On Application of Advanced Algebra Matrix Problem to Sensors

Yu Wang, Niqin Jing and Xin Dai
1 Beijing Polytechnic, Beijing, 100015, P. R. China
2 Information Engineering Department, Wuhan Business University, Wuhan, 430056, P. R. China
E-mail: wyypkm@163.com

Received: 2 November 2013 /Accepted: 22 November 2013 /Published: 30 December 2013

Abstract: In order to reduce energy consumption of the sensor node communication and maximize the life cycle of wireless sensor network, this paper follows the principle of high-dimensional thinking represented by low-dimensional space, puts forward a wireless sensor network energy saving algorithm based on generalized inverse nonnegative matrix decomposition. First of all, it applies generalized inverse decomposition technique to perform initialization operation on the original communication data matrix. Then, non-negative matrix factorization is applied to make dimensionality reduction on characteristic space after the singular value decomposition. The simulation results show that: the algorithm can effectively compress communication data in WSN, solve the high dimension, high redundancy phenomenon of communication data, achieve the energy saving purpose for data transmission. Copyright © 2013 IFSA.

Keywords: Matrix, Generalized inverse decomposition, Wireless sensor network, Energy saving.

1. Introduction

Wireless sensor network is composed of the random distributed integrated sensors, data processing units and the tiny nodes of communication units through a self-organization way. It has the advantages of cheap, self-organizing, dynamic and high fault tolerance, etc. And it is widely used in the field of car business, health care, environmental monitoring, disaster mitigation and relief and so on [1].

WSN sensor node energy is usually supplied by batteries and it is hard to realize timely supplement. Therefore, how to make efficient use of energy to maximize the network life cycle is a major challenge for the sensor network [2]. WSN energy consumption is mainly used in perception, communication and data processing. Studies show that energy consumption of sensor nodes is mainly concentrated in communication and it accounts for more than 80% of the whole sensor life cycle energy. Therefore, effectively reducing the sensor node communication energy consumption is the key to maximize WSN life cycle. Energy consumption of communication is mainly occurred in three kinds of conditions as sensor receiving, sending and sensor free, and there are mainly 3 kinds of methods to reduce energy consumption of communication.

The first method is to reduce energy consumption of communication by improvements on WSN communication protocol [5], and extend the life cycle of WSN. However, this kind of algorithm is lack of generalization ability and is limited to specific applications. Another method is to reduce energy consumption of communication according to the features of sensor environment and apply partially...
Based on the advantages and disadvantages of the present algorithm, this paper, based on generalized inverse non-negative matrix factorization (giNMF), put forward an energy efficient communication algorithm of wireless sensor network. Before using nonnegative matrix factorization (NMF) to make the dimension reduction, adopt the generalized inver
decomposition technique (this paper adopts the singular value decomposition (SVD)) to initialize the original communication data matrix to avoid matrix decomposition uncertainty problem caused by the NMF random initialization. Then, use the NMF to perform dimensionality reduction on characteristic space after the singular value decomposition.

### 2. Non-negative Matrix Factorization

In the actual application environment, sensors tend to transmit information with high-dimensional and sparse features which requires large amount of sensor energy consumption. By the theory of cognitive psychology, we know that the brain only processes the "meaningful" data of the outside world and stays "blind" for other data, i.e., the high dimensional information space can often be represented by low dimensional space to nature. Therefore, establish the map of "meaningful" data from high-dimensional to low-dimensional of the sensor, and thus the sensor will only transmit data after dimension reduction, the energy consumption of communication will be greatly reduced.

Nonnegative matrix decomposition method is to decompose the original high dimensional matrix into the product of two low dimensional nonnegative matrix. Its principle is shown in Fig. 1.

![Fig. 1. Nonnegative matrix decomposition principle diagram.](image)

- **Eq. (1)**
  \[
  \Omega = \sum_{i=1}^{m} \sum_{j=1}^{n} \left( x_{i,j} \log \frac{x_{i,j}}{y_{i,j}} - x_{i,j} + y_{i,j} \right)
  \]

  In the above equation, \( Y = [y_{i,j}] = U \cdot V^T \). If \( X = Y \), then \( D(U \parallel V) = 0 \).

