

# An Algorithm of Wavelet Data Compression Based on Wireless Sensor Network

Luo Xiao  
School of Information and Engineering  
Huzhou Teachers College  
Huzhou, China



**ABSTRACT:** Firstly, this paper gives a data aggregation algorithm based on learning automata to solve the problem that the existing data aggregation algorithm can't solve, the uneven energy cost, and the existing algorithm can't change the gathering path dynamically existing the overhead environment. In the proposed method, nodes can change its gathering path to adjust the overhead environment. All the nodes of WSN equipped with a learning automata. These leaning automata learn all the gathering path of the nodes. In the process of transmit information two kinds of data are transmitted, including data packet, knowledge packet. When the information of the nodes changes, according to the feedback of the nodes, the learning automata gives the reward or punish to the current gathering path, which help to find the best gathering path. Secondly, this paper improved the wavelet data compression algorithm, which was brought out as the correlation between different data. The algorithm do not reduce much of the data relate to the original data. After the wavelet data compression, Huffman coding compression algorithm will improve the data compression ratio.

## Categories and Subject Descriptors

C.2.1 [Network Architecture and Design] Wireless communication E.4 [CODING AND INFORMATION THEORY] Data compaction and compression

**General Terms:** Wireless Communication, Data Compression, Wavelet data compression

**Keywords:** Learning Automata, Gathering Path, Data Compression, Wavelet Transform, Hot Node

**Received:** 11 August 2012, Revised 2 October 2012, Accepted 9 October 2012

## 1. Data Compression Algorithm Based on Wireless Sensor Networks

Shows the characteristics of wireless sensor networks, wireless sensor networks in the information transfer process will consume most of the energy node. The introduction of learning automata in wireless sensor networks, data fusion, to wireless sensor networks carrying data fusion can dynamically select the optimal integration path. After the use of multi-mode data compression, data compression in wireless sensor networks, data fusion. Enhance data fusion and data compression rate, extending the life of wireless sensor networks.

Assume that N sensor network nodes  $S_1, S_2, \dots, S_N$  scattered randomly in a  $L \times L$  rectangular area, used to monitor the particular phenomenon of  $\xi$ , such as temperature, humidity, and so on. To indicate the location of each node  $(x_i, y_i)$ . Each node is detected during the  $\xi$  and the use of multi-hop routing information will be recorded  $\xi$  reported to the base station. A small positive number  $\varepsilon > 0$   $|\xi_i - \xi_j| < \varepsilon$  set up, then and are almost equal,  $\xi_i$  and  $\hat{\xi}_i$  can be aggregated.

Suppose that in a dense sensor network, the value of the measured  $\xi$  in a specific area ( $\Omega \subset A$ ) is basically the same, then the node  $S_i$  for  $\hat{\xi}_i$  and  $\hat{\xi}_i$  can be aggregated,  $\forall S_j \in \Omega$ . A following the performance of  $Y$  is divided into  $M$  regions, respectively  $\Omega_1, \Omega_2, \dots, \Omega_M$ .

$$(1) |\xi_i - \xi_j| < \varepsilon; \forall S_i, S_j \in \Omega_M; k = 1, 2, \dots, M$$

$$(2) \Omega_i \neq \phi; \text{ for } i = 1, 2, \dots, M$$

$$(3) \Omega_i \cap \Omega_j \neq \phi; \text{ for } i \neq j$$

$$(4) \bigcup_{i=1}^M \Omega_i = A$$

Assume that each region  $\Omega_k$  can be approximated by the circular area  $(X_k, Y_k)$  as the center and radius  $R_k$ . At the same time assuming a different location in the network,  $\xi$  changes over time and is constantly changing, and therefore  $\Omega_1, \Omega_2, \dots, \Omega_m$ , is not a static area, the location and radius of each region in the life cycle of the wireless sensor network is constantly changing.

$\Omega_k$  The purpose of the wireless sensor network is the location of the collection of nodes of the cycle, in which  $\Omega_k$  the radius of the area and collect the information  $\widehat{\xi}$ . And periodically collected all kinds of information to pass to the node of the network's base station.

After the route discovery phase, each node in the time interval  $t$  after the node is passed to the base station.

During the routing phase, each node in a data fusion, learning automata to select the best route to the base station to transmit data packets. Each node through automatic machine learning to find the best neighbor of the information transmitted to the base station.

