International Journal of Hybrid Information Technology
Vol. 9, No.7 (2016), pp. 207-216
http://dx.doi.org/10.14257/ijhit.2016.9.7.19

Wireless Sensor Network Application Research Based on Data Fusion

Fengbiao Zan

Qinghai university for nationalities, xi'ning Qinghai Prov.810007, China
zanfb@163.com

Abstract

Due to the high density of the deployment of wireless sensor network, makes a certain range of redundancy data acquisition, inefficient communication.And the data fusion technology to reduce data redundancy, improve the accuracy of information, etc.Therefore, this paper puts forward a kind of wireless sensor network based on data fusion, through simulation experiment, the method to handle well the data redundancy and reduce the energy consumption data transmission, prolong the life cycle of the network.

Keywords*: Wireless sensor； Data fusion； Routing protocol*

1. Introduction

In recent years, the rapid development of network information technology to every corner of the world the connection between the people more and more close, people at any time, any place can easily dialogue and communication.Based on the development of communication technology, sensor technology, sensor size small to can be encapsulated into a millimeter level chip, low cost, high robustness and more powerful.Sensor "wisdom" can help people to accomplish the data collecting, sorting and communication, thus greatly expands the application fields of the sensor.Wireless sensor network will fundamentally alter the interaction between human and the natural world, and make the virtual network world and the physical world and the extension of the field can communicate with people.

Wireless sensor network these characteristics make it in a network, transfer information, processing information, facing great challenges.Traditional wireless network because of the abundant energy, the primary design goal is to provide high service quality and efficient use of bandwidth, next will only consider energy saving;Wireless sensor network due to battery power supply or battery irreplaceability, so energy saving up to the top of the design goals.How the data of sensor nodes in the case of the save energy route to the destination node becomes the hot topics in the study of wireless sensor network.

Most of the current data fusion research has focused on the shortest path tree structure, but the emphasis of the different application scenarios must be different, most of the work focuses on wireless sensor network data fusion of the network data fusion.For fusion query, a certain amount of calculation is acceptable within the network, the sensor sensing data is processed into various parts, and then passed to the gathering node or base station in the process of constantly updated.Each sensor node just pass one or relay a small amount of data to other nodes, thus forming a energy effective method.Current data fusion protocol is mainly divided into three types: the first is based on clusters of data fusion, data fusion method based on cluster area will be divided into several clusters of the entire network, each region elected their own cluster head, sensor node collected data after the data sent directly to the cluster of cluster head, cluster head within the cluster on the data fusion after sent to Sink node.The second is based on data fusion of the tree in WSN, the

ISSN: 1738-9968 IJHIT

Sink node data collected through a reverse in the form of a tree from each node along the trunk to roots, forming a Sink node as root, the source node for the tree structure of the leaf node, the data fusion tree.The structure of the optimal fusion tree problem often into minimum Steiner multicast tree problem.The third kind of data fusion based on multipath the multipath method is suitable for in addition to the ordinary tree topology shapes of arbitrary data fusion.Unlike framework based on tree in a specific time period, in the method based on multipath don't need such a demand.Significant advantages in the multipath method is low communication error[1-4].

2. Related Works

2.1. Wireless Sensor Technology

In a range of monitoring target area for setting a large number of wireless sensor node, the node generated by means of self-organizing, adaptive between a network and the network generally consists of sensor nodes in monitoring area, gathering of Sink node, the Internet and satellite, the console.Works such as the following: a large number of sensor nodes are randomly distributed in a certain range of the target area, according to the specific requirements gathering area data parameters, and USES the single hop or jump way of communication will gather in node collected data sent to the Sink, and then after connected to the console of the Internet and satellite, users can find the data in the console, or give orders to all nodes in the network.

