Energy Consumption Model of WSN Based on Manifold Learning Algorithm

Chunliang Zhou(1), Weiqing Qu(1), Zhengqiu Lu(2), Yangguang Liu(1)

(1) Department of College of Information Engineering, Ningbo Dahongying University, Zhejiang, CHINA
e-mail: 275137531@qq.com
(2) Department of Information & Media, Zhejiang Fashion Institute of Technology, Zhejiang, CHINA

SUMMARY

Energy saving is one of the most important issues in wireless sensor networks. In order to effectively model the energy consumption in wireless sensor network, a novel model is proposed based on manifold learning algorithm. Firstly, the components of the energy consumption by computational equations are measured, and the objective function is optimized. Secondly, the parameters in computational equations are estimated by manifold learning algorithm. Finally, the simulation experiments on OPNET and MATLAB Simulink are performed to evaluate the key factors influencing the model. The experimental results show that the proposed model had significant advantage in terms of synchronization accuracy and residual energy in comparison with other methods.

KEY WORDS: wireless sensor network; energy consumption model; manifold learning; simulation.

1. INTRODUCTION

With the rapid development of Wireless Sensor Network (WSN), the question how to extend the lifecycle of a node and reduce the energy consumption has gradually become a key issue to WSN [1-5]. Compared with the conventional network, WSN is featured with constrained energy and dynamic changes in topology characteristics [6-8]. Therefore, in addition to considering the WSN cluster characteristics, the energy consumption of the node shall be further considered. In order to address the WNS energy issues, the question how to reduce the energy consumption has become an issue that the design of routing protocol, the positioning algorithms, time synchronization algorithms, etc would have to consider. The TPSN (Timing-Sync Protocol for Sensor Networks) adopts the two-way message switching synchronization mechanism, which can obtain a better accuracy. However, as each node interacts with the host node in the group, it will consume a considerable amount of energy which made the foreign and domestic scholars to put forward a great deal of studies. Ref. [9] compared and studied the energy consumed of the cluster host node and of other member nodes within the
cluster, but the proposed clustering algorithm does not take full consideration of the dump energy required by the cluster head. Ref. [10] proposed the creation of a HEED protocol with the primary and secondary parameters, allocating the energy consumption evenly to each node, thus extending the lifecycle of the network. Ref. [11] proposed a self-adaption data collection model based on the node dump energy, which effectively considers its error variance that changes dynamically with time. Ref. [12] adjusted dynamically the data sending according to the node error and established a self-adaption data collection model in relation to the lifecycle-limited PSNs, and such model improved the data accuracy of the base station, and extended the network lifecycle of the WSN nodes. Ref. [13] proposed the calculation of the routing costs and selected optimal forwarding node collections under the adjustable and non-adjustable sending energy mode, hence making the total energy consumption of these data, forwarded to Sink via these nodes, a minimum.

Ref. [14] adopted the segment routing to get a minimum average energy consumption based on the optimal number of nodes in the segment. Ref. [15] reduced the energy consumption of the nodes by regulating the different sleep cycle parameters, thus extending the service life of the node. Ref. [16] considered the energy of the critical nodes in WSN as well as different operating cycles of the network, thus avoiding the tedious gate-boundary function resulting from the repeat creation of shortest path tree. Ref. [17] combined the hybrid-CS technology fusing the joint routing and the compression sensing to minimize the overall energy loss of the network, but such method does not take into account the energy consumption balance. Ref. [18] established a breakdown-based multiple-objective evolutionary algorithm to be applied in the heterogeneous WSNs environment to meet the needs of coverage and energy conservation. Ref. [19] obtained another multi-objective coverage method MOCOAMA, which deals with the multiple objectives like the coverage level, node activity and the dump energy. Ref. [20] proposed the weighted shortest path tree that is more suitable for several-for-one sensor network since it considers the network energy factors, which determines if the node is of great importance for weighted path energy function, and such algorithm prolongs to some extent the network lifecycle, but the shortest path tree is established for the entire network after the death of each node which causes energy waste. Ref. [21] and [22] proposed a spatial-temporal correlativity-based efficient data acquisition policy and dynamically expandable tree based on the spatial correlation of the continuous moments and (or) between the adjacent nodes data, able to utilize the node energy efficiently and to guarantee the data accuracy.

