Chain-routing scheme with compressive sensing-based data acquisition for Internet of Things-based wireless sensor networks

Ahmed Aziz1,2 | Walid Osamy1 | Ahmed M. Khedr3,4 | Ahmed Salim4,5

Abstract
The emerging Internet of Things (IoT)-based systems that integrate diverse types of sensors, mobiles and other technologies to physical world is becoming increasingly popular for use in wide varieties of applications. Compressive sensing (CS)-based information acquisition and in-network compression provide an effective method for accurate data recovery at the base station (BS) with reduced cost of communication. In this study, how CS can be combined with routing protocols for gathering data in IoT-based Wireless Sensor Networks (WSNs) in an energy-efficient manner are examined. A novel chain-routing scheme with CS based data acquisition is introduced that includes the following new algorithms: (1) Seed Estimation Algorithm (SEA) to find the best measurement matrix by selecting the best-estimated seed, (2) Chain Construction Algorithm (CCA) to organise the network nodes during transmitting and receiving process, (3) Compression approach with reduced consumption of energy that improves the lifetime of the network by minimising the local data traffic, and (4) Reconstruction Algorithm (RA) that reconstructs the original data with minimum reconstruction error. Here, extensive simulation and analysis results prove the performance of our proposed method in improving the network lifetime 35% better than the ECST algorithm and 93% better than the PEGASIS algorithm. In addition, the proposed reconstruction algorithm exceeds the other reconstruction algorithms performance.

1 | INTRODUCTION
The gaining popularity of Internet of Things (IoT) makes it an emerging field of research where researchers attempt to connect/interact everyday items to the Internet [1,2]. IoT sensors or computing networks can be embedded in living as well as inanimate objects, including clothing, food, plants, animals and vehicles. The augmentation of these objects with computational capabilities extends their functions and promises to transform numerous fields such as health, logistics, military and entertainment industries [3–6]. In last few years, Wireless Sensor Network (WSN) applications in particular surveillance, transportation and monitoring applications have attracted a number of researchers [7–10]. WSN is considered as the most crucial IoT component in which the IoT devices can communicate wirelessly [11–13]. However, limited power, storage and processor of IoT devices are considered the main challenges that hinder the development of IoT applications. Framing effective routing protocols for IoT network is a crucial aspect. The massive energy constraints of densely integrated sensor nodes motivated the development of several routing protocols for IoT-based Wireless Sensor Networks (WSNs).

Chain-based routing protocols are considered as one of the energy efficient solutions for IoT network, in which the nodes in IoT network are arranged to create a chain. A chain can be built by the WSN nodes on their own by adopting a greedy strategy initiated from any node in the network or can be built by the base station (BS) which will compute and broadcast the chain to all the nodes in the network. Chain-based IoT network can contribute a major part in numerous applications as chain-based routing protocol serves as a major algorithm for organising the sensor nodes transmission process in many IoT applications such as Internet of Vehicles (IoV) [12], health care.
based data acquisition protocols using CS, conserving energy of WSNs with transmission range and energy heterogeneity can be found in [26,29–33].

Effective clustering-based data acquisition protocols using CS, conserving the energy of WSNs with transmission range and energy heterogeneity can be found in [19–36,15]. Effective clustering-based data acquisition protocols using CS, conserving the energy of WSNs with transmission range and energy heterogeneity can be found in [26,29–33].

But the integration between CS method and routing algorithm is considered NP [26] problem and, hence, needs a lot of efforts to adopt CS with an efficient routing technique to achieve energy efficient and high performance. In addition, CS reconstruction process is not an easy process because for CS framework, the number of unknowns is bigger than the input parameters and makes it a very difficult problem.

To address the mentioned above problems, we propose an efficient scheme that makes use of the correlation between sensors readings to compress and correctly reconstruct data using CS. Furthermore, we combine the CS method with suitable chain-routing technique to minimise all-round power consumption that makes improvements to the IoT network lifetime. We adapt CS with routing design for information acquisition in IoT based-WSNs that minimises the energy consumption and thereby extending the WSN lifetime. The main contribution can be summarised as given below:

1. A new seed estimation algorithm (SEA) is proposed to select the best measurement matrix by choosing the best-estimated seed.
2. A new chain construction algorithm (CCA) is proposed for constructing the data aggregation path during the transmitting and receiving process.
3. Finally, to enhance the reconstruction process, an efficient reconstruction scheme (RA) that recovers the original data from the compressed samples is proposed.
4. Extensive simulation and performance analysis of the proposed method are provided and its effectiveness by comparing with existing baseline methods is illustrated.

The definitions of used notations in this paper are provided in Table 1.

The rest of this paper is structured as provided below:

The related work is explained in Section 2. Section 3 provides CS background. Then, Section 4 explains the proposed scheme and Section 5 presents the simulation results of our proposed approach. Finally, Section 6 concludes our research work.

