Acta Technica *62* (2017), No. 4A, 43–54

Wireless sensor network data compression method based on genetic compressed sensing algorithm

Zhang Na[2]

Abstract. In view of wireless sensor network energy, communication bandwidth, computing power and cost limited, not suitable for large-scale data transmission, while the presence of data redundancy, the need for data compression, put forward a new method based on genetic algorithms for compressed sensing reconstruction, applied to wireless sensor network data compression. This paper describes distributed WSNs data compression features, the basic theory of compressed sensing to reconstruct a new method based on genetic algorithm and its application in WSNs CS data compression. Through simulation experiments prove that the compression ratio from four aspects, the average energy consumption of nodes, network lifetime and network delay, etc., compared with DCCM algorithm and CCS algorithm of WSNs data compression algorithm, the proposed algorithm had a higher compression ratio improve the collection of data reconstruction accuracy, reduced data redundancy and network traffic, improved network efficiency.

Key words. Wireless sensor networks, compressive sensing, genetic algorithm, compression ratio, life cycle.

1. Aims and Background

With the development of WSNs unceasingly, the bandwidth of the signal collection needed and the amount of data has been increased as a way for the geometric growth, So new sampling mechanism must be established in order to reduce cost, traffic, delay, energy consumption and information data volume [1]. The energy consumption of WSNs often directly determines the stability and reliability of the whole system. It is very significant for data compression to improve the communication throughput of WSNs [2]. Energy consumption of sensor network is mainly concentrated on three parts, respectively as the signal sampling, signal processing and signal transmission, in which most of the energy consumption is in signal transmission link

[1]Acknowledgement - This work was supported by Foundation of Henan Education Commitee (No.17A520024) and Henan province science and technology hall(No.172102210105).

[2]Workshop 1 - School of Computer and Data Science ,Henan University and Urban Construction , Pingdingshan, Henan, 467036, China; e-mail: znwang1010@126.com

[3]. Compression transmission signals can effectively reduce energy consumption, Two points need to considered in data compression of WSNs design: one is the node processing power is low, the second is limited memory resources. So the compression algorithm designed energy consumption must be as little as possible [4]. Data compression algorithm is the key to compress the energy consumption should be less than the data transfer of energy.

Although genetic compression perception (Compressive Sensing,CS) is field which emerged in recent years[5-6], But it has shown strong vitality and development potential. Many scholars in this field has been to try and hope to contribute and breakthrough. Compressed sensing combine sparse signal compression and sampling of two steps, In order to satisfy a feature: Energy consumption of sensor nodes in the area of the monitoring and calculation ability is very limited, and distal gathering terminals have sustained energy supply and the powerful computation ability. Ying-jie wang[7]put forward wireless sensor network in distributed data compression method , and applied to the wireless sensor network (WSN) data compression algorithm to obtain good results; Xue-jun ren[8] proposed a hybrid coding data compression algorithm for wireless sensor network (WSN); Wang ling[9] proposed wireless sensor network data compression and optimization algorithms based on the time correlation ; Zhang [10] proposed segmentation approximation error is bounded in sensor network data compression algorithm, and obtained good effect. Traditional data compression method which has encoding complex but simple decoding, This means that the use of the limited resources of the sensor nodes in complex coding algorithm, but in the relatively rich resources of cluster head nodes for simple decoding algorithm. Based on the above analysis, this paper proposes a new compression perception reconstruction method based on genetic algorithm, the method is applied to data compression in wireless sensor network (WSN), The experiment result indicates the method can not only make compression part more simple, less transmission of data, but also can make relatively simple decoding algorithm, improve the compression ratio and the collection of data reconstruction accuracy, reduce the data redundancy and network traffic, improve the efficiency of network.

2. Experimental part

(1)Distributed WSNs data compression

In monitoring area of hundreds of thousands of tiny sensor nodes of the wireless sensor network (WSN), which can be able to collaborate to network coverage of objects in the area of information perception, acquisition and processing, finally sent to the base station terminal through processing the information. Typically, WSNs is fully distributed network, And in terms of each sensor node, on the one hand, they have independent perception, processing and communication ability to the information, But on the other hand their communications range, energy, even computing power is limited, So we can not solve the problem of the complexity of the larger. To processing power of sensor node in full use, must be mutual cooperation between nodes, breakthrough the limitation of a single node, common to complete the task, to solve practical problems, thus improve the network performance.

