CHEMICAL ENGINEERING TRANSACTIONS

VOL. 59, 2017

A publication of AIDIC

The Italian Association of Chemical Engineering
Online at www.aidic.it/cet

Guest Editors: Zhuo Yang, Junjie Ba, Jing Pan

ISBN 978-88-95608- 49-5; **ISSN** 2283-9216

DOI: 10.3303/CET1759095

Warning System of Dangerous Chemical Gas in Factory Based on Wireless Sensor Network

Manzeng Ma, Bo Zhang

School of Computer Science & Engineering, Cangzhou Normal University, Cangzhou 061001, China
9527a@163.com

Wireless sensor network (WSN) is a dynamic network topology which is formed in a self-organizing way. In the field of WSN, the perception of information is realized by cooperative work between sensors. The wireless sensor network has the characteristics of simple structure, low cost and high reliability. Usually, most of the energy of sensor nodes is consumed in data transmission. So, reducing the data transmission between sensor nodes becomes an effective way to reduce the energy consumption of sensor networks. Data fusion technology can effectively reduce the amount of data transmission in the sensor network, which can reduce the energy consumption and get more accurate information. In this paper, we propose a hybrid data fusion algorithm based on adaptive weighting algorithm and Calman filter algorithm. In the hybrid algorithm, the Pearson similarity algorithm is used to filter the sensor nodes which do not meet the requirements, and then the average values of the two algorithms are obtained. Firstly, this paper introduces the principle of adaptive weighting algorithm and Calman filter algorithm. Secondly, we propose the improved hybrid algorithm by extracting some nodes with low similarity, and then we get the result by weighted calculation. Finally, the experimental results show that the proposed algorithm has higher accuracy and lower energy consumption.

1. Introduction

Data fusion technology is widely used, and it can be used to create a complete view based on the fragmented information collected from each node. Through data fusion, the sensor network can avoid redundant information transmission and greatly reduce the network communication load.

Many scientists regard data fusion technology as one of their main research directions and they put forward their own solutions through theoretical analysis and experiments. Lin et al., (2008) propose the CHEP method, which can balance the energy load effectively. Xiang et al., (2010) proposes a clustering algorithm by optimizing the position information and the residual energy. Li et al., (2012) proposes a LEACH-New algorithm, which effectively reduces the network energy consumption and ensures the network load balance. Zhang et al., (2001) proposes an optimized Bayesian data fusion method which effectively solves the problem of uncertainty and inconsistency. Qiu et al., (2014) proposes the SAEMDA algorithm based on deep learning and it improves the performance of data fusion in the field of WSN. Yanger et al., (2001) abandons probabilistic method by calculating the support function of each sensor data. Ding Lei proposes SVR algorithm and the experimental results show that the method is superior to the data fusion method of artificial neural network (Ding et al., 2011).

In this paper, we propose a hybrid data fusion algorithm based on adaptive weighting algorithm and Calman filter algorithm. In the hybrid algorithm, the Pearson similarity algorithm is used to filter the sensor nodes, and then the average values of the two algorithms are obtained. Firstly, this paper introduces the principle of adaptive weighting algorithm and Calman filter algorithm. Secondly, the Pearson similarity algorithm is used to extract some nodes with high similarity, and then the average values of the two algorithms are obtained. Finally, the experimental results show that the proposed algorithm has higher accuracy and lower energy consumption.

Please cite this article as: Manzeng Ma, Bo Zhang, 2017, Warning system of dangerous chemical gas in factory based on wireless sensor network, Chemical Engineering Transactions, 59, 565-570 DOI:10.3303/CET1759095

2. Basic theory and method

2.1Adaptive weighting algorithm

Adaptive weighting algorithm is a flexible method in the field of data fusion. In this algorithm, the monitoring values collected by each sensor node are used to determine the weight coefficient, so that the fused result will be close to the true value. Compared with other traditional data fusion algorithms, we can get a better fusion value, which can provide more accurate data for advanced fusion (Qiu, 2012)).

