Open access • Journal Article • DOI:10.1109/JSEN.2019.2893912

An Energy-Efficient Compressive Sensing-Based Clustering Routing Protocol for WSNs — Source link

Quan Wang, Deyu Lin, Pengfei Yang, Zhiqiang Zhang

Institutions: Xidian University, University of Leeds

Published on: 15 May 2019 - IEEE Sensors Journal (Institute of Electrical and Electronics Engineers (IEEE))

Topics: Cluster analysis, Routing protocol, Efficient energy use, Energy consumption and Wireless sensor network

Article:
Wang, Q, Lin, D, Yang, P et al. (1 more author) (2019) An Energy-Efficient Compressive Sensing-Based Clustering Routing Protocol for WSNs. IEEE Sensors Journal, 19 (10). pp. 3950-3960. ISSN 1530-437X

https://doi.org/10.1109/JSEN.2019.2893912

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2019.2893912, IEEE Sensors Journal

An Energy-Efficient Compressive Sensing-Based Clustering Routing Protocol for WSNs

Quan Wang, *Member, IEEE*, Deyu Lin, *Student Member, IEEE*, Pengfei Yang, *Member IEEE*, and Zhiqiang Zhang, *Member, IEEE*

***Abstract*—A novel algorithm which combined the merits of the Clustering strategy and the Compressive Sensing-based (CS-based) scheme was proposed in this paper. The lemmas for the relationship between any two adjacent layers, the optimal size of clusters, the optimal distribution of the Cluster Head (CH) and the corresponding proofs were presented firstly. In addition, to alleviate the "Hot Spot Problem" and reduce the energy consumption resulted from the rotation of the role of CHs, a third role of Backup Cluster Head (BCH) as well as the corresponding mechanism to rotate the roles between the CH and BCH were proposed. Subsequently, the Energy-Efficient Compressive Sensing-based clustering Routing (EECSR) protocol was presented in detail. Finally, extensive simulation experiments were conducted to evaluate its energy performance. Comparisons with the existing clustering algorithms and the CS-based algorithm verified the effect of EECSR on improving the energy efficiency and extending the lifespan of WSNs.**

***Index Terms*—Wireless sensor networks; compressive sensing; the "hot spot problem"; energy efficiency**

I. INTRODUCTION

THE WIRELESS SENSOR NETWORKS (WSNs) is a kind of networks which consist of a large number of tiny sensor nodes[1][2]. Owing to the low deploying cost, WSNs have gained extensive applications, such as the target tracking, enemy detecting, air pollution monitoring, medical caring, and so on[3][4][5]. In general, most of the sensor nodes are powered by the battery, which means their energy supply is limited. Besides, most of WSNs are usually deployed in the area which is out of human's reach. So it is high-cost or unpractical for them to be replenished. In addition, the "Hot Spot Problem" which refers to the fact that the nodes close to the Sink deplete energy faster than the nodes lying in edge area makes matters worse. The lifespan of WSNs terminates when one or some of the key nodes exhaust their energy due to the finite energy budget as well as the "Hot Spot Problem". Therefore how to prolong the lifetime is one of the main challenges facing the wireless sensor networks without being equipped with any energy harvesting technology. In recent years, the schemes aiming at extending the lifespan of WSNs has attracted extensive attentions. Since the energy supply is limited, improving the energy efficiency is one kind of effective ways to prolong the lifetime.

Generally speaking, two aspects are included concerning improving the energy efficiency, which are the reduction and balance of energy consumption. Usually, there exists numerous redundancy in the source data. For example, most of the data acquired by the sensor node exhibit high Spatial-temporal correlation. The Spatial-temporal correlation includes two aspects: spatial correlation and temporal correlation. The sensor nodes are usually densely deployed in most of the applications of WSNs, which results in the similarity of the data. Namely, the data acquired by the adjacent nodes appear to be highly similar in part or whole. This phenomenon is called as the spatial correlation, which leads to the data redundancy and the waste of energy. In addition, the temporal correlation which means the data sensed in the adjacent time slots appear to be similar also exists in the process of data acquisition. In general, the temporal correlation is resulted from the high frequency of data acquisition. For instance, in some scenarios, such as the disaster detection, endangered species rescue, etc., the data need to be collected in a high frequency in order to achieve a high accuracy. Similar to the spatial correlation, the temporal correlation also leads to the data redundancy. Therefore the energy efficiency can be improved if the redundancy is reduced via aggregation before the data are intactly transmitted to the Sink. In general, the intranet data aggregation can be utilized to reduce the data redundancy resulted from the Spatial-temporal correlation. The traditional intranet data aggregation schemes included the distributed source model, distributed transform coding, and the distributed source coding[6]. However, most of them featured that the operation is complicated for the encoder but simple for the decoder. Obviously it is not very suitable for the resource-limited WSNs. On the other hand, the "convergecast" traffic pattern and the "hop-by-hop" routing mode in WSNs result in the uneven distribution of energy consumption, which is the root cause of the "Hot Spot Problem". Recent years have witnessed a lot of researches which aim at improving the energy efficiency via cutting down

This paper was supported by the National Natural Science Foundation of China (61572385 & 6161101371 &, 61702395), the China Scholarship Council, and the Innovation Fund of Xidian University.

Quan Wang, Deyu Lin, and Pengfei Yang are with the School of Computer Science and Technology, Xidian University, Xi'an, Shannxi, China, 710071 (e-mail: Dashing_lin@126.com).

Zhiqiang Zhang is with Leeds University, Leeds, WestYorkshire, LS2 9JT UK.

and balancing the energy consumption[7][8][9], such as the clustering-based routing protocol[10], the topology controlling scheme[11][12], the strategy adopting the mobile Sink or mobile Relay[13][14], and the data aggregation scheme[15]. However, all of them followed the Nyquist sampling theorem, which demands the sample frequency should be at least twice as much as the largest frequency of the source signal to recover it precisely. It means that the redundant data resulted from the Spatial-temporal correlation are still collected, which leads to a waste of energy to some extent. Besides, even though there are some schemes aiming at reducing energy via in-network aggregation, they still followed the acquisition-first-compress-latter pattern. The energy is wasted for the process of raw data acquisition, therefore the energy efficiency still needs to be improved.

Compressive Sensing (CS)[16] theory provides a new kind of data acquisition paradigm for WSNs. Via the CS theory, the signal can be acquired and recovered at the sampling rate which is much lower than the Nyquist sampling rate. The energy consumption decreases with the sampling frequency, therefore the CS-based scheme can be applied to extend the network lifetime. As for CS-based schemes, if a signal is sparse in nature or sparse under a given sparse basis matrix, it can be reconstructed from a small number of linear measurements. Usually the source signal can be recovered by solving an l_1-norm convex optimization problem. CS theory makes it possible that the sampling frequency depends on the characteristics of the signal itself rather than the signal bandwidth[17]. Moreover, unlike the conventional data compression techniques, such as CAG (Clustered AGgregation)[6] and DSC (Distributed Source Coding)[18][19], CS synchronizes sampling and compression, and only need several simple multiplication and addition operations during the process of compression sampling. While the complicated data recovery operations are conducted by the Sink. Obviously, this feature makes it exactly suitable for the wireless sensor networks where sensor nodes possess limited energy and processing capacity but the energy supply and the processing capacity of Sink nodes are relatively unlimited. Through applying the CS to the WSNs, the data can be compressed when they are being sensed. So the energy for sensing and transmitting is cut down remarkably due to the decline of the amount of data sensed and transmitted.

