VOL. 17, NO. 2, JANUARY 2022 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

IMPROVED COMPRESSIVE SENSING PROTOCOL WITH ENERGY-EFFICIENT CLUSTERING AND ROUTING IN WSNS

M. Karthik[1] and S. Mohanapriya[2]
[1]Department of Computer Science, Periyar University, Salem, Tamil Nadu, India
[2]Department of Computer Applications, Sona College of Arts and Science, Salem, Tamil Nadu, India
E-Mail: karthikphd123@gmail.com

ABSTRACT

In the recent digitalized world, Wireless Sensor Networks (WSNs) are highly deployed for data transfer between sensor nodes through wireless channels. Normally, these sensor nodes have limited power resources and so the use of these resources has become the most challenging process during Data Aggregation (DA). To avoid this condition, a Power-aware Clustering and Routing with Compressive sensing Protocol (PCRCP) has been recommended for balancing the power use during transferring and aggregating the data between Cluster Head (CH) and Base Station (BS). In this protocol, Compressed Sensing (CS) method is proposed for aggregating the data from CHs with the aid of Forwarder Node (FN). However, still DA process has high data transmission cost and number of measurements. Therefore in this article, a Power-aware Clustering and Routing with Improved Compressive sensing Protocol (PCRICP) is suggested to ensure the energy efficiency of DA in WSNs. Primarily, a coalition formation-based CS solution is proposed that utilizes the signal's sparsity distribution for assembling nodes into many coalitions and the CS is executed inside each coalition. Also, a 2-stage Belief Propagation (BP)-based restoration strategy is applied to achieve an acceptable data quality during DA process. This BP algorithm is an iterative data transfer method which determines the marginal distribution or discovers the Most Probable Assignment (MPA) in the Bayesian networks. On the other hand, it has a convergence problem and the design accuracy is differed with the graph cyclicity. So, an improved generalized BP-based algorithm is proposed that can guarantee better convergence in Markov Random Fields (MRFs). In this algorithm, a caching method and chessboard transitory policy are employed to speed-up the convergence. Also, the computational difficulty of group information from quadric to cubic is reduced. Finally, the simulation results exhibit that the PCRICP achieves superior effectiveness than the PCRCP in terms of different network metrics.

Keywords: belief propagation, clustering, coalition formation, data aggregation, markov random fields, PCRCP, routing, WSN.

1. INTRODUCTION

WSN is usually a network consisting of a huge amount of minimally efficient sensor nodes. Every node collects the sensory information and transfers it to the sink nodes through a multi-hop transfer protocol. This network is used in a number of applications such as monitoring, cyber warfare, etc. WSN is distributed evenly with system access in hostile conditions.

It is thus required for WSN to operate with a desire to adhoc network implementation in a centrally controlled and systematic fashion [1-3]. A setup is required that allows for minimal transition to the BS and is chosen at the node level since the WSN has limited resources and the transfer of information requires additional energy.

Clustering protocols are developed to divide the entire network into various clusters to avoid high energy usage. Clustering usually uses the coexistence between the data to reduce communication overhead and energy use by combining it [4-6]. Therefore, in any cluster, few nodes are chosen as the CH for diffusing control functions between nodes. The primary responsibility of CH is to collect and organize data for BS broadcast from other cluster members. The Low Energy Adaptive Clustering Hierarchy (LEACH) protocol rotates CH between all nodes to share resources across the network [7-9]. Many Power-Aware Clustering and Routing Protocols (PA-CRPs) were proposed by various researchers in recent years in order to lessen energy usage and boost the system life [10-12].

Darabkh *et al.* [13] created a Balanced PA-CRP (BPA-CRP) that enhances the life of the network by allocating four separate transmission ranges for each node. The network was initially split up by the batch-based clustering approach as equal clusters with their CH. In addition, the FN selection algorithm was used to pick the FN either explicitly or implicitly for the data transfer from CH to BS. In order to keep the depleted nodes from selecting CHs or FNs, "only-normal" function mode has also been implemented. Both CH and FNs are used for DA and compression. Nevertheless, the FN was selected only based on its energy level whereas it requests the node density and its locality for further increasing the energy efficiency. Also, the computational complexity of DA was high because of more number of data in every node.

As a result, a PCRCP [14] was proposed to balance the energy utilization in the network during data transfer between CH and BS. In this protocol, every CH in equal-sized clusters picks their FNs based on the node's residual energy, its spatial locality and density for ensuring the balanced dissemination of power resources to the FNs. After that, the aggregated data from CHs was computed by the FNs and transmitted to the closest BS. Also, the dilemma of how to efficiently aggregate the data using FNs was solved by using CSDA approach. On the contrary, the data transmission cost and number of measurements during DA were high.

Hence in this article, a PCRICP is proposed for achieving energy-efficient DA in WSNs. At first, a coalition formation-based CS solution is introduced that uses signal's sparsity distribution to combine nodes as multiple coalitions and the CS is implemented within the coalitions. Besides, a 2-stage BP-based restoration strategy is proposed for offering an acceptable data quality. BP is an iterative data transfer method which computes the marginal distribution or discovers the MPA in Bayesian networks. Conversely, this algorithm has a convergence problem and the design accuracy is varied with the graph cyclicity.

To avoid these problems, an improved generalized BP-based algorithm is proposed that provides better convergence in MRFs. In this improvised algorithm, a caching and chessboard transitory strategies are applied for increasing the convergence speed. Also, the computational difficulty is decreased for group data from quadric to cubic.

Thus, data transmission cost and number of measurements are minimized for enhancing the DA process in an effective way. The rest of the article is prepared as follows: Section 2 surveys the works related to this research work. Section 3 explains the proposed methodology and Section 4 portrays its efficiency. Section 5 summarizes the research work.