Assuming that in a wireless sensor area, m is the total number of sensor nodes, n is the number of objects which need to be monitored, and t_{ij} is the monitoring information of the object j from sensor i. Where, i=1, 2, 3…m, j=1, 2, 3…n. The matrix T_{ij} can be expressed as:

$$T_{ij} = \begin{pmatrix} t_{11} & \cdots t_{1j} \cdots & t_{1n} \\ \vdots & & \vdots \\ t_{i1} & \ddots & t_{in} \\ \vdots & & \vdots \\ t_{m1} & \cdots t_{mj} \cdots & t_{mn} \end{pmatrix} \tag{1}$$

Then, the final monitoring information of the sensor network to the object j is calculated as follows:

$$t_j = \sum W_i T_{ij} = \sum_{i=1}^{m} w_i t_{ij} \tag{2}$$

In the above formula, w_i represents the weight of a sensor node, and the weights of all nodes satisfy the following formula:

$$\sum W_i = \sum_{i=1}^{m} w_i = 1 \tag{3}$$

The measurement variance of each sensor node is defined as σ_1, σ_2, σ_3…, σ_m, and the total variance is calculated as follows:

$$\sigma = E[(T - \bar{T})^2] = \sum_{i=1}^{m} w_i^2 \sigma_i^2 \tag{4}$$

Through the formula, we can get that the total mean square error σ is a quadratic function. According to the principle of calculating minimum value of variance function, we can get the partial derivative of the above formula.

$$\begin{cases} \dfrac{\partial \sigma}{\partial w_i} = 0 \\ \sum_{i=1}^{m} w_i = 1 \end{cases} \quad (i = 1,2,3\cdots n) \tag{5}$$

The weight under the condition of minimum mean square error is as follows:

$$w_i = \frac{1}{\sigma_i^2 \sum_{l=1}^{m} \frac{1}{\sigma_l^2}} \quad i = 1,2,3\cdots n \tag{6}$$

Finally, we get the minimum mean square error by the following formula.

$$\sigma_{min}^2 = \frac{1}{\sum_{l=1}^{m} \frac{1}{\sigma_l^2}} \tag{7}$$

The advantage of adaptive weighted data fusion algorithm is that it does not need the prior probability distribution of the data and it can quickly determine the weights of each sensor according to their variance. What is more, it has a strong fault tolerance and anti-interference ability

2.2 Calman filtering algorithm

Calman filtering algorithm is a method of valuation based on minimum variance. This method is suitable for the redundant data processing under the condition that the environment around the sensor node changes at any time. The optimal estimate of the present time is obtained based on the optimal estimate of the system state at the last moment and the observed value of the present time (Liu, 2014).
The discrete state equation of linear time-varying system is presented:

$$S(t) = F \cdot S(t-1) + T \cdot U(t) + W(t) \tag{8}$$

In the above formula, S(t) and W(t) represent respectively the state vector and measurement process of moment *t*. *F* represents the state transition matrix, *U*(*t*) represents the dynamic noise of moment *t* and *T* represents the system control matrix.
The corresponding observation equation is:

$$Y(t) = H \cdot X(t) + N(t) \tag{9}$$

Where, *Y*(*t*) represents the observation vector, *N*(*t*) represents the and the observation noise and *H* represents the observation matrix at moment *t*. We assume that they are consistent with the gauss distribution noise formula, and the covariance is respectively *Q* and *R*. In this paper, we assume that the mean square error does not change with the change of the system state.
Based on the assumption above, we can get the vector estimate of the next moment.

$$S(t + \Delta t \mid t) = F \cdot S(t \mid t) + T \cdot U(t) \tag{10}$$

In the above formula, $S(t+\Delta t|t)$ is the estimated value which is calculated based on the vector value of moment *t*, *U*(*t*) is the state control vector of moments t . When the control vector does not exist, the value of *U*(*t*) will be 0.
Using $C(t+\Delta t|t)$ to represent the covariance of $S(t+\Delta t|t)$.

