

# Research on Data Fusion Technology Based on Compressed Sensing

<http://dx.doi.org/10.3991/ijoe.v12i08.5646>

Xianghong Tian

Jinling Institute of Technology, Nanjing, China

**Abstract**—First introduce a traditional compressed sensing data fusion technique is introduced, and with the goal of addressing the shortcoming of the traditional method, a compressed sensing data fusion technology based on temporal-spatial correlation is proposed. Then modeling of the proposed data fusion technique, improves upon the CS reconstruction algorithm and presents the iterative threshold reconstruction algorithm based on temporal-spatial correlation. With the synthetic signals as the research objects, experiments show the effectiveness of the algorithm, further proving that it displays obvious improvement for the proposed data fusion technology to reduce the network load and saves the amount of energy consumption of the network.

**Index Terms**—Compressed sensing, Data fusion, Restriction algorithm, Temporal-spatial correlation.

## I. INTRODUCTION

A wireless sensor network (WSN) is composed of a large number of sensor nodes randomly distributed for gathering the temperature, humidity, seismic waves, sound and image information of the external environment, converging to the sink node in the communication mode of the wireless short distance, with low power consumption, multi-hop [1]. However, sensor nodes have limited power, which is a challenge for the wireless sensor network in collecting and transmitting large amounts of information. At present, many research projects use the correlation of node information to propose a variety of data fusion technologies to eliminate redundant information and reduce the energy consumption of nodes, such as distributed decision data fusion [2], data aggregation based on optimal trees [3], data fusion of mobile agents [4], etc. But these methods still have many challenges for WSN with a complex application environment, limited resources and strong network dynamics.

Hesham et al. [5] propose clustering data fusion technology combined with routing, making the network nodes constitute a minimum spanning tree with the dynamic selection of cluster head nodes. The received packets are fused into a single packet, thus reducing network traffic, reducing energy consumption of the aggregation nodes and prolonging the network lifetime.

Leveraging the spatial-temporal properties in sensory data from real deployments, in-network compression is an essential technique to reduce the amount of data transmission while preserving relatively high reconstruction accuracy in the sink [6,7,8]. The emergence of compressive sensing (CS) theory has opened up a new research avenue for in-network compression. Its basic idea is to process as much information as possible from as few values as possi-

ble, thus easing the pressure on limited hardware resources in large capacity data processing. Wang Xiao et al. proposed the CS thought based on spatial correlation combined with an effective routing strategy according to the network topology [9], which combines transmission and compression, realizing data transmission with low energy consumption.

Starting from the spatial correlation of WSN, Qing et al. [10] developed an optimal decentralized algorithm for sparse signal recovery and demonstrated its application in monitoring localized phenomena using energy-constrained large-scale wireless sensor networks. Capitalizing on the sparse spatial localized phenomena, compressive data collection is enforced by turning off a fraction of sensors using a simple random node sleeping strategy, which conserves sensing energy and prolongs network lifetime.

Combining the GOSSIP algorithm with CS, Bajwa et al. [11] used the high efficiency of the GOSSIP algorithm to obtain global statistics, and at the same time reduced the cost of sampling and transmission through the use of CS. This saves much cost and decreases the complexity for the decentralized large-scale WSN.

Reference [12], through the research of the data collection application of CS in WSNs, aims at minimizing the loss of network energy by the fusion of joint routing and compression. This paper describes the optimal solution to prove the NP-completeness. The proposed mixed integer programming framework and the fusion trees of the optimal and suboptimal solutions of Greedy Heuristics can be obtained. The results confirm the function of Greedy Heuristics and the energy efficiency of the proposed routing and fusion designs.

In practical application, the communication capability of the sensor nodes is limited. Communication between nodes must pass the related routing protocols, and only considering the temporal correlation or spatial correlation of the signal does not make full use of the characteristics of WSN signals to save cost of the network energy. In order to address the shortcomings of the above methods, this paper presents an efficient data fusion technology combining the temporal-spatial correlation of WSN signals and the CS technology and routing. Research on existing CS reconstruction algorithms, and iterative threshold algorithm based on temporal-spatial correlation is proposed to realize signal reconstruction.