Generally, sensor nodes can be regarded as a kind of the batteries of tiny embedded systems, storage capacity and communication capacity is weak, limits the this kind of sensor nodes and Sink nodes of direct communication.From the Angle of the function of network, sensor nodes except for local data collection and integration, but also on the data from other nodes relaying operation, both terminal and repeater dual role.Gather Sink node can be a powerful nodes, have enough energy to support more memory and processing power, also can serve as a gateway device to connect wireless sensor networks and the Internet, realize the conversion between the two kinds of protocol stack.

From the Angle of system properties, sensor nodes including software systems, hardware, system of two parts, including the basic structure of the hardware system mainly includes: power supply module, sensor module, processor module and wireless communication unit, the concrete structure as shown in Figure 1.

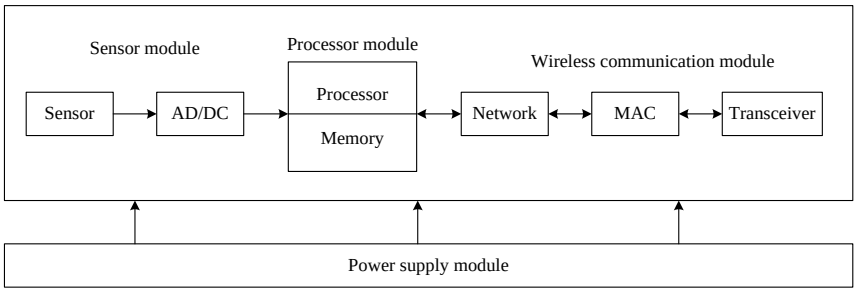

Figure 1. Sensor Hardware System Structure

The normal operation of the power supply, battery module is usually used for the sensor to provide energy.With the improvement of integrated circuit technology, the energy consumption of the sensor module and processor module becomes very low, but on the wireless communication module consumes most of the energy.Sensor nodes transmit information to more energy than when performing calculations, perform 3000 jump instruction is equivalent to transfer 100 meters 1 bit of information consumption power.

Wireless communication module usually have send, receive, idle and sleep four state.Sleep state energy cost the least, sending state energy consumption is more than sleep, receiving and spare time than sending the state only save about 30%.Sleep state node fully closed communication module, natural energy consumption is very low.But in the free state, the node will always monitor channel, check whether there is data to send to come over, so is part of the energy consumption.From energy saving point of view, should as far as possible envoys point in the sleep state.

Wireless communication module is main task with other nodes wireless communication, exchange control message sending and receiving data, and its energy consumption in the direct ratio n times as the square of the distance of transmission.In addition, the data from the adjacent sensor nodes with larger repeatability, so as to reduce network traffic using data fusion technology for processing.

The processor module is made up of microprocessor and memory.With miniature and low power consumption microprocessor embedded CPU, assume control function, calculation data, implement communication protocol, processing sensory data parameters and forward data.Memory is mainly responsible for data storage before and after processing[5-6].

Sensor module is composed of sensor and AD/DC converters, responsible for collecting data in the target area, is responsible for the AD/DC converter to convert the analog and digital data.

Software system is also an indispensable component of wireless sensor network, in addition to including embedded operating system, also contains all kinds of communication protocol, and the fit and unfit quality directly related to the performance of the "life" of the network.In addition, individual special function of sensor node include mobile device, positioning module and part of the solar energy device.

As well as Ad hoc, wireless sensor network infrastructure support, all nodes in a distributed operation, through the interaction with adjacent nodes to transmit data, is adaptive, self-organizing and intelligent network management.However, wireless sensor network compared with the traditional Ad hoc networks have unique characteristics:

1) resource limited

Sensor nodes is a kind of small volume, low cost, less power consumption of tiny embedded devices, storage capacity is small, so it has less energy, computing, processing and communication ability is weak, not suitable for dealing with high energy consumption of complex task.So minimizing energy consumption becomes the wireless sensor network routing protocol design principal and main consideration factors.

