Precision agriculture compressed sensing and data fusion algorithm for wireless sensor networks

Liu Hong-Xia [1]*, Zhang Feng [1, 2]

[1] *School of Information Engineering, Yulin University, 719000, Yulin, China*

[2] *School of Software, Xi'an University, Xi'an 710065, China,*

Received 1 January 2014, www.tsi.lv

Abstract

In order to improve the energy efficiency of WSN nodes and prolong the life of the network, reduce data redundancy. In this paper, proposed the spatial correlation node data compression and fusion algorithm based on the theory of compressed sensing. Firstly, make signal node random projection based on time correlation, then, for random routing instability and network transmission of data fusion technology reconstruction energy effects, proposed the energy consumption of compressed sensing and clustering data fusion technology. The experiment showed that after data fusion, not only effectively removes the redundant information of neighbouring nodes, and the reconstruction error is small, and can accurately realize data decompression, thereby reducing the node communication in wireless sensor network capacity, reduce power consumption of node, provides important support for the field environment in large scale wireless sensor network deployment.

Keywords: compressed sensing, intelligent wireless sensor, agriculture IoT, automatic control, data fusion.

1 Introduction

In recent years, the Internet of Things (IoT) as an international research hotspot, have obtained broad attention [1]. It's represents the future trend of development of the network, and requires sharing interoperability and information, so as to realize human society, the information space, the physical world ternary comprehensive connectivity and integration as the goal. Therefore, the Internet of things is regarded as the third technological revolution in information field [2].

In WSN, due to the large number of sensors, individual sensor nodes with limited resources and processing capacity and have highly redundant data, transmit the data to sink node separate method is not reasonable in the collection of information in the process of using each node [3]. To consider the accurate degree of data, network performance, resource consumption and reliability constraints, solve the associated data between nodes, sharing and integration. The fusion performance in wireless sensor networks, the algorithm should be simple, easy to implement, and reduce resource and energy costs.

An important feature of WSN network is data-centric, data processing is a key problem in WSN, largely determine the performance of the network. WSN nodes are generally battery-powered node energy is limited. How to save node energy and prolong the life of Yes WSN research is a key issue. In this paper, a large-scale and high density WSN as research background, for its data collection, transmission, storage process, the problems, the introduction of compressed sensing theory, the focus of research networks in the process of collecting large amounts of data present in large amount of data problems. In precision agriculture WSN, the general area of relatively is large, relatively large amount of information acquisition.

2 Compressed sensing based spatial correlation of distributed

2.1 COMPRESSED SENSING THEORY

Compressed sensing or compressed sensing (Compressive Sensing, CS) is a theory developed in recent years, is a new data acquisition theory. Innovation is the core of the theory: for compressible data signals can be much lower than the Nyquist sampling via a standard way, is still able to accurately recover the original signal. This theory is largely broadened the scope of the compressing signal, in the theoretical framework for data integration CS, since the data amount is increased, the amount of data transmission can be effectively reduced, thereby reducing the communication energy [4].

Compressed sensing theory mainly includes three aspects: sparse signal representation, measurement selection matrix remodelling achieve optimal algorithm. The theoretical framework can be represented in Figure 1.

* Corresponding author. Tel: +86-189-9220-3580;
E-mail: 75419305@qq.com

First, if the signal $x \in U^N$ is compressible in an orthogonal basis ψ, get $\Theta = \Psi^T X$ is determined Ψ transform coefficients or the equivalent approximation sparse representation;

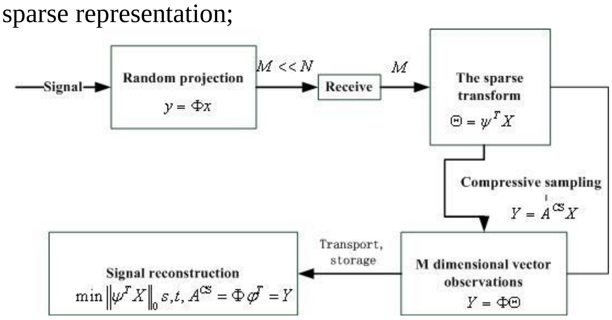


FIGURE 1 Theory framework of CS.