2. LITERATURE SURVEY

Use this document as a template by simply typing your text into it. Use this document as a template by simply typing your text into it. Hwang *et al.* [15] considered the Bayesian CS approach and proposed two efficient algorithms for reducing the amount of energetic nodes when sustaining great efficiency. The main intent was to lessen the computation fault through reducing the determinant of the fault covariance matrix which was comparative to the degree of the belief ellipsoid. The centralized greedy choice method was achieved for a directly optimized result according to the least belief ellipsoid. Additionally, equivalent degree of efficiency was achieved as the combinatorial choice approach. The decentralized choice method was modeled via the determinant of the fault covariance matrix. But, the performance was limited when considering the quality of transmission and sensing medium.

Singh *et al*. [16] proposed CS-based acoustic event detection in wireless multimedia sensors. The main objective of this method was recognizing the tree cutting attempt in the forest region through finding the acoustic model created because of an axe striking a tree hole using sensors. A sequence of processes by the hamming window, wiener filter, Otsu thresholding and numerical morphology were applied to eliminate the redundant clutter from the spectrum acquired from those attempts. Also, a CS-based energy-efficient data collection method was applied by the sparse behavior of the audio signals for accurate event reporting. But, the network lifetime was less.

Nguyen & Teague [17] proposed an integration of CS and Random Walk-based Routing (CSRWR) to lessen the power use in WSNs. In this method, every sensed detail was restored at the BS depending on few amounts of CS metrics than the overall amount of sensors. Every CS metric was gathered via a RW routing having a fixed duration. Every random CS metrics was transmitted to the BS for the CS healing task in either instantly or by forwarding via intermediary nodes. Then, a trade-off between the communication range and the RWs duration was analyzed for achieving the least power use. Additionally, the mean used power of every RW was formulated depending on the communication range. Also, the total power use in various scenarios was computed and the optimized scenario was suggested for prolonging the network lifetime. However, total number of hops was high.

Singh *et al.* [18] proposed a MRF structure framework decomposing the network into various homogeneous regions using efficient back propagation-based in-network inference for data collection. In this framework, sensor node's local measure was updated on the basis of neighborhood information and its local observation. Also, a Decomposition Based CS (DBCS) approach was integrated with the MRF framework by considering the transmission constraints in WSN for globally estimating the state of target region and optimizing the transmission cost. But, the lower bound was not analyzed for the number of measurements needed to guarantee the approximate reconstruction.

Zhang *et al*. [19] suggested a novel information aggregating method depending on CS using the clustering structure of WSN. The sink in the cluster can set the respected seed vector according to the network distribution and transmit it to every cluster head. The respected individual arbitrary spacing sparse matrix was created according to its accepted seed vector and information was gathered via CS approach. Then, these ranges were forwarded to the sink by clusters. However, the overall hops alteration was high and also the computational complexity was high.

Xiao *et al*. [20] proposed a category of distributed CS approaches depending on interval relation. A linear regression method was used by means of interval relation for segmenting the test signals. So, the mutual sparse framework of distributed CS was enhanced and a compression matrix was created for extracting the signal's linear fitting segment. After that, an adaptive CS was employed for compressing the signal. But, the effectiveness was not analyzed.

Sun *et al*. [21] proposed a CS data collection algorithm depending on Packet Loss Matching (CS-PLM). A sparse scrutiny matrix was designed depending on PLM and satisfied the Restricted Isometry Property (RIP) with likelihood randomly nearby one. So, trustworthy compressed information transfer was guaranteed via using the multiple paths backup routing between CS nodes. However, the packet loss ratio was not reduced while the network flow was adequately large or small.

Liu & Li [22] proposed an Improved Discrete Differential Evolution (IDDE) method depending on fuzzy clustering for CS reconstruction wherein signal with

ARPN Journal of Engineering and Applied Sciences

unrecognized sparsity was taken as particle coding. Also, the sparse signal was precisely restored via an iterative growth of inhabitants. But, computational complexity was high.

Lv *et al.* [23] proposed a CS-based sequential data collection method. In this method, the covariance matrix was used for generating the sparsifying source of sensed information. Afterwards, the statistical sparsity was introduced for estimating the sparsity performance. The sparse binary matrix was adopted by the measurement matrix and the amount of metrics was restricted by the statistical sparsity. For all analyses, sensed information was gathered by only few sensors and this information was transmitted to the sink node for data recovery. But, an overall power use was not minimized.

Puneeth & Kulkarni [24] developed a routing protocol with tunable energy ranges and data aggregation using CS. In this protocol, the sensor nodes were adjusted to the acceptable transmit energy ranges for increasing the network lifespan under node-disjoint multi-path routing. But, it does not consider the data loss during transmission. Devi *et al.* [25] designed a Cluster-based Data Aggregation Scheme (CDAS) for latency and data loss minimization in WSN. It has 2 steps: aggregation tree formation and slot scheduling scheme. In the first step, CA was used by every cluster head to receive the data from its members. After, the aggregation tree was formed by the sink using minimum spanning tree. In the second step, the data loss ratio and latency were considered when prioritizing and allocating timeslots to the nodes with aggregated data. Still, the data loss rate and residual energy were not effective.

3. PROPOSED METHODOLOGY

In this section, the PCRICP for achieving effective DA in WSNs is described in brief. The main aim of this protocol is to improve the CS theory for aggregating the data from CHs and transmitted to the BS. For this purpose, a distributed CS method with spatial relation between sensors is introduced for grouping them into coalitions. This coalition creation approach is characterized by the block diagonal metrics matrix where every diagonal element associates with any coalitions. By creating coalitions, the spatial-temporal relation-based CS method is executed within every coalition for scheduling sensors and encoding their observations. The temporal relation between sensor observations facilitates this PCRICP for adjusting the amount of metrics regarding temporally modifying sparsity degree.