(2)Compression perception (CS) theory

Traditional data sampling method which is sampling, compression, transmission, decompression is likely to fail. While compression perception sampling with simple sampling and the characteristics of complex decoding, thus can effectively solve the problem. Compression perception is different from the traditional method which first signal sampling and then eliminate redundant part, it will directly compressed signal sampling, and compressing and sampling at the same time, it saves a lot of useless signal sampling part[11]. It will directly compressed signal sampling, and compressing and sampling at the same time, it saves a lot of useless signal sampling part [11].

Compressed sensing includes three aspects: signal sparse representation, coding measurement and reconstruction method. Signals can be compressed perception is a precondition to sparse representation or compressible, namely in a transformation matrix, the signal can be used to a more concise way to express, its number of nonzero coefficients must be far less than the number of natural zero signal central Africa [12].

For a limited long one-dimensional signal$x \in R^n (n \in N)$,Suppose it is based on an orthogonal basis is sparse of $\psi = \{\psi_i\}(1 \leq P < n,\ and\ P \in N)$,the sparseness is P,so x is expressed as follows

$$x = \sum_{i=0}^{N-1} \theta_i \psi_i = \sum_{p=0}^{P-1} \theta_{i_p} \psi_{i_p} \tag{1}$$

sparse matrix projection coefficient isθ_i ,(1) can be abbreviated to

$$x = \psi\theta \tag{2}$$

θ is N $*$ 1dimensional sparse matrix columnn, ψ is a sparse transformation matrix.

The compression perception theory shows that when the signal x sparse or after sparse transform ψ sparse, you can use an unrelated $m \times n$ observation matrix $\psi(m\}\ nm \in N)\ (m\ \}\ n, m\ \in\ n)$ to linear transformation of x, get the observed value y, namely

$$y^{m\times 1} = \phi^{m\times n} x^{n\times 1} = \phi^{m\times n} \psi^{n\times n} \theta^{n\times 1} \tag{3}$$

It can be seen that the number of elements in the observed value y is less than the number of elements in the x, thus realize the sampling of sensing signal compression. Moreover we can obtain x through reconstituting the observated objects to y by solving optimization problems. The result is as follows

$$\theta = \arg\min \|\theta\|_1, \quad s.t.\ \ y = \phi\psi\theta \tag{4}$$

(3) new reconstruction method about CS based on genetic algorithm

The compression perception theory shows that new reconstruction method about CS based on genetic algorithm is different from the traditional reconstruction algorithm based on matching, It is mainly obtained from the optimization of objective function, the sparse reconstruction results equal to chromosomes in biological. Sup-

pose in a certain size of population, after copy, selection, crossover swaps and mutation genetic process, after the iteration the infinite optimal reconstruction result of the closest node sampling value perception, from sparse reconstruction results after the location information of each non-zero elements. After that the location information of each nonzero elements is obtained from the sparse reconstruction result. The range of the non-zero elements information is obtained by least square method, finally complete reconstruction and perform sensing data compression information transmission process.

1) Non-zero elements location information in the sparse reconstruction results

1.the scheme of population and coding

According to the compression theory of perception, sparse signal $\widehat{\theta}$equivalent into chromosomes (that is, the binary string) to set initial population coding.

2.copy

In the process of parent chromosomes produce offspring, it is required for a given genetic generation gap GGAP specific values, GGAP $\in$ (0, 1), Assuming each generation multiplied by the number of individuals in a population B 1 - GGAP optimal individuals are copied directly to the next generation. And other individuals, produced by the selection operation.

3.selection

Select the target function value is defined as

$$f = \left\| y - \phi\psi^H \widehat{\theta} \right\|_2 \tag{5}$$

The final iteration optimization goal is to make the objective function to obtain the minimum, min$\{f\}$is obtained.

