# A Lossless Distributed Data Compression and Aggregation Approach for Low Resources Wireless Sensors Networks

**Elie TAGNE FUTE** ( eliefute@yahoo.fr )
University of Dschang https://orcid.org/0000-0001-8979-7869
**Hugues Marie KAMDJOU**
University of Dschang https://orcid.org/0000-0002-0728-8655
**Adnen EL AMRAOUI**
University of Artois
**Armand NZEUKOU**
Victor University Institute of Technology, University of Dschang

**Research Article**

**Keywords:** Aggregation, Data compression, Image, Low resources, Network lifetime, Wireless Sensor Networks

**Posted Date:** June 29th, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-583690/v1

# A Lossless Distributed Data Compression and Aggregation Approach for Low Resources Wireless Sensors Networks

Elie TAGNE FUTE [1,2,*], Hugues Marie KAMDJOU[1], Adnen EL AMRAOUI [3], Armand NZEUKOU [4]

[1] Department of Mathematics and Computer Science, Faculty of Science, University of Dschang, Dschang, Cameroon;

[2] Department of Computer Engineering, Faculty of Engineering and Technology, University of Buea, Buea, Cameroon;

[3] LGI2A (« Laboratoire de Génie Informatique et d'Automatique de l'Artois »), University of Artois, UR 3926, F-62400 Bethune, France;

[4] Department of Industrial Systems and Environmental Engineering, Fotso Victor University Institute of Technology, University of Dschang, Cameroon, Faculty of Science, University of Dschang, Bandjoun, Cameroon;

**Corresponding author:** * eliefute@yahoo.fr (Elie TAGNE FUTE)

**Abstract:** Wireless Sensor Networks (WSN) have been as useful and beneficial as resource-constrained distributed event-based system for several scenarios.Yet, in WSN, optimization oflimited resources (energy, computing memory, bandwidth and storage) during data collection and communication process is a major challenge. Most of energy consumption (as much as 80%) for standard WSN applications lies in the radio module where receiving and sending packets are necessary to communicate between stations.This paper proposes an approach to achieve optimal sensor resources by data compression and aggregation regarding integrity of raw data.Data aggregation discarded a certain sensing data packet, which leads to low data-rate communication and low likelihood of packet collisions on the wireless medium. Data compression reduces a redundancy in aggregated data, which leads to save storage and sending only one small data stream in the bandwidthof communication.The performance of the proposed approach is qualified using experimental simulation on OMNeT++/Castalia. Theperformance metricswere evaluated in terms of Compression Ratio (CR), data Aggregation Rate (AR), Peak Signal-to-Noise Ratio (PSNR) and Mean Square Error (MSE) and Energy Consumption (EC).The obtained resultshave significantly increased the network lifetime.Moreover, the integrity (quality) of the raw data is guaranteed.

**Keywords:** Aggregation, Data compression, Image, Low resources, Network lifetime, Wireless Sensor Networks

## 1. Introduction

During the last few years,the application of Wireless Sensor Networks (WSNs) hasbeen an increasing interest in unattended environments[1]. The WSN is composed of hundreds to thousands of wireless nodes. Each node has some computational power and sensing capability, and operates in an unattended mode [2].These devices are able to monitor a wide variety of ambient conditions, like temperature, pressure, luminosity, humidity, composition of the soils, human or vehicular movement,supply chain, noise levels, the presence or absence of certain kinds of objects (like those of the medical imaging field), mechanical stress levels on attached objects and so on. Due to the inaccessibility of the hostile area and also the large number of sensor nodes in the network area, it is not always possible to expect the sensor nodes to be plugged into an electrical outlet or to change their batteries frequently. Therefore, it is crucial to optimize the amount of consumed energy expended by the sensor nodes, since the consumed energy determines the lifetime of a sensor network. However, wireless communicationconsumes more energy than other activities.

The communication radius is generally greater than the range of a single node. Hence, the farther sensor that has to transmit data requires more energy and thereforethe lifetime will be more reduced. To tackle this main issue, resource optimization becomesa crucial problemto design an efficient compression and aggregation approach whichminimizes at the same time packet loss, collisions, congestion, power consumption and the amount of communication required by the sensor nodes.

The autonomous sensors are randomly deployed, thus, various sensor nodes often collect a common phenomenon,which creates redundancyin the data communicatedfrom sources node to a particular Cluster-Head(CH) or sink. It is known that leveraging the correlation between different samples of the observed data will lead to better utilization of sensor resources reserve. However, a large number of sensors periodically collect data and send them to a border node in the network. Resources saving can be archived if different sensor reading can be combined into a single super packet through compression and aggregation, which eliminates redundancy, minimizes the number of transmissions and thus savessensors resources.

Nevertheless, aggregation should only be done if the amount of energy taken to aggregate data byte and transfer is less than just transferring data without aggregation. The approach also examines the complexity of optimal data aggregation, showing that although it is a NP-hard

(Nondeterministic Polynomial time) problem in general. Most of the datacompression and aggregation methodsin literatures investigated on lossy compression and address-centric aggregation routing protocol.The proposed approach focuses on lossless data-centric compression and aggregation to obtain the approximate polynomial solution.

The rest of this article is organized as follows: section 2 presents a related work and focused issues on data compression and aggregation. Section 3 proposes a distributed lossless data compression and aggregation approach in WSNs. Section 4 presents the implementations and discusses the experimental simulation results. Section 5 concludes the paper.

## 2. Related work and focused problems

It should be mentioned that, this section reviews literature on data compression and aggregationin WSNs, different authors have implemented the possible approaches. Although there are researcheswhich depend only on one of the two methods mentioned above.

Most of the aggregation schemes presented in several literatures investigated to save sensor's energy by considering unconstrained data traffic [3-4, 5]. In aggregation, the intermediate nodes can remove data redundancy received from multiple sources in order to transmit the compressed data. The compression approaches can be grouped into two main categories: lossless and lossy data compression. Lossless compression generates a statistical model of the data and maps the data to bit strings based on the generated model. Conversely, lossy compression transforms the data into a new space using appropriate basis functions. In the new space, the data information is usually concentrated on a few coefficients. Hence, compression can be achieved after quantization and entropy coding [6-7-8, 9]. The best known methods in the literatureswill be introducedin the followingsection.