However, these optimization algorithms or policies still exhibit some problems, like large amount of computing or slow convergence rate [23-27]. On the basis of the above studies, this paper first considers the energy consumption model of WSN and then takes advantage of popular learning algorithms to optimize and solve the above-mentioned methods and problems. Meanwhile, OPNET and MATLAB are combined for simulation experiment, making in-depth study on the key factors that affect this model. The structure of this paper is as follows: Section 2 exhibits the energy consumption model of WSN and the computation method; Section 3 clarifies the solution algorithm using the popular learning algorithms; Section 4 explains the simulation experiment; Section 5 summarizes the whole paper.
2. ENERGY CONSUMPTION MODEL

Assuming that the WSN coverage area is an area with the side length \( a \) where a large number of nodes is distributed evenly, and the total number of nodes of \( N \), with the known distance between the adjacent nodes. The broadcasting distance of each sensing node is \( R \). Sink node is located at the centre of the network.

![Fig. 1 Schematic Diagram for Information Exchange](image-url)

After the completion of the hierarchical creation, the Sink node initiates the information synchronization. Layer by layer all the nodes in the sensor network are synchronised and eventually the entire network. At this point, the Sink node broadcasts the synchronization information packet, and selects the node with the shortest distance \( d_{0,i} \) as the response node. As shown in Figure 1, the node \( S \) selects the node \( D \) with the shortest distance \( d_{i,i+1} \) as the response node, and the synchronized information packet received by all nodes from \( S \), records the local time, but only the node \( D \) will reply the response information to \( S \) after delaying a random interval. After the node \( S \) receives the response information packet, it will calculate the time migration \( \delta \) between the node \( S \) and the node \( D \) as well as the propagation delay \( \alpha \) with the two-way message switching synchronization mechanism Eq. (1).

\[
\begin{align*}
\delta &= \frac{(T_2 - T_1) - (T_4 - T_3)}{2} \\
\alpha &= \frac{(T_2 - T_1) + (T_4 - T_3)}{2}
\end{align*}
\]

The one-dimensional feature vector is obtained when dividing and processing sensor data during the recognition. In order to obtain the three-dimensional space trajectories, the accelerated speed has to be known, and then the displacement requested by two integrates can be calculated, the Eq. (2) is shown as follows:

\[
\begin{align*}
v_x &= \int_{t_0}^{t} a_x dt + v_0 \\
S_x &= \int_{t_0}^{t} v_x dt + S_0
\end{align*}
\]

As the acceleration signal acquired is a discrete one, the displacement recursion equation can be derived by determining the curve with the use of the above equation and the method of value calculation:
Whereas, * in the equation represents any one of XYZ, \( \Delta t \) indicates the sampling interval, \( \Delta T = 5\Delta t \).

The entire network synchronization has to go through two stages: level discovery and synchronization. The level discovery phase limits the broadcasting of some nodes, and the area shown in Figure 2 indicates one covered by a single broadcasting, \( S \) indicates the central node, \( R \) indicates the broadcasting distance, and \( d \) indicates the limited distance of graded broadcasting. \( S \) makes the graded broadcasting. The nodes in the shaded area are restricted by graded broadcasting, i.e. the node with the distance less than \( d \) from the central one, is restricted with graded broadcasting, and the area in the shape of circle ring serves as the broadcasting nodes of next grade. If the shaded area is \( \pi d^2 \), then the circle ring is \( \pi R^2 - \pi d^2 \), and the entire sensor network area is \( a^2 \), with the total nodes number \( N \).

\[
S_n(i + 1) = S_n(i) - \left( \sum_{k=i}^{i+4} a_n(k) \Delta t \right) \Delta T + \left( \sum_{k=i+5}^{i+9} a_n(k) \Delta t \right) \Delta T
\]

