Sensors & Transducers, Vol. 23, Special Issue, July 2013, pp. 157-161

www.sensorsportal.com

Sensors & Transducers

http://www.sensorsportal.com

Distributed Source Coding Algorithm Suitable for Clustering Wireless Sensor Network

[1] Geng Zhang, [2] Hao Xu, [3] Hao Deng

[1] Chongqing Technology and Business Institute, No. 3-5, Cuiyuan Community, No. 1, Hualong Avenue, Jiulong Science and Technology Park, Jiulongpo District, Chongqing City, 400052, China

[2] Chongqing Electric Power College, No. 3-5, Cuiyuan Community, No. 1, Hualong Avenue, Jiulong Science and Technology Park, Jiulongpo District, Chongqing City, 400052, China

[3] Chongqing Industry Polytechnic College, No. 3-5, Cuiyuan Community, No. 1, Hualong Avenue, Jiulong Science and Technology Park, Jiulongpo District, Chongqing City, 400052, China

Tel.: 86+18523073848, fax: 86+18523073848

[1] E-mail: qydbzzw@sina.com

Received: 15 April 2013 /Accepted: 20 July 2013 /Published: 30 July 2013

Abstract: Since intensive wireless sensor networks have a lot of redundant information deficiencies, this paper applies clustering wireless sensor networks, distributed source coding algorithm discussed analysis. This algorithm initial reference source as side information source through a correlation between all sources to determine the coding sequence for the reference signal source, and, after the source through the respective reference signal for its corresponding source code, based on the reference signal source and the associated coding sequence of the receiving end decoded. Meanwhile, according to the encoding method modulus value, come up with a less complex decoding algorithm. Simulation and theoretical analysis results confirmed that the algorithm for clustering routing protocol applications, can greatly reduce the number of bits sent by the nodes, so as to realize the reduction of energy consumption, and extend the network lifetime. *Copyright © 2013 IFSA.*

Keywords: Clustering wireless sensor networks, Distributed, Source coding, Algorithm.

1. Introduction

Wireless sensor network (WSN) refers to the large number of low-cost micro sensor nodes are engineering a self-organizing jump more wireless network, the technology in intelligent instrumentation, environmental monitoring, agriculture and health and other fields are widely used in [1]. Normally, the battery is the main source of power of sensor node, and the change cannot be implemented after the deployment completes, therefore, how to ensure that a limited number of micro sensor nodes of wireless sensor network for effective use over a long period of time, and gradually developed into a large scale wireless sensor network, gradually become the focus of the researchers in related field. Through spatial correlation of different data reduces the amount of data sent to help network operation efficiency. In order to improve the accuracy of a single node of the robustness and the demand of perception, the operation of the WSN usually has more precise deployment, so as to ensure the spatial correlation of data between the more abundant [2].

Combined with hierarchical routing and distributed source coding, the acquired data can

implement more efficient compression. The researchers believe that, source code can be used in distributed in chain-like clusters in the scheme, the implementation of source code related to the transmission path through signal. And the researchers think, can be applied in multi-hop clusters distributed source coding scheme, and puts forward method of backward and forward along the signal path fusion two data fusion [3]. However, these two methods are only between two adjacent nodes of the use of the information, but not to maximize compression efficiency. Although it is proposed on this basis to achieve the theoretical maximum Slpeian-Wolf coding algorithm, each node and its neighboring nodes based on the information between the encoding. This theory has neither been put into practice nor explains how to use the information between adjacent nodes [4]. For the above theoretical deficiencies manifested, this paper proposed clustering wireless sensor network design a new coding method, its main features include: First, the targeted application of the modulus value coding method presents a relatively simple decoding algorithm, enabling the decoding of the gradual simplification; Second, to design a new multi-source coding algorithm (MSCC); Third, in the wireless sensor network clustering algorithm using MSCC, thus greatly extending the network use of time [5].