## Table 1 Description of notations

| Notation | Description                  |
|----------|-----------------------------|
| \( \phi \) | Measurement matrix          |
| \( d_i \) | Sensor i's reading          |
| \( \psi \) | Transform matrix            |
| \( \theta \) | \( \theta = \phi \psi (M \times N \text{ matrix}) \) |
| \( y \) | Measurements vector         |
| \( M \) | Number of measurements      |
| \( g \) | Sparse representation       |
| \( s \) | Sparse level                |
| \( \varepsilon \) | Global seed                 |
| \( L \) | Number of rounds            |
| \( a_i \) | Node i's coefficients vector|
| \( y_i \) | The compressed vector for node i |
| \( r^2 \) | The residue after L iterations |
| \( U \) | Candidate set               |
| \( D \) | The approximation solution set |
| \( Z \) | Least square signal approx set |
| \( A \) | Estimated set               |
| \( L_{\text{max}} \) | Maximum number of iterations |
| \( H \) | Selection step size         |
| \( Y \) | Remove step size            |
| \( \delta \) | Termination parameter       |

## 2 RELATED WORK

Low Energy Adaptive Clustering Hierarchy (LEACH) [37] is the widely popular example of cluster-based routing in WSN, where the WSN is partitioned into a set of clusters. One node from each cluster is chosen as a Cluster Head (CH) and the other nodes act as Cluster Members (CMs). The CMs in each individual cluster have to transmit the sensed data to its corresponding CH, and the CH will then forward the aggregated data to the BS. LEACH protocol improves the performance of WSNs in comparison with the direct transmission method. The energy difference between CH and its CMs in LEACH protocol is the main problem in LEACH and it is minimised in many later works such as Intra-Balanced LEACH (IB-LEACH) as presented in [38]. For higher rate of data collection in WSNs, the authors of [39,40] presented a better technique called Power-Efficient Gathering in Sensor Information Systems (PEGASIS). The basic concept behind PEGASIS is to create a chain-list[i] of nodes in WSN. PEGASIS has less energy consumption and better performance than LEACH. However, most of the current routing algorithms do not meet the needs of the huge data traffic in IoT-based
WSNs. For that reason, it is always advisable to use some kind of compression techniques before transmitting data for reduction in size of transmitted data and total power consumption by a sensor node.

The key concept of these protocols is to incorporate the traditional routing protocol with CS to enhance the lifetime, stability period and energy efficiency of homogenous/heterogeneous WSNs. An approach integrating CS with Random Walk (RW) to minimise energy utilisation in WSNs is exploited in [25], in which a predefined length-based RW routing is used to collect each CS measurement and all the random CS measurements are transferred either directly or through intermediate nodes to BS for CS recovery. The reconstruction of all the sensory data at the BS is done based on lesser count of CS measurements compared to the total count of WSN nodes. An improved RW technique for mobile collector is proposed in [27] by exploiting Kronecker Compressive Sensing (KCS). A small subset of nodes is chosen randomly along a random routing path of the mobile collector and the mobile data collector collects the temporal-compressive measurements from them. A hybrid CS method integrating conventional schemes of data gathering with CS is adopted for data aggregation in [28], where the compressed data is rapidly collected over the backbone using the pipelined scheduling approach. Compressive Data Gathering (CDG) [21] is the pioneer research that acquainted CS to WSNs. It blends routing technique with CS so as to reduce the total consumption of energy. The crucial issue with CDG is that it explains the CS idea without analysing. The efficient CS-based routing technique (ECST) is suggested in [18] which prolongs the network lifetime using the CS method to compress sensors reading before sending to BS. The combination of CS and tree routing protocol to minimise the total forwarding energy consumption is proposed in [20,41]; however, it increases the power consumed by leaf and intermediate nodes. A combination of tree routing protocol and CS in a hybrid way is proposed in [19] where only parent nodes do CS operation. It is suitable for small networks, but not for large ones. In [42], the authors proposed the CS scheme for collecting data in large-scale WSNs. A scheme to detect abnormality in sensor readings and to provide the CS reconstruction that is capable of overcoming the abnormality in sensed data adaptively by exploiting the correlation between different nodes data is proposed in [43]. The work in [44] eliminated the need of choosing sparsity weightings priori which leads to improve reconstruction process. A non-uniform compressive sensing method-based heterogeneity of WSN is proposed in [45]. The proposed work in [46] reduced the energy consumption by utilising the sparse random measurement matrix. The work proposed in [47] is the first work that integrates CS with IoT from the viewpoint of data-compressed sampling. It applies CS without organising the sensors to send or receive the data from and to BS. CS technique convenient for data collection in large-scale WSNs is proposed in [20]. Signals data correlation is utilised to reduce the sampling ratio in [48]. By exploiting correlation between different nodes data, a new method to reconstruct lost data is proposed in [49].

All above CS-based schemes have achieved efficient performance in terms of data reduction and extend the network lifetime. However, they have the following limitations:

1. The impact of seed selection process in improving the CS performance is not studied.
2. No one proposed a complete CS scenario, that is, they focussed mainly on adapting CS compression process, however, they did not give careful consideration to the CS reconstruction process.

To address these drawbacks, in this work, a novel CS scheme is proposed in which a new Seed estimation algorithm (SEA), Chain-Based routing algorithm (CCA) and CS reconstruction algorithm (RA) are proposed. The advantages and disadvantages of the CS schemes mentioned above are summarised in Table 2.

## 3 | COMPRESSION SENSING BACKGROUND

As an alternative to conventional compression techniques where sampling and compression are performed in successive steps, CS offers a direct approach that facilitates data compression and sampling in a single step [38]. Moreover, the CS reconstruction algorithm provides successful reconstruction of the original data collected from sensors without any prior knowledge, from the compressed samples [51,52].

