2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC)
Dec 20-22, 2013, Shenyang, China

Research on Multi-sensor Data Fusion Algorithm of Soil Carbon Sink Factors Based on Neural Network

Wu Qiulan
School of Information Science and Engineering,
Shandong Agricultural University,
Tai'an Shandong,271018,P.R. China
wqlsdau@163.com

Liang Yong*
School of Information Science and Engineering,
Shandong Agricultural University,
Tai'an Shandong,271018,P.R. China
yongl@sdau.edu.cn

Geng Xia
School of Information Science and Engineering,
Shandong Agricultural University,
Tai'an Shandong,271018,P.R. China
gx@sdau.edu.cn

Xu Hongli
School of Information Science and Engineering,
Shandong Agricultural University,
Tai'an Shandong,271018,P.R. China
coco2006@sdau.edu.cn

Abstract*—Accurate acquisition of farmland soil carbon sink factors is the key to realize estimation of soil carbon sink, and development of wireless sensor technology provides an important mean to promotion of measurement accuracy of carbon sink factors. Based on wireless communication protocol, this paper combines cluster-based hierarchical structure with the hierarchical structure of neural network, designs a neural network data fusion algorithm based on cluster-based routing protocol, and uses it in acquisition system of farmland soil carbon sink factors. It is indicated by the results that this algorithm can enhance measurement accuracy of carbon sink factors, reduce data transmission quantity effectively, and decrease energy consumption of the network, thus extend the service life of the network.

Keywords*—data fusion; soil carbon sink; neural network; multi-sensor

I. Introduction

With the continuous intensification of global climate change, more and more attentions have been paid to the carbon sink of terrestrial ecosystem. The change in farmland soil organic carbon is an important basis for impact of terrestrial ecosystem on atmospheric CO_2 source sink effect and many researches concerning the estimation model of soil carbon sink have been conducted at home and abroad and made great progresses. The researches show that there are a lot of influencing factors of soil organic carbon, including soil temperature, moisture content, PH value, nutrient index (nitrogen, phosphorus and potassium), texture index,etc. The accurate acquisition of carbon sink factors is the key to realize estimation of soil carbon sink and the development of wireless sensor technology provides an important mean to improve the measurement accuracy of carbon sink factors. The multi-sensor data fusion technology is a very important technology in wireless sensor network and many researchers have been carried out focused on data fusion at home and abroad[1-6]. The multi-sensor data fusion has been applied in the field of agricultural information acquisition in recent years. Jiang Dingguo conducts a research on two-level data fusion based on the greenhouse monitoring system of wireless sensor network by using Bayes's conditional probability formula to improve the accuracy rate in circuit fault judgment of nodes and realize the remote monitoring of agricultural environmental information[7]. According to the measured farmland soil information of each sensor, Feng Minmin etc. designs a fusion algorithm based on BP natural network, builds a reasonable data fusion model and obtains the accurate information of farmland soil by simulation[8].Wu Xiaojia etc. raises a distributed sensor system structure applicable for agricultural resources and environmental monitoring and two-level fusion model based on D-S evidential reasoning method[9].Considering the problem concerning the date uncertainty in sensor information acquisition system, Jiao Zhuqing etc. puts forward a sensor data fusion method based on confidence levels and achieves good results by applying to the data fusion of farmland soil moisture content[10]. Zhang Yulin raises a sensor information fusion model based on wavelet packet transform which is applied to the rapid acquisition and processing of soil moisture and conductivity data[11].

In conclusion, the application of multi-sensor data fusion technology in the agricultural field mainly focuses on information acquisition of farmland environment. There is less research on the accurate determination method of soil carbon sink estimation model parameters and there is no report about the acquisition method of soil carbon sink factors based on multi-sensor data fusion. This paper uses a neural network data fusion (NNDF) algorithm based on cluster-based routing protocol, in the acquisition system of farmland soil carbon sink factors and the results indicates that this algorithm can enhance the measurement accuracy of carbon sink factors, reduce data transmission quantity effectively, and decrease the energy consumption of the network, thus extend the service life of the network.