Combining hierarchical routing and distributed source coding, the data obtained can be more effectively compressed. There is a kind of solution mentioned in the eighth reference that applying distributed source coding into the chain cluster which codes along the signal path. In the ninth reference, the writer has put forward a kind of distributed source coding is applied to the multiple hops cluster solution, and gave two data fusion methods namely along the signal path forward integration and backward integration. But these two solutions can only use the two adjacent nodes asked relevant information, compression efficiency maximization. In the tenth reference, the author put forward a kind of can achieve Slepian Wolf theory limit coding algorithm, each node according to oneself and all relevant information between neighbor nodes to encode, stays at a theoretical level, but how to use the related information between the neighbor node has not been mentioned.

2. The Theoretical Basis

2.1. Slepian-Wolf Coding Theory

According to Slepian-Wolf coding theory, for the two *'X'* and *'Y'* interrelated implementation code, if guide information associated with each other between *'X'* and *'Y'*, take the *'p(x, y)'* joint probability distribution as an example, the *'X'* can not understand the basis of *'Y'* basic situation, and understand the code the efficiency of *'Y'* basic same, different coding scheme is expressed in Fig. 1 to get the same results [6]. However, independent coding scheme without mutual communication between the realization with *'Y'*, so as to minimize the communication cost, improve its flexibility, this feature for the limited energy of wireless sensor networks more attractive and high [7].

According to Slepian-Wolf coding theory, edge information is *'Y'*. We code '*X'* and '*Y'* independently, at the same time, it must be in accordance with Fig. 1 and relationships as in formula (1), (2), (3) the said relationship to implement the joint decoding with very little distortion:

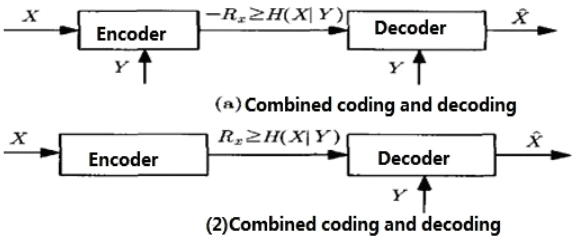

Fig. 1. The two kinds of coding schemes.

$$R_x \geq H(X|Y), \tag{1}$$

$$R_y \geq H(Y|X), \tag{2}$$

$$R_x + R_y \geq H(X,Y), \tag{3}$$

where 'R_y' for *'Y'* coding rate, 'R_x' for *'X'* coding rate, and *'H (X|Y)'* and *'H (Y|X)'* respectively the corresponding conditional entropy, '*H (X, Y)'* represents conditional entropy between '*X'* and '*Y'*. Slepian-Wolf for two related source coding theory, multiple source associated expanded gradually to after [8].

So far, researchers have put forward many practical DSC method, including: the Turbo code method, the modulus value method, method of LDPC codes, and along with the type of distributed source coding method and so on. For application scenarios, including temperature control, chiefly because of its relevance in, in a specific range is continuous and can be used to detect the difference, therefore by means of modulus value can show greater value [9].

2.2. Modulus Value Encoding

In view of the scene temperature control measures, the scope of data set to [*mins, maxs*], the data bits is set to 'n', set the monitoring accuracy as *'△'*, but the specific values visible on the actual situation, its potential monitoring sample collection is as in formula (4) [10]. If source difference between *'X'* and *'Y'* edge information is less than 'M', whose value is as in formula (5), *'X'* can through Kbit

coding for data compression, modulus value encoding of practical application are as in formula (6):

$$\Omega = \{\mathrm{Si/Si} = min\mathrm{s} + i \times \Delta i \in (0,1,...2^n - 1)\}, \quad (4)$$

$$M = 2^{k-1}\Delta(0 \le k \le n), \quad (5)$$

$$f(X) = index(X) \bmod 2^k, \quad (6)$$

where, *index(X)* for '*Ω*' sequence in '*X*', '*f (X)*' for encoding after *'X'* bits of information, its computation is as in formula:

$$index(X) = (X - \min s)/\Delta \quad (7)$$

Divide *'X'* jointly encoded with the same data set to the same, it is called the coset, and *'f (X)'* is an index to the coset of *'X'*, each with concentration has 2^{n-k} elements, coset quantity is 2^k [11].

According to the received by 'f(X)' and the edge information of *'Y'*, known for decoding the *'X'* accurately recover the actual signal and the choice of decoding as in formula (8):

$$\hat{\mathrm{X}} = \min_{\mathrm{ri} \in \mathrm{S}} \left\| \mathrm{Y} - \mathrm{ri} \right\|, \quad (8)$$

where *'ri'* is the elements in the concentration, *'S'* is the coset all given by *'f(X)'*, for the results obtained by decoding.