Once this improved CS solution is applied inside every coalition and the compressed data is forwarded, the BS uses a 2-stage mutual sparsity-based restoration strategy for reconstructing the actual signal. During primary stage, a mutual sparsity framework is used for finding the ordinary sparsity contour between coalitions. During second stage, the ordinary sparsity is computed in every coalition that permits it for realizing information restoration having great precision and less estimations. The entire process in this improved CS procedure is illustrated in Figure-1.

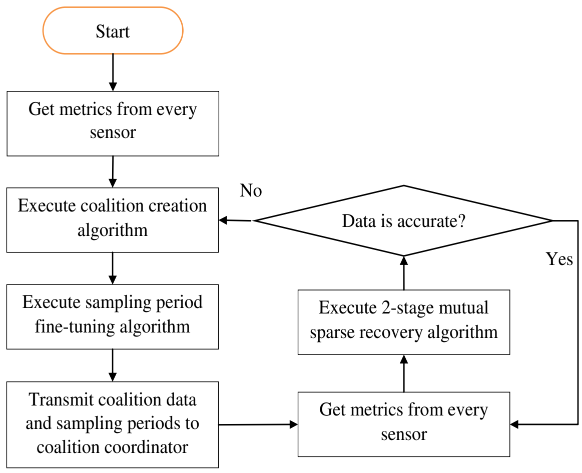

Figure-1. Workflow of the proposed improved CS-based DA and recovery.

3.1 Coalition Formation

The spatial relation between sensors and a support factor distribution model are used for grouping sensor nodes into many coalitions. The utilization of coalitions provides to an energy-efficient CS theory which situates connected sensors near to another in equivalent coalition. Also, coalitions create this CS framework sparse which defines every metric is acquired via the linear mixture of data acquired through some sensors in the same coalition. Therefore, the communication cost is reduced and also the idleness between compressed metrics of various nodes within every coalition is discarded that reduces the amount of data packets.

The sensors notice identical signal having various resolution. In contrast, the signal's sparse illustration is expressed in a sparsity support factor. The support factor distribution is applied in the system for defining the cover factor which expresses the level of support factor enclosed via the sensors.

This factor supports sensors within every coalition create useful metrics resulting in precise information healing through BS. A utility function U is defined depending on cover, communication and sensor relations for providing a concrete and perfect coalition. This factor is applied for evaluating the coalition's efficacy. To achieve this, a trade-off between restoration precision and data communication cost are made by U.

3.1.1 Metric matrix

As the network is split into various coalitions, the information is collected via such coalitions. So, entire data is split into discrete blocks where every block is obtained through the regional metric function. Consider that the WSN is partitioned into N_C coalitions and the signal Z is split into N_C blocks i.e., i.e., $Z_1, Z_2, \dots, Z_{N_C} \in R^N$. Every block denotes coalition j is allocated with a regional metric sub-matrix $\Phi^j: R^N \rightarrow R^{N_j}$. Every metric sub-matrix Φ^j indicates the estimation model in every j. Also, this is allocated to certain coalition matrix C which changes the elements of the actual signal allocated for the particular j. To allocate every j via its metric matrix Φ^j with the changed signal coefficients z_j^C, C is multiplied with Z which creates $CZ = Z^C = \left[z_1^C, z_2^C, \dots, z_{N_C}^C\right]$ and obtain:

$$\gamma = \Phi C Z \tag{1}$$

$$\text{Where } \Phi = \begin{pmatrix} \Phi^1 & & & \\ & \Phi^2 & 0 & \\ & & . & \\ & 0 & . & \\ & & & \Phi^{Nc} \end{pmatrix} \tag{2}$$

The obtained matrix ΦC is the distribution of sensors within every coalition and C is a coalition matrix. According to equation (2) and sparse illustration of $Z = \widetilde{\Psi} a$, get:

$$\gamma = \Phi C Z = \gamma = \Phi\left(C\widetilde{\Psi}\right)a = \Phi \Psi a \tag{3}$$

In Eq. (3), $\Psi = C\widetilde{\Psi}$ is a changed version of $\widetilde{\Psi}$. It is observed that this coalition matrix changes the support factors. So, every row in Ψ is a changed row of $\widetilde{\Psi}$ as:

$$\psi_i^T = \sum_{j=1}^{N} P(i,j)\tilde{\psi}_j^T \tag{4}$$

In equation (4), ψ_i^T and $\tilde{\psi}_j^T$ are the row vectors of Ψ and $\widetilde{\Psi}$, accordingly. If $P(i,j) = 1$, then i^{th} row of Ψ is swapped with j^{th} row of $\widetilde{\Psi}$. After this, an accurate pattern for every Φ^j helps for measuring the information from sensors within every j. This coalition creation process provides to a block diagonal metric matrix with the suitable C corresponds to the sensor's position.

3.1.2 Utility function

C is allocated to any coalition creation strategy. In this task, the optimized trade-off between power conservation and restoration precision is obtained by using the utility function depending on the power, relation and cover level factors.

Energy

The energy factor for coalition creation cases is defined on the basis of transfer, processing and estimation costs. Energy use is defined as:

$$E_i = E_{trans} + E_{process} + E_{measure} \tag{5}$$

In Eq. (5), $E_{trans}, E_{process}$ and $E_{measure}$ are transfer, processing and estimation energy factors, accordingly. The data transfer and estimation costs are focused.

Such costs are normalized and the energy demands are replaced for processing and analyzing the estimation costs. The transfer cost relies explicitly to the remoteness when the estimation cost is affected via the amount of metrics. The energy factor is replaced with the regularized factors of remoteness between any sensors within the coalition and the amount of metrics is expressed as:

$$ECost(i,j) = \frac{D(i,j)}{D_{max}} + \frac{M_i}{M_{max}} \tag{6}$$

In equation (6), $ECost(i,j)$ denotes the energy cost based on the regularized remoteness $D(i,j)$ between nodes i,j and the regularized amount of metrics considered by node $i(M_i)$.

