Data Collection Method Based on Data Perception and Positioning Technology in the Context of Artificial Intelligence and the Internet of Things

Xinbo Zhao[1]*, Fei Fei[2]
School of Management, Liaoning University of International Business and Economics, Dalian, 116052, China[1]
School of Art & Design, Dalian Polytechnic University, Dalian, 116034, China[2]

***Abstract*—Wireless sensor networks are an important technical form of the underlying network of the Internet of Things. The energy of each node in the network is finite. When a node runs out of energy, it can cause network interruptions, which can affect the reliability of data collection. To reduce the consumption of communication resources and ensure the reliability of data collection, the study proposes data collection based on data compression perception positioning technology. This method first uses a Bayesian compression perception method to select nodes, and then adopts an adaptive sparse strategy to collect data. When selecting nodes using this proposed method, wireless sensor networks had the longest network lifespan. In the case of different degrees of redundancy and sparsity, the research method had the lowest reconstruction error, with reconstruction errors of 0.31 and 0.40, respectively. When the balance factor was set to 0.6, the reconstruction error of the research method was the lowest, with a minimum reconstruction error of 0.05. This proposed method has better reconstruction performance, effectively prolongs the lifespan of wireless sensor networks, and reduces the consumption of communication resources.**

***Keywords*—Wireless sensor network; data collection; compression perception technology; Sparse Bayesian Learning; signal reconstruction**

I. Introduction

The development and application of the Internet of Things (IoT) have brought great convenience to people's lives. IoT can connect any object in the physical world with the Internet to achieve real-time monitoring, intelligent control, remote operation and other functions, so as to achieve interconnection between things and people [1-2]. Wireless Sensor Network (WSN) is a key technological component of IoT. Since its inception, WSN has been a hot topic in information research and has been applied to various aspects of society, such as military, agriculture, industry, healthcare, intelligent transportation, and home furnishings [3-4]. However, the energy of nodes in WSN is limited. When energy is depleted, the network will be interrupted, affecting the lifespan of WSN [5]. In addition, there are still issues with anomalies and missing data collected by current wireless sensor network data collection methods, which can result in significant consumption of communication resources. Studying data collection methods can improve the performance of wireless sensor networks, making them more widely applicable in fields such as healthcare, military, and environmental monitoring. Then, compression perception perfectly reconstructs the signal through nonlinear reconstruction algorithms [6]. Therefore, to reduce the consumption of communication resources and ensure the reliability of data collection, the study proposes data collection based on data compression perception positioning technology.

When using data compression perception positioning technology for data collection, the innovative approach is to first use Bayesian compression perception for node selection. Furthermore, an adaptive sparse strategy was adopted for data collection in the experiment. The main contribution of this study is the proposal of data collection based on data compression perception positioning technology, which reduces the energy consumption of WSN and improves the reliability of data collection.

The study will investigate data collection methods based on data compression perception positioning technology from four aspects. Firstly, a review is conducted on the current research status of data collection methods based on data perception and localization technology in the context of artificial intelligence IoT. Next is the research on data collection methods based on data compression perception positioning technology. Then, experimental verification is conducted on the proposed method. Finally, a summary of the research content is provided.

II. Related Works

With the continuous development of IoT, the network scale is getting larger and the network environment is becoming more and more complex. This poses significant challenges to energy conservation, transmission efficiency, security, and other aspects of network communication. Compression perception helps to address these issues in intelligent network communication [7-8]. Wang Y et al. proposed a lightweight method based on deep learning to reduce the computational cost of traditional deep learning methods. The sparsity of the scaling factor was enforced through compression perception. This proposed method effectively reduced the model size and accelerated the calculation speed [9]. Cheng G et al. proposed a hyper chaotic image encryption scheme based on quantum genetic algorithm and compression perception. This eliminated the drawbacks of weak key flow, small key space, and small information entropy in chaotic image encryption schemes. This proposed method was more efficient in resisting statistical attacks and plaintext attacks [10]. Liang P et al. proposed a compression perception technique to address the overwhelming

*Corresponding Author

feedback overhead caused by a large number of antennas in base stations. Secondly, a deep learning-based signal recovery solver framework was used on the base station. This proposed method was superior to existing methods and reduced feedback overhead [11]. The remote sensing image size was relatively large in the new type of hyper chaotic system. Nan S et al. proposed a block compression encryption algorithm that combined a novel hyper chaotic system with block compression perception. This proposed algorithm had good encryption performance, reconstruction accuracy, and anti-attack ability [12]. Xu G et al. proposed a sparse synthetic aperture radar imaging technique based on compression perception and machine learning. This solved the limitation of data bandwidth on the resolution of synthetic aperture radar images. This proposed method further investigated sparse imaging in machine learning and improved image resolution [13].

Data collection is a fundamental and important operation in WSN applications [14]. Aziz A et al. proposed an efficient aggregation scheme for multi-hop clusters based on hybrid compression perception, which effectively combined compression perception and routing protocols to collect data. This reduced the energy consumption of sensor nodes and extended the WSN lifespan. This proposed method was more efficient in collecting data and prolonged the WSN lifespan [15]. Sekar K et al. proposed a data collection method based on compression perception. Random singular value decomposition was used in this experiment to achieve compressed matrix factorization. This reduced the amount of data collected and lowers communication costs. This proposed method had higher accuracy at lower sampling rates, reduced data transmission costs, and extended the lifespan of sensors [16]. Mei Y et al. proposed a sampling algorithm based on compression perception to sample conditioned signal data. This changed the design paradigm of low latency large-scale access requiring random access schemes. This proposed algorithm reduced data sample loss when data samples were transmitted wirelessly and could be used for wind turbine fault detection [17]. Lin C et al. proposed a hierarchical data collection scheme to improve the linear programming effect in agricultural detection and reduce energy consumption in data collection. Data sampling was carried out using precise and greedy methods using mixed compression. This proposed method effectively collected data and planed the path of drones with lower energy consumption [18]. Chang CY et al. proposed a mobile receiver for multi-rate data acquisition to improve receiver speed and obtain fresh data. This proposed method significantly shortened the path length of mobile receivers and collected data comprehensively and efficiently [19].

In summary, with the continuous development of the Internet of Things, wireless sensor networks are being applied in more fields. In the process of collecting data, many scholars and scientists have designed a large number of improved models based on compressed sensing to solve the problem of complex network environments and noise interference leading to shortened lifespan of wireless sensor networks. These models effectively reduce the amount of data collected, but ignore the impact of sensor nodes themselves on wireless sensor networks. Therefore, the study proposes data collection and sensor node selection based on compressive sensing positioning technology.