The contributions of the paper were listed as follow. The lemmas for the relationship between two adjacent layers, the optimal number of CH at each layer and the optimal distribution of CH were presented firstly. In addition, in order to alleviate the “Hot Spot Problem” as well as reduce the energy consumption resulted from the flooding during the process of rotation the roles of CHs, a third role of the Backup Cluster Head (BCH) and the mechanism to rotate the roles between CH and BCH were proposed. Subsequently, the Energy-Efficient Compressive Sensing-based clustering Routing (EECSR) protocol which was a combination of the merits of the Clustering strategy and the CS-based scheme was presented. Finally, extensive simulation experiments were conducted to evaluate its energy performance. Comparisons with the existing Clustering algorithms and the CS-based algorithms verified the EECSR’s effect on improving the energy efficiency and extending the lifespan of WSNs.

The reminders of the paper were organized as follow. Section II introduced the related works. The preliminaries was presented in Section III, which was followed by the theoretical basis of the proposal in Section IV. Subsequently the Energy-Efficient Compressive Sensing-based clustering Routing (EECSR) protocol was proposed in Section V and it was evaluated by simulation in Section VI. Finally the conclusion was drawn and some future directions was pointed out.

II. Related Works

Recent years have witnessed numerous strategies in order to improve the energy efficiency and prolong the lifespan of WSNs. Most of them adopted the Clustering strategies which divided the whole network into logical hierarchical topology. Through clustering, simple aggregation operations can be conducted by CH, which helps to cut down the energy consumption resulted from intra-cluster communication. In addition, the rotation of the roles of the sensor nodes were utilized to balance the energy consumption among the whole network. For example, LEACH (Low-Energy Adaptive Clustering Hierarchy)[20] proposed the Clustering schemes for the first time to improve the energy efficiency. It selected the Cluster Head randomly. In PEGASIS (Power-Efficient Gathering in Sensor Information Systems)[21], the nodes took turns to act as CHs according to a certain network logical topology. In some scenarios, the CHs selection process was controlled according to some predefined thresholds. For example, the similarity among nodes and the node degree are often regarded as parameters for CH election according to the specific application scenarios. There were several related examples. For instance, a double-threshold mechanism, namely, the Hard Threshold (HT) and the Soft Threshold (ST), were used in TEEN (Threshold sensitive Energy Efficient sensor Network protocol)[22] which was applied in the hard real-time scenario. Similar classic Clustering strategies included the DHAC (Distributed Hierarchy Aggregation Clustering) which generated the CHs according to the similarity matrix generated by the node input, LEACH-ERE[23] which considered the expected residual energy through the fuzzy-logic to further improve the energy efficiency, etc. Although the energy efficiency has been improved by the above schemes, the Nyquist sampling theorem pattern has prevented the further improvement of energy efficiency. The redundancy resulted from the data acquisition leads to a waste of energy.

The Compressive Sensing theory provided a new paradigm for the data collection of WSNs. A large number of energy-efficient strategies which were based on Compressive Sensing theory have emerged in recent years. With the help of Compressive Sensing theory, the data redundancy resulted from the Spatial-temporal correlation can be reduced remarkable, which brings in a huge improvement of energy efficiency of WSNs.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2019.2893912, IEEE Sensors Journal

Recent years have witnessed extensive attentions paid to the CS-based strategies which aimed to improve the energy efficiency. Zheng et, al. provided the mathematical foundation to acquire the measurement in a random-walk-based manner[24]. It was obtained that a K-sparse signal can be reconstructed via the l_1-norm minimization decoding algorithm when it took $m = O(k \log(n/k))$ independent random walks whose length was $t = O(n/k)$ steps in a random geometric network with n nodes. The simulation results demonstrated its effectiveness. Similar strategy included another proposal which adopted a random walk-based algorithm and a kernel-based strategy[25]. These schemes were finally proved to be able to not only significantly reduce the communication cost but also combat the unreliable wireless link under various packet losses. Quan et, al. proposed a Neighbor-Aided Compressive Sensing (NACS) scheme for the Spatial-temporal correlated WSNs[26]. The measurements were sent to a randomly and uniquely selected neighbor to distribute the energy consumption equally. Besides, the NACS integrated the Structured Random Matrix (SRM) with the Kronecker Compressive Sensing (KCS) to improve the recovery performance. Xiang, et al. presented a data aggregation scheme which exploited the Compressive Sensing to achieve both the energy efficiency and the recovery fidelity in a arbitrary WSNs[27]. The diffusion wavelet was used for getting a suitable sparse basis which characterized the Spatial-temporal correlation well. The strategy was finally converted into a mixed integer programming formulation solved by a greedy heuristic algorithm. The corresponding performance was evaluated via extensive simulations on both the real data set and the synthetic data set. Yang, et, al. proposed a Compressed Networking Coding based Distributed data Storage (CNCDS) scheme which aimed at achieving high energy efficiency via reducing the total amount of transmissions Nt_{tot} and that of receptions Nr_{tot} during communication[28]. According to the Random Geometric Graphs (RGG) theory, the expressions of Nt_{tot} and Nr_{tot} were obtained to theoretically verify the energy efficiency of CNCDS. The simulation results shown that the values of Nt_{tot}, Nr_{tot}, and the CS recovery Mean Squared Error (MSE) were reduced by 55%, 74%, and 76% respectively. Besides, an adaptive CNCDS was further presented according to the expressions and the values of Nt_{tot}, and Nr_{tot} were reduced by 63% and 32% respectively compared with CNCDS.

Although the CS-based schemes above can improve the energy efficiency, it was not suitable for the large scale WSNs very well. To strengthen the scalablity, the CS-based scheme was integrated with the Clustering algorithm. Luo et, al. proposed a Compressive Data Gathering (CDG) framework for the large-scale wireless sensor network for the first time[29]. In their proposal, the scenario where lots of sensor nodes were densely deployed was taken into consideration. The total communication cost was reduced without introducing much computation or complicated transmission control overhead. Shen et, al. proposed a nonuniform compressive sensing scheme which was applied to the real WSNs data set to improve the energy efficiency via exploiting both data compressibility and heterogeneity[30]. Although the CS-based Clustering strategy improved the energy efficiency, the overhead of Cluster Member increased. To overcome the shortage, Xie et, al. proposed an energy-efficient clustering method which was integrated with the hybrid CS theory[31]. In their scheme, the data acquisition of intra-cluster adopted the Raw Data Gathering (RDG) and that of inter-cluster applied the Compressive Data Gathering (CDG). Zhao et, al. presented a Treelet-based Clustered Compressive Aggregation (T-CCDA) in which the treelet transform was adopted as a sparsification tool [32]. Besides, a novel Clustering routing was proposed to improve the energy efficiency further by taking advantage of the correlation structures which was based on the T-CCDA scheme. To provide a high energy efficiency for the large scale WSNs, the Hierarchical Compressive Data Gathering (HCDG) strategy was proposed. Lan, et, al. proposed a Compressibility-Based Clustering Algorithm (CBCA) [33] where the network topology was converted into a logical chain which was similar to PEGASIS[21]. Li et, al. proposed a novel energy-efficient Compressive-Sensing-based Data Gathering (CS-DG) algorithm to cut down energy consumption in the clustered wireless sensor network and exploit a better reconstruction accuracy[34]. The random sampling and random walks were utilized to get each measurement by summing a few sensory data up. They proved that a constructed $M \times N$ sparse binary matrix was equivalent to the adjacency matrix of an unbalanced expander graph when some conditions were met. Yan et, al. presented an optimal Compressed Data Gathering (CDG) framework [35] which adopted a Diffusion Wavelet Transform Matrix (DWTM) as the sparse representation of the compressed data. Besides, they proposed a novel Measurement Matrix Optimization Algorithm (MMOA) for the process of acquiring the compressed data. Except for the strategies above, other energy-efficient Compressive Sensing-based Clustering schemes have emerged recently, such as the ST-HDACS[36], the chain-based data gathering protocol[37], and the adaptive adjustment of compressed measurements[38].