the second step, a smooth design, and not related to Ψ transform base $M \times N$ dimensional observation matrix ϕ of observations Θ obtained observations $Y = \Phi\Theta = \Phi\varphi^T$ collection, the process may also be expressed as a non-adaptive observation $Y = A^{CS} X$ signal X (where $A^{CS} = \Phi\Psi^T$) @ through the matrix A^{CS}, the information is called CS operator; Finally, the 0 - norm optimization problem solving exact value of X or myopic approach $\hat{X}$:

$$\min\|\psi^T X\|_0 s,t, A^{CS} = \Phi\varphi^T = Y . \tag{1}$$

At present the theory of compressed sensing research is still in the initial stages, but already in some areas the signal and image processing, bio-sensing, analogy information conversion showed great potential. Currently compressed sensing theory research focuses on theoretical and mathematical reasoning to improve the level of the two most be studied. In a study of wireless sensor networks, compressed sensing theory has been the concern of more and more scholars, and related researchers. GHua and U.Chen take advantage of a new source coding techniques to improve the effectiveness of the relevant data. J.Haupt and his colleagues on the use of compressed sensing data acquisition and compression carried out relevant research [5]. Distributed compressed sensing theory concepts through joint sparse data on the correlation between the internal and the data are explored.

2.2 DISTRIBUTED COMPRESSED SENSING

The theory of CS standard is mainly for the design of the internal data structure of the single node, namely the correlation data aware of individual nodes to reduce the node data dimension. In wireless sensor networks, a large number of sensor nodes. For the node space intensive distribution, there is existence of high correlation between adjacent nodes of spatial data. Using this characteristic, natural to think of each node collaboration form measurements are combined, to reduce the overall dimension of data measured by correlation between node data compression effect, will get more. Based on this, the theory of distributed compressed sensing (DCS) emerges as the times require. D. Baron et al. Based on the theory of CS further extends the application of CS, put forward a distributed compressed sensing (DCS, Distributed Compressed Sensing) concept gas to compress a single signal sampling was extended to compress the signal group sampling [7]. The establishment of DCS theory is based on a "joint sparse signals known as group model (JSM, Jointly Sparse Models)" concept [8]. DCS theory pointed out that, if a plurality of signal are8 able to express the sparse in the same medium, and these signals is also has the correlation, then every signal in coding can use another not associated with the sparse matrix based (such as a random matrix) for sensing data observation and coding, which is far less than the original length of the encoded signal.

2.3 SPATIAL CORRELATION IN DISTRIBUTED COMPRESSED SENSING MODEL BASED ON WSN

In the WSN to collect data through various Sink node sensing sensor nodes, and monitoring the event source S to estimate the area, making the distortion estimation results to meet the requirements of WSN applications. Figure 2 shows the compressed sensing based spatial correlation of distributed.

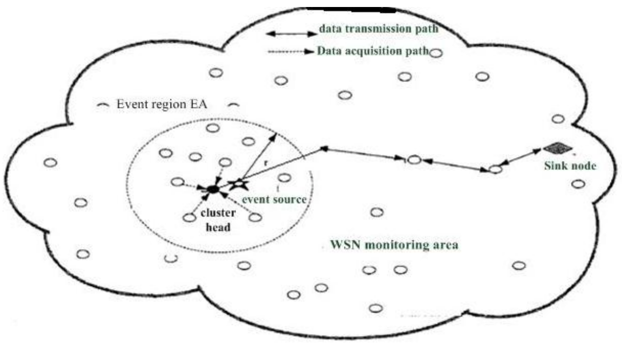


FIGURE 2 Schematic perceptions based on spatial correlation of distributed compression.

As can be seen from Figure 2, because the WSN nodes are densely distributed in S_j and between S_j and S in the presence of varying degrees of spatial correlation. Therefore, to meet the requirements of distortion, by determining the scope of the EA event area, while taking advantage of the perception of spatial data between WSN node correlation data compression and reconstruction, has a very important significance in the WSN.