## II. ORIGINAL DATA AGGREGATION TECHNIQUE BASED ON COMPRESSED SENSING

The distributed wireless sensor network topology is shown in Figure 1:

- (1) N nodes are distributed in the square observa-



The traditional compressed sensing data fusion technology only uses temporal correlation or spatial correlation of data in the WSN without considering the routing mechanism, which causes the node information to consume a large amount of energy in the transmission process. The compressed sensing data fusion technology combined with routings takes into account the energy consumption of a node in long distance transmission, combining the multi-hop routing mechanism, but it does not consider the inner correlation of signals. For distributed sensor networks, nodes in the time dimension and space dimension have a strong correlation. Therefore, currently the focus of much research has been how to effectively use the redundancy of network data to reduce the amount of data communication and transmission costs, while at the same time ensuring reliable communication.

The compressed sensing data fusion technology based on temporal-spatial correlation makes full use of the temporal-spatial correlation of WSN data, and combines random routing, sampling and compression, and transmission and compression. Signals are reconstructed with the fewest observation values. The coding process is shown in Figure 2.

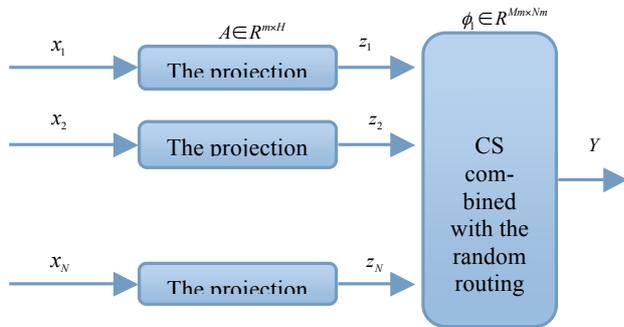


Figure 2. Encoding process of the compressed sensing data fusion technology based on temporal-spatial correlation

It is assumed that there are  $N$  arranged in the observation area regularly, and the network topology is shown in Figure 1. The time signal of a node is  $x_i (x_i \in R^{H \times 1})$ ; namely the sampling signal of node  $i$  in the period of  $T$ .  $x_i$  sparse is in DCT domain, And based on the temporal correlation of signal  $x_i$ , a local random projection is made of  $x_i$ .

$$z_i = Ax_i \quad (4)$$

where,

$A$  is the Gaussian random projection matrix ,

$$A \in R^{m \times H}, \text{ element } a_{ij} : N(0, 1/\sqrt{N});$$

$z_i$  is the projected vector in time dimension of node  $i$ ,  $z_i \in R^{m \times 1}$ ,  $m = H$

For facilitating the calculation, in the algorithm each node adopts the same Gaussian random projection matrix. In the process of encoding, the stronger the temporal correlation of data is, the fewer the number of random projections required by reconstructed signals is; i.e.,  $m$  is smaller, thereby realizing the first compression of network data.

The data in the distributed wireless sensor network has a time correlation. Furthermore, there is strong spatial correlation between nodes due to the different network topologies. For the wireless sensor network with  $N$  nodes, by considering the spatial correlation and random routing, a further projection can be made for the local compressed signal combined with spatial-temporal correlation. The mathematical representation is as follows:

$$y_j = \sum_{i=1}^N a_i z_i \quad (5)$$

where,

$M$  is the path number; namely the number of projection,  $M = N$ ;

$j$  is the path index,  $j = 1, 2, L, M$ ;

$y_j$  is the projection value corresponding to path  $j$ ,  $y_j \in R^{m \times 1}$

In order to explain the encoding process of the compressed sensing data fusion technology based on temporal-spatial correlation, a path in projection is assumed to be  $\{1, 5, 9, 14, \dots\}$ . The projection process of Formula (5) can be expressed as the following matrix form:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix}_{M \times 1} = \begin{bmatrix} a_1 E & 0 & L & a_5 E & 0 & L \\ M & & & & & \\ M & 0 & & 0 & & \\ M & & & & & \end{bmatrix}_{M \times Nm} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix}_{Nm \times 1} \quad (6)$$

where  $E$  is a unit matrix, and  $E \in R^{m \times m}$ .