The above CS-based Clustering strategies have improved the energy efficiency to some extent. However, all of them improved the energy efficiency by reducing the energy consumption resulted from data redundancy. In fact, the energy balance should also be taken into consideration to improve the energy efficiency further. Therefore, in this paper both the energy balance and the energy overhead of cluster formation were considered. Besides, the merits of both the Compressive Sensing and Clustering were utilized to improve the energy efficiency.

III. Preliminaries

In this section, the concept of Compressive Sensing and its components which consist of the K-sparse signal, the Restricted Isometry Property (RIP), and the reconstruction algorithm were presented in detail. Subsequently, the system model and some assumptions were presented.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2019.2893912, IEEE Sensors Journal

A. Concept of Compressive Sensing

Compressive Sensing is a kind of new paradigm which applies the signal sparsity, the mathematical statistics theory, and the optimization theory to the process of signal acquisition. It was firstly proposed by D. Doboho, E. Candes and R. Baraniuk who presented the framework for CS as the foundational achievement[16]. According to the Nyquist sampling theorem, the sampling frequency should be at least twice as much as the maximum frequency of the source signal. On the contrary, the CS theory makes it possible for the sparse signal or the signal which is sparse in a certain transform domain to be precisely recovered with much lower sampling frequency, which brings in a remarkable reduction of energy consumption. Thus the CS theory makes the sampling frequency escape from the dependence on the signal bandwidth. Besides, different from the traditional data aggregation technology, the signal can be sampled and compressed at the same time. Finally, the CS-based data aggregation only need some add and multiply operations and the complex data recovery operation is performed by the Sink whose processing capacity is relatively limitless. All the features above make the CS-based sampling suitable for the wireless sensor networks.

B. Components of the Compressive Sensing Theory

(1) K-sparse Signal

Suppose a wireless sensor network consisting of N nodes, the data acquired by node i are denoted as x_i $(1 \le i \le N)$. Then the source data set collected by the wireless sensor network can be denoted as $X = x\,(x_1, x_2 ... x_n)$. The signal is called to be a K-sparse signal, if for a sparse transform basis Ψ, there exists a coefficient matrix θ which meets the following condition

$$\|\theta\|_0 \le \mathrm{K}, \ \text{s.t.}\ \mathrm{X} = \Psi \cdot \theta \qquad (1),$$

where θ represents the sparse coefficient matrix and $\|\theta\|_0$ denotes the number of non-zero elements in the matrix θ.

(2) Restricted Isometry Property (RIP)

To obtain the valid information from the K-sparse signal, a $M \times N$ measurements matrix Φ is needed. Then the measurements can be represented as $y = \Phi X$, where $X = \Psi\theta$. The foundation of Compressive Sensing lies in the fact that the $M \times 1$ measurement vector is finally obtained through a properly-designed measurement matrix. Since $M << N$ is always established, the rate of compressive sensing is far less than that of the Nyquist sampling theory.

To precisely recover the source data from the measurements, the measurement matrix Φ is needed to satisfy the Restricted Isometry Property (RIP). Namely, for any signal x, if there exists a parameter $\delta_k\,(0 < \delta_k < 1)$ which satisfies the following inequality

$$(1-\delta_k)\frac{M}{N}\|x\|_2^2 \le \|\Phi x\|_2^2 \le (1+\delta_k)\frac{M}{N}\|x\|_2^2 \qquad (2),$$

then the measurements matrix Φ satisfies the RIP. As for the measurement matrix, the row stands for the number of measurements and the column denotes the number of sensor nodes respectively. Generally speaking, the relationship between M and K needs to meet the following condition,

$$M \ge K \log N \qquad (3).$$

It dedicates that the CS theory brings in a remarkable reduction of data acquired and transmitted.

(3) Reconstruction Algorithm

For the measurements collected by the WSNs, RIP is a sufficient condition for them to be accurately recovered. Via the measurement matrix Φ, the measurement y can be obtained as the following expression,

$$y = \Phi \cdot X \qquad (4).$$

The source data X can be precisely recovered by solving the following l_0-norm optimization problem,

$$\min_{\theta} \|\theta\|_0 \ s.t.\ y = \Phi\Psi\theta \qquad (5).$$

Therefore the source signal can be denoted as following expression

$$X' = \Psi\theta' \qquad (6).$$

where θ' stands for the solution of the l_0-norm problem.

However, it has been proven that the solution of l_0-norm is a NP-Hard problem owing to the demand for listing all the possible combinations of vector θ. When the dimension becomes larger, it is harder to obtain the optimal solution. Fortunately Donobo pointed out that the optimal solution of l_1-norm can approximately replace that of the l_0-norm problem if the measurement matrix Φ meets RIP[16]. Therefore the recovery of the source data can be turned into the following l_1-norm convex optimization problem,

$$\theta^* = \arg\min \|\theta\|_1 \quad s.t.\ \ y = \Phi\Psi\theta \qquad (7).$$

Then the following Expression (8) is the approximate value of the source data,

$$x^* = \Psi\theta^* \qquad (8).$$

C. System Model and Some Assumptions

(1) Energy Consumption Model

The sensor node depletes the energy when it senses, receives, and transmits data. There are different operation modes which have large influence on the energy consumption for the sensor node. Generally speaking, there are four modes for the sensor node, namely, the IDLE, TRANSMIT, RECEIVE, and SLEEP modes. However, most of the energy is consumed for transmitting and receiving data. Shih et al. pointed out at the Mobicom meeting in 2002 that most of the sensor's energy was consumed in the communicating module, and the energy consumption in the SLEEP mode was usually negligible[39]. So the energy efficiency can be largely improved if the communication cost is cut down.

In this paper, the first-order radio model[10][11][31] was adopted to describe the energy consumption for a node's transmission. Namely, the energy for a sensor node to transmit a k-bits packet to another node over distance d equals

$$e_{tx} = k(E_{elec} + \varepsilon_{amp} \cdot d^{\alpha}) \qquad (9),$$

where E_{elec} is the energy consumed in the transmitter circuit, ε_{amp} is the transmitter amplifier and $\alpha\ (2 \le \alpha \le 4)$ denotes the propagation loss exponent. In detail, α is 2 for free space and increases to 4 when encountering obstacles. On the other hand, to receive a k-bits packet, the corresponding energy dissipation was shown as Expression (10)

$$e_{rx} = kE_{elec} \quad (10).$$

(2) Network Topology

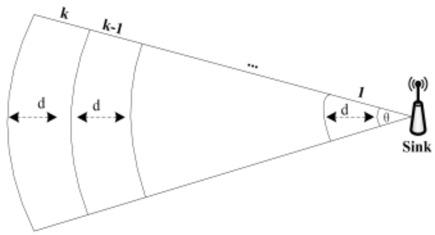

Fig.1 Network topology utilized in this paper

In this paper, assume a sector-shaped network was divided into k layers. The central angle of the sector was denoted as θ. This topology was similar to what was adopted in References [11][37]. The Sink was deployed at the center and each ring-shaped area was d in width, just as shown in Fig.1. Without loss of generality, the sector region can be either an absolute monitor area or just a part of a larger general region. Therefore, the conclusion on the optimal cluster size in section IV can be also applied to the region of any shapes, such as, rectangle, square, triangle, and so on.

(3) Assumptions and Symbols

For the sake of simplicity, some assumptions and symbols were introduced as follow.

All the nodes transmit the data at the fixed power, therefore the transmission radius is also constant. In this paper, the transmission radius was assumed to be a. None of the nodes can be replenished once deployed. On the contrary, the Sink has infinite energy supply and process capacity.

The nodes were independently and evenly distributed in the network topology, generally speaking, the distribution follows the Poisson Point Process[40][41][42][43].

Each sensor node kept stationary once deployed and was accurately aware of its own location and the coordinate with the help of GPS or any other location technologies[44][45]. The location can be used for selecting and announcing the information on the Cluster Head.

