Abstract

Routing is one of the most important technologies in wireless sensor networks. The energy efficiency of the sensor network is realized by combining other technologies with routing protocols in this paper. First, through the analysis of the energy model of the clustering routing algorithm, the local signal to noise ratio (SNR) is used to purposefully select the nodes in the cluster to transmit data. In order to avoid the energy consumption of a large number of nodes to transmit data, a simple architecture is proposed to combine the distributed source code with the cluster routing protocol. The method uses the data on the cluster head as the edge information to decode the compressed information transmitted from the cluster nodes. The simulation result shows that the energy efficiency of the network can be achieved through this combination.

Keywords: Viterbi algorithm; wireless sensor; communication protocol

1. Introduction

With the rapid development and maturity of wireless communication, integrated circuits, sensors, and MEMS, low cost, low power consumption, and multi-function microsensors are stepping into the field of vision [1]. These micro sensors are usually integrated with many functions, such as information acquisition, data processing, and wireless communication [2-3]. Wireless sensor networks (WSN) are multi-hop wireless networks that are deployed in a monitoring area by a large number of sensor nodes in a self-organized form. The purpose is to collaborate to perceive, collect, and process the information of the measured objects in the monitoring area and send it to a unified processing center. Because of its small size and light weight, the micro sensor is also called “intelligent dust” [4]. The emergence of wireless sensor networks (WSN) has attracted worldwide attention. The first wireless sensor network technology was studied by the United States military [5-6]. Since then, the National Natural Science Foundation of the United States has set up a large number of related projects. Intel, Boeing, and other companies have joined the research on wireless sensor networks [7]. With the further development of research and the continuous development of technology, its application has expanded from military defense to environmental monitoring, traffic management, medical treatment, counter-terrorism, disaster relief, and other areas closely related to public life [8]. Wireless sensors measure the information that people are interested in in the surrounding environment by using different kinds of built-in sensors. In communication, sensor nodes usually communicate through multi-hop and peer-to-peer ways, which can avoid signal fading and interference in long-distance wireless signal propagation. Through the gateway and other devices, the wireless sensor network can connect with the existing network infrastructure and publish the collected information to the existing network [9-10].
compatibility of software (TinyOS) and the comparability of hardware platform, most domestic research institutes are also studying foreign famous node hardware platforms as well as foreign research institutions [11-12]. Since 2003, the National Natural Science Foundation has set up less than 20 research projects and key projects that are related to the key technologies of sensor networks, and it has brought the upsurge of sensor network technology research [13]. In 2004, the National Natural Science Foundation of China listed the wireless sensor network as a key research project [14]. The “national long-term science and technology development plan (2006-2020)” has included “sensor network and intelligent information processing” in key areas of support and its priority theme “information industry and modern service industry”, and it focuses on supporting “self-organizing sensor network technolgy” in cutting-edge technology [15]. The National 863 plan and the 973 Plan also began to plan for research support of the sensor network. Overall, however, the research level of domestic research in wireless sensor networks is relatively backward compared with those of other countries because it started late, lacks innovative research on the whole system, and has fewer key independent intellectual property rights [16].

3. Methodology

3.1. Implementation of Distributed Source Coding based on Viterbi Algorithm

The Viterbi algorithm, which was proposed by Andrew Viterbi in 1967, is used for deconvolution in a digital communication link to eliminate noise. The algorithm is widely used in CDMA and GSM digital cellular networks, dial-up modem, satellite, deep space communication, and 802.11 wireless network deconvolution codes [17]. Viterbi algorithm is a dynamic programming algorithm for finding the most likely sequence of observation events - the Viterbi path - implicit state sequence, especially in the context of Markov information sources and hidden Markov models [18]. The terms “Viterbi path” and “Viterbi algorithm” are also used to find the dynamic programming algorithms that are most likely to explain the results of the observation [19]. For example, in statistical syntactic analysis, a dynamic programming algorithm can be used to discover the most likely context independent derived (parsing) strings. Sometimes the research of wireless sensor network nodes known as “Viterbi analysis” is application driven, which is different from the foreign development model. The application developer can detect problems in actual deployment and operation and solve problems of reality by combining production, teaching, and research. It is beneficial to the industrialization of sensor network faster [20]. Wyner-Ziv’s coding problems of continuous source include source coding, channel coding, and estimation. Because the information transmission has a speed limit, the source $X$ should be quantified. The source space $R^L$ is divided into $2^{LR}$ disjoint regions, where $R_s$ represents the rate of the source. That is, mapping $M_1$:

$$M_1 : R^L \rightarrow \{1, 2, \ldots, 2^{LR}\}$$

(1)