3. Design of Algorithms

3.1. Second Source Code

By modulus value coding method to encode processing, formula (9). However, for the specific coding process, it should be executed according to formula (10), (11) and (12):

$$\hat{\mathrm{X}} = \mathrm{Y} + \varepsilon\Delta \quad (9)$$

$$\varepsilon = \begin{cases} \mathrm{X'\text{-}Y'}, & \left|X'-Y'\right| < 2^{k-1} \\ (\left|X'-Y'\right| - 2^k) \times sign(X'-Y'), & \left|X'-Y'\right| \ge 2^{k-1} \end{cases}, \quad (10)$$

$$X' = index(X) \bmod 2^k, \quad (11)$$

$$Y' = index(Y) \bmod 2^k, \quad (12)$$

where *'sign'* is the symbol function.

3.2. Multiple Source Coding

In order to further explore describe accounting method (MSCC) the basic connotation, this paper gives the definition of the following:

Reference source: if the source is *'X_j'* source *'X_i'* when coding the edge information, the information source *'X_j', 'X_i'* is a source of reference source [12].

Correlation matrix: *'C'* matrix is constructed by relevance for the elements, among them, the *'X_j', 'C_{ij}'* refers to the source and source *'X_i'* degree of association between [13].

Relevance: the source is a correlation between the response indicators. The larger the index value, the greater the correlation, the number of bits required for encoding is less [14].

If *'(X_1, X_2, ... X_n, Y)'* are related to a plurality of sources, wherein, *'(X_1, X_2, ... X_n)'* is the encoding process of the source need to be, and *'Y'* is the side information. In order to unify the formula, the need to ensure formula (13) is tenable. Known the correlation matrix *'C'*, '$\overline{\mathrm{S}}$' is the already-encoded source set, *'S'* is the source to be encoded set [15].

$$X_n + 1 = Y \quad (13)$$

'E' is the encoding result obtained, *'R'* is the set of the reference signal source, the specific encoding process, as shown in Table 1.

Table 1. MSCC algorithm coding procedure.

Program 1	Initialize the function $\overline{\mathrm{S}} = \{Y\}$, $\mathrm{S} = \{\mathrm{X}_1, \mathrm{X}_2, ... \mathrm{X}_n\}$, $E = \phi$, $R = \phi$
Program 2	Choose 'X_j' from '$\overline{\mathrm{S}}$' and 'X_i' from 'S', ensure 'X_j' and 'X_i' satisfy the relationship in the function: $\underset{x_i \in S,\ x_j \in \overline{S}}{\mathrm{argmax}}\ \mathrm{C}_{ij}$
Program 3	Delete 'X_i' from S function then add into function of '$\overline{\mathrm{S}}$', add 'X_j' into function of 'R'
Program 4	Implement related information source encoding process on 'X_i' to 'X_j' and add decoding result of 'f(X_i)' in 'E' function
Program 5	Repeat program 2 and 4 again till empty the function'S'
Program 6	Delete 'Y' from'$\overline{\mathrm{S}}$'completely

Finally, *'R'* is the corresponding reference signal source, *'E'* is the final result of all signal source obtained according to the '$\overline{\mathrm{S}}$' order, '$\overline{\mathrm{S}}$' for all source coding sequence, the following figure shows a more detailed image of a coding process, between the source the main measure of correlation as the distance between them, arrow refers reference relationships, that is, the arrow side of the arrow tip source is encoded source reference implementation of the relevant sources, circled numbers refer to the coding sequence, black circle refers to a side information *'Y'*, shown in Fig. 2.

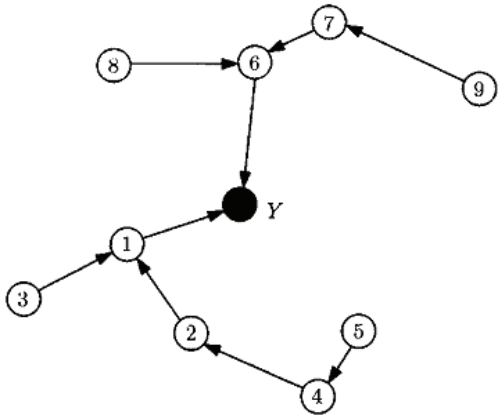

Fig. 2. MSCC algorithm coding procedure.