The data were generated at the same rate by all the nodes, thereby the amount of data which the network generated was proportional to the area of the network topology.

IV. Theoretical Basis

In this section, the theory of the optimal cluster size was proposed. Some lemmas and the corresponding proofs were presented firstly. Subsequently, the role of the Backup Cluster Head (BCH) and the rotation mechanism between the Cluster Head and the Backup Cluster Head were proposed.

A. Lemmas

Lemma 1 Assume the number of clusters at the kth layer is N_k and that of the clusters at the *(k-1)*th layer is denoted as N_{k-1}, they need to meet the condition $N_{k-1} = k\theta d^2/(M-1) + N_k$ in order to keep the energy consumption of the two adjacent layers approximately equal.

Proof: In order to keep the energy consumption of all the layers at an equilibrium state, the following expression should be met,

$$1/2\,(2k-1)\theta d^2 - N_k + MN_k = 1/2\,(2k-3)\theta d^2 - N_{k-1} + MN_{k-1}$$
(11).

Therefore the following relationship between N_k and N_{k-1} can be obtained,

$$N_{k-1} = k\theta d^2/(M-1) + N_k \quad (12).$$

In Expression (12), the energy consumption for both the transmitter and receiver between the two adjacent layers was taken into consideration. Therefore, the energy used for transmitting data from the Cluster Members to their corresponding Cluster Heads also needs to be taken into account in order to maximize the network lifetime. According to Lemma.1, the values of energy consumption of all the layers appear to be equivalent, so the energy consumption of the whole topology is minimized if that of the outermost layer is minimum.

Lemma.2 On the kth layer of the network topology, the mean hop count for each Cluster Member is $\pi a^2/N_k$, where a denotes the constant transmission radius for each sensor node.

Proof: The number of the clusters at the outermost layer is denoted as N_k, then the optimal size of the cluster is established as follow,

$$(1/2(2k-1)\theta d^2)/N_k \quad (13).$$

According to the assumptions, the transmission radius of each sensor node was fixed to be a. Therefore the average number of the transmission range is

$$(1/2(2k-1)\theta d^2)/\pi a^2 \quad (14).$$

The mean hop count for each sensor node is obtained as Expression (15)

$$\frac{(1/2(2k-1)\theta d^2)/N_k}{(1/2(2k-1)\theta d^2)/\pi a^2} \quad (15).$$

After simplification, Lemma.2 is proved.

Lemma.3 The number of clusters at the kth layer to minimize the energy consumption is $N_k^* = da\sqrt{\frac{1}{2}\theta(2k-1)\pi/M}$, where d, a, and M are the width of the kth layer, the transmission radius of each sensor node, and the amount of measurements that each cluster head needs to transmit in each round respectively.

Proof: For the sake of simplicity, the energy consumption of the kth layer is denoted as C_k which consists of the intra-cluster and the inter-cluster energy consumption. According to the energy consumption model in Section III, the energy consumption for per unit packet is proportional to the square or even fourth power of the transmission distance. Besides, according to the assumption, the node transmits data with the constant transmission radius which makes α equal to 2 in this paper. Therefore the hop count is adopted to replace the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2019.2893912, IEEE Sensors Journal

square of transmission distance to make the analysis briefer. Subsequently the value of C_k can be established as follow,

$$C_k = (1/2 \cdot \theta(2k-1)d^2 - N_k) \cdot (\pi a^2 / N_k) + MN_k \quad (16).$$

The derivative of C_k can be obtained as follow

$$C_k^{'} = 1/2\theta(2k-1)\, d^2 \cdot \pi a^2 (-1/N_k^2) + M \quad (17),$$

therefore when Expression $N_k^* = da\sqrt{\frac{1}{2}\theta(2k-1)\pi / M}$ is met, the energy consumption of the kth layer is minimum.

Lemma.4 Let d_{k-Sink} denote the distance from the node lying in the kth layer to the Sink, the energy consumption of each cluster is minimized if condition $d_{k-Sink} = (2k-1)\sin\frac{\theta}{2} d / \theta$ is met, which makes the Cluster Head distribution optimal.

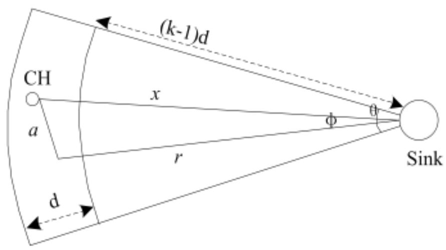

Fig.2 The optimal location of CH

Proof: As shown in Fig.2, assume the area of the kth layer equals to the size of a cluster for the sake of convenience. The Cluster Head lies in the location which is x far away from the Sink. Besides, assume another node which is randomly distributed at the point which is a far away from the Cluster Head. It acts as a Cluster Member. Obviously the location of the CH makes the energy consumption minimal when meeting the condition that the whole distance of all the CMs to the corresponding CHs should be kept minimal to cut down the energy consumption as much as possible. The square of the distance from an arbitrary Cluster Member to the corresponding Cluster Head is denoted as below,

$$a^2 = x^2 + r^2 - 2xr\cos\phi \quad (18).$$

Hence, the sum of the square of the distance from all the CMs to the corresponding CH in one cluster can be obtained as follow,

$$\sum a^2 = \int_{-\theta/2}^{\theta/2} \int_{(k-1)d}^{kd} (x^2 + r^2 - 2xr\cos\phi) d\phi dr \quad (19).$$

Finally, the following Expression can be obtained,

$$\sum a^2 = \theta dx^2 - 2(2k-1)d^2 \sin\frac{\theta}{2} x + 3\theta(3k^2 - 3k - 1)d^3 \quad (20).$$

The conclusion that $\sum a^2$ is minimal when $x = (2k-1)\sin\frac{\theta}{2} d / \theta$ is met can be drawn. According to the energy consumption model, the energy consumption can be minimized when Expression (20) is established.

B. Ratio of Energy to Distance

As for the traditional clustering algorithm, the Cluster Head bears a heavier traffic burden than the Cluster Member. Therefore a CH dissipates energy faster than its corresponding CMs. In order to achieve the energy consumption balance, the mechanism of rotating the role of the sensor node was adopted in the paper. The roles of the sensor node were rotated periodically to alleviate the "Hot Spot Problem". In the phase of rotation, the candidate exchanges message to become CH via broadcast. Apparently, the large amount of the broadcast packet leads to the quick dissipation of energy and the termination of network lifespan.

In order to reduce the energy consumption resulted from the Cluster Heads selection, the concept of a ratio of the residual energy to the distance was proposed in this paper. It represented the ratio of a node's residual energy to the distance from itself to the corresponding Cluster Head. For the sake of briefness, it was denoted as r_{E-d} . In this paper, its mathematical definition was presented as follow.

$$r_{E-d} = E_{res} / d_{s-CH} \quad (21),$$

where E_{res} and d_{s-CH} represented the residual energy of the sensor node and the distance from itself to the corresponding Cluster Head respectively.

C. Backup Cluster Head

In order to avoid exhausting the energy of the Cluster Head ahead of time owing to its heavier energy consumption burden, the mechanism of rotating the role of the Cluster Head was adopted to make the energy consumption more evenly. Compared with the traditional clustering algorithm, a third role which is called the Backup Cluster Head (BCH) was presented in this paper in addition to the roles of the Cluster Head and the Cluster Member. Specifically, the node with the second largest value of r_{E-d} was selected as the Backup Cluster Head by the present Cluster Head. The value and ID of the Backup Cluster Head were kept in the present Cluster Head's memory. Besides, the Backup Cluster Head was informed its role by the present Cluster Head via unicast.