### 3.1 | Mathematical definition

Let the set of sensor nodes reading vector collected be denoted by \( x[n] \), where \( n = 1, \ldots, N \), and \( N \) represents the count of sensor nodes. Any individual signal in \( \mathbb{R}^N \) can be expressed using basis of \( N \times 1 \) vectors \( \{\Psi_i\}_{i=1}^N \) Employing the basis \( N \times N \) matrix \( \Psi = [\Psi_1 \mid \Psi_2 \mid \Psi_3 \mid \ldots \mid \Psi_N] \) together with the vectors \( \Psi_i \) being the columns, we can depict the signal \( x \) as given in Equation (1) [36]:

\[
x = \sum_{i=1}^{N} g_i \Psi_i
\]  

(1)

where the \( \Psi \in N \times N \) corresponds to the orthonormal transform matrix and \( g \) represents the \( N \times 1 \) sparse presentation of \( x \). CS focusses on signals with a sparse representation. That means, the signal \( x \) has \( S \) basis vectors, \( S < N \). That is, \((N - S)\) values are zeros and only \( S \) values of \( g \) are non-zeros. By using Equation (1), the compressed samples \( y \) (compressive measurements) can be derived from Equation (2) as

\[
y = \Phi x = \Phi \Psi g = \theta g
\]  

(2)

Here, the compressed samples vector \( y \in \mathbb{R}^M \), with \( M << N \) and \( \theta \) is \( M \times N \) matrix.
The challenge of deriving solutions of an undetermined set of linear equations has attracted the researchers and, as a result, different practical methods were introduced to deal with this challenge. In the CS approach, the main responsibility is to offer an efficient reconstruction method enabling the recovery of a large and sparse signal with the help of a few available measurement coefficients. The reconstruction of signal using this available incomplete set of measurements is really challenging and relies on the sparse representation of signal. An easiest approach for recovering the original inherent sparse signal using its small set of linear measurements as shown in Equation (2) is to compute the number of non-zero entries obtained by solving \( \|L_0\| \) minimisation problem. The reconstruction problem can, thus, be denoted as

\[
x = \text{argmin} \|x\|_0 \quad \text{subject to } y = \Phi x
\]

The \( \|L_0\| \) minimisation problem works well in theoretical aspects. But, in general, the problem is NP-hard \([53,54]\) and, hence, Equation (3) is computationally intractable for any vector or matrix. Alternatively, two solutions exist in the framework of CS to efficiently solve Equation (2). The first solution is using the Convex Relaxation-based optimisation, leading to \( \|L_1\| \) minimisation \([55]\) and the second solution is using the Greedy Algorithm, for example, Orthogonal Matching Pursuit (OMP) \([56]\), Stage-wise OMP (StOMP) \([57]\) and Regularised OMP (ROMP) \([58]\).
4 | THE PROPOSED SCHEME

Reducing network’s energy consumption, designing effective sensor data aggregation technique and managing large amount of information are considered to be the major challenges faced by IoT-based WSNs, which consist of interconnected sensor nodes. To address these problems, we integrate CS framework with an efficient routing technique. The proposed scheme consists of the following: (1) data compression and aggregation phase and (2) reconstruction phase. Figure 1 shows the block diagram of the proposed scheme. We discuss the detailed description of the proposed method in the following sections.

4.1 | Data compression and aggregation phase

The data compression and aggregation phase consists of three processes: Seed Estimation, Sensors Organisation, and Data Acquisition and Compression.

I. Seed Estimation Process: As already mentioned, the measurement matrix used in the CS method is a random matrix. It is generated using seed and is used in compression and reconstruction process. BS generates the best estimated seed and uses it to generate the random matrix. Every sensor then uses this seed to compress its data; while the BS uses it to reconstruct the original data. The SEA is described as follows:

**SEA:** To create sensing matrix $\Phi$, a global seed $\varepsilon$ is generated and broadcasted to all nodes by BS.

Each node $j$, $j = 1, 2, \ldots, N$ that receives $\varepsilon$ generates the series of corresponding coefficients $a_j = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix}$. BS stores the value $\varepsilon$ and all the nodes identifications, and re-generates these coefficients for reconstruction process. It is clear that the global seed selection is very important factor in building the measurement matrix, which the sensors use to compress their sensed data and the BS uses to reconstruct/

**Algorithm 1 Seed Estimation Algorithm (SEA)**

| **Input:**     | Compressed Samples From the Closest Nodes $P$ |
|---------------|---------------------------------------------|
| **Output:**   | Global Seed                                 |
| **1:**        | BS multi-casts a new random seed to a set of closest nodes denoted by $P$ |
| **2:**        | for each node $i \in P$ do                 |
|               | 1. generate seed using $\varepsilon$       |
|               | 2. generate the corresponding series of coefficients $a_i$ |
|               | 3. compute the measurement $y_i = a_i d_i$ |
|               | 4. transmit the measurement $y_i$ to the BS |
| **3:**        | end for                                     |
| **4:**        | BS collects the measurements $y_j$, $\forall j \in P$, |
| **5:**        | BS generates the measurement matrix $\Phi$ using global seed $\varepsilon$ |
| **6:**        | BS calculates the reconstruction error.     |
| **7:**        | if reconstruction error is not satisfied then |
| **8:**        | Go To Step 1                                |
| **9:**        | Else                                        |
| **10:**       | BS starts the aggregation process based on this seed. |
| **11:**       | **end if**                                  |

II. Sensors Organisation Process: In this process, we explain the Chain Construction Algorithm (CCA). The main objective of CCA is to organise the network sensors into Chain List (CL) and then send CL information to the entire network. The outlines of the CCA is as follows:
Chain construction algorithm:

1. BS initialises CL with the nearest node $c_0$.
2. BS updates CL with the nearest unselected node $c_1$ to node $c_0$.
3. BS then selects the position of next nearest unselected neighbour node $c_i$ according to the following scenario:
   a. $c_i$ will be added to the end of CL, if the distance between $c_i$ and $c_j$ is smaller than the distance between $c_i$ and any consecutive nodes in CL.
   b. Otherwise, $c_i$ will be inserted between the consecutive pair that has minimum distance to $c_j$.