Learning Automata (LA) associated to the node  $S_i$  proposed  $LA_i$  the  $RL_i$  kinds of action, choose to go to every act of the possibility of probability are set to  $1/RL_i$ .  $LA_i$  every action corresponds to the corresponding  $|$  the  $RL_i$  in a neighbor used to send information to the base station. When the node  $S_i$  is the active node, and have measured the relevant information, the node will send a signal to bring their own automata, choose one of its neighbor nodes to the measured information to the base station requests automata. Stage in the automatic machine learning routing data to be aggregated. The choice made by the automaton, the return signal reward or punishment, reward or punishment given by returning signal will cause the current routing path is selected, the probability of change will be given by a neighbor node.

For two random movement of nodes  $S_j$ , from its neighbor node receives two different packages; packet of the packet and knowledge. For packet storage should be measured, such as temperature, as well as the location of the information. Knowledge package is made for the packet issued by  $S_j$  reaction and answer. The information is as follows.

$K_i$  represents the number of packets aggregated into this package.

$\widehat{\xi}_i K$  Represents the use of (1) calculated from the aggregated data.

$$\widehat{\xi}_i K = \frac{\widehat{\xi}_i + \sum_{j=1}^n \widehat{\xi}_j}{n+1} \quad (1)$$

Where  $n$  is the node  $S_i$  received packet number.  $(X_i K_i, y_i K_i)$  represented by the formula (2) calculated from the aggregate location information.

$$x_i, K_i = \frac{x_i + \sum_{j=1}^n x_j}{n+1} \quad y_i, K_i = \frac{y_i + \sum_{j=1}^n y_j}{n+1} \quad (2)$$

Received packet will be temporarily stored and will be the next activity of the nodes of  $S_i$ .

Node  $S_i$ ,  $N$  action will result in the following.

$S_i$  set  $K_i = 1$ .

$S_i$  perceive the surrounding environment and measurement  $\widehat{\xi}_i$ .

$S_i$  Set up  $\widehat{\xi}_i, K_i = \widehat{\xi}_i, x_i K_i = x_i, y_i K_i = y_i$ .

If the next  $n-1$  action in the new packet is received. So,  $S_i$  creates a packet, the contents of the packet is  $\widehat{\xi}_i, K_i, (K_i, y_i K_i), K_i$ .

$$|\widehat{\xi}_i, K_i - \widehat{\xi}_j, K_j| < \varepsilon \quad (3)$$

If a packet was received  $n-1$  action, then each receives a packet, if the formula (3), will be handled as follows:  $\varepsilon$

Is a threshold value represents the maximum difference of the measured data is less than it, the data fusion can To be executed.

(1) Siusing (4) for data integration

$$\widehat{\xi}_i K_i = \frac{K_i \cdot \widehat{\xi}_i + K_j \cdot \widehat{\xi}_j}{K_i + K_j} \quad (4)$$

(2) Siusing (5) the integration of location information.

$$x_i, K_i = \frac{K_i \cdot x_i + K_j \cdot x_j}{K_i + K_j} \quad y_i, K_i = \frac{K_i \cdot y_i + K_j \cdot y_j}{K_i + K_j} \quad (5)$$

(3) Set  $K_i = K_i + K_j$

(4) the rate of fusion by the use of (6) calculated data

$$DAR_{i,j} = U \left[ \frac{\text{Min}(\widehat{\xi}_i, K_i + \varepsilon, \widehat{\xi}_j, K_j + \varepsilon) - \text{Max}(\widehat{\xi}_i, K_i - \varepsilon, \widehat{\xi}_j, K_j - \varepsilon)}{2\varepsilon} \right]$$

$$U(x) = \begin{cases} x; & x \geq 0 \\ 0; & x < 0 \end{cases} \quad (6)$$

(5) The feedback of the environment as a knowledge packet is sent back to the node of  $S_j$ .  $S_i$  creates a data packet, the contents of the packet is  $\widehat{\xi}_i, K_i, (x_i, K_i, y_i, K_i), K_i$ .

Ket data to the new collection from node  $S_i$ , if not satisfied (1-1) will be passed to the neighbor  $K$ .

