AMSE JOURNALS-AMSE IIETA publication-2017-Series: Modelling C; Vol. 78; N°1; pp 71-82
Submitted Jan. 2017; Revised March 15, 2017, Accepted April 15, 2017

Water Quality Monitoring Method based on Data Fusion Technology

* Chunjiang Liu, ** Changlu Qiao

* School of Economics and Management, Chang'an University
China, No. 126 ring road Beilin District Xian Shaanxi 710061, (chunjiangliu@163.com)
** School of Water Conservancy and Architectural Engineering, Shihezi University
China, North 4th Road Shihezi Xinjiang 832003, (574670349@qq.com)

Abstract

In order to effectively monitor water quality, this paper proposes a data fusion method based on Dempster-Shafer evidence theory to detect pollutants in water. Our proposed water quality monitoring system is organized as a hierarchical structure, and the whole monitoring area is divided into several parts. The water quality monitoring system includes an online monitoring module and an offline monitoring module. In particular, each monitoring area has a cluster that contains several wireless sensor nodes to collect data and communicate with other sensor nodes. Furthermore, multiple water quality parameters are detected in our water quality monitoring system, such as PH, conductivity, temperature, dissolved oxygen, turbidity, etc. The final water quality monitoring decisions are made by fusing various types of water quality indexes using the Dempster-Shafer evidence theory. Finally, experimental results prove that the proposed method can detect pollutants in water with higher accuracy by effectively fusing various types of water quality indexes.

Key words

Water quality monitoring, Wireless sensor network, Data Fusion, Dempster-Shafer evidence theory, ROC curve.

1. Introduction

Water is one of the most vital commodities for humans and is essential for socio-economic development of modern society. Moreover, water is not only a life-sustaining drink for humans and all other organisms but also crucial for modern industry and agriculture [1]. However, water pollution, which greatly affects the development and living conditions of persons, is a substantial problem for environment protection [2, 3]. In particularl, the water quality monitoring system and water quality assessment method are two important tools for ensuring and promoting the quality of water resources [4].

Currently, the shortage of water resources and increasing water pollution should be solved carefully in water resource management [5, 6]. The main reason lies in that industrial wastewater, farmland drainage and urban and rural treated sewage are discharged into rivers, lakes and seas, which are greatly deteriorating water quality [7]. Therefore, the most important task we must tackle is to find and control the pollution sources of industrial wastewater, farmland drainage, and so on [8]. With economic development, the water quality monitoring system should be studied in depth. However, most of the existing water quality monitoring works are implemented manually, which makes it difficult to meet the technical requirements.

This paper aims to provide a water quality monitoring system, in which data are collected by a wireless sensor network. As the water quality is affected by various factors, how to effectively collect many different types of data is a crucial problem for water quality monitoring. In this paper, we demonstrate the problem of data fusion in the water monitoring system. Data fusion has been successfully utilized in many areas, such as generating daily synthetic Landsat imagery [9], detection and characterisation of frauds in bovine meat [10], indoor localization under collinear ambiguity [11], prediction of olive oil sensory descriptors [12], olive oil sensory defects classification [13], in-process inspection of freeform shaped parts [14], vibration condition monitoring system for wind turbine bearings [15], exploring the cancer aberration landscape [16], enhancing information retrieval process [17], minimizing energy consumption by selecting sensors for sampling and relaying data [18], adaptive locality weighted multisource joint sparse representation classification [19].

The paper is organized as follows. In the next section, we introduce the water quality monitoring system based on wireless sensor network. Section 3 illustrates how to monitor water quality by multi-sensor fusion using the Dempster-Shafer evidence theory. Section 4 provides experimental results to test the effectiveness of the proposed method, and the whole paper is concluded in section 5.

2. Overview Of The Water Quality Monitoring Using Wireless Sensor Network

The water quality monitoring system is organized in a hierarchical topology, and the monitoring scenario is divided into several areas, such as near the water pump house, near the factory, near farmland, and near residential areas. In particular, in each monitoring area a cluster is set up which consists of several wireless sensor nodes to collect data and communicate with other sensor nodes. In a given cluster, each sensor node is responsible for monitoring water quality parameters, such as PH, conductivity, temperature, dissolved oxygen, turbidity, and so on.

