (12), the ith node determines the level of compression for each packet by comparing $\mathcal{E}_k(i) + \bar{e}_k(i)$ with $\overline{\mathcal{E}_k^m}(i)$. For example, if the energy level of the ith node with energy harvested is greater than the average of the sum of current energy levels and energy harvested for the next m-hop nodes, that is, $\mathcal{E}_k(i) + \bar{e}_k(i) > \overline{\mathcal{E}_k^m}(i)$, it compresses the data packet as much as possible by increasing the compression level until $\mathcal{E}_k(i) + \bar{e}_k(i)$ becomes less than or equal to $\overline{\mathcal{E}_k^m}(i)$ or all of the blocks of the data packet are compressed. Otherwise, if the node has less energy than the average energy level of the next m nodes, it forwards the data packet to the next node without any compression. This process is carried out at each node while the data packet is forwarded from the source node to the destination node. Thus, as the number of transmitted packets increases, the energy levels of nodes will become more and more close to each other with this algorithm.

Algorithm 1 Determining the compression level at the ith node.

1: $\mathcal{E}_{temp} \leftarrow \mathcal{E}_k(i) + \bar{e}_k(i)$
2: **while** $\mathcal{E}_{temp} > \overline{\mathcal{E}_k^m}(i)$ and $\sum_{j=1}^{i} x(j) \leq b$ **do**
3: $\quad x(i) \leftarrow x(i) + 1$
4: $\quad \mathcal{E}_{temp} \leftarrow \mathcal{E}_{temp} - L_b \cdot \epsilon_c + L_b \cdot \alpha \cdot \epsilon_t$
5: **end while**
6: $l(i) \leftarrow L_b \cdot (b - \alpha \cdot \sum_{j=1}^{i} x(j))$
7: $\Delta e(i) \leftarrow l(i-1) \cdot \epsilon_r + L_b \cdot x(i) \cdot \epsilon_c + l(i) \cdot \epsilon_t$
8: $\mathcal{E}_k(i) \leftarrow \mathcal{E}_k(i) - \Delta e(i)$

Algorithm 1 shows the pseudocode for the proposed algorithm that determines $x(i)$ at each node. First, the sum of the current energy levels $\mathcal{E}_k(i)$ and energy harvested $\bar{e}_k(i)$ is stored in the temporary variable $\mathcal{E}_{temp}$ in line 1. As shown in lines 2–5, each node increases its compression level as much as possible until the constraints in line 2 are violated. Once the compression level is determined, $l(i)$ and $\Delta e(i)$ are calculated, as shown in lines 6–7. Note that we assume that $l(0) = 0$ since the source node does not receive any packets. Finally, $\mathcal{E}_k(i)$ is recalculated by subtracting $\Delta e(i)$ from $\mathcal{E}_k(i)$ in line 8.

Performance evaluation

To evaluate the performance of our proposed algorithm, we considered a multihop WSN, in which N sensor nodes are randomly distributed. It is assumed that each sensor node is an IEEE 802.15.4-based MICAz device powered by two AA rechargeable batteries.[24,25] Each node is also assumed to have 5 kJ (2.5 kJ per each AA rechargeable battery) as the initial energy. As reported in the MICAz datasheet,[24] the transmission range of RF transceiver in MICAz is between 75 and 100 m. Here, the average distance between sensor nodes is set to 80 m so that the adjacent sensor nodes are within the transmission range of each other. We also refer to the energy model in De Meulenaer et al.,[25] where the energy consumed for computation, transmission, and reception are 3.5, 600, and 670 nJ/bit, respectively. Each node is assumed to exploit bzip2 to achieve lossless data compression with compression ratio of 0.7 as discussed in Barr and Asanović.[26] For compression and decompression of one bit, it requires 116 and 31 instructions/bit, respectively. Based on the energy model and data compression above, the parameters used in our simulations are chosen as $\epsilon_t = 600$ nJ/bit, $\epsilon_r = 670$ nJ/bit, $\epsilon_c = 3.5 \cdot 116 = 460$ nJ/bit, $\epsilon_d = 3.5 \cdot 31 = 108.5$ nJ/bit, and $\alpha = 0.7$. The original data packet length L is 1000 bits, and the number of blocks in a packet b is 100 (i.e. L_b is 10 bits/block).