Received the knowledge contained in the packet of node

$S_j$ ,  $LA_i$  punish or reward this action. If  $DAR_{k,j}$  is greater than the acceptable efficiency. The action then  $K$  will be in accordance with the formula (7) to the reward.

$$\begin{aligned} pk(n+1) &= pk(n) + \alpha \cdot (DARK, i) \cdot (1 - pk(n)) \\ pl(n+1) &= pl(n) + \alpha \cdot (DARK, i) \cdot pl(n) \end{aligned} \quad (7)$$

$\forall l \neq k$  Learning automata routing collection  $LA_i$  select the action corresponding to the node, to determine the new aggregate data should be sent to which neighbor. The newly created packets, and pac

Otherwise, the punishment will be in accordance with the formula (1-8).

$$\begin{aligned} P_k(n+1) &= (1 - \beta \cdot (1 - DAR_{K,i})) P_k(n) \\ P_l(n+1) &= \frac{\beta \cdot (1 - DAR_{K,i})}{r - 1} + (1 - \beta \cdot (1 - DAR_{K,i})) P_l(n) \end{aligned} \quad (8)$$

$\forall l \neq k$  If the node  $S_j$  did not receive any knowledge of the packet from the node  $S_k$   $LA_i$  (8) to punish the selected action. (8) In the value of  $DAR_{K,i}$  by default to 0. Node  $S_i$  to be re-passed measure to the information, exclude  $RL_i$  in  $S_k$  node.  $RL_i$  in the existence of other neighbor nodes, node  $S_i$  will it all neighbor nodes issue route request, if a neighbor node of  $S_i$   $S_j$  receives the request, and the node  $S_j$  not in the list  $RL_i$  node  $S_j$  reply to route reply information. Node  $S_i$  statistical received route reply message, and add the node to itself  $RL_i$  list. If the node  $S_i$  and did not receive any route reply, so at this time, this node is already a "dead node", and for the entire wireless sensor network, the network lifetime is over.

Secondary route selection algorithm for the TLA-DA, need to consider when selecting a neighbor node, the change in the routing path the node to be the two nodes of the data fusion. Shown in Figure 1, if you use the basic routing algorithm, node  $S_1$  will be randomly selected nodes in the same region  $S_2, S_3, S_4$  to send a packet thus transferred to the sensor network base station and it is. Secondary route selection algorithm will select the node  $S_4$ , because the next neighbor node  $S_4$  to  $S_5$  is exactly within the domain where also node  $S_1$  can be better data integration.

The second routing algorithm: any node  $S_i$   $n$  times in the data fusion process, from its neighbor node  $S_j$  packet is received by (6) and (7) to calculate the value of  $DAR_{i,j}$ , and then again by (9) to calculate the  $E - DAR_{i,j}$ ,  $E - DAR_{i,j}$  is calculated, will be returned as a node  $S_i$  of feedback to node  $S_j$

$$E - DAR_{i,j} = MAX(DAR_{i,j}, MRR_i(n)) \quad (9)$$

Where  $MRR_i$  represents a node  $S_i$  receives the greatest reward.  $MRR_i$  defined in equation (10):

$$MRR_i(n) = \max_{1 \leq i \leq n} (DAR_{k,j})_{k \in RL_i} \quad (10)$$

Where max values represent the beginning to calculate the max value of this time node  $S_i$  received all feedback from the wireless sensor networks.

Through the secondary route selection algorithm, the greater the value in the network shown in Figure 1, node  $S_2$  is  $MRR$  obtained by calculating the value node  $S_3, S_4$ .  $DAR_{i,j}$  for the three nodes is the same node, but the value of  $S_2$  greater than the value of the other two nodes, so the result when the rate of node  $S_1$  choose a neighbor node, the node  $S_2$  is selected to be greater than the node  $S_3, S_4$ , and the node.

When a node in the same area were unable to find the neighbor nodes for data fusion, as in Figure 1 node  $S_5$ , in the same area were unable to find neighbor nodes for data fusion, the  $S_5$  will choose a node  $S_8, S_9$  maximum residual energy of nodes as data integration neighbors. In the second routing algorithm, in a region boundary node to the node to transfer data outside the region will choose the node with the most remaining energy for data transmission. Secondary route selection algorithm, data fusion in the same region  $\Omega_k$  first consideration is the ratio of the level of data fusion between different regions during the data fusion is the first consideration with the highest residual energy node data fusion.

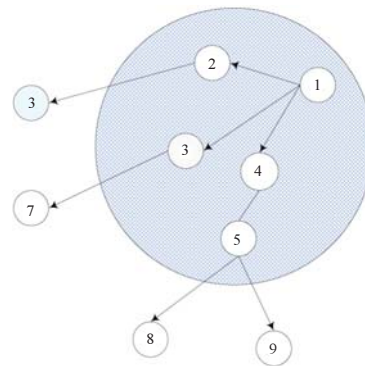


Figure 1. Node environment maps

Second choice routing algorithm, when node  $S_i$  by the neighbor the  $S_j$  passed over packet during feedback the same time, learning automata calculated the value of  $E - DAR_{i,j}$ . If this aggregate is higher than the acceptable data fusion, then the node  $S_i$  will be an acceptable learning automata  $LA_j$  feedback signal. If the adjacent node in a different area, then when the  $S_i$  feedback to the learning automaton  $LA_j$ , it is necessary to take into account the residual energy of nodes. How the feedback signal is given to decision-making according to the number of the node residual energy. If the current energy surplus ratio of  $NRE_i$  is higher than the acceptable rate, then back to the L-reward signal, otherwise the feedback signal to  $LA_j$  punishment