This article uses the linear estimate fitness J_i, first of all, descending order of the objective function value. Place the minimum individual fitness of chromosome in order list of the objective function value of the first position, the most to adapt to the individual placed in position on B_i, B_i is the number of population individuals. Each individual chromosomes according to its position in the sorted populations $W_{0,}$ $J_i$can be launched by calculation fitness values, namely

$$J_i = \min_a\{f_a\} + \frac{W_0 - 1}{B_i - 1} \cdot (\max_a\{f_a\} - \min_a\{f_a\}) \ , \quad a = 1, 2, \cdots A \tag{6}$$

4.crossover

The cross of genetic algorithm choose a single point of intersection, in the individual coding, using crossover probability randomly selected from a series of new intersection, the two matching individual partial chromosomes swap.

5.variation

Mutation of genetic algorithm using the basic variation, the individual coding string according to the process of genetic mutation probability randomly selected a loci on numerical basic mutation operation, execution process is as follows:

1. According to the mutation probability on each individual assigned a loci for mutation point;

2. For every dye the gene value point mutation of invert operation, so as to get

a new generation of genetic individual dyeing.

After a series of operation process, and then through MAXGEN iteration convergence (MAXGEN biggest genetic algebra) can get the optimal chromosome, is the optimal solution. In general, the optimal chromosome is mainly composed of "0" or "1" of coding, including "1" for the sparse results of non-zero elements, and sparse results the location of the zero element information and its corresponding position information.

(2)In the sparse results non-zero elements determine the amplitude information

After knowing location information in each non-zero elements in sparse results , and then in each position using the amplitude is determined by the least square method to do the projection information. Suppose sparse results have a non-zero elements in the first j position, then the amplitude of the nonzero elements.

$$mag = \frac{\langle T_j, y \rangle}{\langle T_j, T_j \rangle} \tag{7}$$

Where $T_j$is expressed the *j column in T, while*

$$T = \phi\psi \tag{8}$$

T is the matrixes of information recovery. By calculating, the range of the non-zero elements in the sparse results are available on the non-zero elements amplitude information.

(4) CS new reconstruction method application in WSNs data compression

Genetic compression perception algorithm combined with a concrete application of WSNs data transmission, able to transmit the data is compressed, the specific process is as follows: First of all, through acquisition of data transmission to the cluster head nodes, to compress the data storage and processing of cluster nodes, as shown in Fig. 3. Cluster nodes and then analyze the data obtained by each sensor node, then sends the compressed data results to gather terminal, Eventually the signal of each node in terminal joint recovery reconstruction. Based on the genetic algorithm for CS steps are as follows:

The remote terminal command. Released by gathering the terminal monitor command transmitted by multiple hops routing to the cluster head nodes, will be in the form of radio of cluster nodes forward command to each sensor node in the cluster.

3. Results and Discussion

This paper uses the Matlab simulation platform for the proposed genetic compressed sensing data compression algorithm for validation, including computer configuration for 2.0 GHz Pentium processor, memory of 2 gb PC ring 4 run under. In wireless sensor network data compression algorithm, use more simple differential coding compression algorithm (differential code compression will method, DCCM) and collaborative compression perception (Cooperative Compressed Sensing, CCS) algorithm[11], etc.,This article will compare with DCCM algorithm and

CCS algorithm[12]for comparative analysis. Genetic algorithm parameters Settings: population size for genetic algorithm (ga), 40 largest MAXGEN = 300 genetic algebra, genetic generation gap GGAP = 0.95, crossover probability p = 0.7, the mutation probability PM = 0.03. Experiment parameters used in the wireless sensor network (WSN) as shown in table 1.

Table 1. Simulation parameters

parameters	values
network size	100m×100m
node number	100
primary energy	1 J
communication range	25 m
node sends power	0.0825W
wireless receive power	0.0891W

Wireless sensor network data compression algorithm performance evaluation index in this paper, from the simple compression effectiveness and sensor data sequence with the other algorithms of the compression ratio, average energy consumption of nodes, survival time and the network time delay, etc.

(1)The effectiveness of the sensor data sequence compression perception

Effectiveness to be improved compression perception algorithm proposed in this paper have a relatively intuitive understanding the sensor, a group of experiment data were collected the data length N = 100, the original data, compression perception to refactor the data and algorithm proposed in this paper after the data after contrast diagram as shown in Fig. 3. As can be seen from the Fig. 3, CCS algorithm reconstruction rate at 83.7%, reconstruction algorithm in this paper rate is 94.8%, root mean square error is small. The proposed algorithm is accurate reconstruction, and reconstruct the signal quality is stable, ensure high precision reducing.