Fig. 2 Schematic Diagram for Broadcasting Area

Assuming that the nodes are evenly distributed, the node density is defined as the number of nodes per unit area, which is represented as \( \rho \), that is:

\[
\rho = \frac{N}{a^2}
\]

The nodes number in the shaded area is \( n_d \):

\[
n_d = \frac{\pi d^2 N}{a^2} = \pi d^2 \rho
\]

The nodes number in the circle ring area (number of nodes in the cluster) is recorded as \( n \), then:

\[
n = \frac{\pi (R^2 - d^2) N}{a^2} = \pi (R^2 - d^2) \rho
\]

The broadcasting information packet number required in this phase \( n(d, \rho) \) equals the nodes number. Namely:

\[
n(d, \rho) = n = \pi (R^2 - d^2) \rho
\]

If the node density is \( \rho \), it will keep \( p-1 \) response message. Traditionally, the required number of broadcasting information packet is \( 2N \), and that required by this algorithm is twice of the level discovery stage, that is \( 2n(d, \rho) \) and \( n_0(d, \rho) \) refer to the broadcasting information packet number for the isolated node. The broadcasting information packet number required by this algorithm is as follows:
The optimal broadcasting limit distance $d$ refers to the value of $d$ when the number of the broadcasting information packet required during the synchronization is at minimum. Then, the partial derivative of $d$ related to total is obtained and made equal to zero. Namely:

$$\frac{\partial (\text{total})}{\partial d} = -6\pi \rho \cdot d + \frac{\partial [n_0 (d, \rho)]}{\partial d}$$

The energy $E_{Tx}$ consumed when the sensor node sends $K$ bits of information through the distance $d$ is as follows:

$$E_{Tx} (k, d) = kE_{elec} \cdot k\xi_{amp} d^\rho$$

And the energy $E_{Rx}$ consumed for receiving $K$ bit of information is follows:

$$E_{Rx} (k) = kE_{elec}$$

Whereas, $E_{elec}$ indicates the energy consumed by transmission circuit or receiving circuit for processing every single bit of data; $\xi_{amp}$ indicates the energy consumed by the amplifier. Then, the total WSN consumption power $E_{total}$ is as follows:

$$E_{total} = E_{CPU} + E_{Radio} + E_{GPRS} + E_{AD} = \sum P_{CPU} \times T_{CPU} + \sum P_{Radio} \times T_{Radio} + \sum P_{GPRS} \times T_{GPRS} + \sum P_{AD} \times T_{AD}$$

Whereas, $P_{ADi}$ indicates the $i^{th}$ state of element $AD$; $T_{ADi}$ indicates the time under the $i^{th}$ state, and this equation does not take into account the energy consumed during the status switch.

### 3. Solution Algorithm

The popular learning is an effective non-linear dimension reduction method that has developed rapidly in recent years. It assumes that the input data are embedded on the low dimensional popular structure of the high dimensional observation space. Its aim is to search for the low dimensional popular structure hidden in the high dimension observation data. The popular learning algorithm based on the conformal mapping is added with increment sample points, and the covariance matrix of the high dimensional sample data, focused on average value of samples calculated in local space, is modified according to the average value of the adjacent, or eliminated with the dimension reduction effect results that are hard to get when the isometric mapping approximates the true mapping.