$$NRE_i = \frac{EL_i}{MaxEnergyLevel} \quad (11)$$

Where existing energy  $EL_i$  on behalf of the sensor nodes,  $MaxEnergyLevel$  represents the sensor network node with all the energy at the outset.

## 2. Wavelet multi-mode compression algorithm

Wireless sensor network node in the practical application of the process, each node usually needs to collect different types of information. Often collected by different data, the data of different data by the variation between the data and performance. In this paper, in order to facilitate description of some of the definition of the concept.

Wireless sensor networks by learning automata by domain, when the nodes in the cluster receives the data of neighbors over  $P$  and  $P$  packets each including measured values of  $Q$  different types of measurement information, sonode to collect the amount of information for  $P * Q$  receives the  $P$  data is divided into  $P$  columns, each column represent the same measurement information signal values, respectively,  $X_1, \dots, X_p$  represents. First of all, pass the  $P$  column of data according to the following algorithm.

```

Input:  $P$  different types of observation collection  $\{X_1, X_2, \dots, X_p\}$ , the relevance of a threshold  $\varepsilon$ 
Output: Handling the collection of the data sequence, the degree of coupling is greater than  $\varepsilon$ 
For  $i$  from 1 to  $p$ 
  int flag [ $i$ ] = 0
  if flag [ $i$ ] = 0
    for each  $X_i \in \{X_1, X_2, \dots, X_p\}$  do
      record corr ( $x_p, x_j$ )
      if (corr ( $X_p, X_j$ ) >  $\varepsilon$ )
        (// If the data is greater than a given threshold
        flag [ $i$ ] = "Z", insert ( $j$ , arry [ $i$ ]) ///  $J i$  where the sequence
        is inserted into
        return  $j + 1$ 

```

The value of the received data, each of which is an array collection. The first item in each collection is the characteristic sequences of other items in the collection of related sequences, then the collection of array processing.

Each set of data on the data processing are strongly correlated, meaning that there is a linear relationship can be used in a straight line representation of data collection between. The straight line is the slope of slash, set the relationship expressed as  $y = ax + b$ . The straight line

$$a = \frac{L_{xy}}{L_x}, b = \bar{y} - \bar{b}x$$

Where

$$L_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - \frac{1}{n} \left( \sum_{i=1}^n x_i \right) \left( \sum_{i=1}^n y_i \right)$$

$$L_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - \frac{1}{n} \left( \sum_{i=1}^n x_i \right)^2$$

Sensor networks usually measured in the practical application of the same kind of data, the temporal correlation and spatial correlation, and processing these data, they usually use the wavelet transform, commonly used wavelet transform Harry, Daubechies, Coiflet Symlet, Meyer, Morlet and Mexican Hat. Relatively speaking, Harr wavelet is a wavelet transform in a simple way. Length  $2^n$  signal  $S^n = \{S_{n,l} | 0 \leq l < 2^n\}$ , Averaging with details of the application to the  $a = S_{n,2^l}, b = S_{n,2^l+1} (l = 0, 1, \dots, 2^{n-1} - 1)$   $\neq$ .

$$\text{Hutchison } S_{n-1,l} = \frac{S_{n,2^l} + S_{n,2^l+1}}{2}$$

$d_{n-1,l} = S_{n,2^l} - S_{n,2^l+1}$  is multi-level Harr wavelet transform decomposition process and the reconstruction process can be used in Figure 2 and Figure 3.

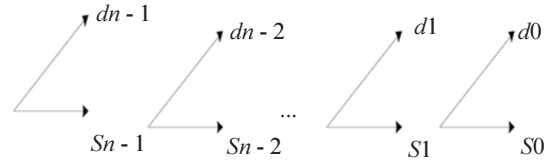


Figure 2. The wavelet decomposition process diagram in

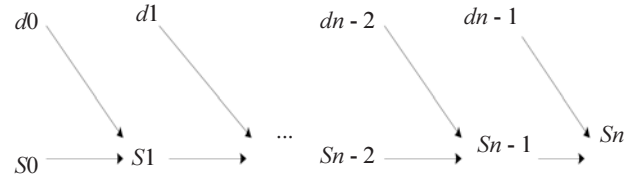


Figure 3. Wavelet reconstruction process diagram in

The Harr algorithm 3 wavelet decomposition of the space - frequency structure shown in Figure 4.