### *2.1. Data funneling*

In compression by funneling approach,local nodes transmit the reading data to a border node which aggregates the data before sending it onto the controller node.The nodes in thearea select a parent node which aggregates the data before sending it onto the base station as the authors present in [10].

### *2.2. Pipeline in-network compression scheme*

In Pipeline approach, the data collected from sensors is buffered in the network aggregation node for a certain lapseof time. Then, the data packets are combined into one data packet by suppressing the redundancies through a pipelined compression scheme as the authors present in [11].

### *2.3. Hardware-Assisted data compression*

Hardware-assisted approach proposed an adaptive compression architecture based on statistical data analysis, for on the fly data compression and decompression whose field of operation is the cache to sensor memory path as the authors presents in [12].

### *2.4. Clustering methods*

In WSNs, clustering methods allow the data aggregation of sensor and improve the scalability of multi-hop wireless networks. Thisapproachdivides the network into subset partitionconsisting of nodes, called clusters.Each partition has one node serving as its Cluster-Head (CH). After the formation of clusters, the nodes transmit their sensing data to the CH for data aggregation. Various clustering protocols have been proposed in literatures [13, 14]. Most of them did not consider data correlation and the assumption of ideal data aggregation, where data are perfectly correlated, such that an arbitrary number of packets within a cluster can be compressed into a single packet.

### *2.5. Routing models*

The routing schemes which use data aggregation in literature are data-centric routing protocol and address-centric routing protocol. In both casesthe sink sends out a query/interest for a certain data collected in which the sensor nodes that have the appropriate data then responds with the corresponding data. However, the difference of the two methods is the way to senddata from the sources to the sink [15]. In address-centric protocol, each source independently sends data along the shortest path to the sink, while in data-centric protocol, the sources send data to the sink, but routing nodes can access the content of the data packet and perform aggregation on multiple input packets [16]. Due to the advantages, in this work, data-centric protocol is consideredto be used.

## 3. Proposedmethod: distributed lossless data compression and aggregation approach

The proposed network architecture approach is to focus on a single network graph that is assumed just for a single cluster attempting to gather data from a certain number of data sources of its cluster [17, 18]. Letusconsider $n$ source nodes ($N_1$, …,$N_n$) and a sink node ($K$). Let the network graph be $G = (N, E, d)$ consisting of all the nodes $N$, with $E$ that is composed of edges between all nodes that can communicate witheach other directly and$d$represents a distance function which maps $E$ into a set of non-negative number.

Let us assume that the number of transmissions from any node in data aggregation node is exactly one. Each sensor $N_i$ sends thesensing data collected to the aggregator (CH), and then the aggregated data is sent to the sink (K). Thus, the aggregation rate is a ratio between discarded packet and the total packet received in the aggregation node. The problem to be sort out is to perform compression and aggregation of the sensing data at a single point of aggregation, before the transmission of the compressed data to the sink. The flow chart of the proposed approach is shown in Figure 1.

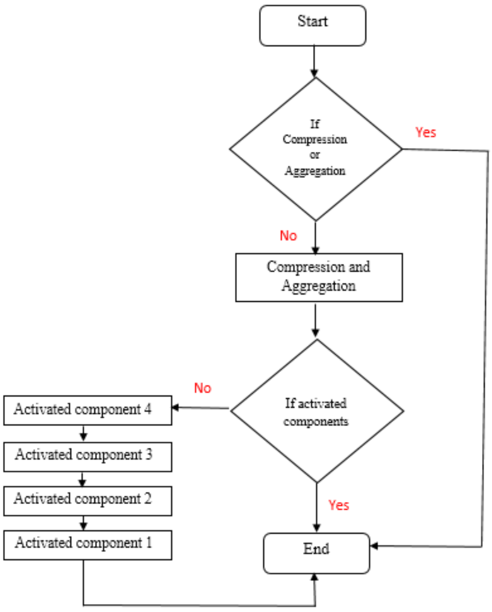


**Figure 1.** Flow chart of the compression and aggregation approach

### *3.1. Aggregation scheme*

The strategy of the aggregation is to use a convenientpacket order in which those packet (pkt) data are sent to convey additional information to the sink.

When the aggregation node receives the sensing data from the neighbors, it explicitly discards some of the data packets, and then the ordering remained data packetsare used to transmitthe information containing the packets that have been discarded. The problemis how many packets can be discarded.

Let $k$ be the range of possible values generated by each sensor, $p$the number of packets present at the aggregation node, $r$the number of discarded packets and $n$ the range of sensornode identification, each node has a unique Identification(Id).

The strategy is to discard$r$ packages and use the appropriate order of the remaining $(p$-$r)$ packages to indicate which values (payload) were contained in the $r$ discarded packets; this induced the number of permutations given by$(p$-$r)!$. Each of the discarded packets contains a payload that can take one of the $k$ possible values and an Id that can be $(n$-$p$+$r)$ value of all valid Ids except those belonging to the packets included in the super packet.

The values (payload and Id) contained within the discarded packets can be considered as symbols of $(n$-$p$+$r) * k$ary-alphabet, resulting in $(n$-$p$+$r)^r *kr$ possible values for the discarded packets. Since, each packet has to be identified with a unique Id, moreover the packets are discarded simultaneously and randomly, then $(n$-$p$+$r)^r$ is more expressed by $C_{n-p+r}^{r}$. Thus, to obtainaggregation rate, the following relation must be satisfied [10]:

$$(p-r)! \geq C_{n-p+r}^{r} * k^{r} (1)$$

However, for large values of $p$ and $n$, this relation becomes a NP-hard problem, which calculation largely exceeds the accuracy of a computer.The problem is critical in WSNs applications, which require a low complexity [19, 20]. Moreover, the inequalityin equation(1) cannot be transformed, so that $r$ is expressed as a function of $n$, $p$and $k$. Therefore, to alleviate this task, numerical's approximation methods and the Stirling's approximation ($x! \cong \left(\frac{x}{e}\right)^{x} \sqrt{2\pi x}$) are used to calculate the optimal value of $r$ satisfying the inequality ofthe relation (2) as follow [21, 22]:

$$\ln(2\pi) + r + (r+1/2) * \ln(r) + (p-r+1/2) * \ln(p-r) + (n-p+1/2) * \ln(n-p) - r * \ln(k) - (n-p+r+1/2) * \ln(n-p+r) - p \geq 0 \qquad (2)$$

Where $ln(.)$ represents a natural logarithm.