Assumed that $C_n$ is the $n$ samples covariance matrix, after adding $m$ sample points $(x_{n+1}, x_{n+2}, \ldots, x_{n+m})$, the recursion equation of the sample covariance matrix $C_{n+m}$ is updated as follows:

$$C_{n+m} = \frac{n}{n+m} C_n + \sum_{i=1}^{m} \frac{1}{n+i-1} W_{n+i}$$

Whereas, $W_{n+i} = (x_{n+i} - \bar{x}_{n+i})(x_{n+i} - \bar{x}_{n+i})^T$, $\bar{x}_{n+i}$ is the mean value of $x_1, x_2, \ldots, x_{n+i}$, and $i=1, 2, \ldots, m$. For each point $x_{n+i}$ in the increment subspace represented by $W$ matrix, $k$ is the subscript for the adjacent point of its $K$. Assuming that the adjacent $X_k$ of $x_{n+i}$ is small enough to meet better the local sub manifold linear approximation. The covariance matrix of the high dimensional sample data centred with average value of samples is turned into one centred with the average.
value of adjacent, and embedding of maximum variance in the local space block is searched for. Assumed that $\bar{x}_j$ is the average vector of $\{x_j : x_j \in X_k\}$, the target function on such block is as follows:

$$\max_{x_j \in X_k} \sum_{i=1}^{n} \|x_{n+1} - \bar{x}_j\|^2$$

(14)

For the local blocks corresponding to all points, the target function should be sum of such target function maximized on each block:

$$\sum_{i=1}^{m} \sum_{x_j \in X_k} \|x_{n+1} - \bar{x}_j\|^2 = \frac{1}{2k} \sum_{i=n+1}^{n+m} \sum_{j=1}^{k} \text{tr}\left[(x_i - \bar{x}_j)(x_i - \bar{x}_j)^T\right]$$

(15)

The weighting factor is defined as the local average distance of each point, namely the average distance between such point and all the adjacent points of $K (x_j \in X_k)$:

$$S(i) = \frac{1}{k} \sum_{j=1}^{k} \|x_i - x_j\|$$

(16)

The projection under a set of basis $Q \in \mathbb{R}^{D \times d}$ in the low dimensional space is as follows:

$$Y = Q^T X, Y = [y_1, y_2, \ldots, y_n], Y \in \mathbb{R}^d, d \in D$$

(17)

The covariance matrix is shown as follows:

$$\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x}_j)(x_i - \bar{x}_j)^T = \frac{1}{n} \hat{X}_n^T \hat{X}_n$$

(18)

A $G$ matrix is defined for each point $x_k$ in the proximity, the column vector is the adjacent centralization point $x_i$:

$$G_i = x_k - \bar{x}_i, \bar{x}_i = \text{Ave}\{x_k : k \in N_i\}$$

(19)

Provided that the basis $Q$ in the low dimensional space is orthonormal, the best model can be obtained in the following way:

$$\max \text{tr}\left(Q^T D Q\right), \text{s.t.} Q^T Q = I$$

(20)

Whereas, $D = \frac{1}{s(i)^2} \left(G G^T \left(I - \frac{1}{n} e e^T\right)\right)^2$.

Based on the above approach, the algorithm, according to the popular learning, sets out the following steps:

**Step 1** Enter the high-dimensional sample points and the new sample points;

**Step 2** Initialize the sample, centralize the sample: $\hat{X}_n = X \left(I - \frac{1}{n} e e^T\right)$, obtain the covariance matrix after the centralization: $C_n = \frac{1}{n} \hat{X}_n \hat{X}_n^T$;

**Step 3** Calculate the updated covariance matrix $C_n + m$ from Eq. (13) after adding $m$ samples, perform Eigen decomposition, and obtain the feature vector corresponding to the eigenvalue of the first $d$ maximum, and add new sample points into the updated feature space;
Step 4 Write \( D \) into the matrix form \( Q = [q_1, q_2, ..., q_d] \) based on Step 3;
Step 5 Determine the low-dimensional output coordinates according to equation \( Y = Q^T \);
Step 6 For each new data point \( x_{\text{new}} \), seek the feature vector \( V \) corresponding to the eigenvalue of maximum of first \( d \) for \( 1/(2MM^T) \);
Step 7 Calculate the low-dimensional output coordinates according to the function \( U = VTX_{\text{new}} \);
Step 8 The end.