For example, if $c_j$ and $c_k$ are consecutive pair of nodes in CL and if $\text{dist}(c_i, c_{\text{last}}) > \text{dist}(c_i, c_j) \quad \text{and} \quad \text{dist}(c_i, c_{\text{last}}) > \text{dist}(c_i, c_k)$, then node $c_i$ will be inserted between $c_j$ and $c_k$; otherwise $c_i$ will be added to the end of CL after node $c_{\text{last}}$, where $c_{\text{last}}$ is the last added node to CL and $\text{dist}(c_i, c_k)$ is the distance between $c_i$ and $c_k$.

4. BS repeats the previous update process to include all nodes in CL.

Unlike ECST [14] that adds the nearest node to the end of CL without considering the other nodes in CL, the proposed algorithm rearranges CL with every added node to the CL which minimises the overall power consumption in transferring data per round.

III. Data Acquisition and Compression Process: This process has two tasks: header selection and adaption of CS method at each sensor node in the chain to compress its data.

a. **Header selection:** In each round $r$, the position of CL header will be randomly selected. Any node in CL can be a CL header whenever the identification number becomes $r \mod N$.

b. **CS Adaptation:** CS reduces the total data traffic and the energy consumption of each node.

To combine sensors readings, every $c_i$ in the CL, $i = \{1, 2, \ldots, N\}$ uses to generate $a_{ci}$, computes its compress vector (measurement) $\mathbf{y}_{ci} = a_{ci} \mathbf{d}_{ci}$, where $\mathbf{d}_{ci}$ is the reading of $c_i$, and transmits $\gamma_{ci}$ to its neighbour node in CL (node $c_{i+1}$).

$c_{i+1}$ computes its measurement $\gamma_{ci+1} = a_{ci+1} \mathbf{d}_{ci+1}$, and then transmits the summation vector $\sum_{i=1}^{N} \gamma_{ci}$ to the next node $c_{i+2}$.

Once $c_{i+2}$ receives this value, it does the same computation as before and sends the joined value to the next node and so on until the CL header.

Finally, the CL header sends the packet containing the collection of all readings in the CL to the BS. The header node can be located at one of the following positions:

**Case 1 CL header located at the end of CL ($c_N$):**
- CL header sends an identification message to the first node $c_1$.
- $c_1$ computes and transmits $\gamma_{c_1}$ to node $c_2$.
- $c_2$ computes $y_{c_2}$ and transmits $\gamma_{c_1} + y_{c_2}$ to node $c_3$ and so on till header.
- Finally, CL header computes $\sum_{i=1}^{N} \gamma_{ci}$ using the received data from $c_{N-1}$ and sends it to BS.

**Case 2 CL header located at the beginning of CL ($c_1$):**
- CL header passes an identification message to the last node of the CL, that is, $c_N$.
- $c_N$ computes and transmits measurement $\gamma_{c_N}$ to node $c_{N-1}$.
- $c_{N-1}$ computes the measurement $\gamma_{c_{N-1}}$ and then transmits $\gamma_{c_N} + y_{c_{N-1}}$ to $c_{N-2}$, and so on till CL header.
- Finally, CL header $c_1$ computes $\sum_{i=1}^{N} \gamma_{ci}$ using the data received from $c_2$ and sends it to BS.

**Case 3 CL header located at position $j$: $1 < j < N$ of CL:**
- CL header passes an identification message to the first node $c_1$ as well as to the last node $c_N$.
- $c_1$ computes and transmits $\gamma_{c_1}$ to $c_2$.
- $c_2$ computes $\gamma_{c_2}$ and transmits $\gamma_{c_1} + y_{c_2}$ to $c_3$ and this process continues till CL header is at position $j$.
- At the same time, node $c_N$ computes and transmits $\gamma_{c_N}$ to $c_{N-1}$. Then node $c_{N-1}$ computes the measurement $\gamma_{c_{N-1}}$ and transmits $\gamma_{c_N} + \gamma_{c_{N-1}}$ to $c_{N-2}$, and this process continues till CL header is at position $j$.
- Finally, the CL header node computes $\sum_{i=1}^{N} \gamma_{ci}$ using the received data from $c_{j-1}$ and $c_{j+1}$ and sends it to the BS.

Once the compressed data from the CL header is received at the BS, the original data is reconstructed from the compressed data at BS making use of the reconstruction algorithm.

4.2 | Reconstruction phase

A new reconstruction algorithm is proposed in this phase that reconstructs the sensors readings. It is a greedy algorithm and it consists of two processes. In the first process (Selection and Estimation), the estimated set $U$ will be updated by adding $\eta$ (where $\eta$ is called the Selection step size) columns from the CS matrix. The columns correspond to the largest (best) entries in the least-squares signal approximation set $Z$ from $\Phi$ as the correct solution of least squares problem, $Z = \Phi^\dagger \psi^{l-1}$.

In the second process (Check and Remove), the estimated set $U$ is clipped by removing $\nu$ columns which were wrongly chosen in the Selection and Estimation process. This accurately identifies the true support of $U$, where $\nu$ is called the Removable step size. Algorithm 2 provides our proposed reconstruction algorithm.

**Algorithm Description:** Our algorithm includes three steps: Selection and Estimation, Check and Remove, and Update.

1. Parameter Initialisation: residue $\rho_0 = y$, approximation $D^0 = 0$, Estimated set $\mathbf{A} = \phi$, candidate set $U = \phi$ and number of rounds $L = 0$ (Algorithm 2, Lines 3-8).
2. **Selection and Estimation**: In this step, set \( U \) is updated by adding the \( \eta \) \( \left( \eta = \frac{M}{2} \right) \) largest elements from the solution \( Z \) of the least squares problem \( (Z = \Phi \hat{\Phi}) \). Then, extend set \( A \) by adding the components of \( U \) to the approximation set \( D_{A}^{-1} \) (Algorithm 2, Lines 11-13).