$$C(t + \Delta t \mid t) = F \cdot C(t \mid t) \cdot F' + Q \tag{11}$$

We use $C(t|t)$ to represent the covariance of $S(t|t)$ and F' to represent the covariance of matrix *F*. So we can get the best estimate value of moment $t+\Delta t$:

$$S(t + \Delta t \mid t + \Delta t) = S(t + \Delta t \mid t) + Kg(t + \Delta t) \cdot (Y(t + \Delta t) - F \cdot S(t + \Delta t \mid t)) \tag{12}$$

Where, $Kg(t+\Delta t)$ represents the Calman gain.

$$Kg(t + \Delta t) = \frac{C(t + \Delta t \mid t) \cdot H'}{H \cdot C(t + \Delta t \mid t) \cdot H' + R} \tag{13}$$

Therefore, we can get the updated covariance:

$$C(t + \Delta t \mid t + \Delta t) = (M - Kg(t) \cdot H) \cdot C(t + \Delta t \mid t) \tag{14}$$

For a single model:

$$M = 1 \tag{15}$$

In Calman filtering algorithm, the noise contained in measurement data is required to obey the Gauss white noise distribution. Otherwise, it will bring greater distortion when they do not obey the Gauss distribution.

3. Improved hybrid data fusion algorithm

The hybrid data fusion algorithm in this paper is based on adaptive weighting algorithm and Calman filtering algorithm. By making Pearson similarity analysis on the data matrix of the two algorithms, we can get the similarity matrix. Then, we extract the nodes with high similarity by setting a threshold. When the similarity of a node is less than the threshold value, the node is abandoned. Conversely, the monitoring information of the corresponding node will be saved for reference. Finally, the effective nodes are re-calculation to get the results. Assuming that W_{ij} is the result matrix of adaptive weighting algorithm, and K_{ij} is the result matrix of Calman filtering algorithm.

$$W_{ij} = \begin{pmatrix} w_{11} & \cdots w_{1j} \cdots & w_{1n} \\ \vdots & & \vdots \\ w_{i1} & \ddots & w_{in} \\ \vdots & & \vdots \\ w_{m1} & \cdots w_{nj} \cdots & w_{mn} \end{pmatrix} \tag{16}$$

$$K_{ij} = \begin{pmatrix} k_{11} & \cdots k_{1j} \cdots & k_{1n} \\ \vdots & & \vdots \\ k_{i1} & \ddots & k_{in} \\ \vdots & & \vdots \\ k_{m1} & \cdots k_{nj} \cdots & k_{mn} \end{pmatrix} \tag{17}$$

According to Pearson similarity algorithm, we can get the similarity of the two algorithms:

$$sim(W,K) = \frac{\sum_{i \in Twk} (W_{ij} - \overline{W})(K_{ij} - \overline{K})}{\sqrt{\sum_{i,j \in Twk} (W_{ij} - \overline{W})^2 (K_{ij} - \overline{K})^2}} \tag{18}$$

Where, T_{wk} is the intersection of W_{ij} and K_{ij}. That is to say, $T_{wk}=W \cap K$.
Assuming that the number of sensor nodes which meet the threshold condition is l, and we can get the estimated value of the latest time.

$$H_{ij} = \frac{W_{ij} + K_{ij}}{2} \tag{19}$$

Therefore, the final result can be get by the following formula:

$$h_j = \sum W_i H_{ij} = \sum_{i=1}^{l} w_i h_{ij} \tag{20}$$

4. Realization of warning system in factory

The early warning system of dangerous chemical gas is composed of three parts. They are data acquisition section, wireless communication sectionand remote monitoring section.The system frame diagram is shown in Figure 1