L2	H2	H1	H0
----	----	----	----

Figure 4. SPACES - frequency chart

According to the above-described fast Harr wavelet transform wavelet analysis - frequency structure shown in Figure 4. For wavelet transform, wireless sensor networks to collect data transformation, compared to the high-frequency part of most of the coefficient value of approximately 0, most of the energy is concentrated in the low frequency coefficients. Shown in Figure 4, after the initial signal after the wavelet transform broken down into three parts h0h1h2 high-frequency part of the L2 part of the low-frequency part. For the above data after wavelet classification, most of its high-frequency part of the coefficient is 0. RLE is characterized by relatively simple, and easy. RLE compression effect and the average run length and the number of pending compressed data will be a recurring character.

Huffman compression by Huffman coding (Huffman coding) in the fifties of last century, Huffman coding is a kind of variable length coding (VLC). Its high coding efficiency, computational speed is fast, flexible way, Huffman coding is widely used on the application of the Huffman coding such as JPEG.

Huffman coding is based on the probability of the source signal is encoded to appear shorter the higher the probability of symbol for the design of the code word, on the contrary the smaller the probability of the symbol, its corresponding code word is the longer, thus average less yards long. Theoretical studies have shown that the Huffman coding method is close to the entropy of the source.

Binary tree with n nodes, the full binary tree or a complete binary tree with minimum path length, the length of the binary tree with the right path:

$$WPL = \sum_{i=1}^m W_i L_i$$

Where:  $n$  is the number of terminal nodes of a binary tree; the  $W_i$  is the weight of the  $i$ -th terminal node; of  $L_i$  for the length of the path from the root to the  $i$ -th terminal node. Suppose there are  $n$  weights  $\{W_1, W_2, \dots, W_n\}$ , construct a binary tree of  $n$  terminal nodes. Each terminal node of the weight of  $W_i$ . Obviously this binary tree can be constructed out of trees, which there must be a weighted path length of the smallest binary tree, the tree binary tree called Huffman tree (Huffman), also known as the optimal tree. The right value node in the Huffman tree, the more close to the root node.

The pseudo code of the binary tree constructed as follows:

```
void CreateHuffmanTree (HuffmanTree T)
{// Construct the Huffman tree, T [m-1] for its root
an int i, p1, p2;
InitHuffmanTree (T); // T-initialization
InputWeight (T); // input the value of the leaves of the
right to T [0. . weight domain of n-1]
for (i = n; i < m; i++)
{// N-1 merger, the new node is in turn stored in T [i]
SelectMin (T, i-1, & p1, & p2);
// T [0. . i-1] to select the smallest root of two weights,
serial numbers, respectively, p1 and p2
T [p1] parent = T [p2, parent = i;
T [i].l child = p1;
// Minimum weight of the root node is the left child of
the new node
T [i].Rchild = p2; // small right to the root of the right
child of the new node
T [i] weight = T [p1]. Weight + of T [p2]. Weight;
} // End for
```

Shown in Figure 5 three binary tree has four leaf nodes, and with the same weights 5, 4, 1, 2. Three binary tree with the right path length, respectively, as follows:

$$WPL = 5 \times 2 + 4 \times (2 + 1) \times 2 + 2 \times 2$$

$$WPL = 5 \times 3 + 4 \times 3 + 1 \times 1 + 2 \times 2$$

$$WPL = 5 \times 1 + 4 \times 2 + 1 \times 3 + 2 \times 3$$

The length of the binary tree  $c$  with the right path, so that  $c$  is the Huffman number, and shows that the minimum weighted path length of the tree is not necessarily a complete binary tree.

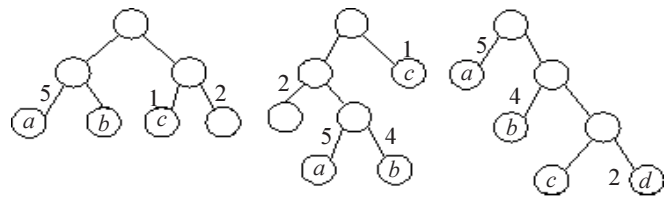


Figure 5. Binary Tree

Huffman coding, the Huffman tree is created, agreed to the left branch, said the character “0”, the right branch of the character “1”. You can use a string of characters from the branch of the root node to leaf nodes in the path as the character encoding of the leaf section. Figure 5 a, b, c, d four nodes can be Huffman coding and encoding of a 0, b encoded as 10, c is encoded as 110, d is encoded as 111. The encoding process shown in Figure 6.