4. SIMULATION EXPERIMENT

In order to validate the effectiveness of the above methods, this paper uses the OPNET and the MATLAB for simulation experiment. The area of WSN is set to \( 60\times60 \text{ m}^2 \) in OPNET, with the total nodes number \( N \), evenly distributed in this area, with the broadcasting distance \( R \) of 10 m. According to the definition, the node density \( \rho \) can be changed throughout the total network node number \( N \). This paper adopts the step-wise approximation to obtain the optimum value of \( d \) with the minimum broadcasting information packet required by synchronizing all the nodes under a certain network node density \( \rho \); and then changes the network node density \( \rho \) by changing the value of \( N \), thus the optimum broadcasting limit distance value of \( d \) is determined under different network node density \( \rho \). Figure 3 below represents the relation schema between the node density \( \rho \) and the optimum broadcasting limit distance \( d \). As it can be read from the figure, with the increase in the node density \( \rho \), the optimum value of \( d \) will gradually increase, and eventually get closer to the broadcasting distance \( R \).

![Fig. 3 Curve Relation between the Node Density \( \rho \) and the Optimum Broadcasting Limit Distance \( d \)](image-url)
Secondly, Figure 4 illustrates the relationship between the total dump energy and the node density $\rho$ for the algorithm proposed by this paper and TPSN (Timing-sync Protocol for Sensor Networks) algorithm upon the completion of one time synchronization [28]. The figure shows that with the increase of density, the energy consumption of TPSN algorithm exhibits a fast decreasing trend, while the remaining energy of the algorithm presented in this paper exhibits a stable decreasing trend, which can save energy better and extend greatly the life of wireless sensor network. Manifold learning can highlight some specific characteristics of energy consumption data between nodes and effectively reduce dimensionality, therefor it performs better in maintaining remaining energy.

In addition, the sampling interval is, to some extent, closely related to the energy consumption. The acceleration sensor embedded within the sensing node can sample and buffer the data independently. The buffer sampling mode reduces the idle waiting time by optimizing the data acquisition process, thus saving the energy effectively, while the consecutive sampling mode can cause greater energy consumption due to extensive unnecessary idle waiting time. Figure 5 and Figure 6 depict the comparison chart for time sequence of consecutive and buffer sampling mode.
At the same time, Figure 7 shows the variation trend of synchronization precision and algorithm network grade presented in this paper and the TPSN under the condition of optimal broadcasting limited distance $d$. As it can be seen from Figure 7, the algorithm presented in this paper is superior to the TPSN in terms of synchronization precision. This is due to the combination of one-way broadcast and two-way exchange synchronization adopted at the synchronization stage, and only one response node selected in TPSN algorithm, while other nodes are synchronized by the use of time deviation and transmission delay between such node and base station, resulting in precision reduction. However, the incremental sample points are added in the manifold learning algorithm, and the covariance matrix of high-dimensional sample data with the sample mean as the centre is calculated in the local space, bringing isometric mapping close to real mapping, and the difficulty of obtaining real dimension reduction results is eliminated with the adjacent taken as the centre. Moreover, the nearest node is selected as the response in the manifold learning algorithm, while random selection is used for TPSN algorithm. As the transmission delay between root node and response node is reduced by manifold learning algorithm, it is more accurate than TPSN algorithm.
In addition, in order to describe the impact of $N_{\text{nodes}}$ and $N_{\text{cluster}}$ on the network energy consumption in the linear segment formed by 36 nodes, Figure 8 the simulated chart is used. As the node that is the closest to Sink node receives and forwards the most amount of data, with the longest working hours and the largest energy consumption, the simulation data analyse the energy consumption of the nearest node under different clusters, which is also the worst in energy consumption in the entire network segment. As it can be read from Figure 6 when $N_{\text{nodes}} = 6$, $N_{\text{cluster}} = 6$ have the better effectiveness under the condition the number of the network segment is 36. When $N_{\text{nodes}} = 4$, the number of nodes within the cluster is little, with low fusion depth, and the proportion of the valid data volume in the transmission data is small. When the data size transferred in the network is large, then the energy consumption increases accordingly; the maximum data value transferred by node is smaller than one when $N_{\text{nodes}} = 6$, which has little impact on the transmission time and the deviation of about 1 ms. When $N_{\text{nodes}} = 8$, the length of data after fusion is relatively large, the working time slot of the node is longer, resulting in a larger amount of energy consumption from the nearest node, while the transferred data size deviates slightly from that of $N_{\text{nodes}} = 6$.