'*D'* indicates a decoding result, if the encoding process information formed *'(E, R,* $\overline{\mathrm{S}}$ *)'* can be obtained, then the decoding process to be successful, the decoding process shown in Table 2.

Table 2. MSCC algorithm decoding program.

Program 1	Initialize $\{Y\} = D$
Program 2	'r' is the first element of 'R', 'e' is the first element of 'E', implement related information source encoding process on 'e' appointed by 'r' to 'D'. Add the decoding results to 'D' in order, delete 'r' and 'e' at the same time
Program 3	Carry out the second step again till empty 'R' and 'E'
Program 4	Delete 'Y' in 'D', then D become the decoding result according to $\overline{\mathrm{S}}$, the specific codes order. Sort 'D' in natural order to get the final results for decoding

4. Algorithm Application

In wireless sensor networks, distributed data compression has obvious advantages, and clustered topologies helps DSC algorithm implementation, therefore, this clustering routing protocol in the application of the MSCC algorithm, enabling network performance improvement. Given this stage most of the sub-cluster routing protocol are based on LEACH protocol, which was the foundation of the implementation of improvement has obvious advantages. This paper briefly describes the basic connotation of LEACH and LEACH protocol proposed by MSCC algorithm based on the practical application.

4.1. LEACH Protocol

LEACH protocol for wireless sensor networks is an early kind of routing protocols, this protocol can transfer the data into multiple rounds program, and in their specific application process is repeated restructuring program clusters, each re-institutions include stable operation phases and cluster building phase, in which dispatch radio, members join the cluster head broadcast and competitive advantage of cluster building stage cluster heads of the four components.

4.2. Practical Application

In cluster sensor networks, since nodes in the cluster in the geographic location is closer, so their choice is also a strong correlation between the data, this correlation is a measure based on the knowledge and expertise mutual distance there between, you can also follow a specific period of statistical results. Known in the relevant characteristics, based on a small number of LEACH protocol to implement improvements MSCC algorithm and practical application, the specific procedures are as follows:

First, at the beginning of each round of the data transfer process before beginning each cluster head node competition. LEACH protocol with the same competitive basis, that is, the individual nodes are randomly formed one at 0-1 between the data and comparative analysis with the threshold. If the threshold is greater than the random number, it is treated as a new round of cluster heads, and shall inform the other neighbors, on the contrary it indicates that the cluster has been elected as the first or competition is unsuccessful, but only as a member node to which, $T_{(n)}$ can be values as in formula (14):

$$T(n) = \begin{cases} \dfrac{p}{1 - p[r \bmod (1/p)]}, & n \in G \\ 0, & others \end{cases}, \tag{14}$$

where *'G'* represents *'1/p'* elected cluster head round all the nodes, *'r'* refers to the number of cycles is, *'p'* represents the percentage of the cluster head.

Second, the success of competitive cluster head itself can be regarded as cluster heads broadcast messages.

Third, the failed node competition, cluster members by selecting with oneself the closest cluster head, and apply for its members.

Fourth, the cluster head with MSCC algorithm, according to the communication between cluster members on *'R'* and '$\overline{\mathrm{S}}$ ' calculated respectively, the order of time slot allocation is the order of each source in '$\overline{\mathrm{S}}$ ', the source of the reference source for related elements of 'R', and regard it as scheduling information, and broadcasting within clusters.

Fifth, relative to a reference source nodes respectively coding and encoding, the data to the cluster head sends in the allocation of time slots, and the cluster head to implement decoding.

Sixth, after the completion of data transmission, the first to the fifth step again.

Improvements idea generally comprises three points: First, the member node in MSCC slot allocation based on slot allocation algorithm. Second, the cluster head node should be members of the corresponding reference information broadcasting source. Third, the various sources without complete information to pass, usually pass encoding the information obtained, the need to decode the associated cluster heads to recover all the information.