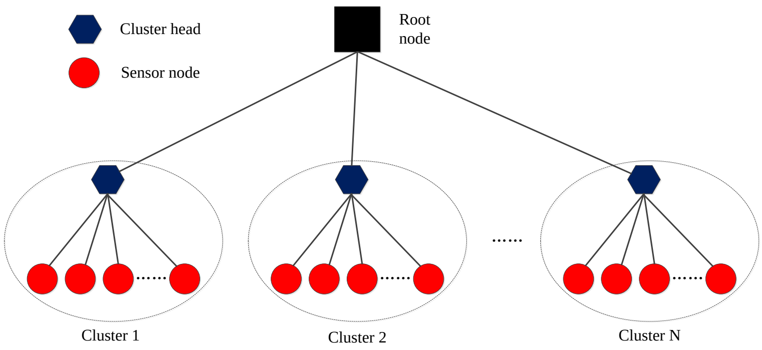


Fig.1. Hierarchical topology of the water quality monitoring system

As is shown in Fig.1, in each monitoring area, each cluster corresponds to a monitoring area, and each cluster is made up of several sensor nodes. In particular, five monitoring points are installed in each cluster at a distance of 10 meters from each other. Furthermore, various types of sensors are used to sense different types of factors which can reflect water quality. As the nodes are installed in an outdoor environment, solar cells are used as the power module to fully utilize renewable energy.

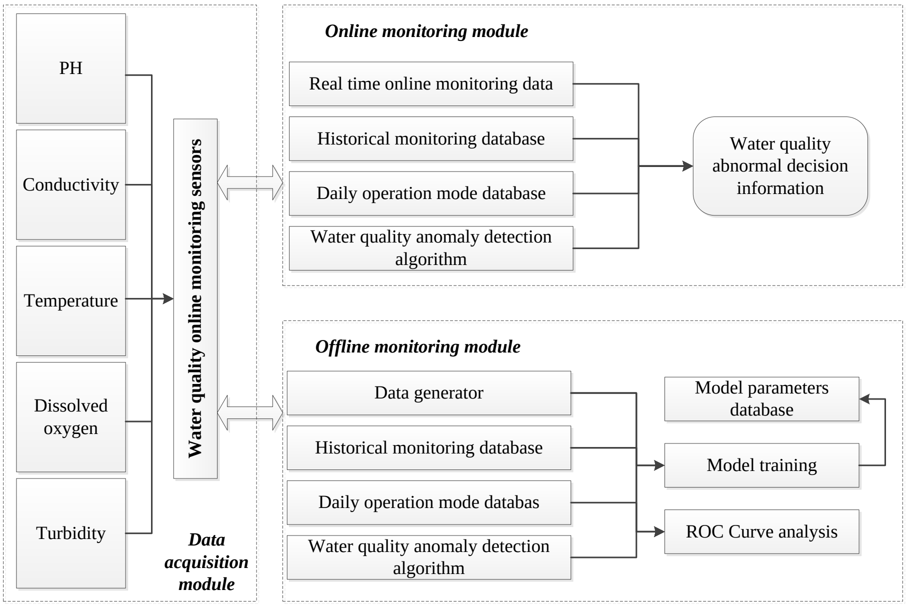


Fig.2. Framework of the water quality monitoring system

As is shown in Fig.2, the framework of the water quality monitoring system is illustrated. The proposed water quality monitoring system made up of an online monitoring module and an offline monitoring module. PH, conductivity, temperature, dissolved oxygen, turbidity of water are all collected from various sensors. The main functions of our system include 1) data acquisition, 2) data transmission and control, 3) real time online detection of water quality and 4) offline water quality analysis. The data acquisition module is designed to control the water quality sensor in accordance with set sampling intervals. In addition to data transmission, the control module is responsible for configuring sensor parameters, and then transmits real-time water quality information and control commands.