III. Data Collection Method based on Data Compression Perception Technology

A. *Data Collection Method Based on Adaptive Sparse Strategy*

Bayesian compression perception is a method that combines Bayesian theory and compression perception to solve signal reconstruction problems. Considering the prior distribution of the signal, Bayesian compression perception can more accurately estimate the original signal and adaptively match the signal sparsity. The main advantage is that it can better handle noise and uncertainty [20-21]. IoT and artificial intelligence technology are constantly evolving. Artificial intelligence IoT has been applied in smart homes, intelligent transportation, intelligent healthcare, and other fields, bringing great convenience to humanity. As the sensing terminal of IoT, WSN mainly consists of three parts, including nodes, sensor networks, and users. Each node in the network needs to reserve energy for long-term use. Energy depletion can cause network interruptions [22-23]. In addition, sensor networks suffer from noise interference, data anomalies, and missing issues, which affect the reliability of data collection. To reduce the consumption of communication resources and ensure the reliability of data collection, the study adopts an adaptive sparse strategy for data collection. The position of the model sensor nodes is fixed, the composition structure is the same, and the energy is sufficient. Fig. 1 shows the WSN structure.

Sparse Bayesian learning is a method for reconstructing sparse signals, which does not require setting regularization parameters. It first considers the observation of a vector and outputs the sample y affected by noise interference, expressed as Eq. (1).

$$y = \Phi\omega + \varepsilon \tag{1}$$

In Eq. (1), ε refers to the observation noise. $\Phi\omega$ refers to the mean. Φ refers to the observation matrix. ω refers to the weights used by model learning to construct each column in Φ. The purpose of Sparse Bayes is to improve the model's generalization ability, interpretability, and efficiency by introducing sparsity, that is, selecting only a few important features. The Gaussian likelihood function model is represented by Eq. (2).

$$p\left(y \mid \omega, \sigma^2\right) = (2\pi\sigma^2)^{-N/2} \exp\left\{-\frac{1}{2\sigma^2}\left\|y - \Phi\omega\right\|_2^2\right\} \tag{2}$$

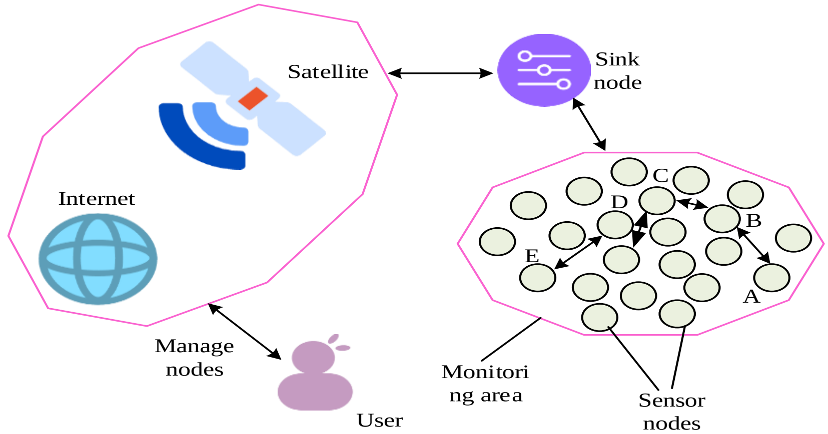


Fig. 1. Wireless sensor network architecture diagram.

In Eq. (2), y refers to the target variable, σ^2 refers to the variance, and N refers to the matrix of the sample. In this case, the task of obtaining maximum likelihood estimation is equivalent to finding the minimum norm solution while obtaining non-sparse solutions. Therefore, to find sparse solutions, the sparse Bayesian learning model estimates parameterized prior weights from the data, represented by Eq. (3).

$$p(\omega;\gamma)=\prod_{i=1}^{M}(2\pi\gamma_i)^{-\frac{1}{2}}\exp\left(-\frac{\omega i^2}{2\pi\gamma_i}\right) \quad (3)$$

In Eq. (3), γ represents M hyperparameter vectors. Secondly, Bayesian inference is performed. A posterior distribution is applied to all unknown variables to obtain the mean and covariance of the weight parameters. Finally, the hyperparameters are updated and estimated using Eq. (4).

$$(\sigma^2)^{new}=\frac{\|y-\Phi u\|_2^2}{N-\sum_i \gamma_i} \quad (4)$$

In Eq. (4), u represents the posterior mean. In order to solve the problem of excessive energy consumption caused by ineffective estimation in noisy environments, a compressed sensing adaptive sparse strategy was adopted to collect data. The steps are as follows: first, the sampling rate for data collection is determined to generate sampling random weights. Next, determine whether the data collection is initiated by a node. If the data collection is initiated by a node, then collect perception data and use the compressed perception adaptive sparse algorithm in the data sorting center to compress the original data. The data collection is complete. On the contrary, if there is no initiating node for data collection, it is determined whether the child node data has been received. If the child node data is received, the perception data is collected, and the compressed perception adaptive sparse algorithm is used in the data sorting center to compress the original data, and the data collection is completed. If no child node data is received, data collection ends. The data collection process is shown in Fig. 2.

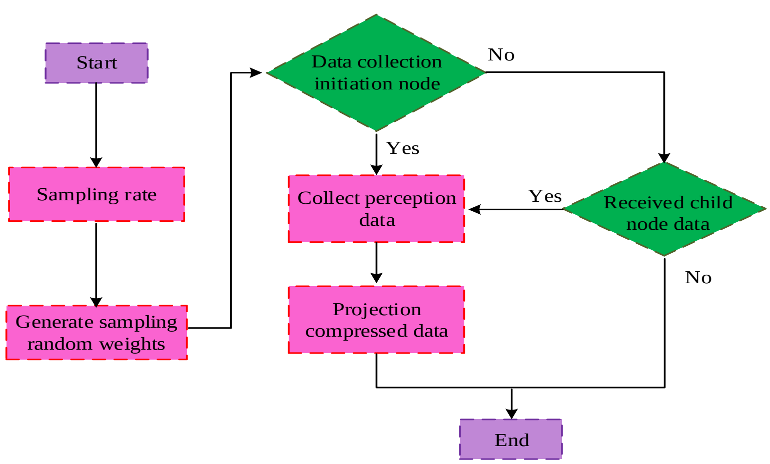


Fig. 2. Data collection flowchart.