3. **Check and Remove**: In this step (correction step), the proposed algorithm removes the incorrect columns indices which were wrongly selected by **Selection and Estimation step**, that is, the algorithm updates the approximation set \( D_{A} = R_{M/3} \) by removing \( u = M/3 \) columns indices which have the smallest values in set \( R \) (Algorithm 2, Lines 14-16).

4. **Update**: The samples are updated as \( r' = y - \Phi D_{A} \) and it will be terminated in two situations: (1) the residue \( r' \left\Vert r' \right\Vert \) is smaller than the termination parameter \( \delta \). The termination parameter selection process depends on the level of noise (Algorithm 2, Lines 17-18). (2) The maximum limit of iterations has reached, that is \( L_{\text{max}} \) (for e.g., \( L_{\text{max}} = M \)). At the end of the algorithm, \( D_{A}^{l} \) contains the corresponding non-zero values.

**Algorithm 2**: Reconstruction Algorithm (RA)

1: **INPUT**: Measurement matrix \( \Phi : m \times N \); \( y \): compressed vector

2: **OUTPUT**: approximate solution \( D_{A}^{l} \) of the original data

Algorithm parameter’s initialisation:

3: \( \rho^{0} = y \) (the residual initialisation)

4: \( D_{A}^{0} = 0 \) (the approximation set initialisation)

5: \( A = \Phi \) (the estimated set initialisation)

6: \( U = \Phi \) (candidate set initialisation)

7: \( Z = \Phi \) (least-squares signal approximation set)

8: \( L = 0 \) (number of rounds)

9: **while** \( \text{true} \) **do**

10: \( L = L + 1 \) (round increment)

11: **Selection and Estimation Step**:

12: \( Z = \Phi \hat{\Phi} \) (solve the least square problem)

13: \( U = \text{supp}(Z) \) \( \left( \text{Select } \eta = \frac{M}{2} \text{columns} \right) \)

14: \( A = U \cup \text{supp}(D_{A}^{-1}) \) (Union of old and new solution)

Check and Remove Step:

15: \( R \mid A = \Phi \hat{\Phi} \) \( y \) (find the first solution)

16: \( R \mid A' = 0 \) (fill the index of non-selected column index with zero)

17: \( D_{A}^{l} = R_{M/3} \) (retaining only \( M/3 \) columns)

Update Step:

18: \( r' = y - \Phi D_{A}^{l} \)

19: \( \text{if } \left\Vert r' \right\Vert_{2} < \delta \) or \( L = L_{\text{max}} \) **then**

20: \( \text{Break} \)

21: **end while**

5 | SIMULATION RESULTS

We provide the simulation results of our algorithm in this section and analyse the overall performance. This section is divided into two parts. In the first part, we evaluate data compression and aggregation phase in terms of average power consumption as well as lifetime of the network (first node dies). In the second part, we analyse our proposed reconstruction algorithm (Reconstruction phase) and evaluate the performance with existing baseline algorithms: OMP [56], ROMP, ROMP [58], Forward-Backward Pursuit (FBP) [59] and Subspace Pursuit (SP) [13].

5.1 | Evaluation: data compression and aggregation

The environment setup is as follows: \( 100 \times 100 \) network region size with sensor nodes range between 50 and 200 with an increment of 50. BS is positioned at location \( (x = 50, y = 50) \). The energy parameters are similar to [37]. The radio-power consumption to send \( l \) – bits message to a distance \( d \) will be

\[
E_{T_{l}}(l, d) = \begin{cases} 
1E_{\text{elec}} + l\epsilon_{l}d^{2} & d < d_{0} \\
1E_{\text{elec}} + l\epsilon_{mp}d^{4} & d \geq d_{0}
\end{cases}
\]

and the energy expended on receiving a message will be

\[
E_{R_{l}}(l) = lE_{\text{elec}}
\]

where \( E_{\text{elec}} = 50 \text{nJ/bit}, \epsilon_{l} = 10\text{pJ/bit}m^{2}, \epsilon_{mp} = 13/10,000 \text{pJ/bit}m^{4}, d_{0} = \sqrt{\frac{G_{b}}{\epsilon_{mp}}} \), and node’s initial energy = 2].

**Performance Metrics**: The given below are the metrics used to evaluate and compare our proposed approach with PEGASIS [39] and ECST [18]:

- **Average Energy Consumption**: gives the average consumed energy by WSN nodes during their operations like forwarding operation, sending and receiving. The average energy-consumption per round is calculated as given below:

\[
E = \frac{\sum_{i=1}^{N} E_{i}(r)}{r}
\]

where \( N \) is the nodes count and \( r \) is the round number.

- **Network lifetime**: measures the lifetime until the first node dies.

Figure 2 shows the average energy consumption per round. We can easily notice that our proposed algorithm has lower...
energy consumption in comparison with that of PEGASIS and ECST. The reasons are as follows. PEGASIS did not consider any data compression scheme in the data aggregation process. Even though ECST used the CS-based data acquisition scheme, it appended the nearest node to the end of CL without considering the other nodes in the CL. However, our proposed scheme rearranges CL with every added node using the CCA algorithm which reduces the total energy consumed during data transfer in each round. Moreover, the proposed scheme uses the best CS matrix based on the proposed SEA which improves the compression performance when compared with the ECST algorithm.

Figure 3 gives the lifetime of the network for all the evaluated protocols. It illustrates that our proposed algorithm extends the network lifetime than ECST and PEGASIS because the proposed algorithm reduces the power consumption for each node by compressing its data before sending, which is not considered in the PEGASIS algorithm. Also, unlike ECST, the proposed CCA algorithm which rearranges CL with every added node minimises the overall power
consumption in transferring data per round, which in turn increases the network lifetime.