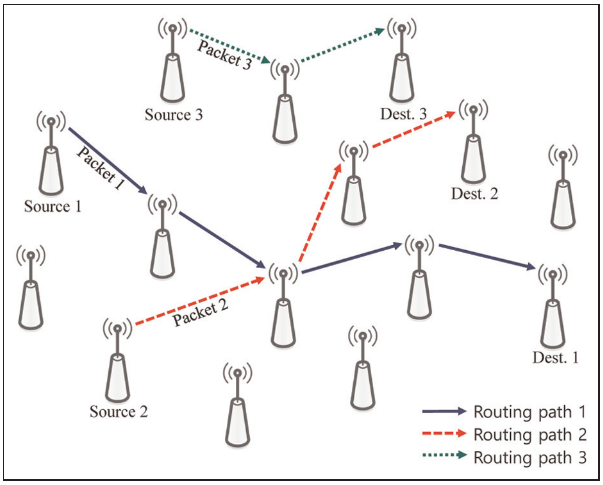

Figure 3. Multihop network topology with 15 nodes and 3 data flows.

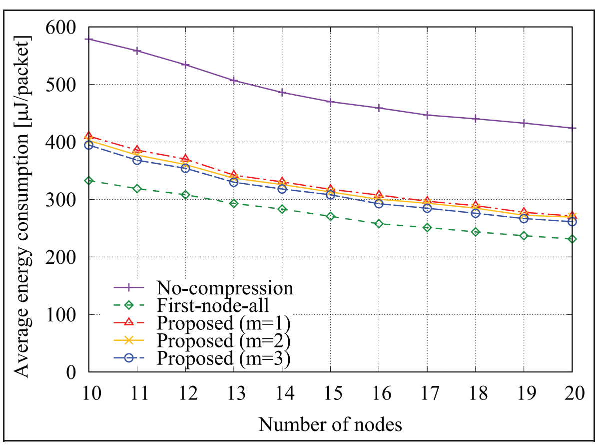

Figure 4. Average energy consumption with respect to the number of nodes.

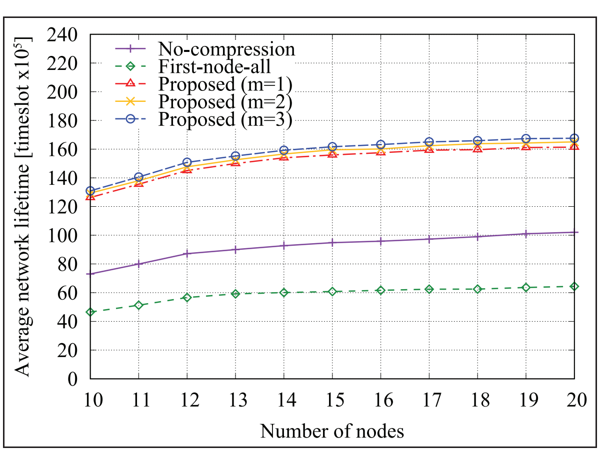

Figure 5. Average network lifetime with respect to the number of nodes without energy harvesting.

Figure 3 shows an example of multihop WSN with 15 nodes and 3 data flows. We assume that one packet transmission from the source node to the destination node is completed within one timeslot, and in each timeslot, each node is recharged by energy harvesting that is determined by a Gaussian random distribution with a mean of 100 μJ and a standard deviation of 20 μJ. At every packet transmission, a source node and a destination node are randomly selected, and then, a multihop path from the source node to the destination node is determined by a routing algorithm. There exist a variety of routing algorithms for WSNs such as location-based routing, data-centric routing, QoS-based routing, and energy-aware routing by Goyal and Tripathy.[27] Here, we adopt the simplest one, shortest-path routing algorithm, to investigate the performance of the proposed data compression algorithm without being affected by any other conditions such as routing. We performed simulations of our proposed algorithm with three cases, $m = 1$, $m = 2$, and $m = 3$, and compared them with a no-compression scheme and a first-node-all-compression scheme, in which the source node in a multihop path compresses all of the blocks, regardless of its energy level, in order to minimize the sum of the energies consumed by the network.