The study adopts an adaptive sparse strategy based on compression perception for data collection. The data collection model is represented by Eq. (5).

$$f = BRx + v \tag{5}$$

In Eq. (5), f represents the observed signal. B refers to a random beta effort matrix. R is a sparse routing matrix. v represents the noise interference of the model. In data collection, there may be insufficient sparsity of the original signal. To obtain suitable sparse bases, the perception data are learned. The optimization process is represented by Eq. (6).

$$\min_{\Psi} \left\| X - \Psi S \right\|_F^2 \; s.t \forall i, \left\| S_i \right\| \leq K \tag{6}$$

In Eq. (6), Ψ refers to the sparse basis in the model that needs to be optimized. K refers to the sparsity of optimizing sparse bases. S refers to a sparse vector matrix. $\left\| \bullet \right\|_F$ refers to the Frobenius norm of a matrix. If only one sparse basis is optimized, the constraint isometry condition needs to be satisfied. If not, multiple measurements need to be taken. Therefore, low coherence is used to compensate for the deficiency. Then, optimization and update are carried out through covariance, represented by Eq. (7).

$$\min_{\Psi} \left\| \left(\Psi^T \Phi^T V \Phi \Psi + A \right) \right\|_F^2 + \eta \left\| \Psi^T \Phi^T V \Phi \Psi - I \right\|_F^2 \tag{7}$$

In Eq. (7), η refers to the equilibrium factor in the model. A refers to the variance matrix of sparse signals. V refers to the inverse matrix of noise variance. $\left\| \Psi^T \Phi^T V \Phi \Psi - I \right\|_F^2$ refers to meeting low coherence. In a successfully deployed WSN, there is a significant amount of redundancy in the perception data collected from sensor nodes due to temporal and spatial correlations. Compression perception is a commonly used data collection technique. Therefore, the study adopts compression perception for data collection. The compression perception performance will improve with the increase of sample size, but the sparsity of sample data is uncertain. Fig. 3 shows a designed data collection framework to continuously update node sparsity.

B. Node Selection Based on Bayesian Compression Perception

The study adopts a method based on Bayesian compression perception adaptive sparsity to collect data. This can reduce the consumption of communication resources and ensure the reliability of data collection. However, this method ignores the performance differences of the nodes themselves. The WSN nodes produced by each company have different characteristics depending on the selected core processor, radio frequency communication chip, and expansion function. The sensor node mainly consists of four parts: sensing unit, processing unit, wireless transceiver unit, and power supply unit [24]. The differences in nodes can lead to a shortened network lifecycle. In this regard, the study adopts a method based on Bayesian compression perception to select nodes. Fig. 4 shows the composition of sensor nodes.

The research on data collection in WSN has been ongoing for many years and has achieved significant research results. However, sensor nodes still face issues such as imbalanced energy consumption, ineffective correlation measurement between nodes, and inability to perform effective spatial clustering. Therefore, when selecting sensor nodes, the study simplifies WSN. Because each sensor node is different, the sink node is affected by noise during the data collection process. Therefore, the model update is represented by Eq. (8).

$$g = \Phi x + a \tag{8}$$

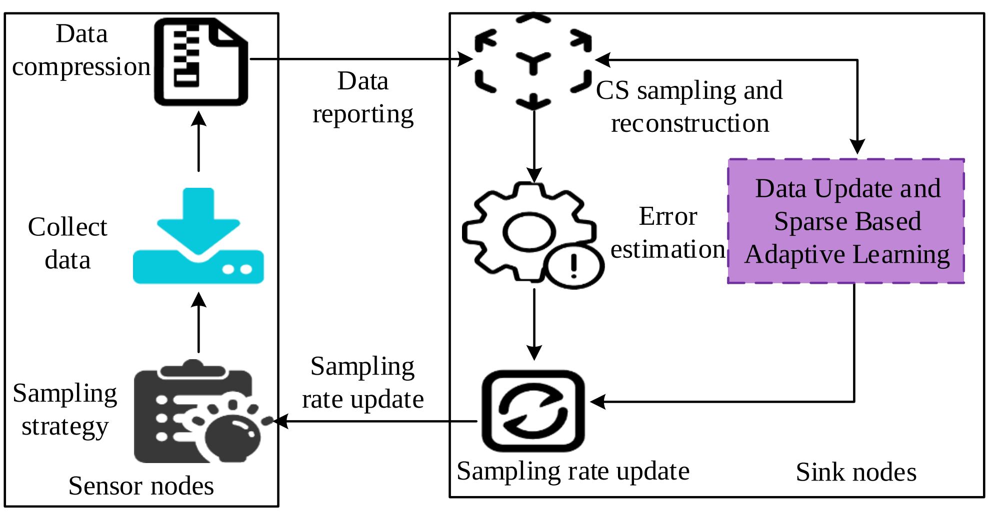


Fig. 3. Data collection framework diagram.

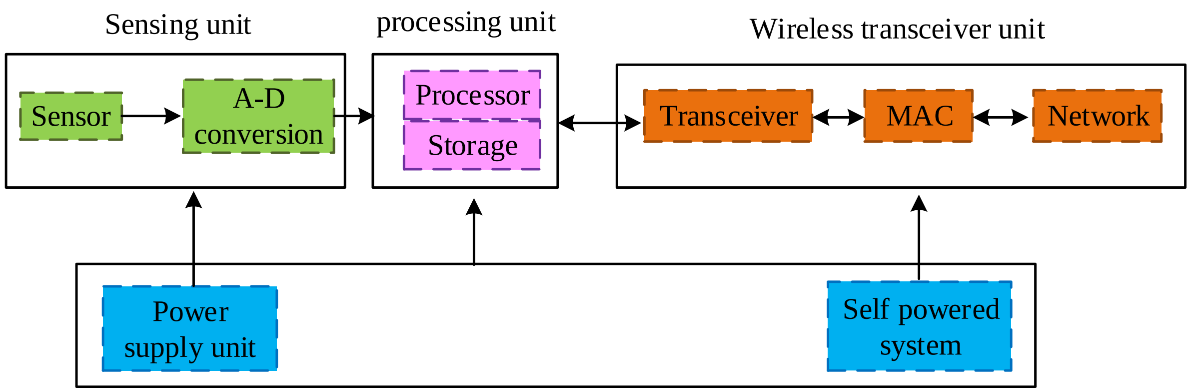


Fig. 4. Composition structure of sensor nodes.