V. Energy-Efficient Compressive Sensing-based Clustering Routing (EECSR) Protocol

A. Introduction of the EECSR Algorithm

The whole network topology was modeled as an undirected graph $G = <V, E>$, where the vertex set V consisted of all the sensor nodes as well as the Sink and the edge set E was composed of the active wireless link which connects any two nodes in set V . For any two nodes v_i and v_j in the set V , the wireless link was active if the following condition was established,

$$|v_i - v_j| \le a \quad (22),$$

where $|v_i - v_j|$ denoted the Euclidean distance between nodes v_i and v_j .

Once the Sink obtains the number of clusters at each layer, it will determine the optimal Cluster Head distribution. Subsequently, the Sink broadcasts the Cluster Head information to all the Cluster Members. Generally speaking, the broadcast contains the optimal location of the Cluster Heads and the corresponding layer ID.

The algorithm in this paper adopted the Compressive Sensing theory and there were two kinds of data collecting modes. Specifically, the intra-cluster data collection and inter-cluster data collection coexisted in the whole network topology. The intra-cluster data were collected via Compressive Data Gathering (CDG) mode and the compressed data were transmitted to the upstream CHs via "hop-by-hop" mode. Therefore, each cluster acquired the source data via CDG mode independently. And the compressed data were transmitted among the CHs without any further process to the Sink. As for the cluster consisted of isolated node, it collected the data via RDG mode. Since the whole large scale topology was divided into small topology, it provided a higher Scalability. When the Sink received the data from each cluster, it conducted reconstruction algorithm.

B. Details of the EECSR Protocol

In this section, the Energy-Efficient Compressive Sensing-based clustering Routing (EECSR) protocol was detailed. On the whole, it included three main phases, namely, the cluster formation phase, the spanning tree construction phase and the data acquisition phase. The generation and the announcement of the Cluster Head, and the process of the Cluster Member's choosing the optimal Cluster Head happened in the clustering phase. The spanning tree construction phase generated a shortest path from the Cluster Heads to the Sink. As for the data transmission phase, the data collected by the sensor nodes were transmitted to the Sink through "hop-by-hop" mode.

(1) Cluster Formation Phase

Once the Sink got the values of the parameters, namely, *d*, *a*, *k*, and *M*, it calculated the optimal number of clusters at each layer basing on Expression (17). Subsequently, the Cluster Heads of each layer were selected according to the optimal number of Cluster Head at the $ith\ (i=1,2,\dots,k)$ layer and the number of Cluster Head was determined to be $N_i\ (i=1,2,\dots,k)$. Specifically, the Sink announced all the information about the optimal number of CH and the optimal Cluster Head distribution to all the sensor nodes in the form of broadcast. Once obtaining the broadcast from the Sink, the sensor node calculated the ratio of the residual energy to the distance and kept it in its memory. Subsequently, each node flooded a type of packet (HELLO) which contained the Cluster Head ID and the value of r_{E-d} to select the optimal Cluster Head. The node compared the ratio with its own when receiving the HELLO packets. If its own ratio was smaller, the node chose the node identified by the packet to be CH and forwarded the packet to its upstream neighbor. Otherwise, it discarded the packet and generated a new cluster head selection packet.

Repeated the above steps until all the sensor nodes joined the clusters. For the node who has not received any HELLO packets, it selected itself to be the Cluster Head to form an isolated cluster which only contains itself.

At the end of each round, the Cluster Head compared its own ratio r_{E-d} with that of the Backup Cluster Head, it would hand over the role of the Cluster Head to the Backup Cluster Head if its own ratio was smaller than that of the Backup Cluster Head. In this manner, the rotation happened between the Cluster Head and the Backup Cluster Head to reduce the rotation overhead. After the Backup Cluster Head received the hand-over message from the Cluster Head, it broadcasted to inform all the Cluster Members. Therefore it turned to the next cluster formation phase. At the same time, the new Cluster Head generated the new Backup Cluster Head basing on the ratio of the Cluster Members.

(2) Spanning Tree Construction Phase

Once all the Cluster Heads have been selected, the backbone routing tree needed to be constructed. The vertexes of the backbone routing tree consisted of all the Cluster Heads and the Sink. For the sake of simplicity, the concept of sub-graph was presented in this section. Suppose all the Cluster Heads constituted the vertex set V_{CH} , then the sub-graph introduced by V_{CH} and the Sink was denoted as $G[V^{'}]=(V^{'},E^{'})$. In the sub-graph $G[V^{'}]$, the set $V^{'}$ was the union of V_{CH} and $\{Sink\}$, namely, $V^{'}=V_{CH}\cup\{Sink\}$, and edge set $E^{'}$ consisted of the logical paths from the Cluster Heads to the Sink. It was obvious that the sub-graph $G[V^{'}]$ was a kind of complete graphs. The Shortest Path Tree (SPT) can be introduced by the sub-graph $G[V^{'}]$ and then the backbone routing tree was established.

(3) Data Acquisition Phase

The CH generated the pseudo-random seed which was used for generating the sub-measurements matrix and sent it to its CMs. Subsequently, the CMs generated the sub-measurement and each cluster acquired the data via the CDG mode according to the sub-measurement. As for the isolated cluster, it collected the data via RDG mode. After intra-cluster data acquisition, all the data were sent to the Sink through the backbone routing tree. Besides, the pseudo-random seed was also sent together with the compressed data to the Sink.

Once the backbone routing tree was constructed, each Cluster Head had the whole information about the backbone routing tree. The intra-cluster data acquisition via CDG mode, each CM processed the data via add or multiply operations and then sent them to its corresponding CH. Subsequently the CHs forwarded the compressed data to their upstream Cluster Head towards the Sink. Once received the measurements from the Cluster Head, the Sink generated the corresponding sub-measurements according to the pseudo-random seed and recovered the compressed data through the data recovery algorithm. For the data transmitted by the isolated cluster, the Sink can simply decapsulate the packet and obtained the source data.

VI. Simulation and Results Analysis

In this section, the performance of EECSR was evaluated via

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2019.2893912, IEEE Sensors Journal

the simulator NS2. EECSR was a kind of algorithm which adopted both of the merits of CS-based algorithm and the Clustering scheme, therefore it needed to be compared with the CS-based algorithm and the Clustering algorithm in terms of energy efficiency and the throughput. As for the traditional Clustering algorithm, there exists much data redundancy compared with the Compressive Sensing technology. The data redundancy leads to a waste of energy. Besides, it also makes the traffic burden of the Cluster Heads much heavier than that of the Cluster Member, which results in the “Hot Spot Problem”. Both of the data redundancy and the “Hot Spot Problem” shorten the lifetime of WSNs to some extent. However, the EECSR integrated the merits of Compressive Sensing theory with that of the Clustering algorithm. The energy consumption resulted from the data redundancy can be alleviated effectively since the CS theory was adopted. Besides, the rotation of roles contributed to the energy balance. In this section, the algorithm EECSR was compared with the LEACH, TEEN, PEGASIS, and LEACH-ERE which belongs to the Clustering algorithm in terms of network lifetime and the throughput.

On the other hand, for the CS-based schemes without considering the energy overhead of leaf nodes which referred to CDG in this paper, the waste of energy is resulted from the data redundancy of the leaf node. In addition, the energy overhead exists in the process of cluster formation for those schemes which integrated CS theory with the Clustering strategy and was denoted as HCDG in some literatures. Although there are some existing CS-based strategies which were combined with the Clustering scheme, they did not take the energy overhead resulted from cluster formation into consideration. Compared with the existing CS-based algorithm, EECSR was integrated with the Clustering algorithm. Besides, it not only considered the energy cost for cluster formation, but also presented the ratio of the residual energy to the distance and the role of the Backup Cluster Head and the corresponding rotation mechanism between CH and BCH. In addition, the optimal cluster size and the optimal Cluster Head distribution were obtained to balance the energy consumption. So it can reduce the traffic burden of the Cluster Heads and balance the energy consumption of WSNs at the same time. Therefore, EECSR also needed to be compared with Clustering strategies as well as the existing CS-based algorithms, such as CDG and HCDG, to evaluate its effect on reducing the data redundancy. Therefore this section also shown the comparisons with PEGASIS, the CDG strategy and the HCDG strategy in terms of the reduction ratio of transmission[31].