Let's consider a running example where there are $n = 8$ nodes with Id's N1, N2, N3, N4, N5, N6, N7 and N8. The number of messages that arrive at an aggregation node is $p = 6$. By considering a black or white color of an image, each sensor generates an independent reading, which is from the set {white, black}, then $k = 2$. Using the above formula (equation 2) scheme allows the encoder to reject $r = 2$ packets. To clearly illustrate the scenario, Table 1 shows the mapping obtained from the shift cursor permutation of Algorithm 1 as shown in Figure 2 [23].

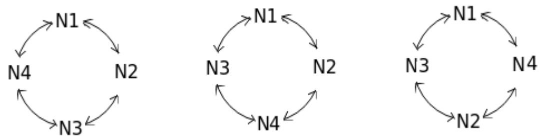


**Figure 2.** Shift cursor permutation

**Table 1.** Mapping from permutations corresponding to discarded packets (w = white and b = black)

| Permutation packets Ids | Discarded packets values |
|---|---|
| **N1,N2,N3,N4** | (N5,w)(N6,w) |
| **N1,N2,N4,N3** | (N5,w)(N6,b) |
| **N1,N4,N2,N3** | (N5,b)(N6,w) |
| **N1,N3,N2,N4** | (N5,b)(N6,b) |
| **N1,N3,N4,N2** | (N5,w)(N7,w) |
| **N1,N4,N3,N2** | (N5,w)(N7,b) |
| **N4,N1,N3,N2** | (N5,b)(N7,w) |
| **N4,N1,N2,N3** | (N5,b)(N7,b) |
| **N2,N1,N4,N3** | (N5,w)(N8,w) |
| **N3,N1,N4,N2** | (N5,w)(N8,b) |
| **N3,N1,N2,N4** | (N5,b)(N8,w) |
| **N2,N1,N3,N4** | (N5,b)(N8,b) |
| **N2,N3,N1,N4** | (N6,w)(N7,w) |
| **N2,N4,N1,N3** | (N6,w)(N7,b) |
| **N4,N2,N1,N3** | (N6,b)(N7,w) |
| **N3,N2,N1,N4** | (N6,b)(N7,b) |
| **N3,N4,N1,N2** | (N6,w)(N8,w) |
| **N4,N3,N1,N2** | (N6,w)(N8,b) |
| **N4,N3,N2,N1** | (N6,b)(N8,w) |
| **N4,N2,N3,N1** | (N6,b)(N8,b) |
| **N2,N4,N3,N1** | (N8,w)(N7,w) |
| **N3,N4,N2,N1** | (N8,w)(N7,b) |
| **N3,N2,N4,N1** | (N8,b)(N7,w) |
| **N2,N3,N4,N1** | (N8,b)(N7,b) |

Let us assume in WSNs of n = $2^8$ sensors, k = $2^4$, the discarded packets (*r*) as a function of the *p* packets in aggregation node isshown in Figure 3, as defined in equation 2. Figure 4illustrates the family function of discarded packets.

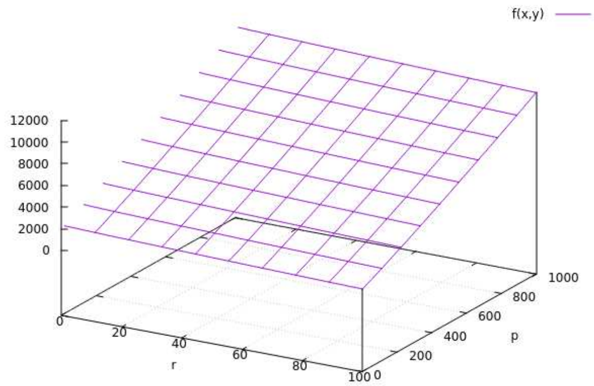


**Figure 3.** Aggregation function

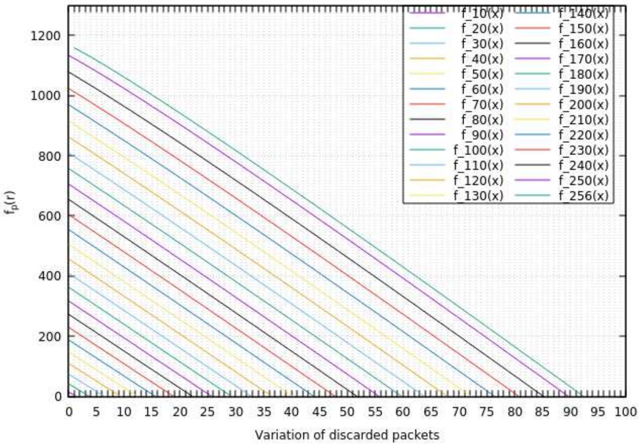


**Figure 4.** Family function of discarded packets

| **Algorithm 1**: Shift cursor permutation; |
|---|
| *Input : p packets;*<br>*Output: ordered packets;*<br>*1. BEGIN*<br>*2. Starting from N1, N2, ... , Np;*<br>*3. If the level order of the cursors is N1, N2, ... , Np, stop;*<br>*4. Determine cursors which level > j and their positions;*<br>*5. Output ordering packets;*<br>*6. Generate and output next integer* $((\frac{j}{e})^j\sqrt{2\pi j})$*-1) permutations directly;*<br>*7. Go to step 3;*<br>*8.End.* |

### *3.2. Compression scheme*

The strategyof the method is to reducethe aggregated data packetto obtain the optimal compression size in WSNs.The approach is designedby four components as illustrated in Figure 5. For now, the four components which are integrated are based on the entropic coding: Arithmetic coding, Run Length Encoding, Move ToFront encoding and Burrows Wheeler Transform encoding [2, 6]. These coding were chosen becauseof their particular performance in applications constrained by resources.Their implementation involvessimple instructions of additions and integer values shifts.