In Eq. (8), a represents the noise vector measured by the model, x represents the data collected by each node, and g represents the measurement data. The model collects raw data without sparsity. Therefore, by utilizing the spatiotemporal correlation of data to transform the original data, sparse data can be obtained. The sparsity of the original data is represented by equation (9).

$$x = \Psi s = \sum_{i=1}^{N} \psi_i \cdot s_i \tag{9}$$

In Eq. (9), s refers to a sparse signal. x refers to raw data. Ψ refers to the sparse basis in the model that needs to be optimized. To better reconstruct the obtained data, assuming that the information of sparse basis and measurement matrix is known, Bayesian algorithm is used to reconstruct the measurement data. The posterior expectation of sparse signal is represented by equation (10).

$$\mu = \sum \Theta^T V^{-1} y \tag{10}$$

In Eq. (10), μ refers to the posterior expectation of the sparse signal. V refers to the inverse matrix of noise variance. Θ refers to the perception matrix. The posterior covariance of sparse signals is represented by Eq. (11).

$$\sum = \left(\Theta^T V^{-1} \Theta + A\right)^{-1} \tag{11}$$

In Eq. (11), $\sum$ refers to the posterior covariance of the sparse signal. A refers to the variance matrix of sparse signals. Θ refers to a perception matrix. In the framework of compression perception, selecting appropriate active sensor nodes is crucial for ensuring reconstruction performance when the data collection quality of each node is different. Bayesian compression perception utilizes prior knowledge to improve reconstruction performance, especially when there are differences in node quality. The node selection steps are as follows: first, all sensor nodes in the wireless sensor network obtain data sources and transmit them to the sink node through node selection. The sampled data is compressed, perceived, and reconstructed at the sink node. Then, the node selection is optimized based on the reconstruction information. Finally, the sink node distributes the optimized node selection information to various sensor nodes in the network through control signals, iteratively until the conditions are met. The architecture diagram for node selection is shown in Fig. 5.

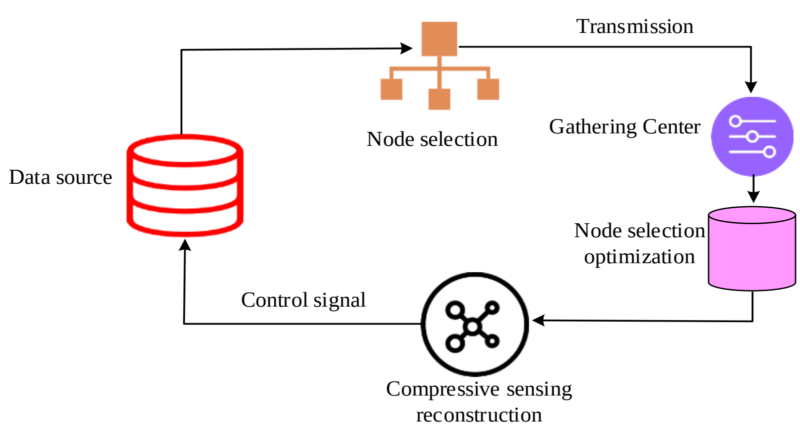


Fig. 5. Node selection framework diagram.

Assuming that the sink node provides complete information for hyperparameters and rewrites the mean and covariance, the observation matrix is an important parameter that affects the reconstruction effect. The reconstruction effect can be controlled by selecting different nodes. Mean square error is a commonly used indicator in machine learning to evaluate the model quality. It reflects the predicted and true values' error. This study uses mean square error to determine the reconstruction error, which is represented by Eq. (12).

$$MSE = \frac{\|\hat{x} - x\|_2^2}{\|x\|_2^2} \tag{12}$$

In Eq. (12), x represents the raw data. $\hat{x}$ represents the reconstructed original signal. The WSN lifecycle generally refers to the energy consumption of a node or a specific area of nodes. In order to extend the lifecycle of WSN, it is necessary to reduce the computational complexity. The average network lifetime is represented by Eq. (13).

$$E(L) = \frac{\xi_0 - E(E_u)}{P_c + \rho E(E_t)} \tag{13}$$

In Eq. (13), P_c refers to constant continuous power consumption. ρ refers to the average rate of reporting data. ξ_0 refers to the starting total energy. $E(E_t)$ refers to the average transmission energy consumption of sensors. $E(E_u)$ refers to the average idle energy during network downtime. The energy limitation of the node itself is represented by Eq. (14).

$$H_{i,i} = \frac{\varepsilon_i}{\vartheta_i} \tag{14}$$

In Eq. (14), ε_i represents the energy stored by a node. ϑ_i represents the energy consumed by a node when transmitting information. The energy consumption of sending data packets from each node to the sink node is represented by Eq. (15).

$$E_{\delta_1}^i = E^e \delta_1 + \bar{E}_{\delta_1} \delta_1 \hat{r}^{\gamma} \tag{15}$$

In Eq. (15), δ_1 refers to the length of the data packet. $\bar{E}_{\delta_1}$ refers to the energy consumption required to achieve the target signal strength. E^e refers to the energy related to the receiving radio wave device. γ refers to the loss in the process of information dissemination. $\hat{r}$ refers to the distance between nodes. To select active nodes, first, sink node sampling is performed. Then, Bayesian compression perception reconstruction is performed. Finally, it is determined whether the energy consumption of the node exceeds the limit. If exceeding this limit, the node selection is redone. If not exceeding this limit, the reconstructed original data are output. Fig. 6 shows the node selection process.

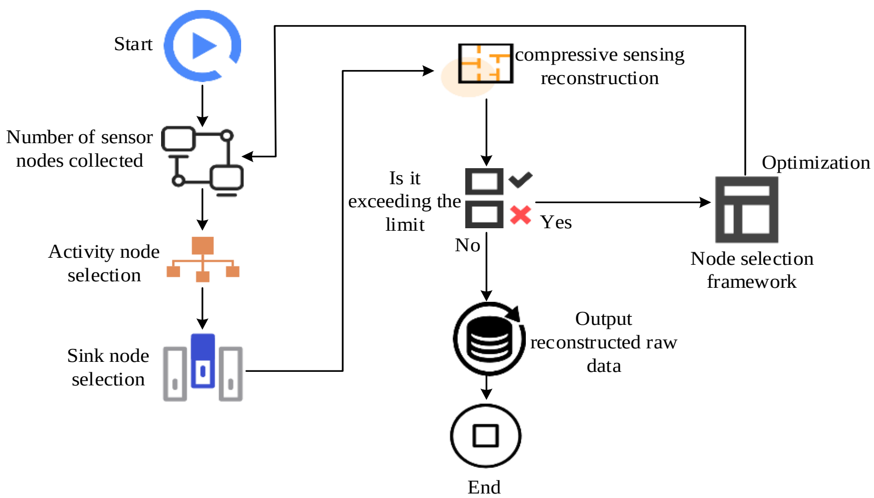


Fig. 6. Node selection flowchart.