Correlation Level

As sensors are situated nearby every other and observe similar signal having various resolutions, spatial relation ranges between them can be found. This improved CS method attempts to discard the idleness between compressed information and so leveraging this spatial relation. To achieve this, the algorithm is introduced in the coalition-formation step for considering this correlation including other factors. While the connected nodes are in similar coalitions, the idleness in the compressed information is eliminated.

The correlation matrix $(Corr)$ between sensors is described as:

$$Corr(i,j) = \frac{Cov(y_i, y_j)}{\sigma(y_i)\sigma(y_j)} \tag{7}$$

Based on this, a binary variable CR is defined which represents whether 2 sensors are adequately connected. As a result, a user-defined correlation threshold (TH) is used.

$$CR(i,j) = \begin{cases} 1, & if\ Corr(i,j) > TH \\ 0, & if\ Corr(i,j) \leq TH \end{cases} \tag{8}$$

Cover Level

Estimated signals are denoted via a sparsity factor distribution over WSN. Such factors are combined into single or multiple coalitions. The efficiency of this coalition process based on restored information precision highly concerns on the characteristics of the sparsity support factors. For efficiency analysis, a Sparsity base Cover Degree (SCD) factor is used to estimate the overlap level among every coalition with support factors Ψ. Basically, it indicates the energy overlap between the support factors and coalitions. The SCD factor between every support factor i and coalition $Coal_j$ is defined as:

$$SCD(j,i) = \sum_{m \in Coal_j} \psi^2(i,m) \tag{9}$$

In Eq. (9), m represents the sensor positioned in j. $SCD(j,i)$ shows that measurements gathered from coalition j have data regarding the metrics of another coalition which enclose similar i. Considering this exposure level between various coalitions, a joint sparse signal healing method is used for recovering the actual signal. But, these are conditions where Ψ is enclosed just via single coalition.

This refers that, the sparsity supports are comprised in single coalition for a K-sparse signal. As it is not recognized prior, it has to collect information from every coalition having less energy effective. But, the surplus metrics from another coalition which doesn't overlap with the support don't add to increase the information precision. If Ψ has overlap between various coalitions, the information healing precision is maximized. The maximum SCD is defined for quantifying the coverage degree of Ψ over coalitions as:

$$SCD_{max}(\Psi) = SCD_{max}(C\widetilde{\Psi}) = max_{j,u} \sum_{\omega} \Psi^{j^2}(\omega, u), SCD(\Psi) \in [0,1] \quad (10)$$

In Eq. (10), SCD_{max} is the highest coverage degree of every coalition with the sparsity support when Ψ^j denotes the support sub-matrix allocated to j.

Creating coalitions such that the sparsity factor is restored using various coalitions enhances the recovered data accuracy. But, the amount of connected coalitions and the amount of metrics should be reduced for minimizing the energy usage.

Utility Function Formulation

A utility function U is calculated for evaluating the candidate coalition model. In the coalition creation step, the main goal is to construct N_C coalitions for achieving the minimum energy use when meeting data-quality demands. Reducing the transmission and measurement costs depends on SCD_{max}, M_i and $D(i,j)$. The utility factor U for every mixture of $(n_i, coal_r)$ is defined as:

$$\mathrm{U(n_i, coal_r) = CR(n_i) \times \left(ECost(n_i, coal_r) + \alpha SCD_{max}(n_i, coal_r)\right), \alpha > 0} \quad (11)$$

In Eq. (11), $CR(n_i)$ is the relation degree of node i with another node in $coal_r$. Concerning $SCD_{max}(n_i, coal_r)$ factor, including the node to various coalitions can modify SCD_{max} range.

3.1.3 Coalition creation strategy

In this framework, the network consists of N sensor nodes $SN = \{n_1, \dots, n_N\}$ and $L = \left\{l_{n_i n_j}\right\}$ is the group of every probable link among sensors. Any 2 sensors are assumed to be linked if they are located in transmission region of every other. Consider that Ψ is recognized and each sparsity support is regularized to one hence that $SCD \in [0,1]$. The main goal is to reduce energy use given that the information accuracy demand is satisfied. Inclusion of a fresh sensor to the coalition is a choice process which analyzes the optimized candidate to add a node. To achieve a better choice, an optimization method is executed on U and the SCD factor should be refined.

In the coalition creation, while including a fresh node to the coalition, SCD analyzes the impact of including this node through allocating a weight for a connection between previous node in the coalition and the fresh node. For analyzing the impact of including the node to every coalition, SCD is represented as $SCD_{max}\left(l_{n_i,n_j}, coal_r\right)$ through taking the connection $l_{n_i,n_j} \in L$ and a considered coalition $coal_r$. This connection doesn't interact 2 sensors in similar coalition which refers $n_i \in coal_r$ and $n_j \ni coal_r$.

The utility function U for every combination of $\left(l_{n_i,n_j}, coal_r\right)$ is redefined as:

$$U(n_i, coal_r) = CR(n_i, n_j) \times \left(ECost\left(l_{n_i,n_j}\right) + \alpha SCD_{max}\left(l_{n_i,n_j}, coal_r\right)\right), \alpha > 0 \quad (12)$$

In Eq. (12), $SCD_{max}\left(l_{n_i,n_j}, coal_r\right)$ is the highest coverage degree while a fresh node is included to coalition $coal_r$ by l_{n_i,n_j}.

While choosing a suitable coalition for a fresh node, the optimization method analyzes the utility of including the fresh node to every coalition. For every coalition, it is represented on the basis of connection linking the fresh node to a previous node in the coalition.