We first evaluated the average energy consumption. Figure 4 plots the average energy consumption with respect to the number of nodes N in the network. The average energy consumption means the total energy consumption divided by N when one packet is transmitted from the source node to the destination node. Our proposed algorithm exhibits an average energy consumption that is roughly 120 μJ lower than that of the no-compression scheme regardless of the number of nodes. The first-node-all-compression scheme shows a much lower average energy consumption than our proposed algorithm because every node, except the first node, receives and transmits the fully compressed packet, which facilitates a lower energy consumption.

In order to evaluate the network lifetime, we next measured the average network lifetime with respect to N for two cases, that is, without energy harvesting (Figure 5) and with energy harvesting (Figure 6). Here, if any one of the nodes runs out of energy, the network lifetime is considered to be finished, as mentioned in section. "Proposed decision scheme for compression." Let us assume that the total number of timeslots means the network lifetime since the number of timeslots is equal to the number of possible packet transmissions. In both Figures 5 and 6, it is seen that our proposed algorithm exhibits the longest network lifetime, which is approximately $5{\cdot}10^5$ timeslots more than the no-compression scheme over a wide range of the number of nodes. On the other hand, the first-node-all-compression scheme shows the shortest network lifetime among the three schemes, even though it exhibits the

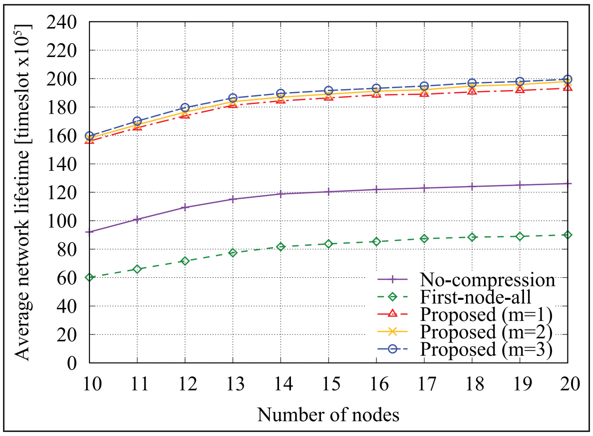

Figure 6. Average network lifetime with respect to the number of nodes with energy harvesting.

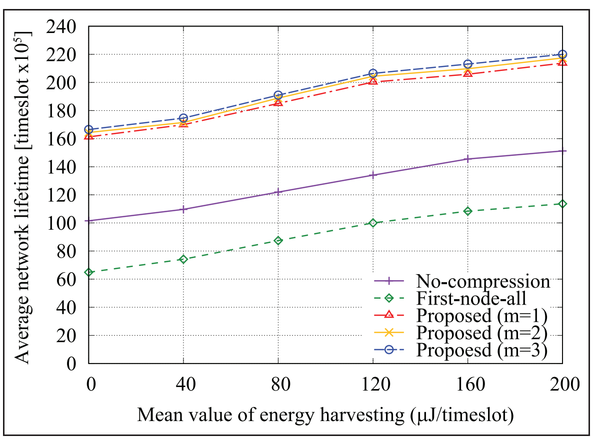

Figure 7. Average network lifetime with respect to the mean value of energy harvesting.

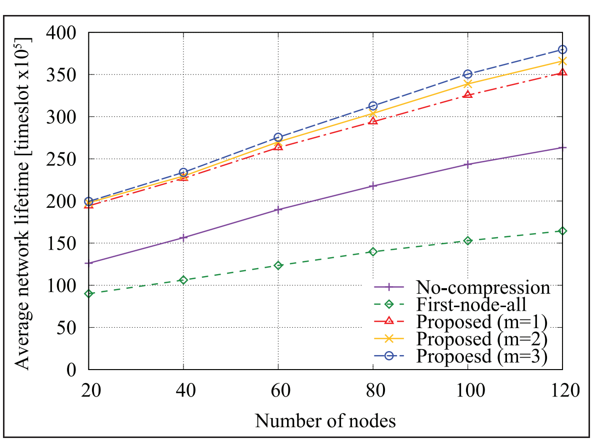

Figure 8. Average network lifetime with respect to the number of nodes in a larger network.

lowest average energy consumption, as shown in Figure 4. This definitely implies that the one node that is frequently chosen to be the first node is exhausted much quicker than any other node.