2) since the organization

Wireless sensor network in the network infrastructure without any cases, using the object (*e.g.,* plane) scatters in remote or regional conditions, nodes may be down and stop work at any time, that will cause the dynamic change of network topology.In addition, for the sake of saving energy, sensor nodes often run in two modes work and sleep, so network topology are subject to change, also requires that the sensor nodes in a network of should possess self-organizing, adaptive ability to realize self-help, intelligently through topology control and distributed algorithms such as a series of actions needed to complete the network data.

3) on a large scale and high density

The number of nodes in wireless sensor network and distribution density is greater than the traditional Ad Hoc networks.In order to obtain accurate information, wireless sensor network often need to large-scale deployment node, and the high density of nodes, a large number of redundant nodes to guarantee the existence of a single node failure cases, data integrity and accuracy, makes the wireless sensor network has strong robustness and accuracy.

4) strong network dynamics

Wireless sensor nodes in a network of more traditional Ad Hoc network nodes in the

network dynamics.Although very little move the nodes in wireless sensor network, but because of the energy of the nodes is used up, changes in the environment, the emergence of a new node or node movement situations, such as topological structure will be changed accordingly.As a result, the dynamic reconfiguration of wireless sensor network must have a strong ability to quickly achieve working state.

5) application of correlation

Wireless sensor network is an application, but because of the different operating mode, the emphasis of the network routing protocol design is also different.Some application requirements, for example, is a single gathering node, and some requirements are more gathering node;Some pay attention to the security of network, and some of security requirement is not high.In short, the design of wireless sensor network road by agreement does not have versatility, must according to each specific application in mind.

6) for the center with the data

Traditional computer network based on address (IP address, for example) as the center of the network, the wireless sensor network because of its is task-based network, the characteristics of the node number more, often do not focus too much on the details.In the target tracking system, for example, the user generally don't care about the source of the data, but pay more attention to data conclusion and spatial location.In wireless sensor network is not necessarily in accordance with the address to choose the path, but according to the point of interest to establish from the path between the sender and the receiver.In addition, the traditional computer network is the "original copy", generally not in the transmission process of data analysis and processing, but to reduce the energy consumption of wireless sensor network, sometimes require in the process of transfer of data redundancy, data fusion is a key technology in WSN.

7) poor safety

Most of wireless sensor network using the wireless channel and distributed control mode is easy to be exposed to the enemy, vulnerable to the enemy or hacker intentionally tease or darkly network attack, such as: denial of service attack, sinkhole attack, Sybil attack and Wormhole attack, *etc*.Combined with the security development is imperfect, so the safety performance than traditional networks.

2.2. The Data Fusion Method

According to the definition of popular now, the data fusion technology, including the various information sources give useful information collection, transmission, comprehensive, filtering, correlation and synthesis, in order to assist people in environmental decision, planning, detection, validation, diagnosis.Sensor data fusion is inner connection to a variety of data acquisition, and comprehensive processing and optimization technology.Sensor data fusion technology to get inner link and law of all kinds of information, combining with the spatial and temporal correlation of sensor nodes, eliminate redundant and false information, keep the key and the right information, finally realizes the information of simplified and optimized.Single sensor can only get the environment or the part of the object information, and the multi-sensor information after fusion can perfectly and precisely reflect the characteristics of the environment.Data fusion technology in the research of intelligent information processing technology has the effect that cannot replace.

In wireless sensor networks, due to the randomness of node deployment and high density, so the induction to the information characteristics of redundancy and complementarity, real time, to obtain low into nature, *etc*.Data fusion can take full advantage of the characteristics of data, effectively reduce redundancy, enhance complementarity, improve the real-time performance and reduce the transmission of low into nature.The use of data fusion can increase the survival ability of the system, expanded the coverage of space, expanded the coverage of time, improves the reliability and reduces the ambiguity of information, improve the detection performance, and

improve the spatial resolution.