5.2 Evaluation: reconstruction performance

Here, we evaluate the overall performance characteristics of our reconstruction technique and compare it with the performance of OMP, ROMP, SP and FBP algorithms (with Selection size = 0.5*M and Removable size = forward step size – 0.3*M).

First, we test the proposed algorithm with the signals obtained from 54 sensor nodes placed in Intel Berkeley Research Lab [61]. The experiments also cover computer-generated signals reconstruction for diverse non-zero coefficient distributions, which include Gaussian and Uniform distributions along with binary non-zero coefficients. The reconstruction performance is evaluated using Gaussian and Bernoulli observation matrices.

Second, we evaluate our reconstruction algorithm's performance in the presence of noisy observations and then the results are compared with FBP, SP, OMP and ROMP algorithms. For the evaluation, we used the metric: Average Normalised Mean Squared Error (ANMSE) which represents the reconstruction accuracy of the algorithms and is measured as the average ratio of the norm of reconstruction error for 500 test-samples.

We make use of specific Observation Matrix following Gaussian distribution having zero mean and standard deviation (S.D.) with 1/N for every test sample.
Figures 6–9 show that our algorithm succeeds to achieve high performance in recovering temperature and humidity signals.

Figure 10 provides the results for relative recovery error distribution meant for the sixth node, for temperature signals.

The results clearly depict that the performance of the proposed algorithm exceeds the other greedy algorithms: FBP, SP, OMP and ROMP because the selection of columns in the proposed algorithm depends on solving least square problem which finds the correct columns in each round.

**Performance evaluation using different coefficient distributions:** Here, we executed three tests: (i) The first test utilises uniform sparse data in which the non-zero values are collected from the uniform distribution $U[-1, 1]$. Figure 11 shows that the proposed algorithm provides lower ANMSE than FBP, SP, ROMP and OMP.

(ii) The second test utilises Gaussian sparse values where the non-zero entries are taken from standard Gaussian distribution. Figure 12 depicts that our proposed algorithm is considerably superior in reconstruction performance than OMP, ROMP, SP and FBP.

(iii) Finally, the third test utilises sparse binary vectors and the non-zero coefficients are chosen according to Equation 1. Figure 13 shows that the proposed algorithm achieves overwhelming success over OMP, ROMP, SP and FBP methods in this case too.

In all the three cases, the proposed algorithm exhibits higher performance than the existing algorithms because the selection of columns in the proposed algorithm depends on solving least square problem which finds the correct columns in each round.

**Performance evaluation under different observation lengths**

In the previous experiments, the observation length $M$ (which is the compressed samples size) is fixed and we change the sparsity level. One more important test is to evaluate the
reconstruction capability by changing the observation length $M$ and fixing the sparsity ratio $S$.

Figure 14 depicts the reconstruction performance over $M$ ($M$ ranges from 50 to 130 and increments by 5 and $S = 25$). $\Phi$ is the Gaussian Measurement matrix. It can be noted easily that the ANMSE of our proposed method is comparatively lower than SP, ROMP, OMP and FBP algorithms. In Figure 15, we replicate the last scenario with $\Phi$ taken from

**FIGURE 11** The results of reconstruction versus sparsity level for uniform sparse signals in proposed, SP, ROMP, OMP and FBP algorithms.

**FIGURE 12** Reconstruction results versus sparsity level for Gaussian sparse signals in SP, ROMP, OMP and FBP algorithms.
**FIGURE 13** Reconstruction results versus sparsity level for binary sparse signals in proposed, SP, ROMP, OMP and FBP

**FIGURE 14** Evaluation of performance under different observation lengths for binary sparse signals ($S = 25$) with single Gaussian Observation Matrix for every $M$
FIGURE 15 Evaluation of performance under different observation lengths for binary sparse signals ($S = 25$) with single Bernoulli Observation Matrix for every $M$.

FIGURE 16 ANMSE results for proposed, OMP, ROMP, SP and FBP algorithms in reconstructing uniform sparse signals under the noisy observations.
Bernoulli distribution. In this case also, our proposed algorithm has lower ANMSE than SP, OMP, FBP and ROMP algorithms.

Reconstruction performance for noisy observations: here, the sparse signals reconstruction has been simulated using noisy observations, where $y = \Phi x + n$, usually acquired through contamination with the White Gaussian Noise component (WGN) $n$ ($n = 10^{-4}$).

Figures 16 and 17 provide the reconstruction error for noisy uniform and binary sparse signals. It shows that our proposed algorithm has less error when compared to OMP, ROMP, SP and FBP.

We can conclude the performance results presented in this section as the following: the proposed reconstruction algorithm improves the forward selection step by solving the least square problem instead of selecting the CS columns based on the inner product process as done in all other algorithms. This helps to improve the selection process and improve the overall reconstruction performance when compared to others.

6 CONCLUSION

It is necessary that a routing technique for IoT-based WSNs should guarantee good performance in handling huge data transmissions in IoT network which comprises a large number of sensor nodes with constrained resources. As a solution, a novel chain-routing scheme with CS-based data acquisition in IoT-based WSNs by combining the advantages of CS in data size reduction and chain-based routing scheme for effective delivery of gathered data to BS are proposed. In contrast to existing schemes, the proposed algorithm operates as follows. The first phase allows the BS to select the best seed of CS matrix by introducing SEA. Then, the proposed Chain CCA is executed in which the CL is formed, and the sensor nodes select the chain header and apply the CS method to compress their data. In the second phase, an efficient reconstruction algorithm is proposed that reconstructs the original data successfully at BS. The results of simulation prove that the proposed method decreases the power consumption, increases the WSN lifetime, and, at the same time, minimises the reconstruction error.