Nodes in the path transmit information in a multi-hop manner. The amount of data for a single node to transmit is  $m$ . After each jump, the message fuses the observation value  $z_i$  of the current node as input value of the next hop. Data converges to the sink node in the end through  $p$  nodes. Each path only needs  $m$  data sizes to transmit.  $M$  paths correspond to  $M$  projections.

$Y = [y_1^T, y_2^T, L, y_M^T]^T$  converges to the sink node as the final data, and the process achieves the second compression of data.

The encoding process of the compressed sensing data fusion technology based on temporal-spatial correlation is shown in Formula (7).

$$Y = \phi_1 \phi_2 X \quad (7)$$

where,

$X$  is the temporal-spatial signal in the wireless sensor network,  $X = [x_1^T, x_2^T, L, x_N^T] \in R^{NH \times 1}$ ,  $x_i \in R^{H \times 1}$ .

$\phi_1$  is the sparse random projection matrix,  $\phi_1 \in R^{Mm \times Nm}$ ;

$\phi_2$  is the local random projection matrix of node,  $\phi_2 \in R^{Nm \times NH}$ ;  $Y$  is the observation vector,  $Y = [y_1^T, y_2^T, \dots, y_M^T]^T \in R^{Mm \times 1}$ ,  $y_j = R^{m \times 1}$ .

In the practical application of distributed wireless sensor networks,  $x_i$ , the sampling signal of node, is the environmental information, such as temperature, humidity and light intensity, or is the physical parameters, such as current, voltage etc. These signals have strong temporal-spatial correlations. Based on the redundancy of signals, the system takes two projections, and eventually the signal  $Y$  with the total length of  $Z$  is sent to the sink node.

As a sparse random projection matrix based on spatial correlation, each row of  $\phi_1$  contains multi-hop routing information of network data transmission. Its matrix representation is shown in Formula (8).

$$\phi_1 = \begin{bmatrix} a_1 E & 0 & L & a_5 E & 0 & L \\ M & & & & & \\ M & & 0 & & 0 & \\ M & & & & & \end{bmatrix}_{Mm \times Nm} \quad (8)$$

$\phi_2$  is the local random projection matrix based on temporal correlation. In this paper,  $\phi_2$  conducts piecewise projection for WSN signals, and each section of signals corresponds to  $x_i$ , the sampling signal.

$$\phi_2 = \begin{bmatrix} A & 0 & L & L \\ 0 & A & & \\ M & & 0 & \\ M & & & A \end{bmatrix}_{Nm \times NH} \quad (9)$$

When  $\phi = \phi_1 \phi_2$ , then Formula (7) is:

$$Y = \phi X \quad (10)$$

Since  $\phi_1$  and  $\phi_2$  are generated randomly,  $\phi$  satisfies the RIP characteristic, applying to the framework of compressed sensing system theory.

The above analysis shows that the original amount of information transmission is  $N \times H$ . The compressed sensing data fusion technology based on spatial-temporal correlation compresses data twice, so that the total amount of information transmission in the network is reduced to  $M \times m$ , realizing lightweight information processing and transmission. The signal recovery is completed in the sink node, which reduces the computational complexity and transmission energy consumption of the sensor node and solves the problem of limited energy so as to prolong the network life span.

#### IV. EXPERIMENT RESULTS AND ANALYSIS

The research object of this paper is the synthesized signal. Taking mathematical modeling for the compressed sensing data fusion technology based on temporal-spatial correlation, and according to the characteristics of the signals in wireless sensor networks, an iterative threshold algorithm based on temporal-spatial correlation is proposed. MATLAB is utilized as the tool of the simulations. The network topology is shown in Figure 1. Nodes are distributed uniformly in a square region of 16 meters per side. The information of the observation area is compounded by three signals obeying two-dimensional Gauss distribution, whose mathematical expression is shown in Formula (11).

$$x(i, j, t) = 10^3 \sum_{k=1}^3 \left( g_{\mu_k, \Sigma_k}(i, j) e^{-\frac{t}{\delta_k}} \cos(2\pi f_k t) \right) \quad (11)$$