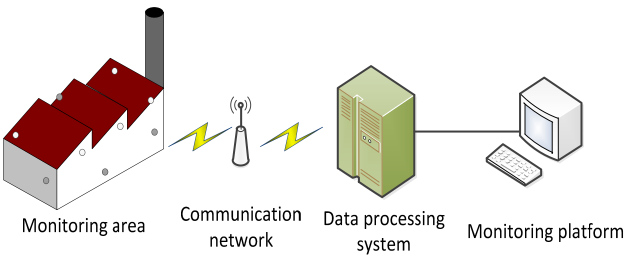

Figure 1: Early warning system of dangerous chemical gas

The data acquisition part is composed of a large number of sensors, which are deployed in every corner of the chemical plant, and they are the sources of monitoring data. The distribution of these sensors can be regular, and the distribution can also be stochastic. The better the performance of the sensor, the more accurate the monitoring data in the later stage.
Wireless communication mainly includes wireless sensor network and network information transmission.In wireless sensor networks, wireless communication is used to transmit between sensor nodes. Wired communication network is used to transmit data between sensor network and monitoring center.
The remote monitoring section receives the signal and sends it to the monitoring platform through network. Three dimensional visualization technology is used to model the factory buildings, which can realize the 3D simulation environment. In addition, the data processing algorithm directly determines the performance of the whole system.

5. Experiment and result analysis

In order to test the performance of the algorithm, in this paper we compare the three algorithms in terms of time consuming, average energy consumption and precision. The experiments were carried out in the region of 1000m*1000m, and the numbers of samples are respectively 80, 160, 280, 440, 640, 880, 1200. Assuming that $\hat{T}_j$ is the actual monitoring value and T_j is the estimated value of the hybrid algorithm, then we can get the accuracy of the monitoring results.

$$\varepsilon_j = \frac{T_j - \hat{T}_j}{\hat{T}_j} \qquad (21)$$

Thus, the monitoring accuracy of the whole sensor network can be get as follows

$$\varepsilon = \frac{T - \hat{T}}{\hat{T}} = \frac{\sum_{j=1}^{n}(T_j - \hat{T})}{\sum_{j=1}^{n}\hat{T}} \qquad (22)$$

As can be seen in the Figure 2, with the increase of sensor nodes, the accuracy of the three algorithms is improved. When the number of nodes reaches a certain threshold value, the change of precision tends to be stable. However, generally speaking, the accuracy of the hybrid algorithm is higher than the first two algorithms

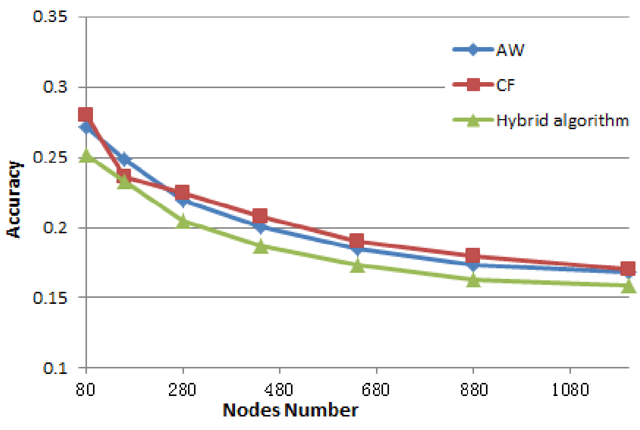

Figure 2: Comparison of accuracy

Figure 3 respectively represents the comparison of energy consumption. With the increase of nodes, the energy consumption of sensor network is gradually improved. As a result of abandoning some nodes with smaller weights, the energy consumption of the hybrid algorithm is lower than the other two algorithms.
Figure 4 respectively represents the comparison of time-consuming. Because the computational complexity of the hybrid algorithm is more complex, it is more time-consuming than other algorithms. Nevertheless, the time is still in the acceptable range.