Then, the different cluster heads are deployed in the simulation system, with the communication radius of the cluster head set to 50 m, and the communication radius of the sensor node 25 m. Figure 9 illustrates the impact of number of cluster head $\text{NH}$ on the network lifecycle. It can be seen from Figure 9 that with the increase in cluster head $\text{NH}$, the network lifecycle increases accordingly. When $\text{NH}$ is less than or equals to 20, the network lifecycle will accordingly increase significantly; when $\text{NH}$ reaches 50 or so, the network lifecycle area is stable. Therefore, the network lifecycle may not be extended by merely increasing the number of the cluster heads.

![Fig. 8 Curve Chart for Low Energy Consumption Uplink Protocol Strategy](image-url)
For the implementation of popular learning algorithms in large-scale networks, a certain number of cluster heads should be deployed in the network, then the network can be divided into smaller clusters. The number of cluster heads will directly affect the nodes energy consumption in data collection. Suppose that the $N_c$ number of cluster heads and 100 sensor nodes are deployed in the network. The communication radius of the cluster head is 50 meters, and that of the sensor node is 25 meters. The influence of cluster head number on the network life cycle is shown in Figure 10. Figure 5 clarifies that with the increase of cluster head $N_c$ and so on, the network life cycle will increase. When $N_c$ is less than or equal to 20, the network life cycle will increase significantly. When $N_c$ is stabilized at about 50, the network life cycle reach its maximum. Therefore, only the increase of the cluster number does not make the network life cycle longer.

In order to verify the recognition rate of the system, in each collected sensor data 30 samples are recorded, and 10 samples are randomly selected from it for training. By writing automatic recognition rate test function, the average value of sensing data recognition rate of the corresponding training sample number is obtained. As it can be seen from Figure 11, when the number of training samples is about 25, the recognition rate of the system is up to 94%, being the highest. If the training sample continues to increase, the recognition rate of the system changes little, so the method of segmentation and processing the collected sensor data to obtain one dimensional feature vector is no longer noticeable at this time.
Dynamic clustering strategy can guarantee network connectivity, but it has the disadvantage of high control cost. The wireless sensor energy consumption model based on popular learning algorithm adopts approximate static clustering strategy, which can balance network overhead and network connectivity. As shown in Figure 12, this algorithm has a better network life cycle than TPSN. The main reason is that it adopts the approximate static clustering strategy, collects data according to the required point, and that each node holds the cluster head in turn. These strategies make the energy consumption balanced. Therefore, this algorithm is superior to the TPSN algorithm.

5. CONCLUSIONS

In order to address the problem in conventional energy consumption model for WSN in low characterization accuracy, this paper proposes a new model based on the popular learning algorithms. The model sets out the composition and calculation equation of energy consumption in WSN, and establishes the target optimization function. Meanwhile, it also solves the computing methods established by leveraging the popular learning. Finally, OPNET and MATLAB are adopted for simulation experiment enabling in-depth study on the key factors affecting this model. The experimental results demonstrate that the model proposed by
this paper has greater advantage over other methods in synchronization accuracy, dump energy, etc. In the follow-up study, the factors like signal interference, energy attenuation, etc. will be considered to further improve the energy consumption model for WSN.

6. ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China, Zhejiang Public Services Projects Fund (2016C31129, 2015 C33236), the Scientific Research Project of Education Department in Zhejiang Province (Y201432666), the Academic Climbing Project for the Young and Middle-aged Subject Leaders in the Colleges and Universities in Zhejiang Province (pd2013443) and Ningbo National Natural Science Foundation Project (2012A610071).

7. REFERENCES