978-1-4799-2565-0/13/$31.00 ©2013 IEEE

II. ACQUISITION SYSTEM OF SOIL CARBON SINK FACTORS

The acquisition system structure of soil carbon sink factors based on wireless sensor network is shown in Fig.1[12].

The system can realize the real time acquisition of soil carbon sink factors including soil temperature, moisture and pH value etc. and has the following features:

(1) The sensor nodes including soil temperature, moisture and PH value etc. are laid in the network and conduct communication in a wireless form;

(2) The sensor nodes adopt solar cells for power supply and wireless communication routing protocol and also selects cluster heads within the whole network to form cluster hierarchical structure and optimal path to transmit data, so as to save energy;

(3) Conduct certain data fusion of each carbon sink factor acquired on cluster head nodes and send the fusion results to the aggregation nodes;

(4) The aggregation nodes collect the fusion results of soil carbon sink factors and give an inquiry command to all or partial sensor nodes and then the corresponding sensor nodes will send data directly to the aggregation nodes according to the command content; meanwhile, the cluster head noses can conduct certain data fusion according to the command content and send the fusion results to the aggregation nodes.

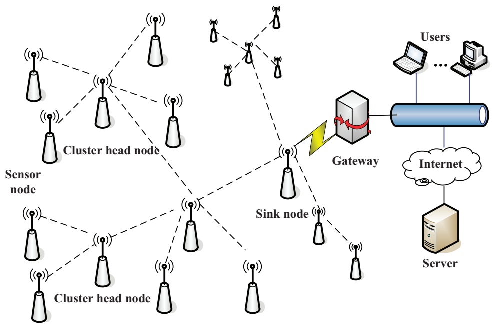

Fig.1. Structure of acquisition system of soil carbon sink factors

III. THE MULTI-SENSOR DATA FUSION ALGORITHM BASED ON NEURAL NETWORK

In the acquisition system of farmland soil carbon sink factors, the sensor nodes continuously collect the soil carbon sink factors and send the results after fusion of data acquired to the aggregation nodes. As the data acquired by adjacent sensors may be the same or the data acquired by same sensor node in a short time may be the same, the data acquired by the sensor nodes continuously have large redundancies. If the information acquired by all sensor nodes is sent to the aggregation nodes directly, problems such as loss of large numbers of data and channel congestion will occur. It is allowed to use the cluster-based routing protocol to conduct clustering of all sensor nodes in the farmland and send the information of soil carbon sink factors by cluster head nodes; considering the data of carbon sink factors acquired with large volume and high redundancy, the data processing can be conducted in each cluster member node and the data with any change will be sent to its cluster head node, which reduces the data transmission effectively and reduce network consumption; the cluster head nodes fuse the information acquired by its cluster member node, such as temperature and moisture to form an effective decision and send the fusion results to the aggregation nodes[13-14].

According to the cluster-based routing protocol and specific application requirement, the structure chart based on NNDF algorithm is shown in Fig.2. The three-level sensor neural network model is adopted to correspond to a cluster in the wireless sensor network where the input layer and first hidden layer are within the cluster member node and the output layer and second hidden layer are within the cluster node.