The online monitoring module in particular is the key component of the water quality monitoring system, and it is made up of the integrated anomaly detection algorithms library, daily operation pattern database, historical database of monitoring and real time water quality monitoring information. The offline monitoring module is comprised of an historical database and a simulation data generator, and it can simulate various situations of water quality anomalies. Moreover, this module can mine useful information from historical patterns, and then establish some basic information databases for the online module.

3. Monitoring Water Quality By Multi-Sensor Fusion Based On Dempster-Shafer Evidence Theory

The Dempster-Shafer evidence theory refers to a general framework for reasoning with uncertainty, with understood connections to other frameworks such as probability, possibility and imprecise probability theories [20, 21]. This theory was proposed by Arthur P. Dempster in the context of statistical inference, and then this theory was extended by Glenn Shafer into a general framework [22]. This theory allows one to integrate pieces of evidence from various sources and then reach a degree of belief considering all the available evidence [23]

In the Dempster-Shafer evidence theory, a fixed set of N mutually exclusive and exhaustive elements is defined as $\Theta = \{H_1, H_2, \cdots, H_N\}$. The frame of discernment Θ defines the working space. Information sources are able to distribute mass values. An information source allocates mass values to those hypotheses, for which it has direct evidence. The mass distribution for all hypotheses should satisfy the conditions as follows.

$$m(\phi) = 0$$

$$\sum_{A \in 2^{\Theta}} m(A_i) = 1$$

(1)

Then, mass distributions m_1, m_2 obtained from various sources are integrated via the Dempster's orthogonal rule. The new distribution $m = m_1 \oplus m_2$, and the following equation should be satisfied:

$$m(A_i) = \frac{1}{1-K} \cdot \sum_{A_p \cap A_q = A_i} m_1(A_p) \cdot m_2(A_q)$$

(2)

where K denotes a measure of conflict between different sources, and it is defined as follows.

$$K = \sum_{A_p \cap A_q = \Phi} m_1(A_p) \cdot m_2(A_q)$$

(3)

In particular, for the N information sources $\{B_1, B_2, \cdots, B_N\}$, its Dempster-Shafer rule is defined as follows.

$$m(A_i) = \frac{1}{1-K} \cdot \sum_{B_1 \cap B_2 \cap \cdots \cap B_N = A_i} m_1(B_1) \cdot m_2(B_2) \cdots m_N(B_N)$$

Subject to

$$K = \sum_{B_1 \cap B_2 \cap \cdots \cap B_N = \phi} m_1(B_1) \cdot m_2(B_2) \cdots m_N(B_N) < 1$$

(4)

Belief and plausibility functions are defined as follows.

$$Bel(A_i) = \sum_{A_j \subseteq A_i} m(A_j)$$

(5)

$$Pls(A_i) = \sum_{A_j \cap A_i \neq \phi} m(A_j)$$

(6)

Subject to

1) $$Bel(\phi) = Pls(\phi) = 0$$

(7)

2) $$Bel(A) \leq Pls(A)$$

(8)

3) $$Bel(A) + Bel(\overline{A}) \leq 1$$

(9)

4) $$Pls(A) + Pls(\overline{A}) \geq 1$$

(10)

We assume that all types of water quality indexes are independent of each other, and the final water quality detection decisions can be made by the above process.

4. Experiment

In this section, the experimental results of the multi-sensor data fusion method are analyzed. We introduce the experimental equipment and the experimental analysis platform in advance, and then provide the level of accuracy of water quality monitoring for different methods.

4.1 Experiment Settings

According to different occasions of usage, the water quality index sensor can be divided into online detection equipment and offline detection equipment. Online testing equipment ensures the real-time operation of the water quality anomaly detection system, which can measure the current water quality index data in a short sampling period. At present, the online water quality monitoring department has a complete set of online water quality testing instruments, including a pH meter, turbidity meter, conductivity meter and a COD rapid measuring instrument. Offline detection equipment is required for manual sample collection, followed by data analysis data in the laboratory. The detection results obtained by offline detection equipment are more accurate than online monitoring devices. The water quality index and related detection methods are listed in Table.1 as follows.