  Because there is only one salient characteristic in \( U \) and \( V \) matrix, therefore, it is difficult to design an algorithm to minimize the global objective function. Lee and Seung [10] proved that Eq. (2) and Eq. (3) can achieve the local minimum of the objective function.

  - **Eq. (2)**
    \[
    u_{i,k} \leftarrow u_{i,k} \frac{\sum_{j=1}^{n} x_{i,j} \cdot v_{j,k}}{\sum_{j=1}^{n} v_{j,k}}
    \]

  - **Eq. (3)**
    \[
    v_{j,k} \leftarrow v_{j,k} \frac{\sum_{i=1}^{m} u_{i,k} \cdot x_{i,j}}{\sum_{i=1}^{m} u_{i,k}}
    \]

  Nonnegative matrix decomposition is an important way to deal with high dimensional data set and it is widely used in pattern recognition, signal processing, data clustering and computer security [4], etc. However, there are also some disadvantages:

  1. For large-scale data, calculation of the NMF is very time-consuming which results in that it cannot be used in real-time systems [7].
  2. Dimensionality of characteristic space \( r \) is hard to predict, so it is difficult to find an optimal \( r \).
  3. Unable to solve the global optimal. Initialization of factor \( U \) and \( K \) is random, which brings uncertainty of each decomposition result.

### 3. Communication Data Compression

#### Algorithm of Generalized Inverse Nonnegative Matrix Decomposition

Based on the existing problems in the nonnegative matrix decomposition, using the fast conversion features of generalized inverse, this paper...
put forward a nonnegative matrix decomposition algorithm based on generalized inverse, and the algorithm is applied to wireless sensor network transmission information compression, to achieve energy efficient communication.

3.1. Building of Generalized Inverse Matrix

Definition 1 [11], for finite dimensional matrix X, Penrose defines the existing of a unique generalized inverse matrix \( X^{-1} \), at the same time satisfy the following:

\[
XX^{-1}X = X \quad (4)
\]
\[
X^{-1}XX^{-1} = X^{-1} \quad (5)
\]
\[
(X^{-1})^T = XX^{-1} \quad (6)
\]
\[
(X^{-1}X)^{\ast} = X^{-1}X \quad (7)
\]

Among which, \( X^{\ast} \) is an associate matrix of \( X \).

Definition 2, make two orthogonal matrix \( U \in \mathbb{R}^{m \times n} \) and \( V \in \mathbb{R}^{n \times r} \), \( X \in \mathbb{R}^{m \times n} \). Singular value decomposition is as follows:

\[
X = U \Sigma V^T \quad (8)
\]

Among which, \( \Sigma \in \mathbb{R}^{m \times n} \) is a diagonal matrix:

\[
\Sigma = \begin{bmatrix}
\sigma_1 & & \\
& \ddots & \\
& & \sigma_r \\
0 & & \\
& & \\
0 & & 
\end{bmatrix} \quad (9)
\]

Among which, \( \sigma_1 \geq \cdots \geq \sigma_r \geq 0 \), \( \sigma_i \) is singular value of matrix \( X \). Simplify Eq. (9) as:

\[
\Sigma = \text{diag}(\sigma_1, \cdots, \sigma_r, 0, \cdots, 0) \in \mathbb{R}^{m \times n}, \quad \text{according}
\]

to definition 1 and 2, building of generalized inverse matrix is as follows:

\[
X^{-1} = V \Sigma^{-1} U^T \quad (10)
\]

Among which,

\[
\Sigma^{-1} = \text{diag}(\sigma_1^{-1}, \cdots, \sigma_r^{-1}, 0, \cdots, 0) \in \mathbb{R}^{n \times m}
\]

3.2. Characteristic Space Transformation

According to the principle of NMF decomposition, \( X_{m \times n} = U_{m \times r} \cdot V_{r \times n} \), of which, \( U \) is a weight matrix, \( V \) is a basis matrix, characteristic space transformation process is shown in (7).

\[
[x_1, x_2, \cdots, x_m]^T = [u_1, u_2, \cdots, u_r]^T \cdot V \quad (11)
\]

Multiply both sides of Eq. (11) by the generalized inverse matrix \( V^{-1} \) of \( V \):

\[
[x_1, x_2, \cdots, x_m]^T V^{-1} = [u_1, u_2, \cdots, u_r]^T \cdot V V^{-1} = [u_1, u_2, \cdots, u_r]^T \quad (12)
\]

If each line of \( V \) is linearly independent, then

\[
VV^{-1} = I, \quad I \text{ is the unit matrix, Eq. (12) can be transformed into:}
\]

\[
[x_1, x_2, \cdots, x_m]^T V^{-1} = [u_1, u_2, \cdots, u_r]^T \quad (13)
\]

Through the above transformation steps, original input matrix can be fastly transformed into weighted feature space.