The components 3 and 4 in Figure 5are based on a process of redundancy identification in raw data to facilitate the compression by the component 2. Once the data has no redundancy, the component 1 compresses it to yield an unintelligible file whose size is smaller. The Algorithm 2 and the Algorithm 3 describe the implementations of components.

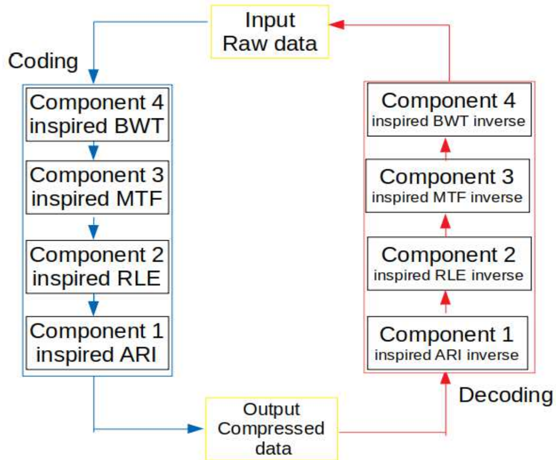


**Figure 3.**Synoptic diagram of the propose compression approach

| **ALGORITHM 2:** Coding; |
|---|
| *Input : raw data;*<br>*Output: compressed data;*<br>*1. BEGIN*<br>*2. Initializations;*<br>*3. Case component 1:*<br>• *range_bound←1; low_bound ← 0; up__bound ← 1;*<br>• *WHILE (Not end of data) DO*<br>• *symbol_pixel ← get symbol_pixel ;*<br>• *up_bound ← low_bound + range_bound * upper_of_current_symbol_ pixel;*<br>• *Low_bound ← low_bound + range_bound * lower_of_current_symbol_ pixel;*<br>• *Range_ bound ← up_ bound - low_ bound;*<br>• *Output_a_code_so_that_low_bound ≤ code <up_bound;*<br>• *END_while;*<br>*4. Case component 2:*<br>• *boolean ← false; meter ← 0;*<br>• *previous_symbol_pixel ← get current_symbol_ pixel;*<br>• *WHILE ((current_symbol_ pixel←next_symbol_ pixel) ≥ 0) DO*<br>• *Add current_symbol_ pixel in_the_buffer;*<br>• *IF (current_symbol_pixel == previous_symbol_pixel) THEN*<br>• *boolean ← false ; count ← 0 ;*<br>• *WHILE ((meter < 255) and ((current_symbol_pixel ←* |

*next_symbol_ pixel) ≥ 0) && (boolean == false)) DO*
- *IF (current_symbol_ pixel ==previous_symbol_ pixel) THEN*
- *meter ← meter + 1 ;*
- *ELSE boolean ← true;*
- *END_if*
- *END_while*
- *Add meter in_the_buffer;*
- *IF ((meter ≠ 255)and (current_symbol_pixel ≥ 0)) THEN*
- *Add symbol_ pixel in_the_buffer;*
- *END_if;*
- *END_if;*
- *previous_symbol_pixel ← current_symbol_pixel;*
- *END_while;*

*5. case component 3:*
- *Initialize the_symbol_pixel_table;*
- *WHILE (Not end of file) DO*
- *S ← get symbol;*
- *Add_position_S_symbol_of_the_symbol_pixel_tablein _the_buffer ;*
- *Move_the_symbol_S_to_the_head_of_the_table_symb ol_pixel ;*
- *END_while;*

*6. case component 4:*
- *L← get_block_string_of_N_symbol_ pixel $C_0 \ldots C_{N-1}$;*
- *sort (L);*
- *output_block_string_composed_of_the_last_character _of_each_of_the_rotations;*

*7. return output;*
*8. End.*

**ALGORITHM 3:** Decoding;

*Input : compressed data;*
*Output: raw data;*
*1. BEGIN*
*2. Initializations;*
*3. Case component 1:*
- *range_bound ← 1; low_bound ← 0; up_bound ← 1;*
- *WHILE (Not end of data) DO*
- *value ← get code_of_symbol_ pixel ;*
- *symbol_pixel ← symbol_succ_that_upper_symbol ≤ ((value – low_bound) / (up_bound – low_bound)) <low_symbol_pixel ;*
- *range_bound ← up_bound – low_bound;*
- *up_bound ← low_bound + range_bound * low_of_current_symbol_pixel ;*
- *Low_bound ← low_bound + range_bound * up_of_symbol_pixel ;*
- *output symbol_ pixel;*
- *END_while;*

*4. Case component 2:*
- *previous_symbol_pixel ← current_symbol_ pixel;*
- *meter ← 0;*
- *WHILE ((current_symbol_ pixel←next symbol_ pixel) ≥ 0) DO*
- *Add symbol_ pixel in_the_buffer;*
- *IF (current_symbol_pixel==previous_symbol_ pixel) THEN*
- *meter ← next_symbol_pixel ;*
- *WHILE ((meter ← (meter – 1)) > 0) DO*
- *Add current_symbol_ pixel in_ the_buffer;*
- *END_while;*
- *END_if;*
- *previous_symbol_pixel ← current_symbol_ pixel;*
- *END_while;*

*5. Case component 3:*
- *Initialize the_symbol_ pixel_table;*
- *WHILE (Not end of data) DO*
- *P ← get symbol_ pixel_position;*
- *Add_symbol_pixel_in_position_P_of_the_symbol_tabl e_in_the_buffer;*
- *Move_the_symbol_S_to_the_head_of_the_table_symb ol_ pixel;*
- *END_while;*

*6. case component 4:*
- *index← get primary index ; L← get buffer_string;*
- *F← sort(L);*
- *Compute_the_transformation_vector_H_such_that L[H[j]] = F[j] for any j;*
  - *FOR i from 0 to length( L) DO*
  - *Output L[index];*
  - *index ←H[index];*

*7. return output;*
*8. End.*

## 4. Implementations and simulations results

The Castalia simulator is used to extend the functionality of the Omnet++ simulator, particularly in the Wireless Channel of transmission module and the energy management module [24-25-26, 27].In the implementation simulations under integrated simulators (Omnet++/Castalia), four scenarios are envisaged as illustrated in Figure 1.