Data fusion can be divided into the pixel level fusion, feature level fusion and decision makers.Pixel level fusion is directly in the raw data were collected on a layer of fusion, before the original data of various sensors without pretreatment on synthesis and analysis of it belongs to the low level of integration;Feature fusion layer on the original data from the sensor feature extraction, the characteristics of comprehensive analysis and processing, it belongs to the middle level of fusion;Policy makers is each sensor in the first place in the local complete preprocessing, feature extraction, recognition and judgment basic processing, set up a preliminary conclusion, and then through the correlation fusion decision, decision makers ultimately joint inference, it belongs to the senior level fusion.When the node deployment of a high density of sensor sensing data is redundant, *i.e.* multiple adjacent sensor nodes to collect to the same information.To save the energy of the sensor nodes, it is necessary to get rid of the redundant information and reduce the data transmission.With tens of thousands of sensor nodes in the network, sending a message consumes more energy than energy calculation, data fusion can very good to remove the redundant information, making smaller transfer information[7-9].

3. Wireless Sensing Technology Based on Data Fusion

In wireless sensor network node, if the similar or continuous period, sensing data or attribute values is similar, or have a regularity of change, even if at some point has a small amount of node failure can not be involved in the induction time, using the similarity of data also can get a set of data and real data have a very small error or gentle change, then we will say these nodes has time correlation.In a wireless sensor network, the node deployment is random, especially in some remote place, use the planes had been planting seed for node, will cause certain node density, the geographical position close to the node at the same time sensor data will have a high degree of similarity, we call these nodes sensor data with spatial correlation.The characteristics of spatial and temporal correlation of data fusion processing provides the convenience, in the case of the failure of some nodes, the advantage of the characteristics of the spatial and temporal correlation estimates, the fusion result has a broader coverage.

3.1. Data Fusion Tree Building

To store and search a convenient, general use a variety of tree structure to store and transmit data, using the regression mathematical model for data analysis in this chapter, to build a complete trigeminal tree, child nodes only transfer coefficient of a group to its parent node, used for data fusion and data query, greatly reduce the energy consumption, effectively extending the life cycle.In a deployed a large number of sensor network, sensor nodes through the optimal path to the base station, the general is the use of minimum spanning tree to solve the problem, commonly used is the greedy algorithm.But greedy algorithm is a kind of local optimum algorithm, just to make the current appears to be the best way to choose, but not from the overall optimal into consideration, it is just a sense to local optimal solution.We construct a completely trigeminal tree, used for data transmission.Completely trigeminal tree hierarchical network is the role of the various nodes of cluster head, only to perform receiving data, transmit data, and the function of data fusion, as shown in Figure 2.

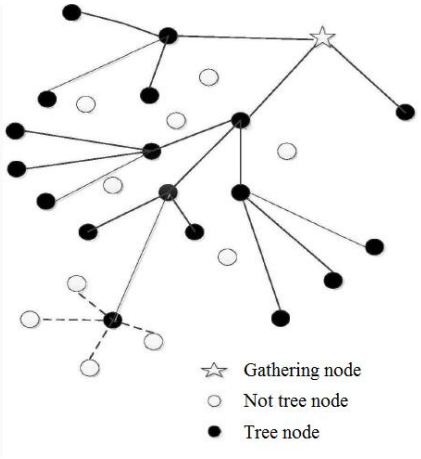

Figure 2. Data Fusion Tree Structure

Said for the convenience of coordinates, we set the wireless sensor network area for the square area, variable length L, assuming that each child area longer for L_s, then the whole network can be divided into L/L_s is the area.Roots must be as close to the adjacent area to satisfy the query multiple areas of information propagation path minimum requirements at the same time.So each query subtree roots is the best deployment in the square subdomain corner position.Each subtree is determined, the root of tree deep analysis and choice.This chapter is to use fully trigeminal tree topology, each parent node should have 3 child node, it's fully guaranteed the tree in the whole area of coverage and the accuracy of data processing[10-12].

The parent node to pass three types of messages: HELLO, PROBE and JOIN.In addition to the HELLO message other message must contain the sender and the receiver's ID.The message exchange between as shown in Figure 3.