3.3. Communication Data Compression

Algorithm of Generalized Inverse Nonnegative Matrix Decomposition

Algorithm thought: first make generalized inverse decomposition on communication data matrix \( X \), initialize \( X \) and quickly get its characteristic space. Then, according to the error constraints or time constraints, perform NMF dimensionality reduction on \( X \) after the generalized inverse decomposition, so as to realize the effective and quick dimension reduction of the original data communication.

Definition 3, in KL divergence function, the error needs to be estimated between \( X \) and \( UV \). Assuming that each iteration leads to an \( \epsilon \) error, which increases the distance between \( X \) and \( UV \), then accumulated error until \( n \) steps is:

\[
\epsilon + (1-\gamma)\epsilon + \cdots + (1-\gamma)^{n-1}\epsilon \leq \gamma \sum (1-\gamma)^i = \frac{\epsilon}{\gamma} \quad (14)
\]

Algorithm 1: Communication data compression algorithm of generalized inverse nonnegative matrix decomposition.

Input: original communication data matrix \( X \in \mathbb{R}^{m \times n} \)

Output: \( U \in \mathbb{R}^{m \times r} \) and \( V \in \mathbb{R}^{n \times r} \)

//Singular value decomposition operation

Step 1: apply the generalized inverse singular value decomposition method to decompose \( X \).

Step 2: select the larger singular values \( \sigma_i \) of \( X \) and its corresponding feature space \( u_i \).

// NMF dimension reduction operation
U = \arg\min_{U \geq 0} \{ D(U \| UV) \}

Step 3: carry out \( U \) = \arg\min_{U \geq 0} \{ D(U \| UV) \}

Step 4: carry out \( V = \arg\min_{V \geq 0} \{ D(X \| UV) \} \)

Step 5: when the error constraints or time constraints is met, calculation termination. Otherwise, looping execution of Step 3 and Step 4.

Step 6: output U and V after the dimension reduction.

Due to the random initialization of U and V in NMF dimension reduction algorithm, each decomposition result is with uncertainty. Sometimes incorrect or unrelated decomposition results are obtained. Time complexity of NMF algorithm is \( O(MN\cdot r) \), space complexity is \( O(MN) \). Whereas giNMF algorithm adopts generalized inverse decomposition (Step 1 and Step 2), and the characteristic space of the original matrix is quickly get, and iteration factor decomposition is avoided. Step 3 and Step 4 calculate the minimal value U and V until circulation condition is met. These two steps are NMF dimension reduction algorithm, because of generalized inverse decomposition, the size of the NMF dimension reduction is reduced greatly, the final dimension reduction results can be obtained by multiplicative update rule. The time complexity of giNMF algorithm is \( O(Nr^2) \), space complexity is \( O(Nr) \).

giNMF algorithm has the characteristic of communication mode independence. Whether it is a single hop or multi-hop, there is no influence on the algorithm. Therefore, giNMF algorithm has strong generalization ability for communication data dimension reduction, and it can be applied to any mode, any node communication data transmission.

4. Simulation Experiment

The subject of this paper is the information of environment temperature and humidity which is collected, sent and received by the sensor. We adopt Crossbow MICA2 Mote sensor and topology model is of single hop. In order to reduce energy consumption, the sensor often adopts concentrating sending mode of multi-information collection, its working principle is shown in Fig. 2.