Executing this method can discover the coalitions which lessen utility factor's connection cost. This optimization is formulated as:

$$\left(n_i, l_{min}, coal_{r_{min}}\right) = argmin\, U(n_i, coal_r) = arg\, min \left[CR(n_i, n_j) \times \left(ECost\left(l_{n_i,n_j}\right) + \alpha SCD_{max}\left(l_{n_i,n_j}, coal_r\right)\right)\right] \quad (13)$$

Subject to $n_i \in SN, n_j \in coal_r, CR(n_i, n_j) \in \{0,1\}$

This Eq. (13) helps to discover the group of connections such that the overall U of the connections is reduced. In this primary phase, consider the group of candidate nodes and connections to be included to the coalitions represented via the group of SN and L. Then, it allocates every coalition coordinator node CC to any N_C coalitions. Also, it represents SN_{coal_r} and L_{coal_r} as a group of nodes and links of coalition r, accordingly.

Additionally, an iterative process is implemented in which it assigns specific node to specific coalitions. To achieve this, initially it discovers the utility factor for every probable links represented in L. Then it operates the optimization factor represented in Eq. (13) and obtains the least utility. The output is a link with the least utility l_{min}. This connection linking (n_i, n_j) includes node n_i to the coalition of n_j. Based on the inclusion of fresh node, the SCD of every connection interacted to the coalitions varies so that the node's connection utility can alter. After that, it eliminates this connection and n_i from the record of

ARPN Journal of Engineering and Applied Sciences

candidate connections and nodes. It is continued until every node is allocated to the coalitions.

Algorithm:
$SN = n_1, \dots, n_N$;
Define $L = l_{ij}$ as the group of every probable connection;
Describe N_C coalitions with any CC;
Describe group of nodes N_{coal_k} and connections for every coalitions L_{coal_k};
$for(P = 1; P \leq (N - N_C); P++)$
$for(Q = 1; Q \leq |L|; Q++)$
Find$L_{prob} = l_{ij}; i \in N_{coal_k}; j \in SN$;
$end\ for$

Compute $U(l_{ii}, coal_k)$;
Insert $n_j$to$coal_{k_{min}}$;
Insert l_{min} to$L_{coal_{min}}$;
Find$\left(n_j, l_{min}, coal_{k_{min}}\right)$;
Eliminate n_j from SN;
Eliminate l_{min} from L;
end for

3.2 Data Collection Inside Coalitions

A distributed CS method is used for the nodes within every coalition to get metrics and restore the actual signal. The energetic nodes within the coalition transmit their observation by using the multi-hop network framework to the CC.

3.2.1 Amount of alive nodes

The amount of nodes needed to be energetic is identified for every coalition. It supports to execute scheduling between nodes and set few nodes into inactive state. Initially, the minimum threshold is defined by the following formula to determine the least information restoration accuracy:

$$\mu(\Phi, \Psi) = \max_{1 \leq k, j \leq N} \left|\langle \Phi_k, \psi_j \rangle\right| \tag{14}$$

If every Φ_j are orthogonal and the sparsifying $\widetilde{\Psi}$ and C are recognized as earlier, then $\mu(\theta)$ is enclosed as:

$$\mathcal{P}\left[\mu(\theta) \leq O\sqrt{SCD_{max}\frac{N_C}{N}\log N}\right] = 1 - O\left(\frac{1}{N}\right) \tag{15}$$

The maximum amount of sensor nodes needed to offer high data quality is defined as:

$$N_{ANodes} = O(SCD_{max} K N_C log^2 N) \tag{16}$$

Adjusting C provides lesser SCD_{max} which creates a smaller amount of nodes when guaranteeing the accuracy of data. The amount of alive nodes related with every coalition is rely on its SCD_{max}. Coalitions which enclose number of sparsity factors are highly useful. Thus, more data should be collected from those coalitions. The amount of alive nodes for every coalition is denoted as:

$$P_j = N_{coal_j} = \frac{SCD_{coal_j}}{SCD_{max}} N_{ANodes} \tag{17}$$

3.2.2 Improved CSDA method

A block diagonal metric matrix is constructed by the spatial-temporal relation between nodes. Consider $SN_{coal_j} = \{1, \dots, N_j\}$ is the group of nodes for j^{th} coalition where P_j of such nodes are allocated to be alive in a random manner. A novel structure is defined for the metric matrix which is well-suited with this coalition creation algorithm. A temporal block diagonal metric matrix Φ_k is used for collecting the information. In every sampling data, spatial analysis of every node is collected at time t and a discrete spatial signal Z_t is generated at t. The combination of temporal analysis of each alive node is a spatial-temporal signal $[Z_1^{tr}, \dots, Z_{ST}^{tr}]$ where ST is a factor denoting the amount of data in every sampling cycle T. Every sampling time has T sampling data identical to the Shannon-Nyquist rate. The amount of sampling periods is adjusted by the BS for minimizing the amount of sampling periods depending on the signal sparsity range.

For every sampling time t, consider Φ_t as a measurement matrix i.e.,$P_j \times SN_{coal_j}$ matrix. The measurement vector γ_{coal_j} has ST_j sub-vector of ST sampling periods such that $\gamma^{coal_j} = \left[\gamma_1^{tr}, \dots, \gamma_{ST_J}^{tr}\right]$ where every γ_i is a$P_j \times 1$ vector.

In every coalition, a block diagram metric matrix is used that efficiently denotes many temporal metric sub-matrices. By fusing such spatial-temporal metrics, obtain:

$$\gamma^j = \gamma_{coal_j} = \Phi^j Z^j \tag{18}$$

$$\gamma^j = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_{ST_J} \end{bmatrix} = \begin{pmatrix} \Phi^1 & & & \\ & \Phi^2 & 0 & \\ & 0 & \ddots & \\ & & & \Phi^{N_C} \end{pmatrix} \begin{bmatrix} Z_1 \\ \vdots \\ Z_{ST_J} \end{bmatrix} \tag{19}$$

Where for every $1 \leq t \leq ST$, Φ^j consists of P_j rows and SN_{coal_j} columns. At last, each sensor node's measurement vector γ^j is transferred to its adjacent node. This data is received by the adjacent node and transmitted it to CC. This temporal block diagonal metric in every coalition can achieve the energy-balanced DA in the coalitions.