REFERENCES

1. Celesti, A., et al.: An IoT cloud system for traffic monitoring and vehicular accidents prevention based on mobile sensor data processing. IEEE Sensor J. (2017)
2. Mieyeville, F., et al.: Wireless sensor networks for active vibration control in automobile structures. Smart Mater. Struct. 7, 075009 (2012)
3. Lay-Ekuakille, A., et al.: Supervised and characterized smart monitoring network for sensing environmental quantities. IEEE Sensor J. 17(23), 7812–7819 (2017)
4. Zheng, J., et al.: The internet of things. IEEE Commun Mag. 49(11), 3031 (2011)
5. Palopoli, L., Passerone, R., Rizano, T.: Scalable offline optimization of industrial wireless sensor networks. IEEE Trans. Ind. Inf. 7(2), 328–329 (2011)
6. Singh, S.P., Urooj, S., Ekuakille, A.I.: Breast cancer detection using PCPCET and ADEWNN: a geometric invariant approach to medical X-ray image sensors. IEEE Sensor J. 16(12), 4847–4855 (2016)
7. Haupt, J., Bajwa, W., Rabbat, M.: Compressed sensing for networking data. IEEE. 603–611 (2012)
8. Akyildiz, I., et al.: A survey on sensor networks. IEEE. 40(8), 5102–5114 (2002)
9. Yick, J., Mulherjee, B., Ghosal, D.: Wireless sensor network survey. Comput. Netw. 52(12), 2292–2330 (2008)
10. Khedr, A.M., Osamy, W., Agrawal, D.P.: Perimeter discovery in wireless sensor networks. J. Parallel Distrib. Comput. 69(11), 922–929 (2009)
11. Soundararajan, R., et al.: Secure and concealed watchdog selection scheme using masked distributed selection approach in wireless sensor networks. IET Commun. 14(6), 948–955 (2020)
12. Tharang Dandala, T., Krishnamurth, V., Alwan, R.: Internet of vehicles (10 V) for traffic management. IEEE International Conference on Computer, Communication and Signal Processing (ICCCSP), 1–4 (2017)
13. Smail, O., et al.: ESR: energy aware and stable routing protocol for WBAN networks. In: 2016 International Wireless Communications and Mobile Computing Conference (IWCMC), Paphos, 452–457, (2016)
14. Banjara, M., Hong, J.: Energy-inefficient unusual chain length clustering for wireless sensor networks in smart cities. Wireless. Commun. Mob. Com. 12, (2017)
15. Omar, D.M., Khedr, A.M.: ERPLBC: energy efficient routing protocol for load balanced clustering in wireless sensor networks. Ad Hoc Sens. Wirel. Net. 42, 145–169 (2018)
16. Orazo, R., Candes, E.J., Sodickson, D.K.: Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components. Magn. Reson. Med. 73(3), 1125–1136 (2014)
17. Abo-Zahhad, M.: Compression of ECG signal based on compressive sensing and the extraction of significant features. Int. J. Commun. Netw. Syst. Sci. 8(5), 97–117 (2015)
18. Aziz, A., Salim, A., Osamy, W.: Adaptive and efficient compressive sensing based technique for routing in wireless sensor networks. Proc. INTHITEN 8, 3–4 (2015)
19. Luo, J., Xiang, L., Rosenberg, C.: Does compressive sensing improve the throughput of wireless sensor networks? Proc. IEEE Intl Conf. Comm. 1–6 (2010)
20. Xiang, L., Jun, L., Athanasios, V.: Compressed data aggregation for energy efficient wireless sensor networks. Aerospace Conf. Proc. IEEE. 3, 1125–1130 (2002)
21. Chong, L., et al.: Compressive data gathering for large-scale wireless sensor networks. In: Proceedings of the 15th Annual International Conference on Mobile Computing and Networking, MobCom ’09, New York, NY, USA. ACM, pp. 145–156 (2009)
22. Lan, K.C., Wei, M.Z.: A compressibility-based clustering algorithm for hierarchical compressive data gathering. IEEE Sensor J. 17(8), 2550–2562 (2017)
23. Luo, C., Wu, F., Sun, J.: Efficient measurement generation and pervasive sparsity for compressive data gathering. IEEE T. Wirel. Commun. 9(12), 3728–3738 (2010)
24. Xiang, L., Luo, J., Vasilakos, A.: Compressed data aggregation for energy efficient wireless sensor networks, 2011 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks, pp. 46–54, Salt Lake City, UT (2011). https://doi.org/10.1109/SAHCCN.2011.5984032
25. Nguyen, M.T., Teague, A.K.: Compressive sensing based random walk routing in wireless sensor networks. Ad Hoc Netw. 54, 99–110 (2017)
26. Li, Y.: Compressed sensing in multi-hop large-scale wireless sensor networks based on routing topology tomography. IEEE Access. 6, 27637–27650 (2018)
27. Zheng, H., et al.: Spatial-temporal data collection with compressive sensing in mobile sensor networks. Sensors 17, 2575 (2017). https://doi.org/10.3390/s17112575
28. Yu, X., Baek, S.J.: Joint routing and scheduling for data collection with compressive sensing to achieve order-optimal latency. International Journal of Distributed Sensor. 13(10), (2018)
29. Zhang, D., et al.: A kind of effective data aggregating method based on compressive sensing for wireless sensor network, EURASIP J. Wirel. Commun. Netw. 159, (2018). https://doi.org/10.1186/s13638-018-1176-4
30. Nikam, K., Mohani, S.P.: IoT Based greenhouse monitoring using data compressive sensing protocol in WSN: a review Kranti. Int. J Innov. Res. Comp. Commun. Eng. 6(2), 1234–1237 (2018)