IV. Effectiveness Analysis of Data Collection Methods Based on Data Perception and Positioning Technology

A. Experimental Parameter Setting and Efficiency Analysis

To verify the effectiveness of the adaptive sparse strategy and node selection strategy, analysis was conducted on a simulation platform. Set a square network monitoring area with a side length of L=10, and divide the monitoring area into 100 cells on average. Randomly deploy a sensor node in each cell, for a total of 100 sensor nodes. Deploy the sink node at the center position (5, 5) of the monitoring area, with an initial observation frequency of M=50. The length of the data packet is set to 20 bytes, the bandwidth is 2M/S, and the sparsity is 16. The simulation parameters and experimental environment settings are shown in Table I.

In the case of different sparsity, the reconstruction error of the adaptive sparsity strategy used in this study was compared with the reconstruction error of other methods. The balance factor was set to 0.6. Fig. 7 shows the statistical results. The reconstruction error decreased with increasing sparsity. The research method had the lowest reconstruction error. When the sparsity was 25, the reconstruction error curve tends to stabilize, and the minimum reconstruction error was 0.05. The reconstruction error of the overcomplete dictionary design method for sparse representation was slightly greater than that of the method used in the study. When the sparsity was 30, the reconstruction error curve tends to stabilize, and the minimum reconstruction error was 0.06. The wavelet transform method's reconstruction error was the largest. When the sparsity was 30, the reconstruction error curve tended to stabilize, and the minimum reconstruction error was 0.49. The reconstruction error of the discrete cosine transform method was slightly smaller than the wavelet transform method's. When the sparsity was 35, the reconstruction error curve tended to stabilize, and the minimum reconstruction error was 0.42.

TABLE I. SAMPLE PARAMETER SETTINGS

Adaptive sparse strategy		Node selection		Experimental environment	
Parameters	value	Parameters	value	Project	Environment
Number of sensor nodes	100	Sparsity	16	CPU	11th Gen Intel(R) Core(TM) i7-11800H @ 2.30GHz
Sink point location	(5,5)	Sink node sampling number	50	GPU	NVIDIA GeForce RTX 3070 Laptop GPU
Monitoring area	10*10	Sink point location	(5,5)	Memory	16GB
Bandwidth	2M/S	Number of sensor nodes	100	Video Memory	8GB
Packet length	20bytes	/	/	Operating system	Windows 10
Observation frequency	50	/	/	/	/

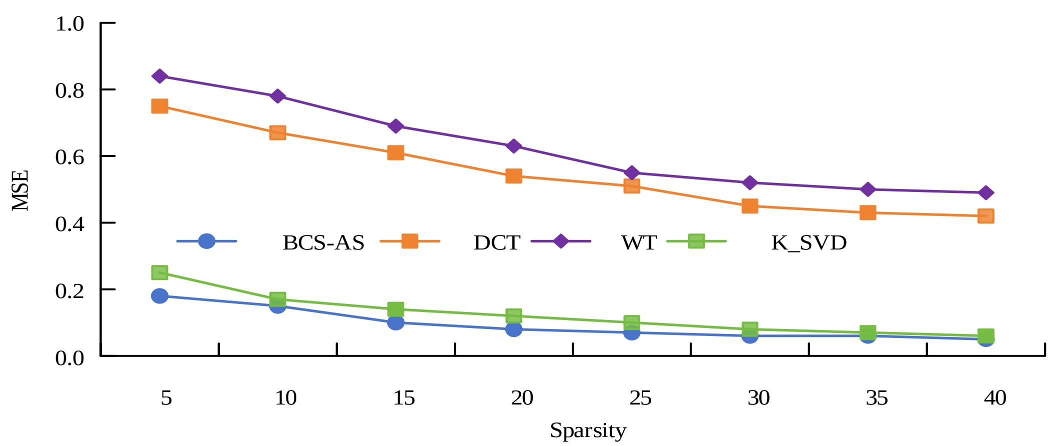
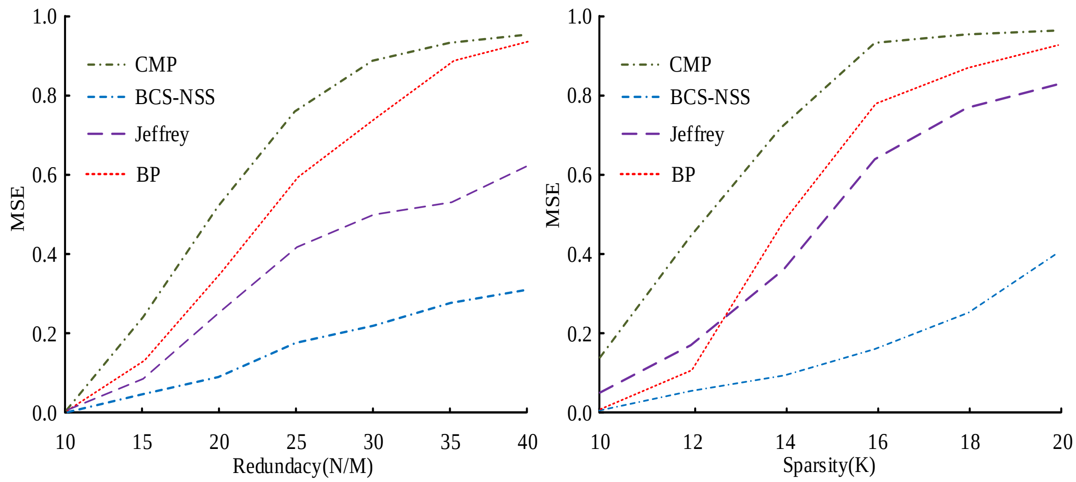


Fig. 7. Reconstruction performance of four sparse bases.