Set $\Gamma = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_{2^{LR}}\}$ to represent the collection of these $2^{LR}$ disjoint regions. Each area is represented by a code word, called the codebook $S$; that is, mapping $M_2$:

$$M_2 : \{1, 2, \ldots, 2^{LR}\} \rightarrow R^L$$

(2)

The goal of source coding is to design mapping $M_1$ and $M_2$. The source is quantified as a code word in the $S$, and the code is transmitted to the decoding end with a rate $R_s$ bit/symbol. The code word quantized as the source $X$ is expressed as an active code word. The random variable describing the activity code is $W$. The partition of $\Gamma$ is based on the edge distribution of $X$ and uses the Lloyd algorithm. It is the best estimate. The decoder optimizes the estimation of the $X$ based on the output $Y$ of the side information and the elements in the $\Gamma$ containing the $X$, making the distortion minimal:

$$\hat{x} = \arg \min_{x \in R} E \left[ \rho(X, a) \right]_{X \in \Gamma_i, y = y}$$

(3)

Among them, $i$ is a received message and $y$ is the output of side information. Therefore, the mapping $M_3$ is:

$$M_3 : R^L \times \{1, 2, \ldots, 2^{LR}\} \rightarrow R^L$$

(4)

The estimated error is a function of $R_s$, where the appropriate $R_s$ is selected and the distortion is maintained within a
given range. In channel coding, the system requires a transmission rate of $R_s$ to ensure a given distortion. By using the correlation between $X$ and $Y$, the decoder resumed the serial number of the $W$. Since the quantized $W$ has a certain correlation with the source $X$, there is also a correlation between $W$ and $Y$, which is represented by the conditional distribution of $P(Y|W)$. Based on this condition distribution, a virtual channel is introduced to simulate the correlation between $W$ and $Y$. $W$ is the input signal of the virtual channel, and the $Y$ is the output signal. The capacity of this channel is more than 0. The target of channel coding is to establish the channel code $C$ applied to the virtual channel, so that the reduction of transmission bit rate is as close to $I(W;Y)$ as possible. Let $2^{L_R}$ represent the number of code words in a well-designed channel code, in which $R_s$ is the channel rate. The quantized output becomes a member of the channel code. Then, the number of cosets is transmitted to the decoder, and the decoder uses it to find $W$. Mapping $M_4$ refers to the coset number of the channel code corresponding code word $W$, which is quantized by the calculation of the encoder:

$$ M_4 : \{1, 2, \ldots, 2^{L_R}\} \rightarrow \{1, 2, \ldots, 2^{L_R}\} $$

Then, the serial number is transmitted to the decoder with $R = R_s - R_c$, R bit/symbol transmission rate. Setting the $Y$ side information, the decoder finds the most likely code word in coset code word by channel decoding scheme, so the mapping $M_5$ is:

$$ M_5 : R^c \times \{1, 2, \ldots, 2^{L_R}\} \rightarrow \{1, 2, \ldots, 2^{L_R}\} $$

In this coding method, the decoding error probability is not zero because the edge information is decoded into a wrong code word. However, the probability of decoding error can be reduced by designing effective channel codes. For a given domain in the $\Gamma$, $I(W;Y)$ is decided by choice of the description code, that is, $R_s$.

3.2. The Energy Model of the Protocol

The energy consumption of the node when the data model passes the data is:

$$ E_{tx}(l, d) = \begin{cases} lE_{elec} + lE_{friss-amp}d^2 : d < d_{\text{onwer}} \\ lE_{elec} + lE_{two-ray-amp}d^4 : d \geq d_{\text{onwer}} \end{cases} $$

(7)

The energy consumed when receiving data is:

$$ E_{rx}(l) = lE_{elec} $$

(8)

In which $E_{elec}$ is the radio energy consumption of 50 NAR/bit, $E_{friss-amp}$ is the transmission and amplification of energy consumption (free space) for 10 Pico/bit/square meters, and $E_{two-ray-amp}$ is the amplification and transmission energy consumption (two) for 0.0013 Pico/bit/four meters. $l$ is the size of the packet, the unit is the bit, the $d$ is the transmission distance, and the unit is the meter. We carry out the analysis of the energy consumption of the LEACH protocol.