We also examined the effect of energy recharging by energy harvesting on the network lifetime. As shown in Figures 5 and 6, energy harvesting with a Gaussian random distribution $\mathcal{N}$ (12.5 mJ, 2.5 mJ) has an effect on the extension of the network lifetime. Specifically, the overall network lifetime of every scheme with energy harvesting is roughly $3{\cdot}10^5$ timeslots higher than the lifetime without energy harvesting. Also, Figure 7 simply shows that the network lifetime is extended as the mean value of energy harvesting, which follows a Gaussian random distribution, is increased.

Finally, we evaluate the network lifetime performance of the schemes in a larger network. Figure 8 shows the average network lifetime with respect to the number of nodes varying from 20 to 120. The performance results verify that the proposed algorithm provides the longest network lifetime in the larger network environment regardless of the value of m. It is also observed that the network lifetime of every scheme proportionally increases as the number of nodes increases. This is because the probability that each node is involved in routing paths decreases as the number of nodes in the network increases. As a result, each node consumes less energy for receiving, compressing, and transmitting data packets.

Through our numerical simulations, it is obvious that our proposed algorithm using data compression extends the network lifetime while reducing the energy consumption of WSNs. Furthermore, how to choose an appropriate value of m has been investigated. As shown in Figures 4–8, the performance differences between different values of m are not remarkable, even though the algorithm with a larger m, for instance, $m = 3$, results in slightly less energy consumption and a longer network lifetime than those with $m = 1$ or $m = 2$. However, it becomes more difficult for each node to get the energy-level information of the other nodes in real multihop networks as the number of hops between nodes increases. Therefore, a value of m of 1 or 2 is sufficient since our proposed algorithm ensures a sufficiently long network lifetime, even with a small m value.

Conclusion

In this article, we have considered energy savings for wireless sensor nodes that suffer from a limited battery capacity. For energy-harvesting WSNs that can be slowly recharged, we have proposed a data compression algorithm to decrease the average energy consumption while maximizing the network lifetime. As a data packet is relayed along the multihop path in the network, our proposed m-hop averaging compression algorithm determines the amount of packet to be compressed at each node by mainly considering the average energy levels of the next m nodes within m hops. Our extensive simulation results have verified that our proposed data compression algorithm achieves a considerable reduction in the energy consumption with a significantly extended network lifetime.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by the Electronic Warfare Research Center at Gwangju Institute of Science and Technology (GIST), originally funded by the Defense Acquisition Program Administration (DAPA) and Agency for Defense Development (ADD).