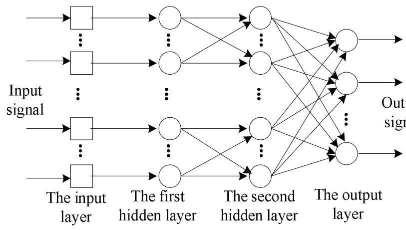

Fig.2. Structure of data-fusion algorithm

In the information acquisition of farmland soil carbon sink factors, assuming that there are n cluster member nodes in a cluster in the wireless sensor network and m different data are acquired in each cluster member node, there are n×m nodes of input layer in the neural network model and n×m neuron of first hidden layer. The quantity of neuron of second hidden layer and output layer can be adjusted according to the specific application and has no relation with quantity n of the cluster member nodes. For different data, the quantity of second hidden layers can be different. There is no full join between input layer and first hidden layer and between first hidden layer and second hidden layer and different data can have different fusion processing; there is full join between second hidden layer and output layer and different data can have comprehensive fusion processing. According to the above three-level sensor neural network model, after each clustering, the fusion algorithm of neural network data can have simple processing of all data acquired by each cluster member node according to the neuron function of first hidden layer, and then send the processing results to its cluster head node, which will have further fusion processing according to neuron function of second hidden layer and output layer; finally, the cluster head node will send the fusion results to the aggregation nodes.

Each neuron can conduct information processing and the neuron processing model is shown in Fig.3.

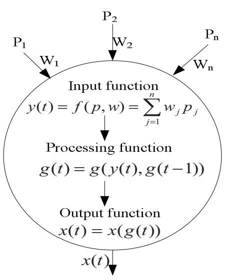

Fig.3. The neuron processing model

The neuron processing model can be divided into three areas: input area, processing area and output area, which respectively correspond to input function, processing function and output function. The input function is one that adopts weighted sum form, in which pj is the input signal and wj is the weight of each input signal. The weighted sum form is a first-approximation of the integrative function of neuron input signals and a kind of space integration. In the processing area, the processing function g(t) can process and deal with the input after the integration.

The flowchart of NNDF algorithm is shown in Fig.4.

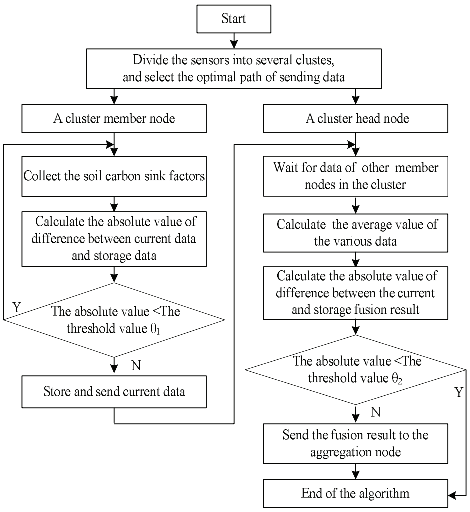

Fig.4. The flowchart of NNDF algorithm

The processing functions of the cluster member node and the cluster head node are introduced in detail below.

(1) The processing function of the cluster member node

The processing function of cluster member node is shown in Fig.5.

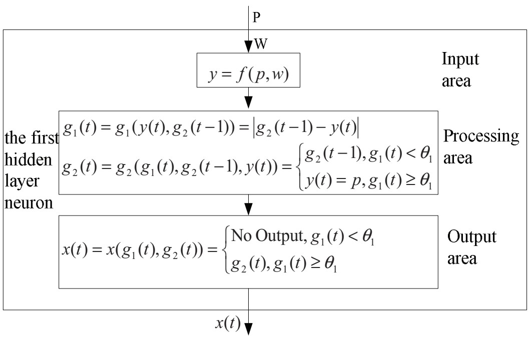

Fig.5. The processing function of the cluster member node

In the data fusion algorithm based on the neural network, the cluster member node corresponds to the first hidden layer neuron, the number of which is decided by the data type acquired by the cluster member node. In the item, there are three types of data type acquired by the sensor node, namely soil moisture, temperature and PH value, hence there is a first hidden layer neuron with three single input in each sensor node, the processing function of the processing area in the first hidden layer neuron adopts weighted sum method; as there is only one input signal in each first hidden layer neuron, the processing function of the input area becomes an identity function.

The function of the processing area is to calculate absolute value of the differential value between current data and last stored data, and to judge whether the current data is needed to be stored; the processing area adopts two processing functions.

The processing function of the output area adopts one special threshold function, the practical significance of which is that when g1 (t) is less than the threshold value θ1, the neuron has no output result; when g1(t) is more than or equals to the threshold value θ1, the output result of the neuron is g2(t).