In this section, all the sensor nodes were randomly distributed in a 100×100 square area and the Sink was deployed at the center of the area with the coordinate $(0,0)$. The value of d was set to be 20, therefore the value of k equaled 3. Since the Sink located at the center of the network topology, the value of θ was 2π. In order to comprehensively evaluate the energy efficiency of EECSR, two kinds of experiment scenarios were adopted, namely the scenario with changeable number of the sensor nodes and that with constant number of the sensor nodes. For the sake of convenience, they were denoted as scenario 1 and scenario 2 respectively. In scenario 1, the number of the sensor nodes varied from 100 to 400 with the step 100 and the results of it were compared with CDG and HCDG. In scenario 2, the number of the sensor node were fixed to be 200 and the results were compared with the traditional clustering algorithm, such as LEACH, PEGASIS, TEEN, and LEACH-ERE. To make the comparison more conspicuous, the compression ratio $\rho = N/M$ was proposed in this section. In the experiment, the value of ratio was set to be 5 or 10 and it met the demand of the precision for data recovery[25][43].

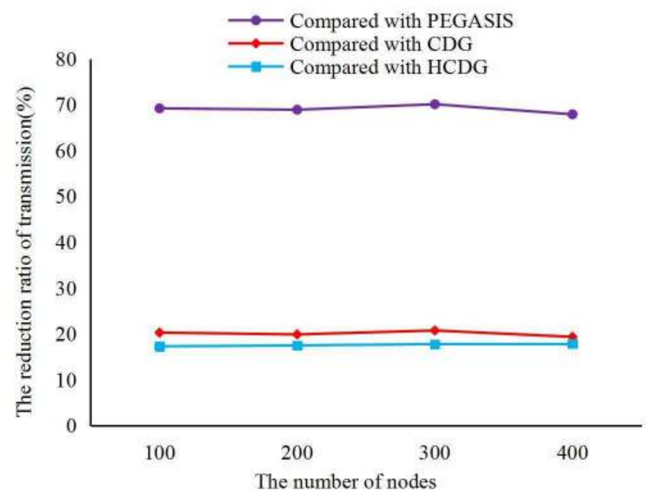

Fig.3 The reduction ratio of transmission when $\rho = 5$

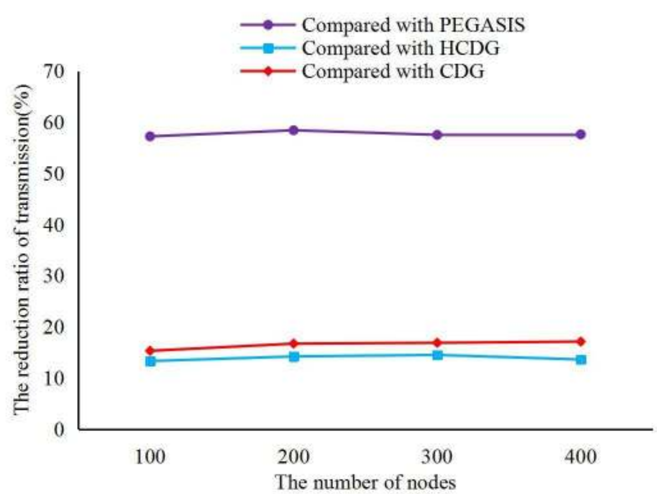

Fig.4 The reduction ratio of transmission when $\rho = 10$

Figs.3-4 shown how the reduction ratio of transmission varied with the number of the sensor nodes when $\rho = 5$ and $\rho = 10$ respectively. Compared with the Clustering strategy without compressive sensing technology which was PEGASIS in this paper, the amount of data transmission was reduced by 69% and 58.3% when the transmission ratio were 5 and 10 respectively. When compared with the CDG scheme, the proposal in the paper can achieve the reduction ratio to be 19.5% and 15.4% when ρ equaled to 5 and 10 respectively. In addition, the amount of data transmission was reduced by

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2019.2893912, IEEE Sensors Journal

17.4% and 14.2% compared with the HCDG scheme when ρ equaled to 5 and 10 respectively. Therefore the conclusion that EECSR algorithm can effectively reduce the energy consumption caused by the cluster formation phase can be drawn. Besides, the comparisons of Figs.3-4 shown that the reduction ratio of data transmission of EECSR can keep steady and be free from the influence of the increase of the compressive ratio. Therefore the EECSR algorithm is steady and has good performance in improving the energy efficiency.

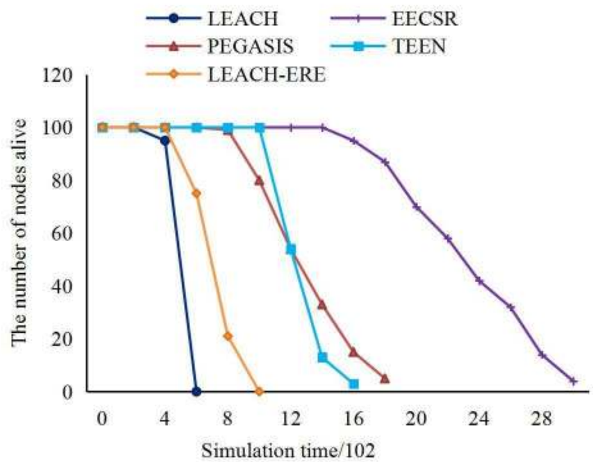

Fig.5 The number of nodes alive when $\rho = 5$

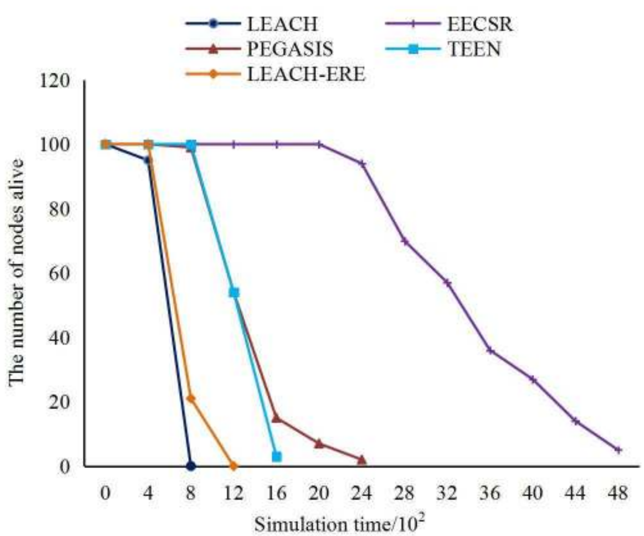

Fig.6 The number of nodes alive when $\rho = 10$

Figs.5-6 shown the comparisons of the number of nodes alive when $\rho = 5$ and $\rho = 10$ respectively. The simulations were conducted with the constant number of sensor node being 100. EECSR was compared with the traditional clustering protocol, such as the LEACH, the PEGASIS, and the TEEN. Besides, to verify the advantage of the proposal of the paper, EEREG was also compared with the LEACH-ERE which was one of the state-of-the-art energy-efficient strategies. AS was shown in Figs.5-6, the curves of LEACH, PEGASIS, TEEN, and LEACH-ERE kept the same with different compressive ratios because the CS was not adopted. On the contrary, the compressive ratio played an important role in the network lifetime for EECSR because the CS theory was utilized to reduce the data redundancy. Besides, it can be easily obtained that the amount of traffic was reduced with the increase of the compressive ratio. Therefore the network lifespan extended with the increase of the compressive ratio. Besides, it was also easily obtained that the network lifetime did not linearly vary with the compressive ratio strictly. Since the hybrid compressive sensing theory was adopted in EECSR, the data transmission among the intra-cluster did not decrease linearly with the compressive ratio.