3.3 Joint Sparse Signal Recovery Using Improved Generalized Back-Propagation Algorithm

The BS has the responsibility of restoring the actual signals from the metrics accepted from the coalitions. Assuming many coalitions, signals collected by CC are connected in spatial and temporal domains. Consider two kinds of correlations such as spatial-temporal relation within every coalition and spatial relation between coalitions. In this joint sparse signal healing, this relation is represented depending on position and amplitude of signal's non-zero coefficients. The relation between various coalitions or nodes is represented

VOL. 17, NO. 2, JANUARY 2022 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

as a similarity level ($SimL$) and described according to the position of non-zero coefficients as:

$$SimL = \frac{comm}{X} \tag{20}$$

In equation (20), $comm$ is the total amount of ordinary non-zero coefficient positions between various coalitions or nodes and X is the overall amount of signal's non-zero coefficients. At last, the common sparsity profile CSP_{coal} among various coalitions is computed by the BS. Coalitions that fulfill the least resemblance demands are taken in the joint sparse signal healing; or else, its healing proceeds independently.

By defining $SimL$, a 2-stage mutual sparse signal healing is developed which utilizes the spatial-temporal previous data for reconstructing the actual signal. Based on this process, the amount of measurements is reduced when maximizing the accuracy. This signal recovery process is carried out both within and between the coalitions. In the primary phase, this algorithm is executed by the BS among coalitions and their CSP_{coal} are obtained. In the secondary phase, a joint sparse recovery is executed in every coalition by using CSP_{coal} as an input for completing the CSP_{coal} of every coalition $\left(CSP_{coal_j}\right)$. As the nodes in every coalition are greatly connected, executing the second cycle provides other factors of CSP_{coal}.

Based on obtaining CSP_{coal_j}, the BS executes the individual recovery algorithm for finding the independent elements. For entire healing, the BS introduces the two-level BP-based recovery algorithm which is an iterative data transitory strategy which computes the marginal distribution or the MPA in Bayesian networks. But, it has convergence problem and design accuracy is varied with the graph cyclicity.

Therefore, an improved generalized BP-based recovery algorithm is introduced that provides better convergence in MRFs and achieves acceptable data quality. The main aim of this algorithm is computing highly informative data among areas other than nodes. It facilitates a random amount of nodes for DA as a clique and encompasses the clique data to the entire transmission which provides enhanced approximation to the posterior likelihood. Since the clique data encompassed in the transmission, the search ability for the least of an energy factor is modified. The modification criteria of the canonical generalized BP are:

$$\mathrm{m_{rs}} \leftarrow \mathrm{k}\frac{\sum_{\mathrm{z_{r\backslash s}}} \varphi_{\mathrm{r\backslash s}}(\mathrm{z_{r\backslash s}}) \prod_{\mathrm{m_{r''s''}} \in \mathrm{M(r)\backslash M(s)}} \mathrm{m_{r''s''}}}{\prod_{\mathrm{m_{r's'}} \in \mathrm{M(r,s)}} \mathrm{m_{r's'}}} \tag{21}$$

$$\mathrm{b_r} \leftarrow \mathrm{k}\varphi_{\mathrm{r}}(\mathrm{z_r}) \prod_{\mathrm{m_{r's'}} \in \mathrm{M(r)}} \mathrm{m_{r's'}} \tag{22}$$

In Eqns. (21) & (22), r is the areas and s is their respective sub-region, m_{rs} refers to the data transmitting from r to its sub-region s, $\varphi(z)$ denotes the local evidence of node z, $M(r)$and $M(s)$ are the group of data transmitting from outside of r or s to few nodes inside r or s, respectively, $M(r,s)$ is the group of data transmitting from few nodes in r but not in s to few nodes in s and b_r is the belief of r.

Algorithm

Discover SL among various coalitions;
Discover SL_{min};

Execute spatial-temporal joint recovery in coalitions with SL_{min};

$\mathrm{for}(\mathrm{j}=1; \mathrm{j} \le \mathrm{N_{coal}}; \mathrm{j}++)$
$\mathrm{if}\left(\mathrm{SL_{coal_j}} > \theta\right)$
Utilize CSP_{coal} from coalition with SL_{min} as input;
end if
Execute spatial-temporal joint recovery algorithm;
end for

Thus, this PCRICP can reduce the transmission cost and the amount of measurements for achieving energy-efficient DA in WSNs.

4. SIMULATION RESULTS

This part simulates the PCRICP using Network Simulator version 2.35 (NS2.35) and compares its efficiency with the existing protocols: CSRWR [7], CS-PLM [11], CDAS [15] and PCRCP [4] in the aspect of network lifespan, energy consumption, packet loss and the number of nodes alive. In this simulation, 250 nodes are deployed over 1200m×1200m. The simulation parameters are given in Table-1.

VOL. 17, NO. 2, JANUARY 2022 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

Table-1. Simulation parameters.

Parameters	Range
Network dimension	1200×1200m^2
Network topology	Flat grid
Antenna category	Omni antenna
Channel category	Wireless channel
MAC layer	802.11
Routing protocols	PCRCP and PCRICP
Number of nodes	250
BS position	(50,45)
Data packet size	8500 bits
Control packet size	60 bytes
Traffic type	Constant Bit Rate (CBR)
Initial energy	2 J
dis_χ	87.71
Energy used for DA	5 nJ/bit
Energy used in the transmitter	50 nJ/bit
Energy used by the free space transmission	10 pJ/bit.m^2
Energy used by the multipath transmission	0.0013 pJ/bit.m^4
Transmission radius	100 m
Simulation time	1000sec

4.1 Network Lifespan

It is the time taken to build the network up to the initial node dies. Figure-2 portrays the network lifespan (in No. of rounds) for different protocols with different amount of nodes. For 150 nodes in the network, the network lifespan of PCRICP is 8.59% increased than PCRCP, 24.66% increased than the CDAS, 45.55% increased than the CS-PLM and 63.53% increased than the CSRWR.