In conclusion, if the current data acquired by the cluster member node has no big variation compared with the data stored last time, the cluster member node will not send data to its cluster head node; only if the differential value between the current data and the data stored last time is more than or equal to the threshold value θ1, the cluster member node will then send the newly acquired data to its cluster head node; according to the change of its newly acquired data, each cluster member node decides whether to send data to its cluster head node or not, which can avoid many repetitive and highly similar data being sent to the cluster head node, thus saving energy consumption of each cluster member node.

(2) The processing function of the cluster head node

The cluster head node corresponds to the second hidden layer neuron and the output layer neuron and its processing function is closely related to the practical application.

The quantity of second hidden layer neuron and the processing function of each second hidden layer neuron rely on the practical application and the number of hidden layer in the cluster head node varies with the practical application.

Taking the soil carbon sink factors—soil moisture for example, the neuron processing function to calculate the average soil moisture is shown in Fig.6.

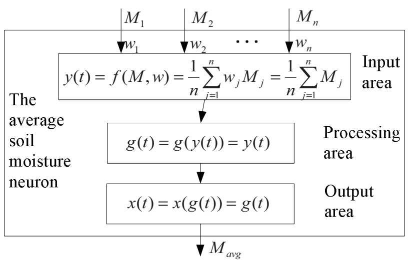

Fig.6. The processing function of the average moisture neuron

The average soil moisture neuron is that having "n" input signals (the soil moisture value (Mj) acquired by each cluster member node) and one single output signal (the average soil moisture value Mavg), in which n means the number of the cluster member node that changes dynamically with the routine protocol. Since the weight of each input signal is the same, actually the processing function of the input area is calculating the average value of the input data. Because the function needed to be achieved by the neuron is to calculate the average value of the soil moisture, there is no need to deal with the processing area and the output area; identity function can be adopted to output the average soil moisture value in the cluster.

The cluster head node requires not only fulfilling the practical application but also sending the fusion result to the aggregation node. The output layer neuron is used to deal with the result of the application requirement, whose quantity is decided by the practical application requirement.

The processing function of the output layer neuron is shown in Fig.7.

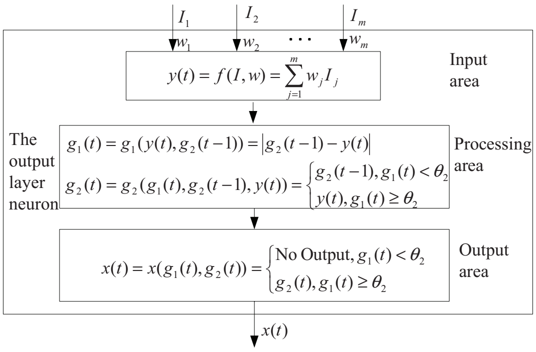

Fig.7. The processing function of output layer neurons

The processing model of the output layer neuron is quite similar to the processing method of the cluster member node. The biggest different point goes to the number of the input signals; the first hidden layer neuron is single input neuron while the output layer neuron is multi-input neuron. In the input area of the output layer neuron, the input data of the output layer neuron is the output result of the second hidden layer neuron. Considering the difference among the soil carbon sink factors, the weight of the farmland soil carbon sink factors must be taken into consideration, hence weighted sum processing function is adopted.

Similar to the cluster member node, the processing area and the output area of the output layer neuron need to conduct comparative judgment concerning the weighted sum result of the input area; If the absolute value of the differential value between the current weighted sum result and that stored last time is more than the threshold value θ2, the output result of the neuron is the current weighted sum, namely, the cluster head node will send the newly fused result to the aggregation node; otherwise, the neuron has no output result, that is to say, there is no big change on the information in the internet. The weight W is of great significance to the output layer neuron, which serves as an important parameter for the data fusion algorithm. Since there are too many factors influencing the farmland soil carbon sink, any consideration of one single factor cannot form a reasonable and effective judgment. At the same time, the factors that influence the soil carbon sink are interactive and interrelated; if one factor changes, other related factors will change accordingly. Therefore, if the acquired data of various types can receive a comprehensive and correct judgment through effective data fusion, then the farmland soil carbon sink can be judged effectively. In the neural network, those weights can be learned, trained and adjusted through the neural network itself; while in the wireless sensor network, because of the restriction of system resources such as communication, CPU and storage space, the training of weight cannot be conducted in the wireless sensor network, but finished in the outside of the wireless sensor network[15].