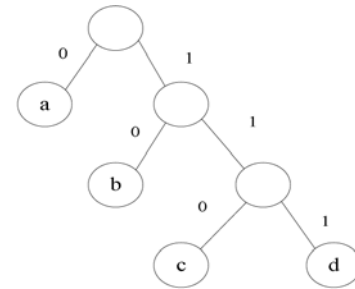


Figure 6. Huffman coding schematic

If the source file is aaabbbbaaacddabdd “need 18Byte of storage space, the use of Huffman compressed binary encoding is: 11111101011111 000101010000110110 the just 4Byte + 1Bit storage space, and Huffman coding of the source file can save nearly 14Byte of storage space.

In terms of data compression for wireless sensor networks, through optimization of the algorithm and the algorithm computation time to save energy consumption in wireless sensor networks, data transmission process is a very big impact. From the compressed data, decompress the reconstruction quality, make the wavelet coefficients of the energy loss are kept to a minimum. Therefore, based on the actual situation on the wavelet transform coefficient of the allocation of the “adaptive” threshold. Harr transform the wavelet data compression based on low frequency data after compression, Huffman coding, under the premise of ensuring data quality, improve the data compression rate.

Formal language based on a combination of both data compression method is described as follows:

Input: P different types of observation collection  $\{X_1, X_2, \dots, X_p\}$ , the correlation threshold.

Output: After Huffman coding Harr transform the low-frequency data compression data.

Corr ( $X_i, X_j$ ),  $1 \leq i \leq p$ ;  $1 \leq j \leq p$ ;

// Calculate the data of the input data sequence correlation

if Corr ( $X_i, X_j$ ) >  $\epsilon$ , insert ( $j$ , array [ $i$ ]);

// Preprocess the data, the data of the data is greater than  $\epsilon$  is divided into a data sequence

the if lenth (array [ $i$ ]) > 1, Harr, (array [ $i$ ]);

```

// If the number of elements and process data
collection is greater than 1, the data memory Harr
transform
    The RLE (h0, h1, h2); / RLE / high-frequency
data
    Huffman, (L2); // Huffman coding low
frequency data
end;

```

### 3. Simulation experiments and results analysis

The experiment mimics the six different temperatures, position and movement speed of the external environment. Mobile simulation environment limited the boundaries of wireless sensor networks, when the environment mobile wireless sensor network boundary, then the environment will be toward its original direction of movement in the opposite direction and move in order to ensure that all nodes in a simulation environment. Moving speed of the environment according to  $\mu = 0.001$ ,  $\sigma = 0.0001$  normal distribution random selection. In experiments based on learning automata data aggregation algorithm parameters alpha, beta is set to 0.1, and the acceptable range of compression settings of a quilt is set to 5. Q-Learning algorithm of the parameter  $\hat{\alpha}$  is set to 0.1, beta is set to 0.01. Nodes every 10s to the base station sends a data rate of data reception is set to 0.85. In this experiment, called LA-DA algorithm, known as the secondary route selection algorithm based on LA-DA algorithm to improve the algorithm for the TLA-DA algorithm based on learning automata for wireless sensor networks data fusion algorithm.

The packet number of the comparison. Figure 7 shows the experimental simulation environment based on learning automata for wireless sensor networks data fusion algorithm described in this article to the other three algorithms comparison of the number of packets received.

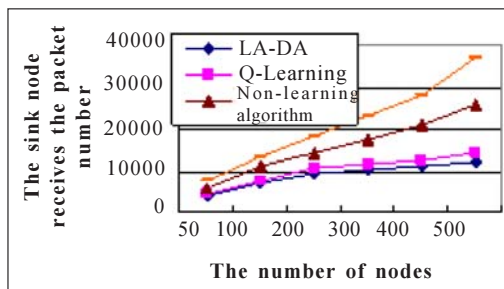


Figure 7. Comparison chart of number of packets of base station receives

The number of packets shown in Figure 7, with the increase of the number of nodes, four algorithms receiver to an uptrend, but the amount of the received data is based on data fusion algorithm for learning automata with the increase of the number of nodes, data the amount of increase is far less than other nodes.

Figure 8 500-node sensor network when  $\mu$  region [0-0.05] within the region changes, the base station receives the total amount of data comparison.