Assuming the event source S WSN location (0,0), the distribution area in the event of a node $EAn(x,y)$, $n(x_r, y_r)$ coordinates are $n(x,y)$, $n(x_r, y_r)$, its perception data is $S(x,y)$, $S(x_r, y_r)$ variorums is defined as:

$$\gamma(x)=\frac{1}{2}\left[\left(S(x,y)-S(x_r,y_r)\right)^2\right], \quad (2)$$

where $(x-x_r)^2+(y-y_r)^2=r_2$, if the value of the variorums $\gamma(x)$ is smaller, then the perceived correlation between the data will be stronger. In polar coordinates can be defined within a node $n(r,\theta)$ perception data to $S(r,\theta)$, EA within the scope of the event area in the monitored area WSN triggered the event source S, between the perception of sensory data node $n(0,0)$ data and information to other nodes around the following correlation:

$$\begin{aligned}&S(0,0)=I_{(U=T)}+I_{(U=H)}\int_\theta\int_\gamma\left(S_{(r,\theta)}+Z\right)\delta(R=r)\\&\delta(\Theta=\theta|R=r)\end{aligned}, \quad (3)$$

wherein $U=T$ represents sensing data $S(0,0)$ is obtained by a random variable γ, the probability is β; $U=T$ data through the sensing means adjacent to the $S(0,0)$ sensing data of the node $n(r,\theta)$ is worth to $S(r,\theta)$, the probability is $1-\beta$. Random variable Y and Z are the probability density function $f_Y(y)$, $f_Z(z)$ and a random variable and random variables Y and Z are each independently sensing data $S(r,\theta)$.

3 Spatially distributed compressed sensing information acquisition method

3.1 GREEDY ALGORITHM

The basic steps of the greedy algorithm are:

1. Given initial estimates of $\theta^0=0$;
2. At each iteration, according to $A(\theta-\theta^0)=A\theta-A\theta^0$ determine the estimated value Δ of $\theta-\theta^0$;
3. Leaving only the larger Δ value items, other items will be set to zero. Update $\theta^0=\theta^0+\Delta$.

There are many ways about greedy algorithm, mainly has the orthogonal greedy algorithm (Orthogonal Matching Pursuit, OMP), regularized orthogonal greedy algorithm (Regularized Orthogonal Matching Pursuit, ROMP), piecewise orthogonal greedy algorithm (Stage wise Orthogonal Matching pursuit STOMP) and the gradient of the greedy algorithm (Gradient pursuit method)

3.2 ITERATIVE THRESHOLD ALGORITHM

Iterative threshold algorithm is a recursive method to determine a solution of $y=A\theta$, specifically, is defined by the following equation sequence θ^n to close to the optimal solution $y=A\theta$:

$$\theta^{(n+1)}=S\tau\left[\theta^{(n)}+A^*A\theta^{(n)}\right], \quad (4)$$

where $S\tau(x)$ is: $S\tau(x)=\begin{cases}x-\tau,\ x>\tau\\ 0.x\le\tau\\ x+\tau,\ x<\tau\end{cases}$.

Gets the iterative sequence can be obtained by two methods: the operator splitting and two function approximation method. Firstly, by introducing a regularization factor K that set (P1) problem is transformed into unconstrained problems as follows:

$$\min L(\theta,\lambda)=\|\theta\|_1+\lambda\|y-A\theta\|_2^2. \quad (5)$$

Because $\|\theta\|_1$ is non smooth convex separable function, $\|y-A\theta\|_2^2$ is a smooth non separable convex function, so if set $T_1=\partial\|\theta\|_1$, $T_2=\partial\|x-D\theta\|_2^2$, so we can write

$$\begin{aligned}&\hat\theta=\arg\min L(\theta,\lambda)\\&\Leftrightarrow\hat\theta\in(T_1+T_2)(\theta^*)\\&\Leftrightarrow\hat\theta\in(I+\tau T_1-I+\tau T_2)(\theta^*)\\&\Leftrightarrow(I-\tau T_2)(\theta^*)\in(I+\tau T_1)(\theta^*)\\&\theta^*\in(I-\tau T_2)^{-1}(I+\tau T_1)(\theta^*)\end{aligned}. \quad (6)$$

The $(I-\tau T_2)$ is a step τ of the gradient operator, and $(I-\tau T_2)^{-1}$ is $S\tau(\bullet)$.