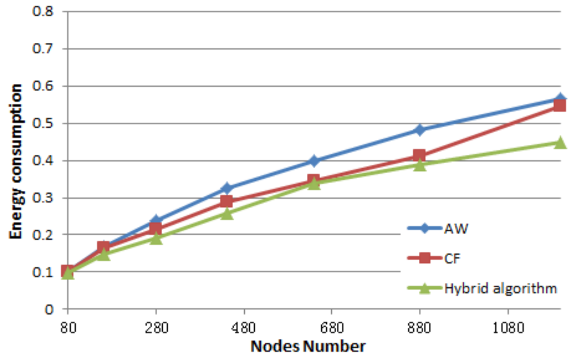

Figure 3: Comparison of energy consumption

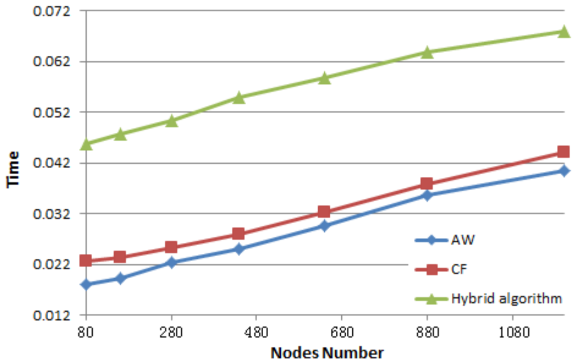

Figure 4: Comparison of time-consuming

In general, the improved hybrid algorithm proposed in this paper has more advantages in energy consumption and precision. Although the complexity of computational logic leads to long computation time, the overall time is still within acceptable limits.

6. Conclusion

Chemical enterprises play an important role in the economy of our country, and the products of chemical enterprises bring great convenience to human life. However, there is a great safety risk in the production of chemical enterprises, which need to attract enough attention. In view of the potential danger of chemical industry production, this paper proposes an early warning system of dangerous chemical gas based on wireless sensor network, which can monitor the production process in real time by sensors in all corners of the plant. In the aspect of sensor data fusion technology, a hybrid fusion algorithm based on adaptive weighting algorithm and Calman filtering algorithm is proposed, which not only improves the monitoring accuracy, but also reduces the energy consumption.

References

Ding L., Liao T.Q., Tao L., 2011, The Method of Sensors Data Fusion Based on SVR. Chinese Journal of Sensors and Actuators, 24(5), 710-712

Li D., Liu T.J., Lin N., Huang C.Z., 2014, Data Aggregation in Wireless Sensor Network Based on Deep Learning Model. Chinese Journal of Sensors and Actuators, 27(12), 1705-1709

Li L., Wang L., Zhang F.G., Wang X.Z., 2012, Lower energy adaptive clustering hierarchy routing protocol for wireless sensor network. Journal of Computer Applications, 32(10), 2700-2703

Lin K., Zhao H., Yin Z.Y., Luo D.D., 2008, A Clustering Hierarchy Arithmetic Based on Energy Prediction for Wireless Sensor Networks. Acta Electronica Sinica, 36(4), 824-828

Qiu C., 2012, Research on data fusion algorithms in wireless sensor network for indoor environment monitoring system. 23-27

Liu X.R., 2014, Research of The Data Fusion Technology in Wireless Sensor Networks. 27-29

Xiang M., Shi W.R., Jiang C.J., Zhang Y., 2010, Energy efficient clustering algorithm for maximizing life time of wireless sensor networks. AEU-International Journal of Electronic & Communication, 64(4), 289-298

Zhang P., Dong W.J., Gao D.D. 2014, An Optimal Method of Data Fusion for Multi-Sensors Based on Bayesian Estimation. Chinese Journal of Sensors and Actuators, 27(5):244-247

Yager R.R., 2001, The Power Average Operator. IEEE Transactions on Systems, Man, and Cybernetics, 31(6), 724-731.