The first scenario is to compress data at the source node before transmission to the CH where it will be aggregated. The second scenario is to aggregate the data collected at the CH before performing the compression.The third scenario is to send the sensing data without compression and aggregation. The last scenario is just to send the sensing data after compression or aggregation.

### *4.1. Compression and aggregation*

In the aggregation process,the aggregation node collects data from its cluster and then applies the aggregation in a first step.In the next step it performs the four compression components. The illustration of the processescan be shown in Figure 6 and Algorithm 4.

The compression process as illustrated in Figure 7 and Algorithm 5, consists to add a table of four components to the super packet and initialize by zero (not active component). Each node can activate at most one compression process component among the four, depending on the data received.

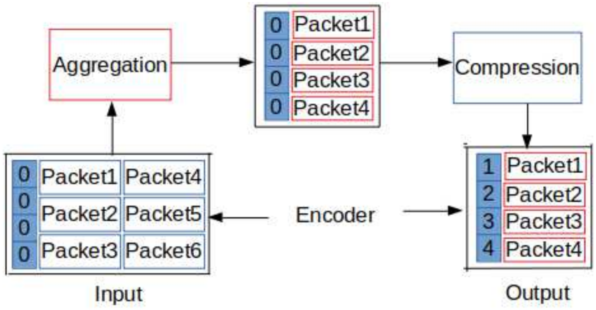


**Figure 6.**Synoptic diagram of aggregation and compression

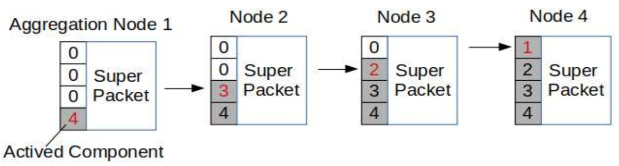


**Figure 7**.Component activated in super packet

| **ALGORITHM 4:** Aggregation; |
|---|
| *Input : super packet;* |
| *Output: compressed packet;* |
| *1. BEGIN* |
| *2. P ← get super packet;* |
| *3. WHILE ( (node(P)=CH ) and (component(P) = 0)) DO* |
| *4. Activete all the component;* |
| *5. END_While;* |
| *6. Send packet to the sink;;* |
| *8. End.* |

| **ALGORITHM 5:** Distributed compression; |
|---|
| *Input : super packet;* |
| *Output: compressed packet;* |
| *7. BEGIN* |
| *8. P ← get super packet;* |
| *9. WHILE ( (node(P)≠sink ) and (component(P) = 0)) DO* |
| *10. Activate the next component;* |
| *11. END_While;* |
| *12. Send packet to the sink;;* |
| *8. End.* |

### *4.2. Performance metricsand simulation parameters*

The performances of the proposed approach are measured in terms of the following metrics:

- Aggregation Rate (AG) as defined in equation 3.

$$AG = \frac{r}{p} \quad (3)$$

Where *r* represents the number of discarded packets and *p* is the number of packets in aggregation node.

- Compression Ratio (CR)as defined in equation 4.

$$CR = (1 - \frac{FSC}{OFS})\ x100 \quad (4)$$

Where FSC is the Frame Size after Coding and OFS is the Original Frame Size.

- Peak Signal-to-Noise Ratio (PSNR) as defined in equation 5.

$$PSNR = 10log_{10}\left(\frac{(Pic)^2}{MSE}\right) \quad (5)$$

Where the pixel values of image are integers that range from 0 (black) to 255 (white), thus Pic represents the max pixel value 255 and MSE is a Mean Square Error.

- Mean Square Error (MSE) as defined in equation 6.

$$MSE = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(x_i - y_j\right)^2 \quad (6)$$

Where NM is the size of image that represents the N x M pixels, N is the number of rows, M the number of columns.

- Energy Cost (EC) as defined in energy consumption model of equation (7) [29, 30]:

$$EC = E_{cap} + E_{comp} + E_{agg} + E_t + E_r \quad (7)$$

Where $E_{cap}$, $E_{comp}$, $E_{agg}$, $E_t$ and $E_r$ respectively represent the consumed energy to capture image, the consumed energy to compress *s*bytes, the consumed energy to aggregate *s* bytes (equation 8), the consumed energy to send*s* bytes at a distance *d*, and the consumed energy to receive *s* bytes (equation 9).

$$E_t(s,d) = (E_{elec} * s) + (E_{amp} * s * d^2) \quad (8)$$

$$E_r(s,d) = (E_{elec} * s) \quad (9)$$

Where $E_{elec}$ and $E_{amp}$ respectively represent the signals electrification energy and amplification energy.

To sort out the different performance metrics, the simulation setup parameters are definedin Table 2.

**Table 2.** Simulation parameters

| | |
|---|---|
| Simulator | Omnet++/Castalia |
| Simulation time | 1860 seconds |
| CPU time | 1860 seconds |
| Simulation area | 50x50 square meters |
| Mac protocol | IEEE 802.11 |
| Protocol Radio (Rx, Tx, Sleep) | CC1000 |
| Routing protocol | GPSR |
| Number of nodes | 10 |
| Image size (Figure 8 (a)) | 320x320 pixels, 325 kilobytes |
| Image size (Figure 8 (b)) | 360x360 pixels, 233 kilobytes |
| Initial energy $E_i$ | 19440 Joules (J) |
| Captured energy $E_{cap}$ | 110 milli-joule (mJ) |
| Compression energy $E_{comp}$ | 10 milli-joule per packet (mJ/pkt) |
| Aggregation energy $E_{agg}$ | 20 mJ/pkt |
| Transmission energy $E_t$ | 150 mJ/pkt |
| Received energy is $E_r$ | 150 mJ/pkt |

The test images from Kaggle dataset [28]are shown in Figure 8. Within the 10 nodes of the network, the sources nodes are node 1, node 2, node 3, node 4 (CH), and node 5 as shown in the simulation environment in Figure 9. The simulation time limit to receive the complete image is 1860 seconds.