The focus of the $d = 86$ m in the figure is the size of the $d_{\text{onwer}}$. From Figure 1, we can get the following conclusions. When $d < d_{\text{onwer}}$, if $E_{elec} = E_{friss-amp}d^2$, then $d = 70$ m. When $d \geq d_{\text{onwer}}$, if $E_{two-ray-amp}d^4 = E_{elec}$, then $d_{\text{onwer}} = 78$ m. However, in the actual network operation process, because the distribution of nodes is very dense and the role of clustering is to reduce the transmission distance of cluster nodes, the actual transmission distance is much less than the calculated values above. This shows that when the two nodes are transmitted peer to peer, both transmission and acceptance will consume a large amount of energy that is independent of distance. Then, we analyze the simulation model given by LEACH according to a certain amount of data. The value of the amount of data is given in Table 1. In the simulation, we make a separate statistics on the different types of energy in the LEACH energy model based on the data of the table. At the same time, we also make statistics separately on the energy consumed by the different identities of the nodes.
Figure 1. A schematic diagram of the varying parts of the energy consumption per unit bit with distance

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation area size</td>
<td>100 Square meter</td>
</tr>
<tr>
<td>Number of nodes</td>
<td>100</td>
</tr>
<tr>
<td>Sink node location</td>
<td>(50, 175)</td>
</tr>
<tr>
<td>Cluster head number</td>
<td>5</td>
</tr>
<tr>
<td>Initial energy of node</td>
<td>2J</td>
</tr>
<tr>
<td>Radio energy consumption</td>
<td>50 nanocad/bit</td>
</tr>
<tr>
<td>Energy consumption (free space) of the box amplifier</td>
<td>10 skin coke/bit/square meter</td>
</tr>
<tr>
<td>Energy consumption of transmission amplifier (two ways)</td>
<td>0.0013 skin coke/bit/four meters</td>
</tr>
<tr>
<td>Broadcast packet size</td>
<td>25 bytes</td>
</tr>
<tr>
<td>TDMA set the size of the packet</td>
<td>44 bytes</td>
</tr>
<tr>
<td>Data Baotou size</td>
<td>25 bytes</td>
</tr>
<tr>
<td>Data size</td>
<td>300 bytes</td>
</tr>
<tr>
<td>Decode energy consumption</td>
<td>20 nanocad/bit</td>
</tr>
</tbody>
</table>

It is assumed that $N$ acoustic sensors are placed in a two-dimensional region. The output of these sensors is the amplitude of the sound signal. The output $X_i$ of the $i$ sensor is defined as:

$$X_i = \theta / \left\| y_i - y \right\|^a + \omega_i$$  \hspace{1cm} (9)

$\theta \in R$ is the state parameter of the target, $\alpha$ is a known attenuation factor of sound waves with distance (in this article, it is assumed to be 2). $\left\| \cdot \right\|$ is the Euclidean distance. The additional noise $\omega$ is the variance of a Gauss random signal with zero mean and a discussion. $y$ represents the location of the target, and $y_i$ represents the location of the sensor node $i$. The network composed of these $N$ sensor nodes is divided into $K$ clusters. $K$ is an optimized value calculated according to the size of the $N$ and the sensor distribution area. If we know the distance of the target distance to the node, then the three anchors can be used to determine the position of the target in accordance with the traditional method. The specific position is the intersection point with three anchors as a circle center and three distances as a radius. If the distance’s proportion information is used to determine the location of the target, four anchors are required. The two anchors can form a circle according to the proportion of information. At least three such independent circles are needed to determine the location of the target. Therefore, we need at least four anchors to determine the location of the target, as shown in Figure 2.

Figure 2 is a diagram of determining the position of a target using a distance ratio. In practical applications, the position information that is detected will be disturbed by the noise, so the position of the estimated target will be near the actual position.
4. Result Analysis and Discussion

Based on the above discussions and the simulation results listed, we can clearly see that the simulation results of the improved DISCUS algorithm, which we call the extended DISCUS method, are significantly better than those of the original method. Additionally, the result of DISCUS simulation under the binary channel is better than that of the DISCUS in the multilevel case. Below, we will make a detailed analysis of the above conclusions.

Figure 3 is a statistical graph of unequal cases of sources X and Y in binary symmetric channel. In the diagram, the difference of the highest level is that the 3-bit letter source code X and Y are not equal. For binary Gauss channels, even at lower signal-to-noise ratio (SNR), the ratio of the number of X and Y to the total simulation code word is about 1/20 of the ratio of DISCUS to higher SNR. Therefore, under the low SNR of the binary Gauss channel, a very good decoding effect can be obtained, that is, a lower bit error rate.