Table 1. Water quality indexes and detection method

Water quality index	**Detection method**
Turbidity (NTU)	Colorimetry method
Conductivity (us/cm)	Electrode method
Residual chlorine (mg/L)	DPD method
Total chlorine (mg/L)	DPD method
Sulfate (mg/L)	Surfa Ver method
ammonia nitrogen (mg/L)	Nessler method
Dissolved oxygen (mg/L)	Fluorescence method
pH	Electrode method
TOC(mg/L)	Direct method
COD(mg/L)	Reactor Digestion method

4.2 Experimental Results

Firstly, we compare three different concentrations of potassium ferricyanide and ferric ammonium sulfate abnormity detection rate, and experimental results are shown in Table.1. Fuzzy c-mean (FCM) [24] is used to make a performance comparison with ours. The FCM algorithm proposed by Dunn and improved by Bezdek is widely utilized in the clustering methods. FCM is comprised of data points divided into c clusters and utilizing some criteria for the optimization of an objective function [25, 26].

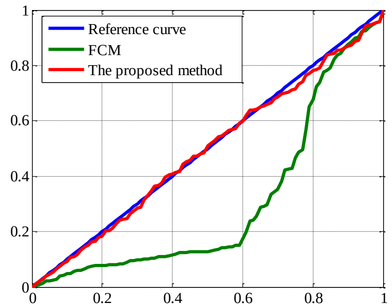
Table 2. Pollutant detection rate for different methods

Pollutant Name	**Method**	**pollutant concentration**	**Detection rate**
Potassium ferricyanide	FCM	0.4	0.21
		0.6	0.76
		0.8	0.83
	The proposed method	0.4	0.22
		0.6	0.81
		0.8	0.88
Ammonium ferric sulfate	FCM	0.8	0.72
		1.0	0.84
		1.2	0.96

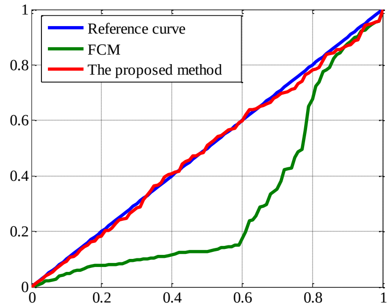
	The proposed method	0.8	0.73
		1.0	0.89
		1.2	0.96

Table. 2 reveals that both methods perform better when the pollutant concentration increases. Moreover, the proposed method can achieve a higher detection rate than FCM for the given two types of pollutants.

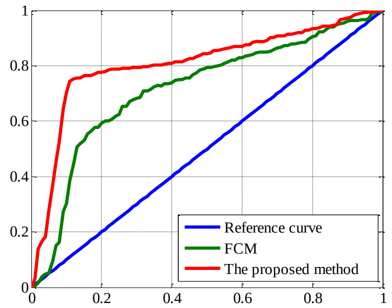
Secondly, we provide a ROC curve for different methods with different abnormal intensity (0.8, 1.0 and 1.2). Fig. 3 shows the ROC curve using the two water quality indexes of turbidity and conductivity.



(a) Abnormal intensity is 0.8



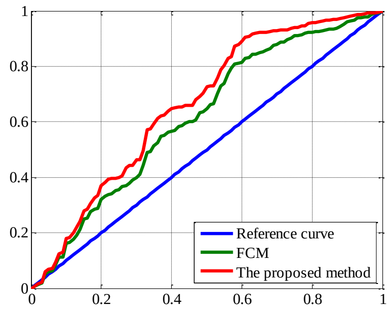
(b) Abnormal intensity is 1.0



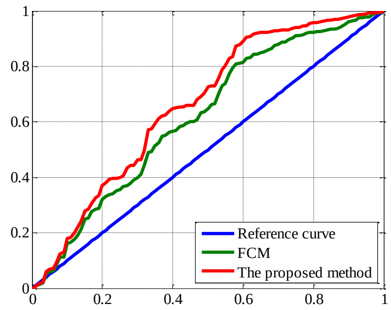
(c) Abnormal intensity is 1.2

Fig.3. ROC curve with two water quality indexes

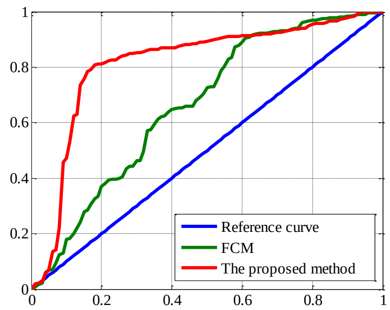
To test the water quality monitoring capability more accurately, we test the performance of water quality monitoring using all the indexes in Table.1. Fig. 4 shows the experimental result.