$x(i, j, t)$  denotes the sampling signal of the node with the observation area location  $(i, j)$  at the moment  $t$ . The frequencies corresponding to cosine function respectively are  $f_1=0.1\text{Hz}$ ,  $f_2=0.2\text{Hz}$  and  $f_3=0.051\text{Hz}$ . The sampling frequency is 5Hz. The means of two-dimensional Gauss distribution are respectively  $\mu_1=[3 \ 2]$ ,  $\mu_2=[8 \ 14]$  and  $\mu_3=[13 \ 4]$ .  $\Sigma_1=[10 \ 2; 2 \ 9]$ ,  $\Sigma_2=[8 \ 2; 2 \ 10]$ , and  $\Sigma_3=[9 \ 1; 1 \ 8]$  are the corresponding covariance matrices. The parameter  $\delta_k$  controls the damped speed of the signal along with the time, and  $\delta_1=10$ ,  $\delta_2=15$  and  $\delta_3=8$  are the corresponding values.

In the simulation, 64 nodes in the observation area are arranged using gridding with the size of  $8 \times 8$ . Nodes gather the signal at the speed of 5Hz within 20 seconds. In the reconstruction algorithm the corresponding threshold processing parameters are set at  $\beta_1=2.88$  and  $\beta_2=2.9$ . The maximum iterations *maxiter* is 100.

In order to compare with the traditional compressed sensing data fusion technology which considers only temporal correlation or spatial correlation, use the reconstruction error is used to measure the performance of the proposed algorithm in the experiment. The reconstruction error is defined as follows:

$$\varepsilon = \frac{\|X - \hat{X}\|_2}{\|X\|_2} \quad (12)$$

where,

$X$  is the the original signal;

$\hat{X}$  is the the reconstruction signal.

The energy consumption of the signal per hop in the process of transmission is assumed to be the same. For the proposed compressed sensing reconstruction technique based on temporal-spatial correlation, because each path corresponds to the same number of nodes, the experiment does not consider the energy consumption error of each measurement. The energy consumption of networks is measured by the total data quantity in transmission; namely  $I=m \times M$ .  $m$  denotes the number of observation based on temporal correlation.  $M$  expresses the number of random projections based on spatial correlation. For the compressed sensing data fusion technology based on the temporal correlation, the number of nodes determines the

value of  $M$ . For the compressed sensing data fusion technology based on the spatial correlation, the value of  $m$  is determined by the sampling number during the period of  $T$ .

Figure 3 shows the curve representing the relationship between the reconstruction errors of the different compressed sensing data fusion technologies and the required number of projections. The reconstruction errors of different number of projections is the average value obtained through 100 experiments. The original network traffic is 6,400. It can be seen from the figure, with an increase in the number of projections, the signal reconstruction error decreases; i.e., increasing the number of observations improves the quality of the reconstructed signal. However, by comparing the same reconstruction error, the required projection number of the proposed technology is significantly lower than that of the traditional technology, and the accurate reconstruction signal can be obtained when the sampling rate of the signal is 20%. The results show that, in order to ensure reliable signal transmission, compared with the traditional compressed sensing data fusion technology, the proposed technology can further reduce the load of WSN communications, thus decreasing the energy consumption of nodes. At the same time, the hard threshold processing adopts the wavelet transformation. The decided threshold requires that the corresponding signal dimension is large enough, so as to effectively remove noise, which is especially suitable for large-scale wireless sensor networks.

To more accurately express the effectiveness of the algorithm, Figure 4 shows the reconstruction results of different methods when the sampling rate mM/NH is 20%. The gray graphs represent the spatial signals of the sampling of a node at the first moment. The curve graphs represent the sampling signals of node 1 during T period. Figure 4(d) is the reconstruction results of the proposed algorithm, which clearly shows that the proposed algorithm can reconstruct temporal-spatial signal of WSN under the low sampling rate. Figure 4(b) is the reconstruction result of the algorithm while only considering the spatial correlation. Compared with the result shown in Figure 4(d), its effect of signal recovery quality is poorer than that of the proposed algorithm. Figure 4(c) is the reconstruction results of temporal dimension signal of nodes, whose signal recovery effect of a single node is far less than that of the proposed algorithm. In summary, with the same energy loss the visual effect of reconstruction using the iterative threshold method based on temporal-spatial correlation is obviously better than that the one that only considers temporal correlation or spatial correlation. It further shows that the proposed compressed sensing data fusion technology based on temporal-spatial is effective.