(a) (b)

**Figure 8.**Overview of the test images

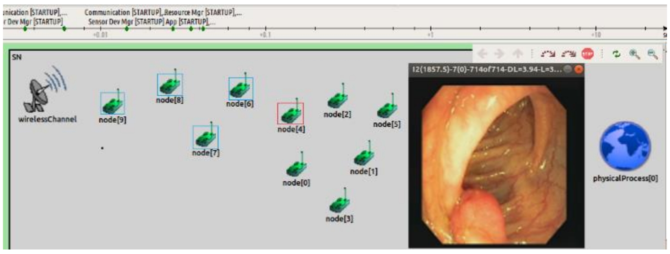


**Figure 9.**Overview of the simulation environment Omnet++/Castalia

### *4.3. Results and Analysis*

The images of Figure 8 were used for the experiments, each camera sensor capture image that the pixel values are integers from 0 (black) to 255 (white). Thus,k = $2^8$is the possible gray intensity values.Figure 10 presents the data Aggregation Rate (AR) of the proposed approach with low-complexity. It can be observed that the gain increases exponentially when the number of packets to be aggravated grows. Which induces a logical reduction of collisions, congestion, data-rate communication and produce various trade-offs among some network related performance metrics such as compression rate, energy, latency, accuracy, fault-tolerance and security.

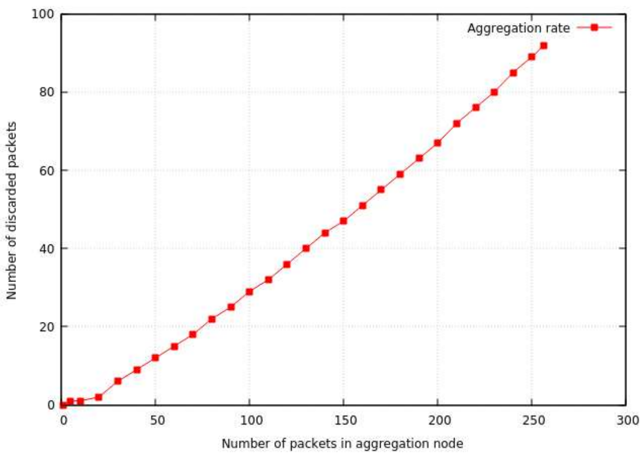


**Figure 10.** Discarded packets as a function of packets at the aggregation node

The proposed compression with aggregation approach was compared with 2D-DCT (two-Dimensional Discrete Cosine Transform) presented in [31]. Figure 11 shown the comparative CR. It shows that 2D-DCT presents a better CR than the proposed approach. It can be justified by the fact that 2D-DCT is a lossy compression approach, whereas the proposed approach is a lossless compression.In contrast, the lossless propose approach presents the best PSNR in Figure 12 and the best MSE in Figure 13 compared to 2D-DCT.

The overall energy consumption of the proposed approach, used for the operation of capture, communication, compression and aggregation processes of each sensor node is shown in Figure 14 and Figure 15. The results reveal that after 1860 seconds of simulation, the first image is received by the sink.

Figure 16 and Figure 17present the comparative network lifetime of remaining energy between the proposed method and 2D-DCT. Thus, generally speaking, the proposed approach of compression with aggregation has the best remaining energy.

In light of these encouraging results, the performance characteristics of the proposed approach are satisfactory.