4 Spatial correlation of distributed compressed sensing algorithm

First of all, need to get the distribution range of event region EA. At the position of the hypothetical event sources have a virtual node is $n(0,0)$, the information data of the node $S(0,0)$, $S(r,\theta)$ data boundary nodes in $n(r,\theta)$ aware S the events triggered by the event source region of the $|S(r,\theta)-S(0,0)|\le\mu$, μ is in line with the conditions, the error threshold, it is used to indicate the difference in perception data information and the event source at node in different position of the distribution of radius r represents the event region EA. According to equation (2) can be obtained:

$$r(r)=\frac{1}{2}E\left[\left(S(r,\theta)-S(0,0)\right)^2\right]=\frac{1}{2}E\left[Z^2\right]$$

$$=\frac{1}{2}\int_{-\varpi}^{\mu} z^2\frac{1}{\sqrt{2\pi\sigma_z}}e^{\frac{z^2}{2\sigma_z}}dz \quad , \qquad (6)$$

$$\frac{1}{2}\sigma_z^2 erf\left(\frac{\mu}{\sqrt{2\pi\sigma_z}}\right)-\frac{1}{\sqrt{2\pi\sigma_z}}e-\frac{z^2}{2\sigma_z^2}=\psi(\sigma_z,\mu)$$

where $\psi(\sigma_z,\mu)$ is a function of σ_z and μ , $erf(x)=\frac{2}{\sqrt{\pi}}e^{-t^2}dt$.

In summary, the distributed spatial correlation of compressive sensing encoding and decoding algorithm based on WSN can be expressed as follows:

Step 1: according to the formula of $T^J(\lambda_j)=\{\lambda_0,\gamma_0,\gamma_1,\ldots,\gamma_{J-1}\}$, the distribution range of the sink node computing event region EA, forming a cluster by multicast routing way node $n_i(i=1,2,\ldots,N)$ activation events in area EA.

Step 2: cluster head node n_h to generate the observation matrix $\Phi=R(S_M,\tau_N)$, where $R(\bullet)$ is a pseudo random number generating function $\tau_N:\{\tau_1,\tau_2,\ldots,\tau_N\}$, τ_i is the number of n_i.

Step 3: the cluster head node of the coded data Y_M transmission by wireless multi hop to the sink node.

Step 4: Sink node and cluster head node generates the same observation matrix $\Phi=R(S_M,\tau_N)$, then get a small wavelet transform matrix ψ according to the network topology, run the distributed compressed sensing decoding algorithm.

5 Simulation and performance analysis

5.1 EXAMPLES OF DATA FUSION

Introduced in front of cluster Limited ad hoc networking mode, it is in accordance with the region of the cluster, the cluster head using ad hoc network between the nodes in the cluster structure, with star, tree and chain structure.

Assuming a field of soil moisture content within the region more than 45% of the area number for inquiries first sends a query request. Suppose there are four field using tree routing, the sensor position and the communication path shown in Figure 3.

In Figure 3, each sensor node has prepared a data, and (Filed, Soil moisture) is expressed in the form of. Need to do the following:

1) Check each sensor node is in accordance with the upload request, to decide whether to participate in the transmission;

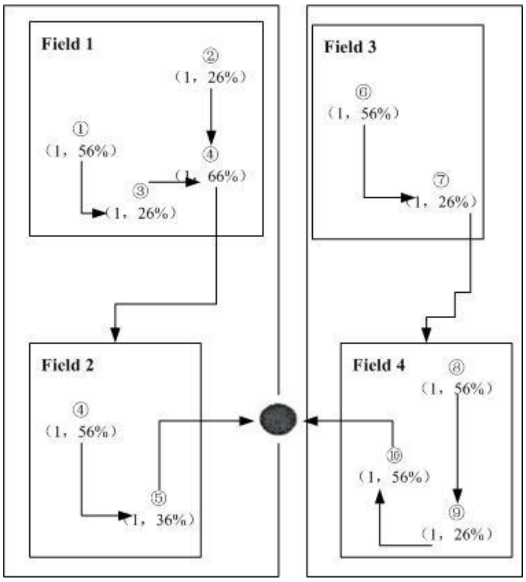


FIGURE 3 Instance of data fusion.

2) Each node receives the data sent by other nodes, for local operations, results of operations continue to submit to the upstream node;

3) intermediate nodes if in a certain period of time has not received the neighbour node to send data, then the default from neighbour nodes below, did not submit the data nodes request.

According to the above operation, the appeal processes a total of 5 copies in the network transmission of data. As shown in Figure 3, the Node10 sent 3 copies of the data, Node 8 to send 3 copies of data, Node 3 to send 1 copy of data. If you do not have any data fusion method, each individual node send data to the sink node, and the sink node concentration calculation results, the network data transmission quantity is 21, far greater than 6, the corresponding node energy consumption will significantly increase.