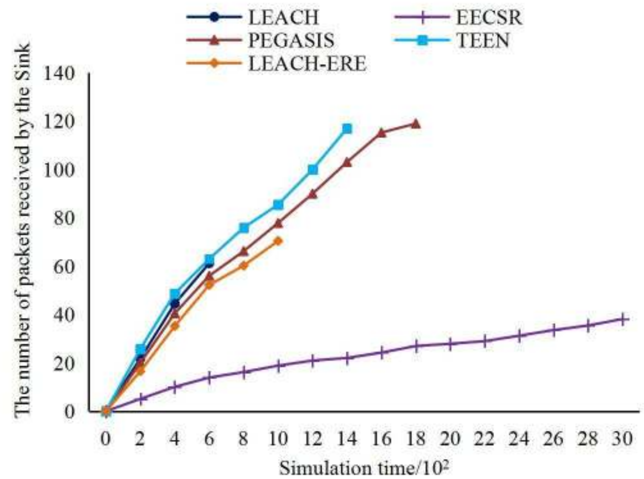

Fig.7 The throughput comparison when $\rho = 5$

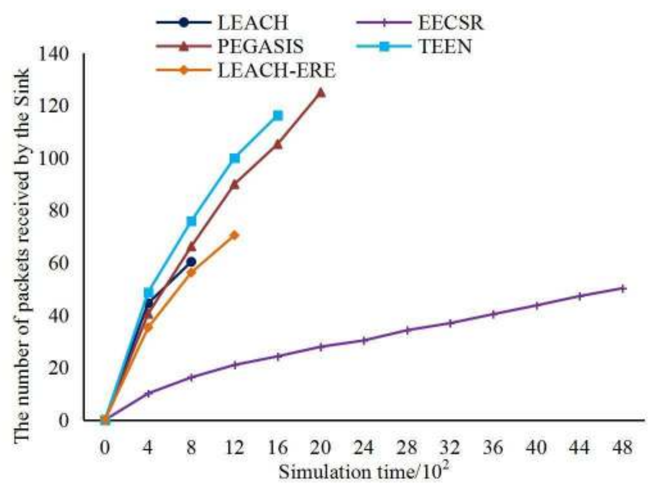

Fig.8 The throughput comparison when $\rho = 10$

Figs.7-8 shown the throughput comparisons when $\rho = 5$ and $\rho = 10$ respectively. It was clear that EECSR can largely reduce the amount of the traffic flow compared with the Clustering protocol, such as LEACH, PEGASIS, TEEN, and LEACH-ERE. It is obvious that the compressive sensing mode contributed to cutting down the data redundancy. Besides, Fig.8 further confirmed that the energy consumption burden decreased with the increase of the compressive ratio. The amount of data transmission decreased, therefore the energy consumption for transmission was reduced and the network lifetime was extended.

VII. Conclusions and Future Research Directions

A. Conclusions

The energy constraint has been one of main challenges facing the traditional WSNs for a long time. The "Hot Spot Problem" leads to the early termination of the network lifetime. The broadcast used for rotation of CHs also leads to a waste of energy. On the other hand, the Spatial-temporal correlation during the process of data acquisition makes matters worse. In this paper, the Clustering strategy was integrated with the CS theory to reduce the energy consumption resulted from the Spatial-temporal correlation. Besides, the Backup Cluster Head and the rotation mechanism was proposed to cut down the energy depletion caused by the broadcast during Cluster Head rotation. In addition, the optimal cluster size, the optimal Cluster Head distribution were proposed to further improve the energy efficiency. Subsequently, the Energy-Efficient Compressive Sensing-based clustering Routing (EECSR) protocol was presented. Finally, the EECSR was evaluated via the network simulator NS2 in terms of reduction ratio of transmission, network lifetime, and the throughput. The simulation results shown that the EECSR algorithm can effectively save energy and extend the network lifetime.

B. Future Research Directions

Since the temporal correlation exists in the data collected by the Sink, the data acquired in the adjacent time slots presents some similarity. Therefore, the predictive coding theory can be combined to reduce the energy consumption further. Via adopting predictive coding, the data acquired previous can be utilized to predict the later data, thus reduce the amount of data acquired and achieve high energy efficiency. As far as we know, there have not been any papers aiming at improving energy efficiency of WSNs via integrating with the predictive coding theory. Furthermore, the predictive coding can also be integrated with the CS theory to reduce the energy consumption resulted from the Spatial-temporal correlation.

References

[1] D. Lin and Q. Wang. "A game theory based energy efficient clustering routing protocol for WSNs," *Wirel. Netw*., vol. 23, no. 4, pp. 1101-1111, 2017.

[2] B. Erkal and O. Kaynak. "Design of a fuzzy variable structure controller for controlling satellite attitude suffering from sensor data delay," *Proceedings of 5th International Conference on Recent Advances in Space Technologies-RAST2011*, Istanbal, Turkey, 2011, pp. 566-569.

[3] Z. Sheng, S. Yang, Y. Yu, et al. "A survey on the ietf protocol suite for the internet of things: standards, challenges, and opportunities," *IEEE Wirel. Commun.*, vol. 20, no. 6 pp. 91-98, 2013.

[4] L. Parra, S. Sendra, J. Lloret, et, al. "Low cost wireless sensor network for salinity monitoring in mangrove forests," *Proceedings of the 2014 IEEE on Sensors*, Valencia, Spain, 2014, pp. 126-129.

[5] J. Lloret, S. Sendra, M. Garcia, et, al. "Group-based underwater wireless sensor network for marine fish farms," *Proceedings of the 2011 IEEE GLOBECOM Workshops (GC Wkshps)*, Houston, TX, USA, 2011, pp. 115-119.

[6] S. Yoon and C. Shahabi. "The clustered aggregation (CAG) technique leveraging spatial and temporal correlations in wireless sensor networks," *ACM T. Sensor Network*, vol. 3, no. 1, pp. 3, 2007.

[7] N. A. Pantazis, S. A. Nikolidakis, and D. D. Vergados. "Energy-efficient routing protocols in wireless sensor networks: a survey," *IEEE Commun. Surv. Tut.*, vol. 15, no.2 pp. 551-591, 2013.

[8] A. Mehmood, S. Khan, B. Shams, and J. Lloret. "Energy-efficient multi-level and distance-aware clustering mechanism for WSNs," *Int. J. Commun. Syst.*, vol. 28 , no.5, pp. 972-989, 2015.

[9] J. Lloret, M. Garcia, D. Bri, et, al. "A cluster-based architecture to structure the topology of parallel wireless sensor networks," *Sensors*, vol. 9, no.12, pp. 10513-10544, 2009.

[10] D. Lin, Q. Wang, D. Lin, et al. "An energy-efficient clustering routing protocol based on evolutionary game theory in wireless sensor network," *Int. J. Distrib. Sens. N.*, vol. 2015, no. 2015, pp. 1-12, 2015.

[11] D. Lin, Q. Wang and J. Liu. "An energy efficient sink deployment scheme aiming at extending the lifespan for wireless sensor networks," *Int. J.Fut. Gene. Commun. N.*, vol. 9, no. 8, pp. 77-86, 2016.

[12] X. Wu, G. Chen and K. D. K. Sajal. "Avoiding energy holes in wireless sensor networks with nonuniform node distribution," *IEEE T. Parall. Distr.*, vol. 19, no. 5, pp. 710-720, 2008.