(2)compression ratio

compression ratio compression calculate formula for

$CR=S_{CP}/S_{OR}$

(9)

Where SOR is for initial data, the SCP is compression after the amount of data. Another CR value is smaller, so the amount of data after compression which accounts for the smaller amount of raw data, compression performance is optimal. Three kinds of algorithm of compression ratio such as shown in Fig. 4.

As can be seen from the Fig. 4, with the increment of node data acquisition, compression ratio of the three algorithms are gradually decline, the size is not very big, In comparison, the CCS compression ratio of compression ratio below DCCM algorithm, has a better compression performance,And GA - CS compression algorithm proposed in this paper the compression ratio than CCS algorithm compression

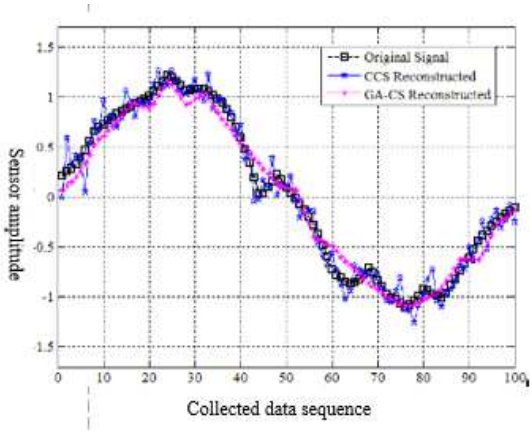

Fig. 1. Sensor data refactoring rendering sequence compression

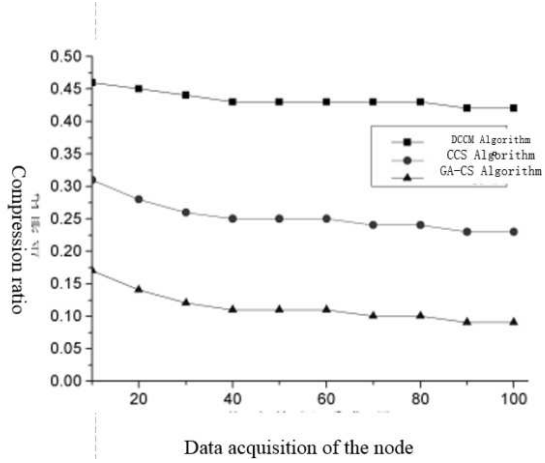

Fig. 2. Compression ratio compared three kinds of algorithm

ratio is low, the compression performance is the best, Because the algorithm is very good mining data-centric wireless sensor networks, data, the characteristics of the correlation between the coding process in addition to the maximum level of redundant information between nodes, obtained the good compression effect, and with the increase of the amount of data collected node, its compression ratio is more and more small, gradually become stable. In addition the convergence and convergence of genetic algorithm is given time as shown in Fig. 5. Genetic iteration algorithm in 180 generation data flat, achieve convergence, at the same time, the simulation time consumption 4.83s.

(3)Average Energy Consume

Node energy consumption is an important parameter of wireless sensor network performance parameters, also suitable for data compression algorithm performance evaluation, data transmission will increase in the process of data compression in the

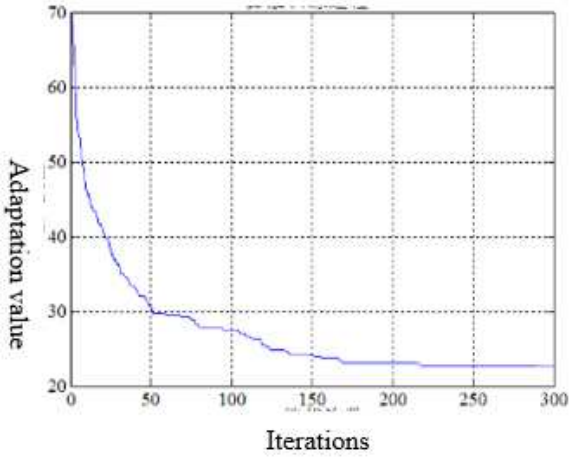

Fig. 3. Genetic algorithm iteration convergence curve

calculation of energy consumption, but in the perception of node energy consumption of the whole process of communication, data compression to calculate the energy consumption can be neglected. Only consider perception nodes in the simulation of energy consumption of wireless communication, set the initial energy, sending and receiving power of simulation nodes, packet sending rate, at the same time, according to the data packet length calculation node residual energy after each sending data. When calculating the network to send and receive the remaining energy difference then know that the network energy consumption. The average energy consumption of three kinds of algorithm is shown in Fig. 6.