References

1. Akyildiz I, Su W, Sankarasubramaniam Y, et al. Wireless sensor networks: a survey. *Elsevier Comput Netw* 2002; 38(4): 393–422.
2. Uysal-Biyikoglu E, Prabhakar B and El Gamal A. Energy-efficient packet transmission over a wireless link. *IEEE/ACM T Netw (TON)* 2002; 10(4): 487–499.
3. El Gamal A, Nair C, Prabhakar B, et al. Energy-efficient scheduling of packet transmissions over wireless networks. In: *Proceedings of IEEE INFOCOM*, New York, 23–27 June 2002.
4. Pal V, Singh G and Yadav R. Balanced cluster size solution to extend lifetime of wireless sensor networks. *IEEE Internet Th J* 2015; 2(5): 399–401.
5. Sankarasubramaniam Y, Akyildiz I and McLaughlin S. Energy efficiency based packet size optimization in wireless sensor networks. In: *Proceedings of the IEEE sensor network protocols and applications (SNPA)*, Anchorage, AK, 11 May 2003.
6. He L. Energy-efficient multi-path routing with short latency and low overhead for wireless sensor networks. In: *Proceedings of IEEE software engineering, artificial intelligence, networking, and parallel/distributed computing (SNPD)*, Qingdao, China, 30 July–1 August 2007.
7. Oh H and Chae K. An energy-efficient sensor routing with low latency, scalability in wireless sensor networks. In: *Proceedings of IEEE multimedia and ubiquitous engineering (MUE)*, Seoul, Korea, 28 April 2007.
8. Seah W, Eu Z and Tan H. Wireless sensor networks powered by ambient energy harvesting (WSN-HEAP)—survey and challenges. In: *Proceedings of Wireless VITAE*, Aalborg, 17–20 May 2009, pp.1–5. New York: IEEE.
9. Gunduz D, Stamatiou K, Michelusi N, et al. Designing intelligent energy harvesting communication systems. *IEEE Commun Mag* 2014; 52(1): 210–216.
10. Nasir A, Zhou X, Durrani S, et al. Relaying protocols for wireless energy harvesting and information processing. *IEEE T Wirel Commun* 2013; 12(7): 3622–3636.
11. Anastasi G, Conti M, Di Francesco M, et al. Energy conservation in wireless sensor networks: a survey. *Elsevier Ad Hoc Netw* 2009; 7(3): 537–568.
12. Medeiros HP, Maciel MC, Demo Souza R, et al. Lightweight data compression in wireless sensor networks using Huffman coding. *Int J Distrib Sens N*, 2014; 2014: 1–11.
13. Marcelloni F and Vecchio M. A simple algorithm for data compression in wireless sensor networks. *IEEE Commun Lett* 2008; 12(6): 411–413.
14. Marcelloni F and Vecchio M. An efficient lossless compression algorithm for tiny nodes of monitoring wireless sensor networks. *The Comput J* 2009; 52(8): 969–987.
15. Capo-Chichi EP, Guyennet H and Friedt JM. K-RLE: a new data compression algorithm for wireless sensor network. In: *Proceedings of IEEE sensor technologies and applications (SENSORCOMM)*, Athens, 18–23 June 2009, pp.502–507. New York: IEEE.
16. Alsalaet JK and Ali AA. Data compression in wireless sensors network using MDCT and embedded harmonic coding. *Elsevier ISA T* 2015; 56: 261–267.
17. Tavli B, Bagci IE and Ceylan O. Optimal data compression and forwarding in wireless sensor networks. *IEEE Commun Lett* 2010; 14(5): 408–410.
18. Incebacak D, Zilan R, Tavli B, et al. Optimal data compression for lifetime maximization in wireless sensor networks operating in stealth mode. *Elsevier Ad Hoc Netw* 2015; 24: 134–147.
19. Ali A, Khelil A, Suri N, et al. Adaptive hybrid compression for wireless sensor networks. *ACM T Sens Netw (TOSN)* 2015; 11(4): 53.
20. Kimura N and Latifi S. A survey on data compression in wireless sensor networks. In: *Proceedings of IEEE information technology: coding and computing (ITCC)*, Las Vegas, NV, 4–6 April 2005.
21. Srisooksai T, Keamarungsi K, Lamsrichan P, et al. Practical data compression in wireless sensor networks: a survey. *J Netw Comput Appl* 2012; 35(1): 37–59.
22. Razzaque MA, Bleakley C and Dobson S. Compression in wireless sensor networks: A survey and comparative evaluation. *ACM T Sens Netw* (TOSN) 2013; 10(1): 5.
23. Chen Y and Zhao Q. On the lifetime of wireless sensor networks. *IEEE Commun Lett* 2005; 9(11): 976–978.
24. MICAz Datasheet. http://www.memsic.com/userfiles/files/Datasheets/WSN/micaz_datasheet-t.pdf
25. De Meulenaer G, Gosset F, Standaert FX, et al. On the energy cost of communication and cryptography in wireless sensor networks. In: *Proceedings of IEEE international conference on wireless and mobile computing, networking and communications*, Avignon, 12–14 October 2008, pp.580–585. New York: IEEE.
26. Barr KC and Asanović K. Energy-aware lossless data compression. *ACM T Comput Syst (TOCS)* 2006; 24(3): 250–291.
27. Goyal D and Tripathy MR. Routing protocols in wireless sensor networks: a survey. In: *Proceedings of IEEE international conference on advanced computing & communication technologies*, Rohtak, India, 7–8 January 2012, pp.474–480. New York: IEEE.