The training steps of the multilayer perception network are as follows:

(1)Data collection

In combination of the subject—multi-scale farmland information acquisition and fusion technique, the main producing area of wheat in Yanzhou City of Shandong Province is chosen as the experimental base. The soil carbon sink factors are acquired including soil temperature, moisture, PH value, etc. Some experimental data is shown in Table I.

(2)Selecting the type and structure of network

The three-layer perception model is adopted according to the hierarchical structure of wireless sensor networks and data acquired using acquisition system of soil carbon sink factors. The cluster member node is the first hidden layer neuron, each type of input data has corresponding two neurons in the second hidden layer neurons, output layer has only one neuron.

(3)Training of the multilayer perception network

Applying the sample data as training data, the essential relationship between the input data and output data in the sample data can be obtained through several training, thus the optimum weight can be got and sent to the lower node using the aggregation nodes.

TABLE I. SAMPLE DATA OF SOIL CARBON SINK FACTORS

Number of sample	***Soil temperature (℃)***	***Soil moisture***	***PH value***
1	18.5	18.41	7.55
2	18.6	18.22	7.56
3	19.5	18.35	7.55
4	25.6	18.35	7.54
5	26.8	18.25	7.52
6	25.5	18.30	7.53

The convergent curve of the multilayer perception network is shown in Fig.8.

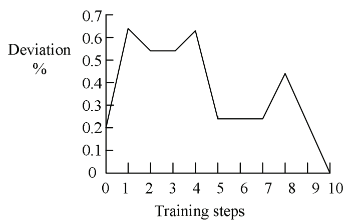

Fig.8. The convergent curve of the multilayer perception network

IV. TEST

In combination of the subject—multi-scale farmland information acquisition and fusion technique, the main producing areas of wheat in Yanzhou City of Shandong Province is chosen as the experimental base to test the efficiency of NNDF algorithm under different network scale. Four kinds of wireless sensor networks of the same node density but different scale are adopted in the test site.

Ten sensor nodes are distributed in the area of $120*120m^2$; twenty sensor nodes are distributed in the area of $150*150m^2$; thirty sensor nodes are distributed in the area of $180*180m^2$; forty sensor nodes are distributed in the area of $200*200m^2$. In the design of the wireless communication routing protocol, the period of the cluster head election is forty minutes and many tests have been conducted. The impact of NNDF algorithm under different scale on the service life of the sensor node and the measurement accuracy of the carbon sink factors is shown in Fig.9, which shows the contrast of the service life and measurement accuracy of sensor nodes obtained by using or without using NNDF algorithms under different network scales and then the following conclusion are obtained:

(1)The amount of data transferred in the wireless sensor network can reduce by an average of about 20%， thus effectively reduce energy consumption, so the service life of the sensor node obtained by using NNDF algorithms is obviously higher than that obtained without using NNDF algorithm.

(2)Whether NNDF algorithms are adopted or not, with the increase of the number of sensor nodes, the service life of the sensor nodes decreases, which is faster without using NNDF algorithms;

(3)After NNDF algorithm is adopted, the measurement accuracy of carbon sink factors has increased.