V. CONCLUSIONS

The proposed compressed sensing data fusion technology based on temporal-spatial correlation makes full use of temporal-spatial correlation in wireless sensor networks. It combines routing and transmits information in a multi-hop fashion, and finally reconstructs signal in the sink node. The process of data fusion significantly reduces the sampling cost and computational complexity of nodes in the WSN and reduces the amount of data transmitted over the network, solving the deficiency of limited energy of nodes in wireless sensor networks. In this paper, by

modeling the compressed sensing data fusion technology based on temporal-spatial correlation, the iterative threshold reconstruction algorithm based on temporal-spatial correlation is proposed. The results prove that with the same sampling rate, the proposed algorithm is superior to the traditional compressed sensing data fusion technology.

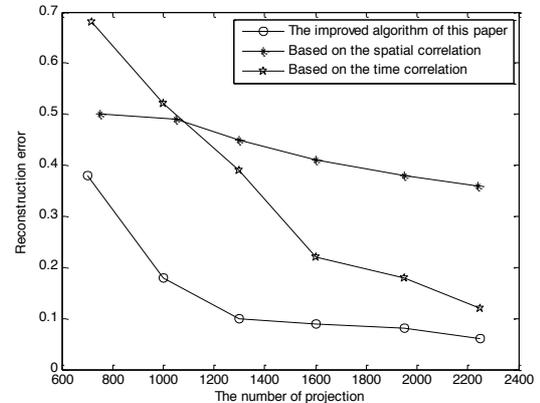


Figure 3. Comparison of the relationship of reconstruction errors of different compressed sensing data fusion technology and the required number of projections

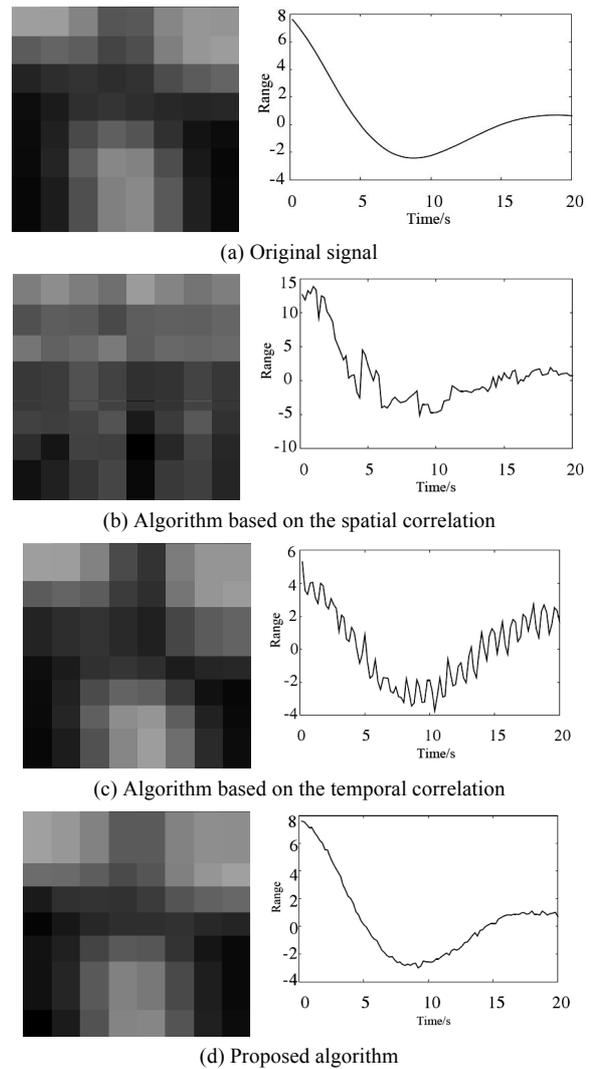


Figure 4. Comparison of reconstruction results of different algorithms at a sampling rate of 20%