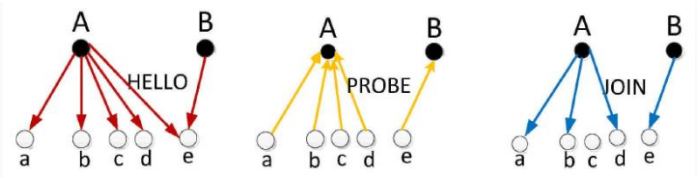

Figure 3. The Message Exchange between Processes

Algorithm description is as follows:

1) HELLO message

J layer node of the tree are to spread a HELLO message packets to give it a jump distance of adjacent nodes j+1, j+1 layer of a node is likely to receive more than HELLO message, it need according to probability p'> p''randomly select a node as its parent node, and sends a PROBE message to the parent node.E is shown in Figure 3 nodes at the same time to accept A, B, e according to the probability of the random choice B as its parent node.

2) the PROBE message

J layer is selected in the parent receive j+1 layer node sends the PROBE of the message, because according to probability conditions chose to return to the PROBE, so a node does not also send messages to two parent node.As shown in Figure 3 A sends A HELLO message to A, b, c, d, e, but only A, b, c, d send the PROBE A message to A node.

3) the JOIN message

Received the PROBE message, have belong to many child nodes, completely trigeminal tree only choose three child nodes, follow the rule is: first select 120 within the boundary of two adjacent nodes, and then according to the received received a, d node of

the size of the intensity of transmitted power, choose between two nodes of the node as the third child node.As shown in Figure 3 A receiving A, b, c, d the PROBE message, but only the node A only chose A, b, d as its child nodes.

Nodes in wireless sensor networks is the spread of according to the above process to build completely fork of the tree.Because completely trigeminal tree than involved fully binary tree node, the messages exchanged between the nodes is more, we can advance design depth of the tree is 3, and the experimental results prove that on the scale of hundreds of nodes, wireless sensor, depth is 3 completely trigeminal tree can achieve very good effect.

In a wireless sensor network, the distribution of a large number of nodes constantly sensing the environment data, and transmitted to other nodes.If when the sudden changes in the surrounding environment or abnormal sensor nodes received interference signal, we need timely to identify abnormal data to analyze abnormal or discarded directly, it is very important in some applications.In forest fire monitoring, for example, deployed a large number of sensor nodes, each sensor node will continue to pass information in active or passive way to detect base station, usually sensor induction temperature is in a fixed range, in the event of fire, the sensor will send out abnormal data values, touches the early warning signal, which requires timely identification to control the spread of fire.But sometimes due to human factors, such as outdoor enthusiasts in the forest fire to cook, just near the furnace frame in a sensor will also because of the abnormal data values, this would require the use of spatio-temporal correlation to judge, if the nodes in a short time back to normal, and nearby the value of the sensor is normal, we can not do any processing.

3.2. Performance Analysis

1. The percentage error The sensor data and the error between the real data in the form of percentage of said:

$$PE = (\frac{z - \hat{z}}{z} \times 100\%) \leq \varepsilon_m \quad (1)$$

One set of threshold for the error ε_m=5%。

2. The number of query tree node output data compression ratio and received other query tree nodes and the query tree node sensing, according to the ratio between the total of the ratio between the output and input.Compression ratio reduces the proportion is less, the greater the redundancy of data is more, the smaller the compression ratio of data redundancy is less, generally we expect compression ratio as small as possible, but also want to consider the completeness and accuracy of the data, is not as small as possible.

3. The fusion data output size We expect the tree and node density varies with different cases, the data output is a fixed value, is advantageous to the network scalability and energy saving[13-15].