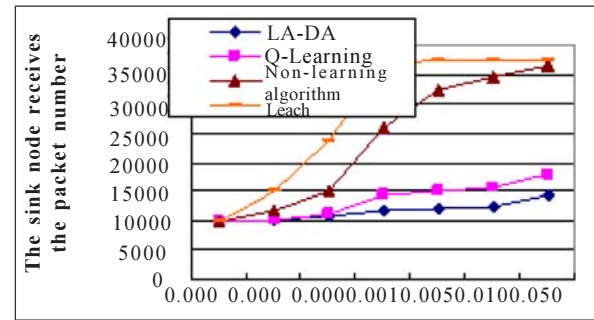


Figure 8. Comparison chart of number of packets of base station receives when  $\mu$  value changes

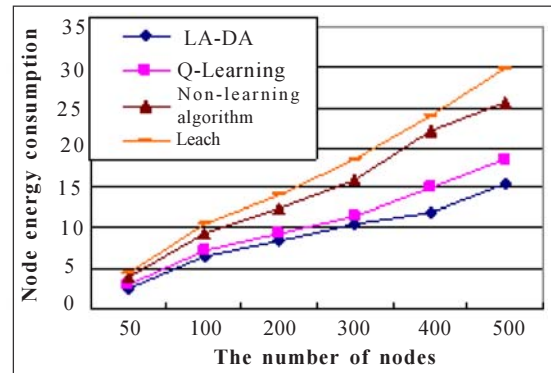


Figure 9. Energy consumption

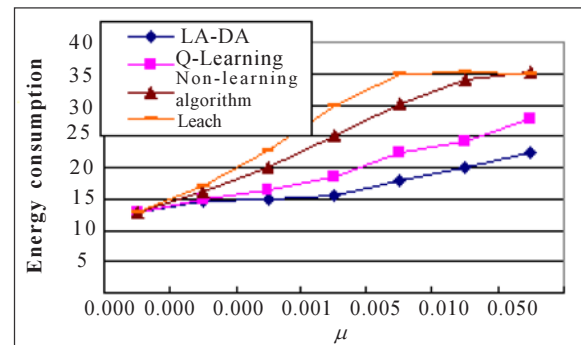


Figure 10. Energy consumption when  $\mu$  value changes

Shown in Figure 8, the packets and the other three methods when the values of  $\mu$  0, the received data fusion algorithm based on learning automata is basically no difference in fusion rate. With the increase of the value of  $\mu$ , learning automata-based data fusion algorithm received packet is significantly less than the received data packet by using the other three algorithms

Experiments compare wireless sensor network node energy consumption, based on learning automata wireless sensor network data fusion algorithm, the knowledge package in the transfer process also need to consume some energy, knowledge package is consumed in the delivery process energy is also being considered in the overall energy consumption within the scope. After averaging the experimental data found for process-based learning automata for wireless sensor networks data fusion algorithm for packet delivery process, a knowledge-based package of energy consumption accounts for each node

packet transmission 30% of the total energy consumption. Figure 9 shows based on learning automata for wireless sensor networks data fusion algorithm, the comparison with the energy consumption of the other three algorithms in the data transfer process.

Shown in Figure 9. Learning automata-based wireless sensor network data fusion algorithm knowledge package is consumed in the process of passing 30 percent of the overall packet delivery, but in the process of passing the overall number of packets less consumption the total energy is less than the other three algorithms.

In order to prove the impact of external environment on the energy consumption in several algorithms, so that  $\mu$  values vary between 0 to 0.5. The energy consumption of the test 500 nodes in this process. The experimental results shown in Figure 10.

Shown in Figure 10. When  $\mu = 0$ , due to the external environment of the network is essentially no change in the energy consumption of the four algorithms are basically the same. With the increase of the value of  $\mu$ , due to the external environment can transform rapidly month, learning automata-based wireless sensor network data fusion algorithm for the other three algorithms, the energy savings more, reflected in the more obvious advantages.

#### 4. Conclusions

Wireless sensor networks are widely used in medical information gathering, military information collection, environmental protection, monitoring, traffic monitoring, and other fields. Wireless sensor network nodes due to limited energy carried, resulting in limited sensor networks, energy-saving life. In this paper, the wireless sensor networks, data fusion, data compression field, overheating node for the existing network data fusion algorithm, and cannot meet the dynamic changes of the external environment; and existing network data compression algorithm is not satisfied with the problems were analyzed.