The reconstruction error of node selection based on Bayesian compression perception was compared with the reconstruction error of other methods under different degrees of redundancy and sparsity. Fig. 8 shows the statistical results. In Fig. 8(a), the reconstruction error increased with the increasing redundancy. The research method had the lowest reconstruction error and the maximum reconstruction error was 0.31. The orthogonal matching tracking algorithm had the highest reconstruction error, with a maximum reconstruction error of 0.96. The reconstruction error of the error backpropagation algorithm was slightly lower than the orthogonal matching tracking algorithm's, with a maximum reconstruction error of 0.93. The reconstruction error of Jeffrey's prior algorithm was higher than the research method's, with a maximum reconstruction error of 0.61. In Fig. 8(b), the reconstruction error increased with the increase of sparsity. The research method had the lowest reconstruction error and the maximum reconstruction error is 0.40. The orthogonal matching tracking algorithm had the highest reconstruction error, with a maximum reconstruction error of 0.98. The reconstruction error of the error backpropagation algorithm was slightly lower than that of the orthogonal matching tracking algorithm, with a maximum reconstruction error of 0.96. The Jeffrey's prior algorithm had higher reconstruction error than the research method, with a maximum reconstruction error of 0.82.

B. Quality Analysis of Data Collection Methods based on Data Perception and Positioning Technology

IoT signals are often affected by other noises during transmission, but this research does not consider noise interference in sensor networks. The reconstruction performance of the research adaptive sparse strategy was compared with other methods under different sampling quantities in Fig. 9. When the samples quantity reached 70, the reconstruction performance of different algorithms was basically the same in the two routing scenarios. The reconstruction performance decreased with increasing sampling. The adaptive strategy had the best reconstruction performance. When the sampling quantity was 70, the reconstruction error curve tended to stabilize, and the minimum reconstruction error was 0.04. The reconstruction performance of the discrete cosine transform method was the worst. When the sampling quantity was 60, the reconstruction error curve tended to stabilize, and the minimum reconstruction error was 0.42. The reconstruction performance of the overcomplete dictionary design method for sparse representation was slightly inferior to the research method. When the sampling quantity was 60, the reconstruction error curve tended to stabilize, and the minimum reconstruction error was 0.13.



(a) Reconstruction performance under different degrees of redundancy

(b) Reconstruction performance under different sparsity levels

Fig. 8. Comparison of reconstruction performance of four algorithms.

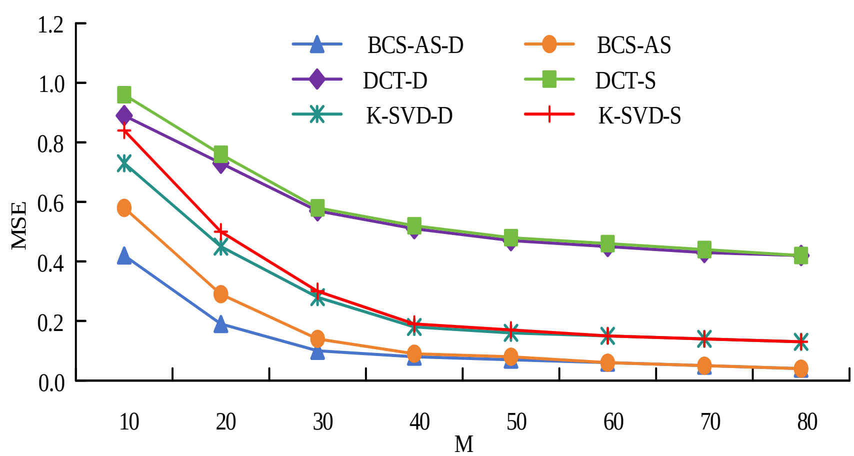
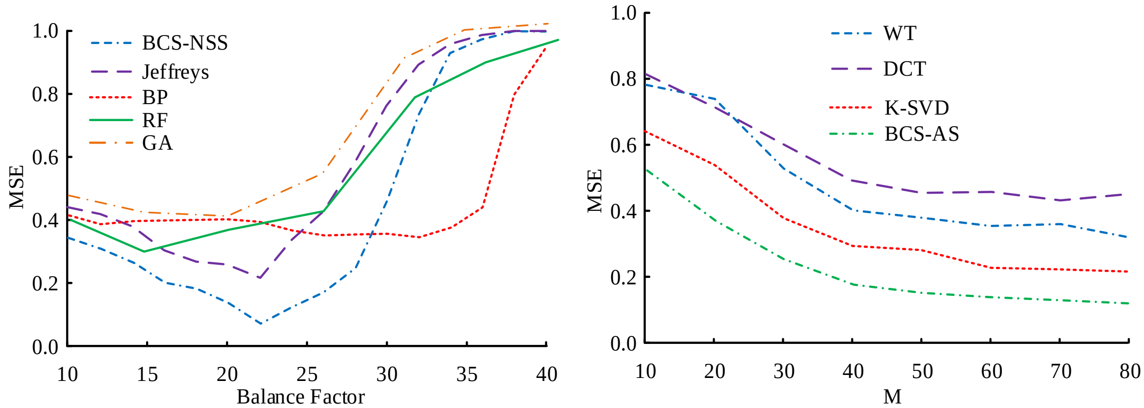


Fig. 9. Reconstruction performance of sparse routing and dense routing.

The reconstruction performance of node selection based on Bayesian compression perception was compared with other methods under different balance factors. In the case of unreliable links, the reconstruction performance of the adaptive sparse strategy was compared with that of other methods in Fig. 10. Observing Fig. 10(a), it can be seen that when the balance factor is small, the reconstruction performance of node selection based on Bayesian compressive sensing is the best, with a minimum reconstruction error of 0.09. When the balance factor is small, the reconstruction performance of genetic methods is the worst. The reason is that for some complex, high-dimensional, and nonlinear problems, genetic algorithms may have difficulty effectively searching for the optimal solution. At this time, the minimum reconstruction error is 0.47. When the balance factor is large, the reconstruction performance stability of the five algorithms is insufficient. In

Fig. 10 (b), the adaptive sparse strategy had better and more stable reconstruction performance, with a reconstruction error of 0.16. The reconstruction performance of the overcomplete dictionary design method for sparse representation was slightly inferior to the research method, with a reconstruction error of 0.22.