4. The Experiment Results Analysis

The conclusion based on MATLAB simulation platform, in order to be comparative method, the advantages of using the simulation data are shown in Table 1 below:

Table 1. Parameter Setting

Symbol	Numerical	Unit
Z	600*600	m^2
R	35	m
S	2500	
Z_s	350*350	m^2
p	3	
n_s	12	

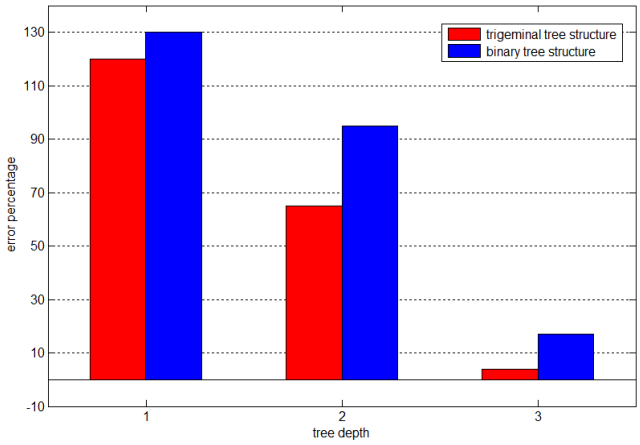

Figure 4. Fully Binary Tree and Fully Trigeminal Tree Error Percentage and Deep Comparison

Can be seen from Figure 4, the tree was deeply 1, since each subdomain with only four nodes is used to process the entire area values of the parameters, in dense nodes distribution receives the sensor data is limited, sometimes not very accurate reflection of real data values, so the percentage error of 110%, when the tree depth is 2, trigeminal tree node to 11, relative to the situation of the four nodes, increase in the number of nodes involved in data fusion, the data fusion precision value increase, but the simulation diagram can also be seen as high as 54%, but when the tree deep 3, belongs to the fusion effect is greatly improved, error rate is only 5%.All these effects are better than complete binary tree, even the binary tree tree deeply 4, the percentage of its error rate is 5.8%.

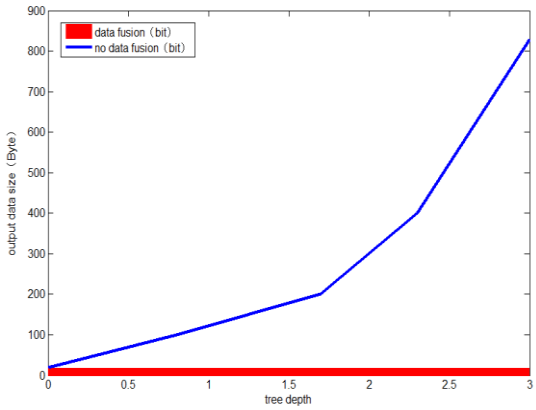

Figure 5. Completely Trigeminal Tree, No Data Fusion Effect

Data output simulation can be seen in Figure 5: the size of the trigeminal tree model based on regression approximation, passed on from one node to another node and base station, is a constant value of the final output data, it also confirms the data values include data fusion coefficient and the coordinate range, because theoretically coefficient is a fixed number of coordinates range is a value, they add up to the length of the data is considered to be a constant.By the simulation results can also see that the output of the data has nothing to do with the number of nodes and trees are deep, this greatly increased the network scalability, which is helpful for the application of this method in a wider range.

5. Conclusion

By establishing the trigeminal tree completely based on multiple variables, we proposed a novel data fusion method, the trigeminal tree leaf nodes within the accepted around the tree node transmission of data, each node has passed this set of coefficients and range of packets to the parent node coordinate, the parent node is using the packet of information and to further data fusion based on the received data.This kind of data fusion strategy to reduce the number of data transfer, saves energy, can also be a quick query.Obtained from the simulation results can be seen that the percentage of error rate in the acceptable range, node length is a constant output data, the data compression ratio is also the basic is a fixed value and the influence of the depth of the trigeminal tree they extremely small, appropriately increase can increase the accuracy of the sensor node density.The roots of the packet can contain all the tree node information, convenient query, and reduce the time of data transmission, to reduce the delay.

Acknowledgments

This work is supported by the Chunhui planning project of Ministry of Education of China under Grant Z2010073.