This analysis indicates that the PCRICP achieves the highest network lifespan than all other protocols.

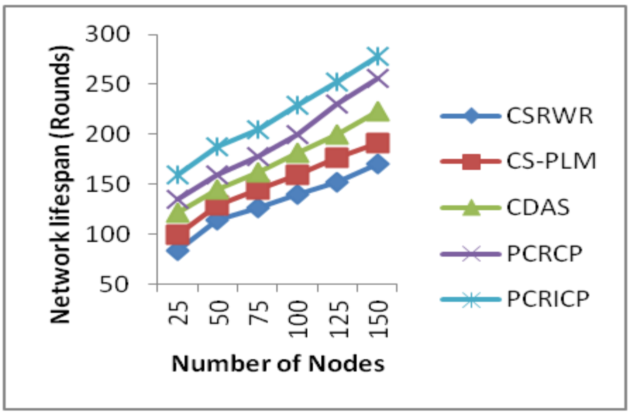

Figure-2. Network lifespan vs. no. of nodes.

4.2 Energy Consumption

It is the sum amount of energy used by the network during the given time. Figure-3 illustrates the energy consumption (in Joules (J)) for different protocols with varying amount of nodes.

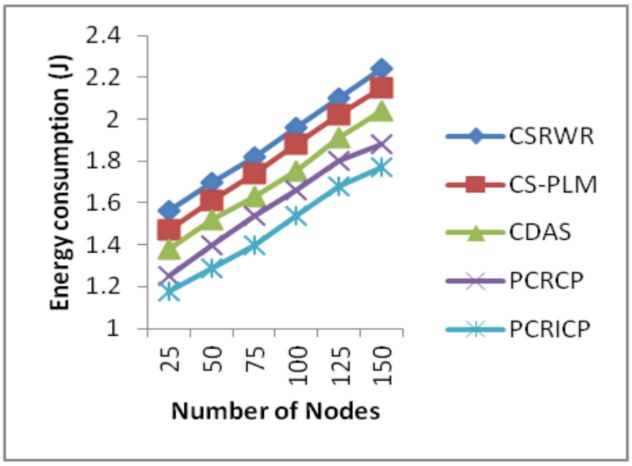

Figure-3. Energy consumption vs. no. of nodes.

When the amount of nodes is 150, the energy consumption of PCRICP is 5.85% reduced than the PCRCP, 13.24% reduced than the CDAS, 17.67% reduced than the CS-PLM and 20.98% reduced than the CSRWR protocols. This analysis indicates that the PCRICP

ARPN Journal of Engineering and Applied Sciences

achieves the reduced energy consumption compared to the all other routing protocols.

4.3 Packet Loss

The number of packets missed during transmission between the BS and CHs or FNs is considered as Packet loss between the BS and CHs or FNs. Figure-4 demonstrates the packet loss (in %) for various protocols with different amount of nodes.

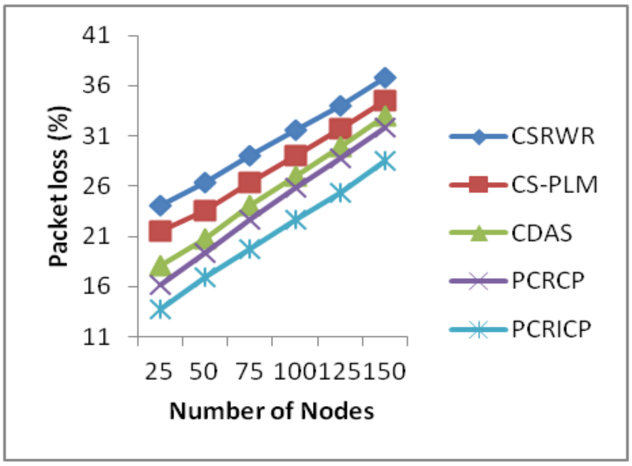

Figure-4. Packet loss vs. no. of nodes.

If the amount of nodes is 150, the packet loss of PCRICP is 10.38% less than the PCRCP, 13.64% less than the CDAS, 17.39% less than the CS-PLM and 22.55% less than the CSRWR. Through this analysis, it is noticed that the PCRICP has the minimum packet loss than the other protocols.

4.4 Amount of Nodes Alive

It provides the amount of nodes active in the network after data transmission. Figure-5 depicts the amount of nodes alive for different protocols with varying amount of rounds.

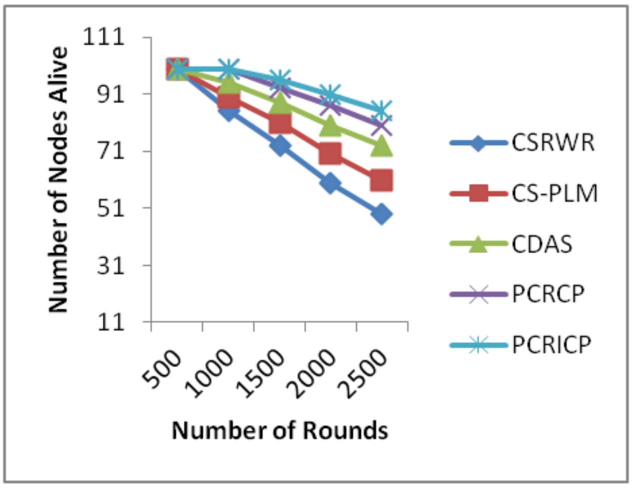

Figure-5. No. of nodes alive vs. no. of rounds.

If the amount of rounds is 2500, then the amount of nodes alive for PCRICP is 6.25% increased than the PCRCP, 16.44% increased than the CDAS, 39.34% increased than CS-PLM and 73.47% increased than the CSRWR. This analysis indicates that the PCRICP achieves the highest amount of alive nodes compared to the other protocols.