In the binary Gauss channel, the decoding error caused by the difference of the highest level occupies 30% ~ 40% of the total error at low SNR. With the increase in SNR, due to the reduction of the error code of the end of the sequence and the decrease of the continuity of the common error, the highest error finally became the only insoluble error. As we have analyzed earlier, only this part of the error is unrecoverable due to the characteristics of the decoder. For an ordinary error, the decoder can effectively correct with the decrease in the probability of its appearance. For the sequence end code word
error, because it has a certain probability of its own appearance, the decoding error gradually disappears with the increase in the SNR. For binary symmetric channels, the decoding error caused by the highest difference occupies more than 80% of the total error number. With the decrease in P, it gradually increases. Under the correlation of binary symmetric channel, the number of common errors also occupies some proportion. Therefore, it is also difficult to decode the decoder, and the bit error rate is still at a high level.

This algorithm has been implemented in the C++ language, and the specific simulation setting is shown as follows. 100 sensor nodes are evenly distributed in the area of 100m × 100m. The occurrence of the uniform distribution of the target nodes in the simulation area is that the X and Y coordinates of the target nodes are uniformly valued within 0 to 100. In order to simplify the process of simulation, \( \sigma_i = 0.005 \) and \( \theta = 1 \) were set up. Although the LSAS algorithm needs to transmit data to the cluster head twice, it can still save a lot of energy according to the simulation data. In the LSAS algorithm, the packet length used to transmit the local SNR is much less than the length of the transmitted state data. In wireless sensor networks, energy consumption is mainly concentrated on the process of data transmission, and this consumption is proportional to the length of the data. Because sensor nodes do not transmit long state data when the local SNR is not high, nodes save a large amount of energy with the simulation. It is shown on the map that in the 500 round, the LSAS algorithm only uses about half the energy of the LEACH-C.

Figure 4 shows that there is no obvious performance gap between the LSAS algorithm and the LEACH-C algorithm through the comparison of the results analysis to the node state. This shows that the Kalman filtering algorithm can be used more effectively by using the data on a node with high SNR. Although the number of iterations is less, the results are more inclined to the true value in each filtering process.

In the process of simulation, the sensor nodes are evenly distributed in the area of 100m × 100m. For the first application, the target appears randomly in the simulation area, and the horizontal ordinate and the ordinate of the target position are evenly distributed between 30 and 70. This ensures that the number of nodes that can participate in each valuation is more than 4. The detection range of the set node is 10m.

Figure 5 compares the balance of performance and energy consumption at different coverage densities. More node participation means more energy consumption. In the case of large density, increasing the number of participating nodes can achieve a better performance-price ratio. In the case of small density, because of the long distance among nodes, the newly added nodes may have the opposite effect on the estimation of the position of the mouth. Therefore, the participation of 4 nodes is a better choice. In this application, the fast descent method is better than the Newton method in the number of iterations and the performance of the estimation. For Newton's method, when the variance of Gauss’s noise changes, its performance is not very stable; this is because the starting point of this method falls faster in the vicinity of the optimal value. When the initial value is far from the optimized value, its performance will not be very stable. For the fast descent method, because the optimized model is an ellipse model, the optimal value can be obtained in one iteration.
5. Conclusion

With the further development of wireless communication technology, large-scale integrated circuits, and MEMS, it is possible to generate micro sensor nodes with communication function. In the development of wireless sensor networks, it is found that technology is urgently needed to compress relevant information collected by densely distributed sensor nodes. In this paper, the theory and implementation of distributed source coding and the optimization of cluster routing protocols are mainly discussed. With the implementation of a distributed source coding scheme based on concomitant and convolutional codes for binary Gauss channel and symmetric channel, the reason why the decoder can obtain better performance only in the case of higher SNR is analyzed. In addition, it is pointed out that a more efficient channel coding method should be introduced to obtain better decoding effect because of the performance limit of the convolutional decoder. The distributed source coding method based on coding and syndrome scheme is simulated and implemented, and the source of the method is decoded by combining multilevel code and syndrome. The simulation result shows that a better performance can be obtained for multilevel code under the lower compression rate. By analyzing the different types of energy consumed differently on the network lifetime in the energy model, a node selection mechanism based on local SNR is proposed.

Acknowledgement

This research was founded as part of the project “Research of Broadband Stationary Near-Field Source Localization Algorithm (Educational, Scientific and Cultural Organization of Jilin Province, 2016 (57))”.

References