References

[1] A. Atassi, N. Sayegh and I. Elhajj, “Malicious Node Detection in Wireless Sensor Networks”, Proceedings of the 2013 27th International Conference on Advanced Information Networking and Applications Workshops, IEEE Computer Society, **(2013)**, pp. 456-461.
[2] W. R. Pires, P. F. T. H. De and H. C. Wong, “Malicious Node Detection in Wireless Sensor Networks”, Proceedings of the 2013 27th International Conference on Advanced Information Networking and Applications Workshops, IEEE Computer Society, **(2013)**, pp. 456-461.
[3] W. B. Pöttner, H. Seidel and J. Brown, “Constructing Schedules for Time-Critical Data Delivery in Wireless Sensor Networks”, Acm Transactions on Sensor Networks, vol. 10, no. 3, **(2014)**, pp. 209-245.
[4] G. Han, H. Xu and T. Q. Duong, “Localization algorithms of Wireless Sensor Networks: a survey”, Telecommunication Systems, vol. 52, no. 4, **(2013)**, pp. 2419-2436.
[5] M. M. Mohiuddin, I. Adithyan and P. Rajalakshmi, “EEDF-MAC: An energy efficient MAC protocol for wireless sensor networks”, Advances in Computing, Communications and Informatics (ICACCI), 2013 International Conference on. IEEE, **(2013)**, pp. 1323-1329.
[6] D. Karaboga, S. Okdem and C. Ozturk, “Cluster based wireless sensor network routing using artificial bee colony algorithm”, Wireless Networks, vol. 18, no. 7, **(2012)**, pp. 847-860.
[7] R. D. Pietro, D. Ma and C. Soriente, “Self-healing in unattended wireless sensor networks”, Acm Transactions on Sensor Networks, vol. 9, no. 1, **(2012)**, pp. 421-470.
[8] R. C. Luo and O. Chen, “Mobile Sensor Node Deployment and Asynchronous Power Management for Wireless Sensor Networks”, Industrial Electronics IEEE Transactions on, vol. 59, no. 5, **(2012)**, pp2377-2385.
[9] M. K. Maggs, S. G O'Keefe and D. V. Thiel, “Consensus Clock Synchronization for Wireless Sensor Networks”, Sensors Journal IEEE, vol. 12, no. 6, **(2012)**, pp. 2269-2277.
[10] J. Gang, J. Lei and L. H. Xie, “Fluctuation control for many-to-one routing in wireless sensor networks”, Journal of China Universities of Posts & Telecommunications, vol. 19, no. 06, **(2012)**, pp. 35-44.
[11] B. Khaleghi, A. Khamis and F. O. Karray, “Multisensor data fusion: A review of the state-of-the-art”, Information Fusion, vol. 14, no. 1, **(2013)**, pp. 28–44.
[12] S. Wu, “Data fusion in information retrieval”, Adaptation Learning & Optimization, vool. 36, no. 2, **(2014)**, pp. 2997–3006.
[13] W. Zhang, A. Li and H. Jin, “An Enhanced Spatial and Temporal Data Fusion Model for Fusing Landsat and MODIS Surface Reflectance to Generate High Temporal Landsat-Like Data”, Remote Sensing, vol. 5, no. 10, **(2013)**, pp. 5346-5368.
[14] D. Lillis, F. Toolan and R. Collier, “Extending Probabilistic Data Fusion Using Sliding Windows”, Lecture Notes in Computer Science, vol. 4956, **(2014)**, pp. 358-369.
[15] L, Sorber, M. Van Barel and L. De Lathauwer, “Structured Data Fusion”, IEEE Journal of Selected Topics in Signal Processing, vol. 9, **(2014)**, pp. 586-600.

Author

Fengbiao Zan, he received the M.S. degree in college of computer science and engineering at University of Electronic Science and Technology of China. He is a professor with school of computer, Qinghai University for Nationalities. . His research interests include wireless sensor network, communication and computer network.