Simulation results show that learning automata-based data algorithms for data aggregation, while the location of the node of polymerization; and learning automata response to the dynamic adjustment of the aggregate path for the data in the environment is a high degree of transformation vary greatly improve its data fusion rate. Improved wavelet multi-mode data compression algorithm, which greatly improved the compression ratio of the algorithm, and no loss of accuracy of data, to extend the life of sensor networks. For improved wavelet multimode data compression algorithm, when the data coupling threshold value set is small, the data compression rate increases, the accuracy of the data will fall. This will serve as the direction of future research efforts.

#### 5. Acknowledgments

The work is supported by Zhejiang Provincial Natural

Science Foundation of China under Grant Nos. LY12F03008.

#### References

- [1] Kasirajan, P., Larsen, C. (2010). A New Adaptive Compression Scheme for Data Aggregation in Wireless Sensor Networks. *Wireless Communications and Networking Conference*. p. 1-6.
- [2] Sun LM, Li JZ, Chen Y, Zhu HS. (2005). *Wireless Sensor Network*. Beijing: Tsinghua University Press. p. 422-431
- [3] DARPA Sensor Information Technology Program [D/OL]: <http://www.Darpa.mil/ito/research/sensit/index.Html>
- [4] Savarese, C., Rabaey, J. (2001). Robust positioning algorithms for distributed ad-hoc wireless sensor networks. *In: Park Y, ed. Proceedings of the USENIX Technical Annual Conference*. Monterey: USENIX, p. 317-328
- [5] Ratnasamy, S., Karp, B. (2002). GHT: A geographic hash table for data-centric storage. *In: ReghavendrvCS, ed. Proceedings of the 1<sup>st</sup> ACM International Workshop on Wireless Sensor Networks and Applications*. New York: ACM Press. p. 94-103.
- [6] University of California at Los Angeles. WINS: Wireless integrated network sensors. <http://www.janet.ucla.edu/WINS/biblio.htm>
- [7] Ganesan, D., Govindan, R., Shenker, S., Estrin, D. (2002). Highly-Resilient, energy-efficient multipath routing in wireless sensor networks. *Mobile Computing and Communications Review*, 1 (2) 295-298.
- [8] Girod, L., Bychkovskiy, V., Elson, J., Estrin, D. (2002). Locating tiny sensors in time and space: A case study. *In: Manoli Y, Kim KS, eds. Proceedings of the International Conference on Computer Design*. Piscataway: IEEE Press. p. 195-204.
- [9] Cerpa, A., Estrin, D. (2002). ASCENT: Adaptive self-configuring sensor networks topologies. *In: Kermani P, ed. Proceedings of the 21<sup>st</sup> International Annual Joint Conference of the IEEE Computer and Communications Societies*. Piscataway: IEEE Press, p.101-111.
- [10] Heidemann, J., Silva, F., Intanagonwivat, C. (2001). Building efficient wireless sensor networks with low level naming. *In: Marzullo K, ed. Proceedings of the 18<sup>th</sup> ACM Symposium on Operating System Principles*. New York: ACM Press, p. 146-159.
- [11] Intanagonwivat, C., Govindan, R., Estrin, D., Heidemann, J., Silva, F. (2002). Directed diffusion for wireless sensor networking. *ACM/IEEE Transactions on Networking*, 11 (1) 2-16.
- [12] Liu, J., Cheung, P., Guibas, L., Zhao, F. (2002). A dual-space approach to tracking and sensor management in wireless sensor networks. *In: ReghavendrvCS, ed. Proceedings of the ACM International Workshop on Wireless Sensor Networks and Applications*. New York: ACM Press, p. 162-173.

[13] Guibas, L.J. (2002). Sensing, tracking, and reasoning with relations. *IEEE Signal Processing Magazine*, 19 (2) 73-85.

[14] Xie, M. D. (2011). The Survey of Latest Researches on Online Code Dissemination in Wireless Sensor Networks, *IEIT Journal of Adaptive & Dynamic Computing*, (1) 23-28, Jan. DOI=[10.5813/www.ieit-web.org/IJADC/2011.1.4](http://dx.doi.org/10.5813/www.ieit-web.org/IJADC/2011.1.4)

[15] Wu, C. C. (2011). A Study of Synchronous and Bucket Trading Behavior of Institutional Investors, *IEIT Journal of Adaptive & Dynamic Computing*, (2) 14-25, Apr. DOI=[10.5813/www.ieit-web.org/IJADC/2011.2.3](http://dx.doi.org/10.5813/www.ieit-web.org/IJADC/2011.2.3)