[13] S. Basagni, A. Carosi, E. Melachrioudis, et al. "Controlled sink mobility for prolonging wireless sensor networks lifetime," *Wirel. Netw*., vol. 14, no. 6, pp. 831-858, 2007.

[14] D. Lin, Q. Wang and J. Liu. "Energy-saving strategy by combining mobile and static sinks schemes for wireless sensor networks," *Journal of Harbin Institute of Technology*, vol. 48, no. 11, pp. 162-168, 2016.

[15] J. Yang, D. Zhang, Y. Zhang, et, al. "A cluster-based data aggregation and transmission protocol for wireless sensor networks," *Journal of Software*, vol. 21, no. 5, pp. 1127-1137, 2010.

[16] D. L. Donoho and M. Elad and V. N. Temlyakov. "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE T. Inform. Theory*, vol. 52, no. 1, pp. 6-18, 2006.

[17] W. Chen and I. J. Wassell. "Energy-efficient signal acquisition in wireless sensor networks: a compressive sensing framework," *IET Wireless Sensor Systems*, vol. 2, no.1, pp. 1-8, 2012.

[18] D. Slepian and J. Wolf. "Noiseless coding of correlated information sources," *IEEE T. Inform. Theory*, vol. 19, no. 4, pp. 471-480, 1973.

[19] G. Hua and C. W. Chen. "Correlated data gathering in wireless sensor networks based on distributed source coding," *Int. J. Sens. Netw*., vol. 4, no. 1, pp. 13-22, 2008.

[20] M. J. Handy, M. Haase and D. Timmermann. "Low energy adaptive clustering hierarchy with deterministic cluster-head selection," *The 4th International Workshop on Mobile and Wireless Communications Network,* Stockholm, Sweden, Sweden, 2002, vol. 1, pp. 368-372.

[21] S. Lindsey and C. S. Raghavendra. "PEGASIS: power-efficient gathering in sensor information systems," *Proceedings of the IEEE Aerospace Conference Montana. IEEE Aerospace and Electronic Systems Society*, Big Sky, MT, USA, 2002, pp. 1125-1130.

[22] A. Manjeshwar and D. Agrawal. "TEEN: a routing protocol for enhanced efficiency in wireless sensor networks," *Proceedings of the 15th International Parallel and Distributed Processing Symposium (IPDPS'01) Work-shops*, California, USA, 2001, pp. 2009-2015.

[23] J. Lee and W. Cheng. "Fuzzy-logic-based clustering approach for wireless sensor networks using energy predication," *IEEE Sens, J.*, vol. 12, no. 9, pp. 2891-2897, 2012.

[24] H. Zheng, F. Yang, X. Tian, et al. "Data gathering with compressive sensing in wireless sensor networks: a random walk based approach," *IEEE T. Parall. Distr.*, vol. 26, no. 1, pp. 35-44, 2015.

[25] H. Zheng, W. Guo and N. Xiong. "A kernel-based compressive sensing approach for mobile data gathering in wireless sensor network systems," *IEEE T. Syst. Man. Cy-S.*, vol. 99, pp. 1-13, 2017.

[26] L. Quan, S. Xiao, X. Xue, et al. "Neighbor-aided spatial-temporal compressive data gathering in wireless sensor networks," *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 578-581, 2016.

[27] L. Xiang, J. Luo and C. Rosenberg. "Compressed data aggregation: energy-efficient and high-fidelity data collection," *IEEE/ACM Transactions on Networking*, vol. 21, no. 6, pp. 1722-1735, 2013.

[28] X. Yang, X. Tao, E. Dutkiewicz, et al. "Energy-efficient distributed data storage for wireless sensor networks based on compressed sensing and network coding," *IEEE T. Commun*., vol. 12, no. 10, pp. 5087-5099, 2013.

[29] C. Luo, F. Wu, J. Sun, et al. "Compressive data gathering for large-scale wireless sensor networks," *Proceedings of the 15th annual international conference on Mobile computing and networking*, Beijing, China, 2009, vol. 14, no. 1, pp. 145-156.

[30] S. Yiran, H. Wen, R. Rajia, et al. "Nonuniform compressive sensing for heterogeneous wireless sensor networks," *IEEE Sens. J.*,, vol. 13, no. 6, pp. 2120-2127, 2013.
[31] R. Xie and X. Jia. "Transmission-efficient clustering method for wireless sensor networks using compressive sensing," *IEEE T. Parall. Distr.*, vol. 25, no. 3, pp. 806-814, 2014.
[32] C. Zhao, W. Zhang, Y. Yang, et al. "Treelet-based clustered compressive data aggregation for wireless sensor networks," *IEEE T. Veh. Technol.*, vol. 64, no. 9, pp. 4257- 4267, 2015.
[33] K. Lan and M. Wei. "A compressibility-based clustering algorithm for hierarchical compressive data gathering," *IEEE Sens. J.*, vol. 17, no. 8, pp. 2550-2562, 2017.
[34] X. Li, X. Tao and G. Mao. "Unbalanced expander based compressive data gathering in clustered wireless sensor networks," *IEEE Access*, vol. 5, no. 99, pp. 7553-7566, 2017.
[35] W. Yan, Y. Dong, S. Zhang and Z. Wang. "An optimal CDG framework for energy efficient WSNs," *Chinese Journal of Electronics*, vol. 26, no. 1, pp. 138-144, 2017.
[36] X. Xu, R. Ansari and A. Khokhar. "Spatio-temporal hierarchical data aggregation using compressive sensing (ST-HDACS)," *Proceedings of the 2015 International Conference on Distributed Computing in Sensor Systems*, Fortaleza, Brazil, 2015, pp. 91-97.
[37] D. Xie, Q. Zhou and J. Liu. "A chain-based data gathering protocol under compressive sensing framework for wireless sensor networks," *Proceedings of the 2013 International Conference on Computational and Information Sciences*, Washington, DC, 2013, pp. 1479-1482.
[38] Z. Nie, D. Wang, D. Wang, et al. "Adaptive adjustment of compressed measurement for energy-efficient data gathering in WSNs," *Proceedings of the 2016 IEEE Conferences on Internet of Things and IEEE Green Computing and Communications and IEEE Cyber, Physical and Social Computing and IEEE Smart Data*, Chengdu, China, 2016, pp. 437-440.
[39] E. Shih, S. Cho, N. Ickes, et, al. "Physical layer driven protocol and algorithm design for energy-efficient wireless sensor networks," *Proceedings of the ACM Mobicom*, Rome Italy, 2002, pp. 272-287.
[40] O. Younis and S. Fahmy. "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE T. Mobile Comput.*, vol. 3, no. 4, pp. 366-379, 2004.
[41] S. Bandyopadhyay and E. Coyle. "An energy efficient hierarchical clustering algorithm for wireless sensor networks," *Proceedings of the Joint Conference of the IEEE Computer and Communications IEEE Societies*, San Francisco, CA, USA, 2003, pp. 1713-1723.
[42] D. Wang and L. X. Lin L. "A study of subdividing hexagon-clustered WSN for power saving: analysis and simulation," *Ad Hoc Netw.*, vol. 9, no. 7, pp. 1302-1311, 2011.
[43] S. Chen, Y. Wang, X. Y. Li and X. Shi. "Data collection capacity of random-deployed wireless sensor networks," *Proceedings of the IEEE Global Telecommunication Conference*, Honolulu, HI, USA, 2009, vol. 17, no. 5, pp.1-6.
[44] K. Yedavalli and B. Krishnamachari "Sequence-based localization in wireless sensor networks," *IEEE T. Mobile Comput.*, vol. 7, no. 1, pp. 81-94, 2008.
[45] A. Nasipuri and K. Li "A directionality based location discovery scheme for wireless sensor networks," *Proceedings of the First ACM Int'l Workshop Wireless Sensor Neworks and Applications (WSNA '02)*, Georgia, USA, 2002, pp. 105-111.