31. Khedr, A.M.: Effective data acquisition protocol for multi-hop heterogeneous wireless sensor networks using compressive sensing. Algorithms. 8, 910–928 (2015). https://doi.org/10.3390/a8040910
32. Omar, D., Khedr, A.M., Agrawal, D.P.: Optimized clustering protocol for balancing energy in wireless sensor networks. Int. J Commun. Network Inf. Secur. 9(3), 367–375 (2017)
33. Khedr, A.M., Omar, D.M.: SEP-CS: Effective routing protocol for heterogeneous wireless sensor networks. Ad Hoc Sens Wirel Net. 26, 211–232 (2015)
34. Osamy, W., et al.: Cluster-tree routing scheme for data gathering in periodic monitoring applications. IEEE Access, 6, 77372–77387
35. Osamy, W., Salim, A., Khedr, A.M.: An information entropy based-clustering algorithm in heterogeneous wireless sensor networks. Wireless Network (2018) https://doi.org/10.1007/s11276-018-1877-y-0123456789
36. Aziz, A., et al.: Effective algorithm for optimizing compressive sensing in IoT and periodic monitoring applications. J Netw. Comput. Appl. 126, 12–28 (2019)
37. Handy, M.J., Haase, M., Timmermann, D.: Low energy adaptive clustering hierarchy with deterministic cluster-head selection, 4th International Workshop on Mobile and Wireless Communications Network, pp. 368–372. Stockholm, Sweden (2002). https://doi.org/10.1109/MWCN.2002.1045790
38. Salim, A., Osamy, W., Khedr, A.M.: IBLEACH: intra-balanced LEACH protocol for wireless sensor networks. Wireless Netw. 20(6), 1515–1525 (2014)
39. Lindsey, S., Raghavendra, C.: PEGASIS: power-efficient gathering in sensor information systems. IEEE. 3, 1125–1130 (2002)
40. Ali, S.A., Refaay, S.K.: Chain-based routing protocol. Int. J. Comp. Sci. Iss. 8, 694–8104 (2011)
41. Luo, C., Wu, F., Sun, J: Efficient measurement generation and pervasive sparsity for compressive data gathering. IEEE T. Wirel. Commun. 9(12), 3728–3738 (2010)
42. Haupt, J., Bajwa, W., Rabatt, M.: Compressed sensing for networked data. IEEE Signal Process Mag. 25(2), 92–101, (2008)
43. Jin, W., et al.: Data gathering in wireless sensor networks through intelligent compressive sensing. Digit. Signal Process. 23, 1539–1548 (2013)
44. Hector Basevi, R.A., et al.: Compressive sensing based reconstruction in bioluminescence tomography improves image resolution and robustness to noise. Biomed. Opt. Express 3(9), 2131–2141 (2012)
45. Shen, Y., et al.: Non-uniform compressive sensing in wireless sensor networks: feasibility and application. IEEE Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP). (2012)
46. Abbasi-Daresati, S., Abouei, J.: Toward cluster-based weighted compressive data aggregation in wireless sensor networks. Ad Hoc Netw. 36, 368–385 (2016)
47. Li, S., Xu, L.D., Wang, X.: Compressed sensing signal and data acquisition in wireless sensor networks and internet of things. IEEE Trans. Ind. Inf. 9(4), 2177–2186 (2013)
48. Burak, N., Erdogan, H.: Compressed sensing signal recovery via forward-backward pursuit. Digit. Signal Process. 23, 1539–1548 (2013)
49. Duarte, M., et al.: Joint Sparsity Models for Distributed Compressed Sensing. In: Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS), Rennes, France, November 2005
50. Salim, A., Osamy, W.: Distributed multi chain compressive sensing based routing algorithm for wireless sensor networks. Wireless Network 21(4), 1379–1390 (2015)
51. Donoho, D.: Compressed sensing. IEEE Trans. Inform. Theory 52(4), 1289–1306 (2006)
52. Baraniuk, R.G.: Compressive sensing, IEEE Signal Process Mag. 24(4), (2007)
53. Mallat, S.: A Wavelet Tour of Signal Processing. Academic Press (1999)
54. Candès, E., Tao, T.: Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE. Trans. Inform. Theory 52(2), 145–156 (2006)
55. Venkataramani, R.C.: Sub-Nyquist multicoset and mimo sampling: Perfect reconstruction, performance analysis, and necessary density conditions, Thesis, University of Illinois at Urbana-Champaign, (2001)
56. Truett, J., Gilber, A.: Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans. Inform. Theory 53(14), 4655–4666 (2007)
57. Donoho, D., et al.: Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit. IEEE Trans. Inform. Theory 58(2), 1094–1121 (2012)
58. Deanna, N., Roman, V.: Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit. Found Comput. Math. 9(3), 317–334 (2009)
59. Bandyopadhyay, D., Sen, J.: Internet of things: applications and challenges in technology and standardization. Wirel. Pers. Commun. 58, 49–69 (2011)
60. Wei, D., Olgica, M.: Subspace pursuit for compressive sensing signal reconstruction. IEEE Trans. Inform. Theory 55(5), 2230–2249 (2009)
61. Intel Berkeley Research Lab. http://dlcslcs.mit.edu/labdata/labdata.html (2017)

How to cite this article: Aziz A, Osamy W, Khedr AM, Salim A. Chain-routing scheme with compressive sensing-based data acquisition for Internet of Things-based wireless sensor networks. IET Netw. 2021;10:43–58. https://doi.org/10.1049/ntw.12002