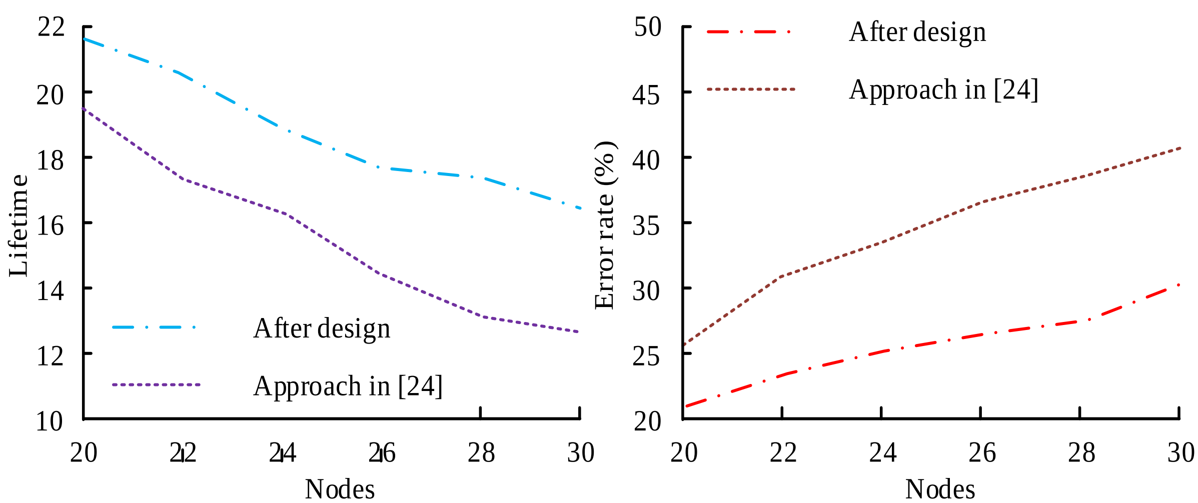
In the case of different numbers of nodes, the study compared the network lifetime and error rate of data collection using methods with the method proposed in study [24]. The statistical results are shown in Fig. 11. Observing Fig. 11(a), it can be seen that the network lifetime of the wireless sensor network data collected by the method used in the study has been extended. When the number of nodes is 28, the network lifetime is 17, and the network lifetime has been extended by 30.8%. Observing Fig. 11 (b), it can be seen that the error rate of wireless sensor network data collection has been reduced in the study. When the number of nodes is 30, the error rate of data collection has been reduced the most, with an error rate of 29%. The error rate of data collection has been reduced by 12%.



(a) Comparison of Reconstruction Performance under Different Balance Factors

(b) Reconstruction performance under unreliable links

Fig. 10. Comparison of reconstruction performance of different methods.



（a）Comparison of network lifespan for data collection using different methods

（b）Comparison of error rates in data collection using different methods

Fig. 11. Comparison of error rates and network lifespan of data collected by different methods.

The adaptive sparse strategy was compared with the discrete cosine transform method in terms of lifespan in sparse and dense routing. The reconstruction error and network lifetime of node selection based on Bayesian compression perception were compared with those of other methods under different numbers of nodes. Table II shows the statistical results. The lifespan of WSN with adaptive sparse strategy was longer than that of discrete cosine transform method. WSN had the longest lifespan under sparse routing. When the nodes were the same, the reconstruction error of node selection based on Bayesian compression perception was minimized. When nodes were 28, the minimum reconstruction error of the strategy used in this study was 0.04. Under the same reconstruction error, the network lifespan of node selection based on Bayesian compression perception was longer. When the reconstruction error was 0.6, the longest network lifetime of the strategy used in this study was 28.

TABLE II. COMPARISON OF RECONSTRUCTION ERROR AND NETWORK LIFESPAN

Same number of nodes			Same reconstruction error		
method	Number of active nodes	Reconstruction error	method	Reconstruction error	Network lifespan
BCS-NSS	22	0.35	BCS-NSS	0.3	25
BCS-DSSR	22	0.55	BCS-DSSR	0.3	6
BCS-MA-DR	22	0.62	BCS-MA-DR	0.3	22
BCS-NSS	24	0.15	BCS-NSS	0.4	26
BCS-DSSR	24	0.46	BCS-DSSR	0.4	20
BCS-MA-DR	24	0.32	BCS-MA-DR	0.4	23
BCS-NSS	26	0.08	BCS-NSS	0.5	27
BCS-DSSR	26	0.24	BCS-DSSR	0.5	22
BCS-MA-DR	26	0.37	BCS-MA-DR	0.5	24
BCS-NSS	28	0.04	BCS-NSS	0.6	28
BCS-DSSR	28	0.36	BCS-DSSR	0.6	26
BCS-MA-DR	28	0.22	BCS-MA-DR	0.6	27

V. DISCUSSIONS

The proposed digital sensing technology has unique advantages in data collection in wireless sensor networks. The study in [15] uses an efficient aggregation method for multi hop clusters based on hybrid compressive sensing to collect data. By combining compressive sensing and routing protocols, although the effect is significant, it requires multiple compression and decompression operations on the signal, thus requiring a large amount of computing resources and time, and the computational cost is high. The method in study [16] achieves compressive matrix decomposition through random singular value decomposition, thereby achieving higher accuracy at lower sampling rates and reducing data transmission costs. However, stochastic singular value decomposition is based on random sampling, and its results are sensitive to the randomness of the initial sampling matrix, which may lead to instability and uncertainty in the results. The method proposed in study [17] can reduce the loss of data samples when transmitted wirelessly, but its applicability is limited. In contrast, when studying the collection of wireless sensor network data based on digital sensing technology, the selection of sensor nodes reduces the amount of data collected and extends the lifespan of the wireless sensor network. In the experiment, as the amount of data continued to increase, the minimum reconstruction error of the research method was 0.04, which was 0.42 lower than DCT and 0.14 lower than SVD. The proposed method achieves significant performance improvement while maintaining low error, making it suitable for large-scale wireless sensor network data collection.

In summary, the research on data collection based on data aware positioning technology reduces the amount of data collected and extends the lifespan of wireless sensor networks by selecting sensor nodes. The research has provided theoretical support and practical guidance for wireless sensor network data collection, further improving the efficiency and accuracy of wireless sensor network data collection. And this method requires high-performance computer resources, so future research directions will focus on how to reduce the demand for computer resources in Bayesian compressive sensing methods.