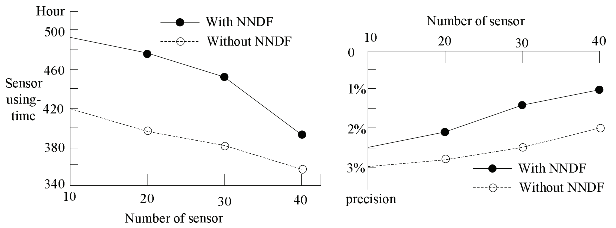

Fig.9. Influence of NNDF algorithm to sensor using-time and measure precision

V. CONCLUSION

Based on the wireless communication routing protocol, this paper combines cluster-based hierarchical structure with the hierarchical structure of neural network, designs a neural network data fusion algorithm based on cluster-based routing protocol, and uses it in acquisition system of farmland soil carbon sink factors. The test results shows that NNDF algorithm of Soil Carbon Sink Factors not only can increase the measurement accuracy, but also greatly reduce the amount of transferred data in the wireless sensor network, thus effectively reduce energy consumption and extend the service life of the network. The NNDF algorithm has good data fusion efficiency, which applies to the wireless sensor network collecting information periodically. This paper still needs to be improved. During the process of data fusion, how to improve the influence caused by fault tolerance and information loss as well as node failure on the network is the content that needs to be studied in the next step.

ACKNOWLEDGMENT

This study was funded by the National High Technology Research and Development Program of China (863 Program) （Grant No.2013AA102301） and the Science and Technology Development Program of Shandong Province (Grant No.2011GGB01308）,and also supported by Shandong Agricultural University. Sincerely thanks are also due to Agricultural Bureau of Yanzhou City for providing the experimental base for this study.

REFERENCES

[1] FU J C,WAN S.Multisensor information fusion based on D-S evidence theory and BP neural network[J].Automation & Instrumentation,pp.22-24, February 2011(in Chinese).

[2] JIANG Y J.Research on key technologies of multi-sensor data fusion[D]. Harbin:Harbin Engineering University,2010.

[3] MALER R. Statistical multisource-multitarget information fusion [M]. Boston:Artech House Publishers, 2007,pp.32-44.

[4] XIAO L H.Research and implementation of data fusion technology in wireless sensor network[D].Guang Zhou:South China University of Technology,2011.

[5] ZHANG Q,QUE P W,CHEN T L.A new multi-sensor data fusion algorithm based on EMD-MMSE[J]. Journal of Shanghai Jiaotong University (Science),vol. E-12pp. 153-158, February 2007.

[6] SHEN G R,GE X Y,HUANG X M,Spatial digital distribution of topsoil organic carbon density and assessment of its reserves in Chongming Island[J].Journal of Shanghai Jiaotong University (Agriculture Science), vol.29,pp.61-66, June 2011(in Chinese).

[7] JIANG D G.Agricultural information monitoring system design and data fusion research based on wireless sensor networks[D].Wu Xi:Jiangnan University,2010.

[8] FENG M M,ZHU P Y. Research on soil electrical conductivity data fusion based on BP neural network[J].Journal of Changshu Institute Technology(Natural Sciences),vol. 23,pp. 93-95, October 2009(in Chinese).

[9] WU X J,YANG P H.The research of application of data fusion technology in agriculture expert system[J],Journal of Taiyuan University of Technology, vol.36,pp.282-285,March 2005(in Chinese).

[10] JIAO Z Q,XIONG W L,ZHANG L.Multi-sensor data fusion method based on belief degree and applications[J],Journal of Southeast University (Natural Science Edition),vol.38,pp.253-257,January 2008.

[11] ZHANG Y L.Research and application on computational intelligence in data fusion of soil information [D].Wu Xi:Jiangnan University,2009.

[12] ZHANG R R, ZHAO C J,CHEN L P.Design of wireless sensor network node for field information acquisition[J].Transactions of the CSAE, vol.25,pp.213-218,November 2009(in Chinese).

[13] HALL. Manual of multi-sensor data fusion[M],Beijing: Electronic Industry Press,2008.

[14] OTHMAN Sidek, QUADRI S A. A review of data fusion models and systems [J]. International Journal of Image and Data Fusion,vol.3,pp.3-21,January 2012.

[15] ZHANG Y J.Research of wireless sensor networks and data fusion algorithm based on farmland soil information [D].Wu Xi:Jiangnan University,2009.