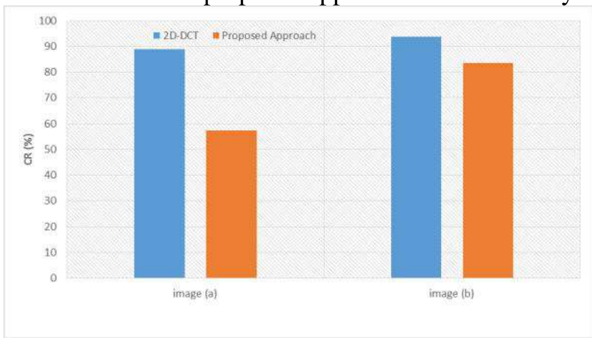


**Figure 11**. Comparative histogram of Compression Ratio (%)

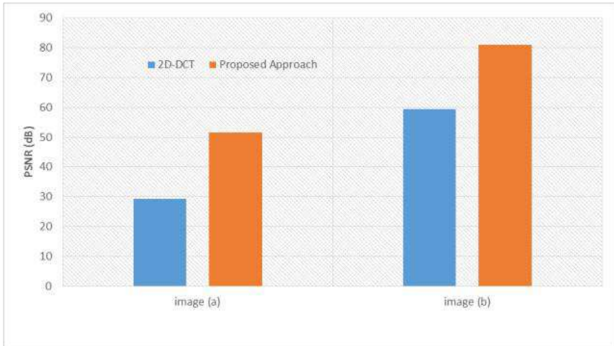


**Figure 12**. Comparative histogram of Peak Signal to Noise Ratio (dB)

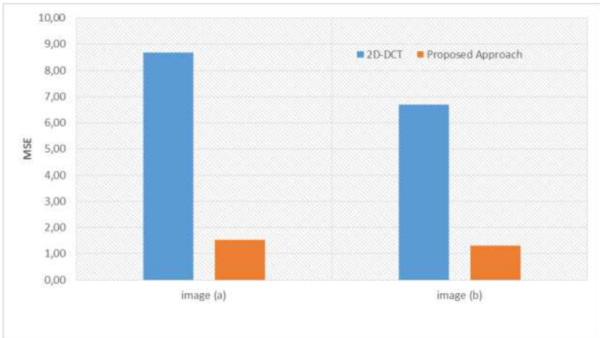


**Figure 13**. Comparative histogram of Mean Square Error

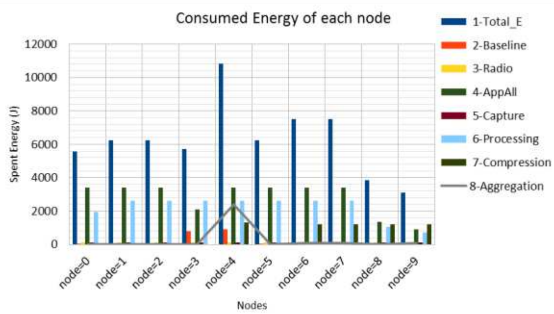


**Figure 14**. Consumed energy of each node of image (a)

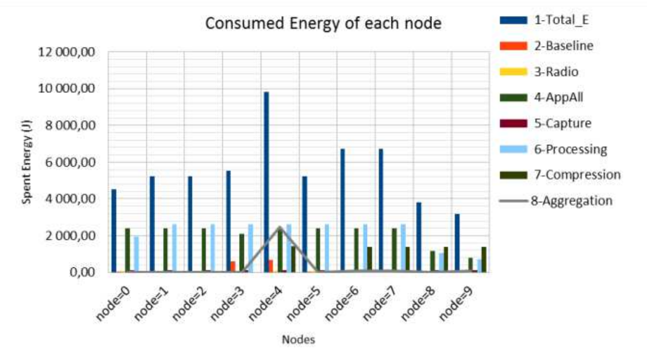


**Figure 15**. Consumed energy of each node of image (b)

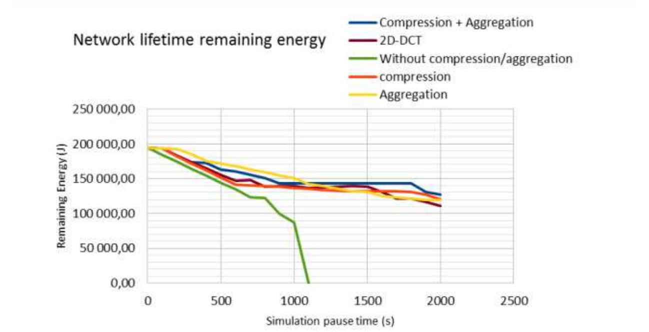


**Figure 16**. Network lifetime remaining energy of image (a)

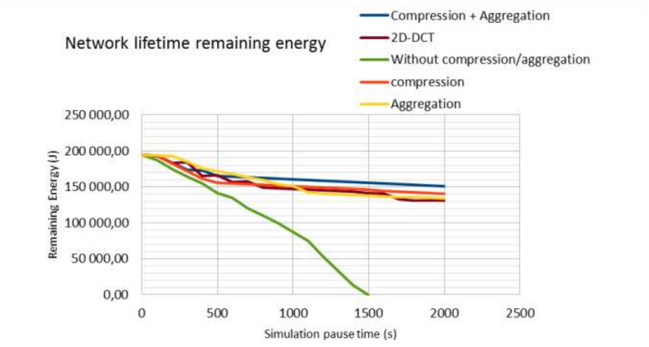


**Figure 17**. Network lifetime remaining energy of image (b)

## 5. Conclusion

This paper has described an approach of lossless data aggregation and distributed compression in low resources network platformsin WSNs. Clustering, aggregation and compressionwere used to provide an architectural framework for exploiting data correlation. The results of the proposed approach were evaluated qualitatively and quantitatively, using performance metrics such as the Compression Ratio (CR), data Aggregation Rate (AR), Peak Signal-to-Noise Ratio (PSNR) and Mean Square Error (MSE) and Energy Cost (EC). The simulation results show that the proposed approach is better than the existing methods. The advantagescan be used to handle some problems when the number of source nodes increases and when the source nodes are located relatively close to each other and far from the sink. The simulation results, though, also seems that the compression and aggregation latency could be non-negligible and should be taken into consideration during the design process.


**Acknowledgements**

We gratefully acknowledge that this work has been partially supported by the Directorate of Central Africa and Great Lakes Region (DCAGLR) of “Agence Universitaire de la Francophonie” (AUF) for the financial support of my research in the "Laboratoire de Génie Informatique et d’Automatique de l'Artois - UR 3926", University of Artois, France.

**Funding**

This work was partially supported by AUF (“AgenceUniversitaire de la Francophonie”).


**Conflicts of interest/Competing interests**

The authors declare that they have no competing interests.

**Availability of data and material**

The datasets used and/or analyzed during the current research are available from the corresponding author on reasonable request.

**Code availability**

The custom code is available from the corresponding author on reasonable request.

**Authors' contributions**

All authors participated during the design, implementation and writing of the manuscript. All authors read and approved the final manuscript.

**References**


[1] Elie T. Fute, Alain B. Bomgni & Hugues M. Kamdjou (2016). An approach to data compression and aggregation in wireless sensor networks. *International Journal of Computer Science and Telecommunications*, vol. 7, pp. 13-19.

[2] N. Kimura & S. Latifi (2005). A survey on data compression in wireless sensor networks. *International Conference on Information Technology: Coding and Computing (ITCC'05),*Volume II, Las Vegas, NV, 2005, vol. 2, pp. 8-13, doi: 10.1109/ITCC.2005.43.

[3] Kamdjou, H.M., Tagne Fute, E., El Amraoui, A. et al. (2021). The Transferable Belief Model for Failure Prediction in Wireless Sensor Networks. *SN Computer Science*,2(269), pp. 1-9.https://doi.org/10.1007/s42979-021-00654-0.

[4] M. Nava Barathy & Dejey (2020). Two-level data aggregation for WMSNs employing a novel VBEAO and HOSVD. *Computer Communications*, vol. 149, pp. 194-213.

[5] SaifZahir & Arber Borici (2012). An Efficient Block Entropy Based Compression Scheme for

Systems-on-a-Chip Test Data.*Journal of Signal Processing Systems*, vol. 69, pp.133-142.

[6] Tagne, E., Kamdjou, H., Bomgni, A. & Nzeukou (2019). An Efficient Data Compression Approach based on Entropic Coding for Network Devices with Limited Resources.*European Journal of Electrical Engineering and Computer Science*, 3(5), doi:https://doi.org/10.24018/ejece.2019.3.5.121.

[7] Uthayakumar, J., T. Vengattaraman & P. Dhavachelvan (2018). A survey on data compression techniques: From the perspective of data quality, coding schemes, data type and applications.*Journal of King Saud University - Computer and Information Sciences*, https://doi.org/10.1016/j.jksuci.2018.05.006.

[8] R. Gravina, P. Alinia, H. Ghasemzadeh & G. Fortino (2017). Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges. Information Fusion. 35(5), pp. 68-80, https://doi.org/10.1016/j.inffus.2016.09.005.