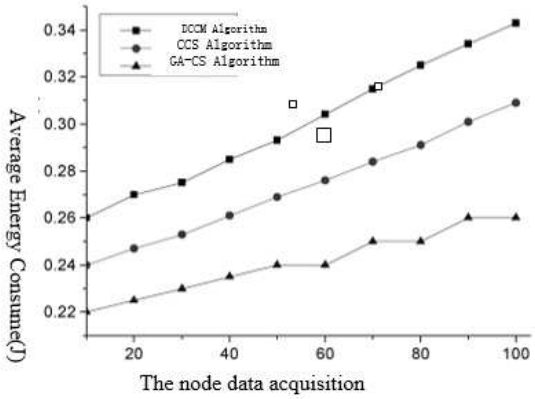

Fig. 4. The average energy consumption contrast three kinds of algorithms

As can be seen from the Fig. 6, node data acquisition phase at the same time, the amount of CCS algorithm the average energy consumption is lower than DCCM

algorithm, the algorithm proposed in this paper the average energy consumption is lower than DCCM algorithm. With the increase of perceptual node sends a packet number, three algorithms of the average energy consumption is increased, but the energy consumption of the proposed algorithm increase the narrowest margin. This is mainly because the proposed genetic compressed sensing data compression algorithm can better dig sensing data correlation between nodes, maximize reduce the redundant nodes information, improve the compression precision of the algorithm.

(4) network lifetime

Similarly network survival time is an important performance index, wireless sensor network (WSN) the length of the reaction network life directly. This article respectively to the three network survival time of the algorithm are simulated, the simulation results are shown in table 2.

Table 2. Mode death time round number

Algorithm	DCCM	CCS	GA-CS
10% exhaust	2067 r	2367 r	2681 r
25% exhaust	2619 r	2846 r	3162 r
50% exhaust	6135 r	6504 r	6943 r
75% exhaust	8534 r	9249 r	10957 r
100% exhaust	9831 r	11357 r	12835 r

On the whole, the proposed GA - CS algorithm node death time round number is longer than DCCM algorithm and CCS, when network awareness node 10% depleted, the GA - CS algorithm nodes death rate of DCCM algorithm is 77% and 88% of CCS algorithm; When network awareness node 50% depleted, the GA - node death rate of CS algorithm DCCM algorithm is 88.3% and 93.7% of CCS algorithm; When network awareness node 75% depleted, the GA - node death rate of CS algorithm DCCM algorithm is 77.9% and 84.4% of CCS algorithm; In 100% of the node runs out of energy consumption, GA - CS algorithm of node network polling time respectively 3004 and 1478. As you can see, in this paper, based on the GA - CS algorithm for wireless sensor network in the network life-time can life longer than the other two algorithms. This is mainly because after data compression, reduces the data redundancy within the network, reducing the energy consumption in the process of data transmission, prolong the network lifetime. Therefore, this article puts forward the algorithm to calculate the cost of limited and is suitable for energy co., LTD., co., LTD. Of the wireless sensor network (WSN).

(5)network delay

Network delay T_Y includes two parts of time, from the perception of nodes to send data to the receiving node receives the packet transmission delay T_S and the node

receives the packet data compression algorithm of calculation time T_C, calculation formula is:

$$T_Y = T_S + T_C \tag{9}$$

Three kinds of algorithm of network time delay is shown in Fig. 7

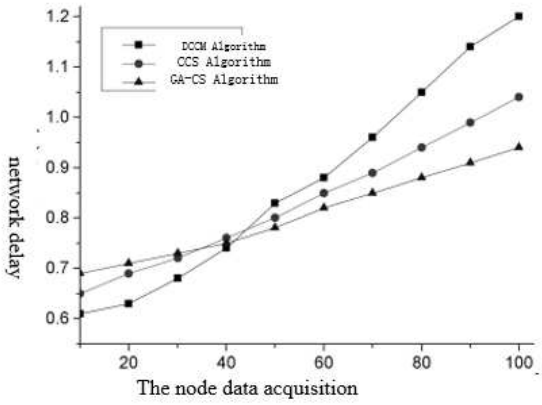

Fig. 5. Network delay compared three kinds of algorithm

As can be seen from the Fig. 7, three algorithms of network time delay is increased with the increase of the node data acquisition, when a node collection capacity is small, DCCM algorithm of network time delay shorter than CCS algorithm, short 0.05 s, and than GA - short 0.1 s of CS algorithm of network time delay, this is mainly because the data is compressed and the consumption of computation delay, but with the increased amount of node data acquisition, the packets sent large DCCM algorithm of network delay is longer than CCS algorithm obviously, network delay is longer than GA - CS algorithm, here basically is not data compression of time consumption calculation time delay, mainly is the data acquisition to increase after transmission time delay time, it can be seen that the proposed data compression algorithm based on GA - CS algorithm, effectively reduces the need to preach in the network

4. Conclusions

Wireless sensor network (WSN) due to energy, communications, computing power and bandwidth costs is limited, not suitable for large-scale data transmission, at the same time data redundancy, the need for data compression processing, this paper proposes a new compression perception reconstruction method based on genetic algorithm, is applied to the wireless sensor network in the process of data compression. For GA - CS algorithm performance analysis, from the compression ratio, average energy consumption of nodes, network survival time and network time delay and so on four aspects, and DCCM algorithm and CCS comparison of WSNs data com-

pression algorithm, experimental verification show that the proposed algorithm has higher compression ratio, and improve the collection of data reconstruction accuracy and reduces the data redundancy and network traffic, improve the efficiency of network.

References

[1] N. Erratt, Y. Liang: *Compressed data-stream protocol: an energy-efficient compressed data-stream protocol for wireless sensor networks.* Communications, IET *5* (2011), No. 18, 2673–2683.

[2] X. Liu, L. Jun, C. Rosenberg: *Compressed Data Aggregation: Energy-Efficient and High-Fidelity Data Collection.* Networking, IEEE/ACM Transactions on *21* (2013), No. 6, 1722–1735.

[3] S. C. Li, L. D. Xu, X H. Wang: *Compressed Sensing Signal and Data Acquisition in Wireless Sensor Networks and Internet of Things.* Industrial Informatics, IEEE Transactions on *9* (2013), No. 4, 2177–2186.

[4] F. W. Ceng: *Adjustable sensing data compression algorithm related to time of the reconstruction error* . Journal of computer applications and software *28* (2011), No. 11, 279–282.

[5] M. Y. Zhang: *Snake model based on improved genetic algorithm in fingerprint image segentation.* Int. J. BIOAUTOMATION *20* (2016), No. 4, 431–440.

[6] W. Xi: *The application of genetic algorithms in the biological medical diagnostic research.* Int. J. BIOAUTOMATION *20* (2016), No. 4, 493–504.

[7] X. J. Ren, Y. Ding: *A hybrid coding data compression algorithm for wireless sensor networks* . Small microcomputer system *32* (2011), No. 6, 1055–1058.

[8] W. Ling: *Wireless sensor network based on time correlation data compression and optimization algorithm.* Journal of computer applications *12* (2013), 3453–3456.

[9] J. M. Zhang, Y. P. Ling, M. Fu: *Sensor network segmentation approximation error bounded data compression algorithm.* Journal of software *12* (2010), 2149–2165.

[10] Y. G. Yin, G. G. Ding: *Data compression techniques in wireless sensor network research.* Journal of computer applications and software *27*, (2010), No. 7, 1–4.

[11] X. H. Fan, S. N. Li, P. L. Du: *An adaptive lossless data compression in WSN system.* Computer measurement and control *19* (2010), No. 2, 463–465.

[12] P. Jiang, F. W. Jian: *The optimal zero suppression based on adaptive wireless sensor network data compression algorithm research* . Journal of communications *2* (2013), 1–7.

Received November 16, 2016