5. CONCLUSIONS

In this article, a PCRICP is proposed for achieving energy-efficient CSDA in WSNs. At first, a coalition formation-based CS solution is proposed that utilizes the signal's sparsity distribution for assembling the sensors into many coalitions and the CS is executed inside each coalition. Also, an improved generalized BP-based algorithm is proposed that can guarantee better convergence in MRFs. In this algorithm, a caching and chessboard transitory strategies are applied for improving the convergence speed. Also, the computational difficulty of group information from quadric to cubic was minimized. At last, the investigational outcomes revealed that the PCRICP has enhanced efficiency compared to the PCRCP in terms of different network metrics.

REFERENCES

[1] M. Carlos-Mancilla, E. López-Mellado and M. Siller. 2016. Wireless sensor networks formation: approaches and techniques. Journal of sensor. 2016: 1-18.

[2] P. Singh, O. P. Gupta and S. Saini. 2017. A brief research study of wireless sensor network. Advances in Computational Sciences and Technology. 10(5): 733-739.

[3] D. Kandris, C. Nakas, D. Vomvas and G. Koulouras. 2020. Applications of wireless sensor networks: an up-to-date survey. Applied System Innovation. 3(1): 1-24.

[4] A. A. H. Hassan, W. M. Shah, M. F. Iskandar, and A. A. J. Mohammed. 2017. Clustering methods for cluster-based routing protocols in wireless sensor networks: comparative study. International journal of applied engineering research. 12(21): 11350-11360.

[5] X. Liu. 2012. A survey on clustering routing protocols in wireless sensor networks. Sensors. 12(8): 11113-11153.

[6] C. Gherbi, Z. Aliouat and M. Benmohammed. 2017. A survey on clustering routing protocols in wireless sensor networks. Sensors Review. 37(1): 1-21.

[7] F. Zhao, Y. Xu and R. Li. 2012. Improved LEACH routing communication protocol for a wireless sensor

network. International Journal of Distributed Sensor Network. 8(12): 1-6.

[8] R. K. Gill, P. Chawla and M. Sachdeva. 2014. Study of LEACH routing protocol for wireless sensor networks. In International Conference on Communication Computing Systems. pp. 3-5.

[9] S. R. Ahmed, M. A. Kadhim and T. Abdulkarim. 2019. Wireless sensor networks improvement using LEACH algorithm. In IOP Conference Series: Materials Science and Engineering, IOP Publishing. 518(5): 1-6, 2019.

[10] N. Elsaka, N. El-Shennawy and M. M. Fahmy. 2019. An energy-aware clustering/routing a* algorithm for wireless sensor networks. In 14th IEEE International Conference on Computer Engineering and Systems (ICCES). pp. 413-418.

[11] G. Sumalatha and P. Anil kumar. 2016. Applications of energy aware routing protocols in WSN. Oriental Journal of Computer Science and Technology. 9(1): 24-29.

[12] C. R. Dongarsane, D. Maheshkumar and H. S. Arvind. 2020. A relative simulation study of energy aware routing protocols for wireless sensor network. International Journal Future Generation Communication and Networking. 13(3): 4060-4068.

[13] K. A. Darabkh, M. Z. El-Yabroudi and A. H. El-Mousa. 2019. BPA-CRP: A balanced power-aware clustering and routing protocol for wireless sensor networks. Ad Hoc Networks. 82: 155-171.

[14] M. Karthik and S. Mohanapriya. 2020. Power-aware clustering and routing with compressive sensing protocol for wireless sensor networks. International Journal of Advanced Science and Technology. 29(6): 4584-4592.

[15] S. Hwang, R. Ran, J. Yang and D. K. Kim. 2016. Multivariated Bayesian compressive sensing in wireless sensor networks. IEEE Sensors Journal. 16(7): 2196-2206.

[16] V. K. Singh, G. Sharma and M. Kumar. 2017. Compressed sensing based acoustic event detection in protected area networks with wireless multimedia sensors. Multimedia Tools and Applications. 76(18): 18531-18555.

[17] M. T. Nguyen and K. A. Teague. 2017. Compressive sensing based random walk routing in wireless sensor networks, Ad Hoc Networks. 54: 99-110.

[18] V. K. Singh, S. Verma and M. Kumar. 2018. DBCS: A decomposition based compressive sensing for event oriented wireless sensor networks. Wireless Personal Communications. 99(1): 351-369.

[19] D. G. Zhang, T. Zhang, J. Zhang, Y. Dong and X. D. Zhang. 2018. A kind of effective data aggregating method based on compressive sensing for wireless sensor network. EURASIP Journal on Wireless Communications and Networking. 2018(1): 1-15.

[20] S. Xiao, T. Li, Y. Yan and J. Zhuang. 2019. Compressed sensing in wireless sensor networks under complex conditions of Internet of things. Clustering Computer. 22(6): 14145-14155.

[21] Z. Sun, R. Tao, N. Xiong and X. Pan. 2018. CS-PLM: Compressive sensing data gathering algorithm based on packet loss matching in sensor networks. Wireless Communication and Mobile Computing. Vol. 2018.

[22] Z. Z. Liu and S. N. Li. 2019. WSNs Compressed Sensing Signal Reconstruction Based on Improved Kernel Fuzzy Clustering and Discrete Differential Evolution Algorithm. Journal of Sensors. Vol. 2019.

[23] C. Lv, Q. Wang, W. Yan and J. Li. 2019. Compressive sensing-based sequential data gathering in WSNs. Computer Network. 154: 47-59.

[24] D. Puneeth and M. Kulkarni. 2020. Data Aggregation using Compressive Sensing for Energy Efficient Routing Strategy. Procedia Computer Science. 171: 2242-2251.

[25] V. S. Devi, T. Ravi and S. B. Priya. 2020. Cluster based data aggregation scheme for latency and packet loss reduction in WSN. Computer Communications. 149: 36-43.