[9] Vincent Lecuire, Cristian Duran-Faundez& Nicolas Krommenacker (2007). Energy-Efficient Transmission of Wavelet-Based Images in Wireless Sensor Networks. *EURASIP Journal on Image and Video Processing,* 047345.

[10] D. Petrovic, R. C. Shah, K. Ramchandran & J. Rabaey (2003). Data Funneling: Routing with Aggregation and Compression for Wireless Sensor Networks. *Proceedings of the First IEEE International Workshop on Sensor Network Protocols and Applications*, Anchorage, AK, USA, pp. 156-162, doi: 10.1109/SNPA.2003.1203366.

[11] T. Arici, B. Gedik, Y. Altunbasak & L. Liu (2003). PINCO: a pipelined in-network compression scheme for data collection in wireless sensor networks. *Proceedings, 12th International Conference on Computer Communications and Networks (IEEE Cat. No.03EX712)*, Dallas, TX, USA, pp. 539-544, doi: 10.1109/ICCCN.2003.1284221.

[12] L. Benini, D. Bruni, A. Macii & E. Macii (2002). Hardware-assisted data compression for energy minimization in systems with embedded processors. *Proceedings 2002 Design, Automation and Test in Europe Conference and Exhibition*, Paris, France, pp. 449-453, doi: 10.1109/DATE.2002.998312.

[13] O. Younis & S. Fahmy(2004). Distributed clustering in ad-hoc sensor networks:a hybrid, energy-efficient approach. *IEEE INFOCOM 2004*, Hong Kong, pp. 640, doi: 10.1109/INFCOM.2004.1354534.

[14] W. B. Heinzelman, A. P. Chandrakasan & H. Balakrishnan (2002). An application-specific protocol architecture for wireless microsensor network. *IEEE Trans. on Wireless Communications*, vol. 1, pp. 660-670.

[15] Bhushan B., Sahoo G. (2019). Routing Protocols in Wireless Sensor Networks. In: Mishra B., Dehuri S., Panigrahi B., Nayak A., Mishra B., Das H. (eds) *Computational Intelligence in Sensor Networks*. Studies in Computational Intelligence, vol. 776, Berlin, Heidelberg,http://doi-org-443.webvpn.fjmu.edu.cn/10.1007/978-3-662-57277-1_10.

[16] Chandana Rani Kandru & Ravi SankarSangam (2019). A Survey on Routing Protocols of Wireless Sensor Networks: A Reliable Data Transfer Using Multiple Sink for Disaster Management.*In Next-Generation Wireless Networks Meet Advanced Machine Learning Applications 2019 IGI Global*, DOI: 10.4018/978-1-5225-7458-3.ch004.

[17] Sovannarith Heng, Chakchai So-In & Tri Gia Nguyen (2017). Distributed image compression architecture over wireless multimedia sensor networks. *Wireless Communications and Mobile Computing*, vol. 21.

[18] Seong-Won Lee & Ho-Young Kim (2018). An energy-efficient low-memory image compression system for multimedia IoT products. *EURASIP Journal on Image and Video Processing*, vol. 87, pp 1-15.

[19] Ian F. A., Tommaso M. & Kaushik R. C. (2007). A survey on wireless multimedia sensor networks. *Computer Networks*, 51, pp. 921-960.

[20] Robert Sedgewick (1977). Permutation Generation Methods. *Computing Surveys*, 9(2), pp. 137-164.

[21] P. Diaconis& D. Freedman (1986). An Elementary Proof of Stirling's Formula. *The American Mathematical Monthly*, 93(2), pp. 123-125, doi: 10.1080/00029890.1986.11971767.

[22] The Publications & Writings of Herbert Robbins (2003). The Annals of Statistics. *JSTOR*, 31(2), pp. 407-413, https://www.jstor.org/stable/3448399.

[23] Gao J. and Wang D. (2003). Permutation Generation: Two New Permutation Algorithms, pp.1-7, http://cds.cern.ch/record/620331.

[24] C. Pham (2015). Low cost Wireless Image Sensor Networks for visual surveillance and intrusion detection applications. *2015 IEEE 12th International Conference on Networking*, *Sensing and Control*, Taipei, pp. 376-381, doi: 10.1109/ICNSC.2015.7116066.

[25] Varga, A. (2001). The OMNeT++ Discrete event simulation system. *In Proceedings of the European Simulation Multiconference (ESM'01)*, Prague, Czech Republic, pp. 319-324.

[26] A. Kopke, M. Swigulski, K. Wessel, D. Willkomm, P. T. Klein Haneveld, T. E. V. Parker & O. W.Visser (2010). Simulating Wireless and Mobile Networks in OMNeT++ The MiXiM Vision. *1st*

*International ICST Workshop on OMNeT++*, doi: 10.4108/ICST.SIMUTOOLS2008.3031.

[27] Athanassios Boulis (2011). Castalia A simulator for Wireless Sensor Networks and Body Area Networks Version 3.2 User's Manual.*NICTA*.

[28] Koulaouzidis, et al. (2017). KID Project: an internet-based digital video atlas of capsule endoscopy for research purposes. *Endoscopy International Open*, 5(06), pp. E477-E483.

[29] Najmeh Kamyabpour & Doan B. Hoang (2011). Modeling overall energy consumption in Wireless Sensor Networks. *International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT 2011),* pp. 240-243, DOI: 10.1109/PDCAT.2010.65.

[30] N. B. Rizvandi, J. Taheri, A. Y. Zomaya & Y. C. Lee(2010). Linear Combinations of DVFS-Enabled Processor Frequencies to Modify the Energy-Aware Scheduling Algorithms. *2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing*, Melbourne, VIC, pp. 388-397, doi: 10.1109/CCGRID.2010.38.

[31] Med Lassaad Kaddachi, Adel Soudani, Vincent Lecuire, Kholdoun Torki, Leila Makkaoui, et al. (2012). Low power hardware-based image compression solution for wireless camera sensor networks. *Journal of Computer Standards and Interfaces*. 34(1), pp. 14-23, doi:10.1016/j.csi.2011.04.001.