VI. CONCLUSION

The development of IoT enables wireless communication between objects and automates data exchange. The foundation of IoT is wireless sensor technology. There is noise interference and abnormal or missing data in WSN, which leads to excessive consumption of communication resources and insufficient data reliability. The study proposes a data collection method based on data compression perception positioning technology. Firstly, a Bayesian compression perception-based method is adopted to select nodes to address the shortened network lifecycle caused by differences in node performance. Secondly, an adaptive sparse strategy based on Bayesian compression perception is adopted to collect data and reduce communication resource consumption. When nodes were 28, the minimum reconstruction error based on Bayesian compression perception node selection strategy was 0.04. When the reconstruction error was 0.6, the longest lifespan of the network based on Bayesian compression perception node selection strategy was 28. Under different sampling quantities, the adaptive strategy had the best reconstruction performance with a minimum reconstruction error of 0.04. When the balance factor was small, the reconstruction performance of node selection based on Bayesian compression perception was the best, with a minimum reconstruction error of 0.09. Compared with existing algorithms, the proposed method has been effectively applied in data acquisition, reducing energy consumption of wireless sensor networks, extending their lifecycle, and improving the reconstruction performance of compressed sensing. However, Bayesian compressive sensing requires a prior distribution of known signals and has a high computational complexity, requiring a large amount of computing resources. Subsequent research will adopt other methods to optimize the Bayesian compressive sensing method, in order to avoid problems such as long computation time and large storage space.

ACKNOWLEDGMENT

The research is supported by Scientific Research Project of Liaoning University of International Business and Economics (No.2023XJLXZD01); Scientific research project of Liaoning Provincial Department of Education (No.JYTMS20230406).

REFERENCES

[1] Zhou I, Makhdoom I, Shariati N, Raza M A, Keshavarz R, Lipman J. Internet of things 2.0: Concepts, applications, and future directions. IEEE Access, 2021, 9: 70961-71012.

[2] Ghaffari K, Lagzian M, Kazemi M, Malekzadeh G. A comprehensive framework for Internet of Things development: A grounded theory study of requirements. Journal of Enterprise Information Management, 2020, 33(1): 23-50.

[3] Liu J, Zhao Z, Ji J, Hu, M. Research and application of wireless sensor network technology in power transmission and distribution system. Intelligent and Converged Networks, 2020, 1(2): 199-220.

[4] Luo J, Chen Y, Wu M, Yang Y. A survey of routing protocols for underwater wireless sensor networks. IEEE Communications Surveys & Tutorials, 2021, 23(1): 137-160.

[5] Antony S M, Indu S, Pandey R. An efficient solar energy harvesting system for wireless sensor network nodes. Journal of Information and Optimization Sciences, 2020, 41(1): 39-50.

[6] Zhang J, Zhao C, Gao W. Optimization-inspired compact deep compressive sensing. IEEE Journal of Selected Topics in Signal Processing, 2020, 14(4): 765-774.

[7] Ye G, Liu M, Wu M. Double image encryption algorithm based on compressive sensing and elliptic curve. Alexandria engineering journal, 2022, 61(9): 6785-6795.

[8] Bhuiyan M N, Rahman M M, Billah M M, Saha D. Internet of things (IoT): A review of its enabling technologies in healthcare applications, standards protocols, security, and market opportunities. IEEE Internet of Things Journal, 2021, 8(13): 10474-10498.

[9] Wang Y, Yang J, Liu M, Gui G. LightAMC: Lightweight automatic modulation classification via deep learning and compressive sensing. IEEE Transactions on Vehicular Technology, 2020, 69(3): 3491-3495.

[10] Cheng G, Wang C, Xu C. A novel hyper-chaotic image encryption scheme based on quantum genetic algorithm and compressive sensing. Multimedia Tools and Applications, 2020, 79(39): 29243-29263.

[11] Liang P, Fan J, Shen W, Qin Z, Li G Y. Deep learning and compressive sensing-based CSI feedback in FDD massive MIMO systems. IEEE Transactions on Vehicular Technology, 2020, 69(8): 9217-9222.

[12] Nan S, Feng X, Wu Y, Zhang H. Remote sensing image compression and encryption based on block compressive sensing and 2D-LCCCM. Nonlinear dynamics, 2022, 108(3): 2705-2729.

[13] Xu G, Zhang B, Yu H, Chen J, Xing M, Hong W. Sparse synthetic aperture radar imaging from compressed sensing and machine learning: Theories, applications, and trends. IEEE Geoscience and Remote Sensing Magazine, 2022, 10(4): 32-69.

[14] Chinthamu N, Karukuri M. Data Science and Applications. Journal of Data Science and Intelligent Systems, 2023, 1(1): 83-91.

[15] Aziz A, Osamy W, Khedr A M, El-Sawy A A, Singh K. Grey Wolf based compressive sensing scheme for data gathering in IoT based heterogeneous WSNs. Wireless Networks, 2020, 26(5): 3395-3418.

[16] Sekar K, Devi K S, Srinivasan P. Compressed tensor completion: A robust technique for fast and efficient data reconstruction in wireless sensor networks. IEEE Sensors Journal, 2022, 22(11): 10794-10807.

[17] Mei Y, Gao Z, Wu Y, Chen W, Zhang J, Ng D W K, Di Renzo M. Compressive sensing-based joint activity and data detection for grant-free massive IoT access. IEEE Transactions on Wireless Communications, 2021, 21(3): 1851-1869.

[18] Lin C, Han G, Qi X, Du J, Xu T, Martínez-García M. Energy-optimal data collection for unmanned aerial vehicle-aided industrial wireless sensor network-based agricultural monitoring system: A clustering compressed sampling approach. IEEE Transactions on Industrial Informatics, 2020, 17(6): 4411-4420.

[19] Chang C Y, Chen S Y, Chang I H, Yu G J, Roy D S l. Multirate data collection using mobile sink in wireless sensor networks. IEEE Sensors Journal, 2020, 20(14): 8173-8185.

[20] Lin Z, Chen Y, Liu X, Jiang R, Shen B. A Bayesian compressive sensing-based planar array diagnosis approach from near-field measurements. IEEE Antennas and Wireless Propagation Letters, 2020, 20(2): 249-253.

[21] Jiang X, Li N, Guo Y, Yu D, Yang S. Localization of multiple RF sources based on Bayesian compressive sensing using a limited number of UAVs with airborne RSS sensor. IEEE Sensors Journal, 2020, 21(5): 7067-7079.

[22] Singh J, Kaur R, Singh D. Energy harvesting in wireless sensor networks: A taxonomic survey. International Journal of Energy Research, 2021, 45(1): 118-140.

[23] Wei X, Guo H, Wang X, Wang X, Qiu M. Reliable data collection techniques in underwater wireless sensor networks: A survey. IEEE Communications Surveys & Tutorials, 2021, 24(1): 404-431.

[24] Xue Z. Routing optimization of sensor nodes in the Internet of Things based on genetic algorithm. IEEE sensors journal, 2021, 21(22): 25142-25150.