## REFERENCES

- [1] J. Yick, B. Mukherjee, and D. Ghosal. Wireless sensor network survey. *Computer networks*, 2008, vol. 52(12), pp. 2292-2330. <http://dx.doi.org/10.1016/j.comnet.2008.04.002>
- [2] Y. Yuan, M. Kam. Distributed decision fusion with a random-access channel for sensor network applications. *Instrumentation and Measurement, IEEE Transactions on*, 2004, vol. 53(4), pp. 1339-1344.
- [3] D. C. Hoang, R. Kumar, and S. K. Panda. Optimal data aggregation tree in wireless sensor networks based on intelligent water drops algorithm. *IET wireless sensor systems*, 2012, vol. 2(3), pp. 282-292.
- [4] Q. Wu, N. S. V. Rao, and J. Barhenl. On computing mobile agent routes for data fusion in distributed sensor networks. *Knowledge and Data Engineering, IEEE Transactions on*, 2004, vol. 16(6), pp. 740-753.
- [5] A. Hesham, S. H. Yang. Dynamic cluster head for lifetime efficiency in WSN. *International Journal of Automation and Computing*, 2009, vol. 6(1), pp. 48-54. <http://dx.doi.org/10.1007/s11633-009-0048-0>
- [6] T. Srisooksai, K. Keamarungsi, and P. Lamsrichan. Practical data compression in wireless sensor networks: A survey. *Journal of Network & Computer Applications*, 2012, vol. 35(1), pp. 37-59. <http://dx.doi.org/10.1016/j.jnca.2011.03.001>
- [7] G. Anastasi, M. Conti, and M. D. Francesco. Energy conservation in wireless sensor networks: A survey. *Ad Hoc Networks*, 2009, vol. 7(3), pp. 537-568. <http://dx.doi.org/10.1016/j.adhoc.2008.06.003>
- [8] Liang, et al. Towards energy saving and load balancing data aggregation for wireless sensor networks. *Information Technology Journal*, 2011, vol. 10(2), pp. 409-415. <http://dx.doi.org/10.3923/ijtj.2011.409.415>
- [9] X. Wang, Z. F. Zhao, and X. Yu. Compressed sensing for efficient random routing in multi-hop wireless sensor networks. *International Journal of Communication Networks and Distributed Systems*, 2011, vol. 7(3-4), pp. 275-292. <http://dx.doi.org/10.1504/IJCNSD.2011.042380>
- [10] L. Qing, T. Zhi. Decentralized sparse signal recovery for compressive sleeping wireless sensor networks. *IEEE Transactions on Signal Processing*, 2010, vol. 58(7), pp. 3816-3827. <http://dx.doi.org/10.1109/TSP.2010.2047721>
- [11] J. Haupt, W. U. Bajwa, and M. Rabbat. Compressed sensing for networked data. *IEEE Signal Processing Magazine*, 2008, vol. 25(2), pp. 92 - 101. <http://dx.doi.org/10.1109/MSP.2007.914732>
- [12] L. Xiang, J. Luo, and A. Vasilakos. Compressed data aggregation for energy efficient wireless sensor networks//Sensor, mesh and ad hoc communications and networks (SECON), 2011 8<sup>th</sup> annual IEEE communications society conference on. *IEEE*, 2011: 46-54.
- [13] M. Balouchestani, K. Raahemifar, and S. Krishnan. Compressed sensing in wireless sensor networks: survey. *Canadian Journal on Multimedia and Wireless Networks*, 2011, vol. 2(1), pp. 1-4.
- [14] G. Quer, D. Zordan, and R. Masiero. WSN-control: Signal reconstruction through compressive sensing in wireless sensor networks. In *IEEE 35<sup>th</sup> Conference of Local Computer Networks(LCN)*, 2010, pp. 921-928.
- [15] D. L. Donoho. De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 1995, vol. 41(3), pp. 613-627. <http://dx.doi.org/10.1109/18.382009>

## AUTHOR

**Xianghong Tian** is a member of the Jinling Institute of Technology, Nanjing, Jiangsu, 211169, China. (e-mail: txh@jit.edu.cn).

This work was supported in part by the Natural Science Foundation of China (No. 61375121); Natural Science Foundation of Jiangsu Province (No. BK20130096). Submitted 14 March 2016. Published as resubmitted by the authors 16 June 2016.