5. Conclusions

In order to effectively get rid of the energy problem in wireless sensor network, and avoid the information redundancy, we have discussed a kind of distributed source coding algorithm suitable for clustering wireless sensor network and we also have done further verification by use of LEACH protocol to this kind of algorithm. And then we improved the LEACH protocol in some details which can be taken for am example for other clustering routing protocols in MSCC algorithms. Simulation results confirmed that this algorithm can greatly reduce the number of bits sent to save energy and prolong the service life of the network. But the information between members of the algorithm needs to be known as the basic premise. So how to obtain accurate information is the key of the algorithm can be successfully applied, which need us to do further research in the future. In addition, MSCC algorithm only from physical layer to send bits, how it might apply to other layers, such as MAC layer, network layer for joint optimization is also worth to continue to explore and research.

References

[1]. Fan Zhi-Ping, Jong-Chol, Xie Dong-Qing, Clustering algorithm for wireless sensor networks based on energy efficient, *Mini-Micro Systems*, Vol. 3, Issue 3, 2003, pp. 535-536.

[2]. Yang Jun, Zhang Yun, Zhang Yun-wing and etc., Cluster-based wireless sensor network data aggregation and transmission protocol, *Journal of Software*, Vol. 21, Issue 5, 2010, pp. 1127-1137.

[3]. Su Yan Jun, Wireless multimedia sensor networks based on distributed processing technology of image, *Shandong University Master's Degree Thesis*, Vol. 1, Issue 1, 2011, pp. 12 - 13.

[4]. Zhijie Han, Ruchuan Wang, Gogh Juan Fna and etc., WSN based on ARMA unbalanced clustering routing algorithm, *Journal of Electronics*, Vol. 38, Issue 4, 2010, pp. 865 - 869.

[5]. Zhao Zenghua, Dangao Tao, Zhang Yebin, New wireless image sensor network node hardware design, *Computer Science,* Vol. 34, Issue 9, 2007, pp. 22-25.

[6]. Xiaoguang Niu, Cui Li, Changcheng Huang, Mobile ad hoc networks based on optimized clustering hybrid routing protocol, *Journal of Communication*, Vol. 31, Issue 10, 2010, pp. 58-67.

[7]. Juan Li, Yan Li, Sina Li and Xinyi Chen, Discussion on multi-sensors data fusion technology, *Journal of Yunnan University*, Vol. 30, Issue 1, 2008, pp. 241-246.

[8]. Jiefeng Ou, Xinghua Liu and etc., Study on wireless sensor network (WSN) networking based on IEEE802.15.4, *Journal of Zhejiang University of Science*, Vol. 1, Issue 1, 2006, pp. 17-36.

[9]. Zhengdao Wang, Jing Zhang and etc., A optimization algorithm of JPEG2000 arithmetic encoder, *Journal of Hunan University (Natural Science Edition)*, Vol. 34, Issue 1, 2009, pp. 41-44.

[10]. Jie Li, Zhijie Han, A uneven clustering routing algorithm for WSN based on forecasting, *Journal of Natural Gas Chemistry*, Vol. 47, Issue 8, 2010, pp. 1459-1465.

[11]. Xin Li, Xiujiang Huang, Zaiping Nie, Discussion of MIMO wireless transmission technology, *Electromagnetic Field and Microwave*, Vol. 36, Issue 8, 2006, pp. 42-47.

[12]. Chao Ma, Research of the key technology of nodes in wireless medium sensor network, *Journal of National University of Defense Technology*, Vol. 1, Issue 1, 2008, pp. 10-12.

[13]. Bo Shen, Shiyong Zhang, Yiping Zhong, Clustering routing protocol in wireless sensor networks, *Journal of Software*, Vol. 17, Issue 7, 2006, pp. 1588-1600.

[14]. Guangxue Wang, Kai Liu, Cross-layer routing protocols in wireless sensor networks, *Journal of Beijing University of Aeronautics and Astronautics*, Vol. 36, Issue 6, 2010, pp. 732-740.

[15]. Xinyue Zhang, Shuqun Shen, The technology and its development prospect of UWB ultra-wideband wireless communication, *Data Communication*, Vol. 2, Issue 1, 2004, pp. 9-12.