Transmit N bags with M-bit length, the total communication energy consumption is the sum of energy consumption of data collection, data compression, data transmission and data receiving stage.

\[
E_{\text{total}} = E_a + E_c + E_s + E_r
\]

(15)

According to the literature [4], energy consumption \( E_a \) in data collection stage and \( E_c \) in data compression stage are far less than that of data transmission and receiving. Therefore, this paper ignored \( E_s \) and \( E_r \). considered only \( E_s \) and \( E_r \), the calculation is as follows:

\[
E_a = N \cdot M \cdot 10^{-6}
\]

(16)

\[
E_r = \begin{cases} 
N \cdot M \cdot 10^{-6} + N \cdot M \cdot d^2 \cdot 22.0 \cdot 10^{-12}, & d \leq 90m \\
N \cdot M \cdot 10^{-6} + N \cdot M \cdot d^4 \cdot 16.0 \cdot 10^{-10}, & d > 90m 
\end{cases}
\]

(17)

Among which, \( d \) is the distance between node sensor and cluster head sensor.

This paper made experimental analysis respectively from three angles:

Comparative analysis with algorithm in literature [9], compare dimension reduction effect and time consuming of giNMF algorithm;
Contrastive analysis on energy consumption of communication between giNMF and LEACH algorithm;
Contrastive analysis on network life cycle between giNMF and LEACH algorithm.

4.1. Dimension Reduction Effect Analysis

Dimension reduction effect comparison is shown in Fig. 3. According to Fig. 3, giNMF algorithm adopted by this paper is more efficient than lraNMF algorithm. For convergence rate of KL divergence, giNMF algorithm is faster. For the same KL divergence, dimensionality obtained by giNMF algorithm is the minimum. And for same dimension reduction, minimum KL divergence means minimum error.

![Fig. 3. Dimension reduction effect comparison.](image)

4.2. Communication Energy Consumption Analysis

Assuming that receiving and sending unit (bit) energy consumption for 50 nj/bit, transmit data matrix $M=2000$ bits, $N=50$. As shown in Fig. 4, the communication energy consumption contrast experiments between giNMF algorithm and LEACH algorithm. From Fig. 4, we learn that energy consumption of giNMF algorithm is less than that of LEACH algorithm. The average energy consumption of communication of LEACH algorithm is above 2 times of that of giNMF algorithm. It can be concluded that the giNMF dimension reduction algorithm can effectively reduce the communication data size, remove the white noise in communication data, achieve the goal of energy saving.

![Fig. 4. Communication energy consumption contrast experiments.](image)

4.3. Network Life Cycle Analysis

The simulation experimental environment is a 100 m×100 m area, which monitors temperature and humidity information of the environment. The number of sensor nodes is $N=100$, with one hop communication between sensor nodes and cluster heads. White noise of sensor nodes obeys the Gaussian distribution $N(0, 2.25)$, the initial value of each sensor node is 0.5J, compared with LEACH algorithm in literature [5]. One of the important indicators for evaluation of wireless sensor network (WSN) temperature and humidity monitoring system is the system life cycle. The longer the life cycle, the better the system. From Fig. 5, we know that system life cycle obtained by giNMF algorithm is about 50 % higher compared with that of LEACH algorithm.

As shown in Fig. 5, the network lifetime comparison between giNMF algorithm and LEACH algorithm. It shows the change of live nodes in the network. With the increase of number of death nodes, node death rate presents a speeding up trend. Dead nodes first appeared in LEACH algorithm, and the first dead node appeared in giNMF is at about 1000 rounds. LEACH algorithm at about 1300 rounds, all network nodes are dead while GiNMF algorithm at around 1800 rounds, all nodes are dead. The algorithm adopted in this paper can obtain the longest network life cycle. Comparison results show that: algorithm adopted in this paper can effectively prolong network lifetime.

![Fig. 5. Contrast experiment of network life cycle.](image)
5. Conclusions

This paper put forward a wireless sensor network communication energy saving algorithm based on generalized inverse nonnegative matrix decomposition which can effectively reduce the amount of communication data and realize energy consumption reduction. The characteristics of the algorithm are: it has strong generalization ability and communication protocol independence, as a result, it has wide application range. Secondary, it can achieve accurate high-dimensional data compression with low error, and compressed data can describe the real original state and it has the quality of distortionless of dimensionality reduction. Moreover, the simulation results show that GiNMF algorithm can effectively compress the communication data, so as to realize the purpose of reducing communication energy consumption and extending network lifetime.

References