5.2 EXPERIMENTAL AND SIMULATION RESULT

Simulation research and focus on the relationship between the number of observation data and reconstruction error of DCS encoding and decoding algorithm of the time correlation analysis based on WSN. The simulation uses Matlab as tools, the spatial correlation of data source in WSN by using two-dimensional Gauss distribution to model, the node in WSN random uniform distribution of events in the 60mx60rn area. In the system, all the nodes of the sensor that produced in WSN sensing data at the same time and send it to the cluster head node encoding, the cluster head nodes using DCS coding algorithm on the perception of the received data compression, and the observation data generated by multi hop wireless transmission to the decoder, accurate reconstruction of the sink node using the DCS decoding algorithm on each node sensing data in the area of the WSN event.

As shown in Figure 4, the relationship between the number of simulation graph reconstruction error and observation data, assuming that perceived value number is 3000. Can be seen from the chart, the observed values of M increases, the reconstruction error is more and more and more small, suitable for multiple observations data acquisition system.

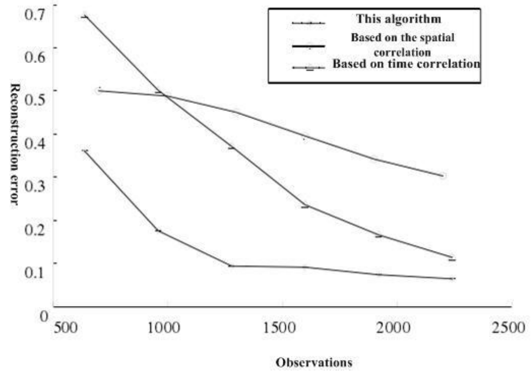


FIGURE 4 the relationship between the observed number and reconstruction error.

6 Conclusions

Compressed sensing is a theory developed in recent years, firstly, the theory of compressed sensing do a simple introduction, focusing on the spatial correlation in WSN based on distributed compressed sensing theory. WSN in precision agriculture, the amount of information needed to capture the scene is relatively large, and therefore the number of sensor nodes required range. Is also a great amount of data transmission, data compression is an effective way to reduce the amount of data transmission, according to the practical application of this article, the choice of the spatial correlation based on compressed sensing theory of distributed network layer data compression.

Acknowledgments

This work is partially supported by the Funding Project for Department of Education of Shaanxi Province in 2012 #12JK0828, 2013 #2013JK1151 and 2013JK1167. Research and Cooperation Project for Department of Yulin city of (**#Sf13-08，#Sf13-23, #Ny12-19, #Ny13-10,#2012cxy3-26**). Thanks for the help.

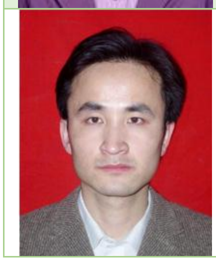
References

[1] Pierce F J, Elliott T V 2008 *Computers and Electronics in Agriculture* **61**(1) 32-43
[2] Wetter M 2011 *Journal of Building Performance Simulation* **4**(3) 185-203
[3] Rodriguez R I, Jia Y 2011 *International Journal on Smart Sensing and Intelligent Systems* **4**(2) 325-37
[4] Haifeng D, Xiaolong Z 2012 *International Journal on Smart Sensing and Intelligent Systems* **5**(4) 824-42
[5] Geziei S 2008 A survey on wireless position estimation, *Wireless personal Communications*, **44**(3) 263-80
[6] Zhibin S, Chao T, Jun C, et al 2010 *IEEE Transaction on Communications* **58**(2) 511-20
[7] Candès E 2006 Compressive sampling. *Proceedings of the International Congress of Mathematicians, Madrid, Spain,* **3** 1433-52
[8] Candès E 2008 The restricted isometry property and its implications for compressed sensing *C.R.Math.Acad.Sci.* Paris **346**(9-10) 589-92.

Authors

Liu Hong-Xia, born 24.10.1981, in Shannxi Yulin

Current position, grades: lecturer in Yulin University.
University studies: MS degree in specialty of educational technology from Shaanxi Normal University in 2007.
Scientific interest: Information technology education, application of Internet of things.



Zhang Feng, born 26.06.1980, in Shannxi Yulin

Current position, grades: associate professor in Yulin University.
University studies: MS degree in Computer science from Xidian University in 2009.
Scientific interest: Cloud integrated manufacturing technology, the modeling of complex systems, the Internet of Things applications.