(a) Abnormal intensity is 0.8



(b) Abnormal intensity is 1.0



(c) Abnormal intensity is 1.2

Fig.4. ROC curve with multiple water quality indexes

Integrating all the experimental results together reveals that the proposed method can more accurately detect pollutants in water. The reason lies in that our proposed data fusion method based on Dempster-Shafer evidence theory can effectively fuse the information collected from different wireless sensors, and all water quality indexes can make contributions to the final water quality monitoring results.

5. Conclusion

This paper focuses on the problem of the water quality monitoring system, which is of great importance in water resources protection. This paper introduces the Dempster-Shafer evidence theory in water pollutants detection. The proposed water quality monitoring system contains an online monitoring module and an offline monitoring module. The water quality monitoring decisions can be derived from fusing various types of water quality indexes. In the end, experimental results demonstrate the effectiveness of the proposed method.

References

1. W.W Yan, J.L Li, X.H Bai, Comprehensive assessment and visualized monitoring of urban drinking water quality, 2016, Chemometrics and intelligent laboratory systems, no. 155, pp 26-35.
2. G. Durrieu, Q.K Pham, A.S Foltete, V. Maxime, I. Grama, L.T Veronique, H. Duval, J.M Tricot, C.B Naceur, O. Sire, Dynamic extreme values modeling and monitoring by means of sea shores water quality biomarkers and valvometry, 2016, Environmental monitoring and assessment, vol. 188, no. 7, pp 1-8.
3. K.E McCracken., S.V Angus., K.A Reynolds, J.Y Yoon, Multimodal Imaging and lighting bias correction for improved mu PAD-based water quality monitoring via smartphones, 2016, Scientific reports, no. 6, pp 27529.
4. M. Kim, Y. Kim, H. Kim, W. Piao, C. Kim, Enhanced monitoring of water quality variation in Nakdong River downstream using multivariate statistical techniques, 2016, Desalination and water treatment, vol. 57, no. 27, pp 12508-12517.
5. F. Ge, Y.N Wang, Energy efficient networks for monitoring water quality in subterranean rivers, 2016, Sustainability, vol. 8, no. 6, pp 526.
6. D.C Barrett, A.E Frazier, Automated method for monitoring water quality using landsat imagery, 2016, Water, vol. 8, no. 6, pp 257.
7. I.B. Roll, R.U Halden, Critical review of factors governing data quality of integrative samplers employed in environmental water monitoring, 2016, Water research, no. 94, pp 200-207.
8. J.F. Griffith, S.B. Weisberg, B.F Arnold, Y.P Cao, K.C. Schiff, J.M.J Colford., Epidemiologic evaluation of multiple alternate microbial water quality monitoring indicators at three California beaches, 2016, Water research, no. 94, pp 371-381.
9. M.Q Wu, C.Y Wu, W.J Huang, Z. Niu, C.Y Wang, W. Li, P.Y Hao, An improved high spatial and temporal data fusion approach for combining Landsat and MODIS data to generate daily synthetic Landsat imagery, 2016, Information fusion, no. 31, pp 14-25.
10. K.M. Nunes, M.V Andrade, A.M Santos Filho, M.C. Lasmar, M.M Sena, Detection and characterisation of frauds in bovine meat in natura by non-meat ingredient additions using data fusion of chemical parameters and ATR-FTIR spectroscopy, 2016, Food chemistry, no. 205, pp 14-22.
11. S. Kumar, R.M., Hegde, Multi-sensor data fusion methods for indoor localization under collinear ambiguity, 2016, Pervasive and mobile computing, no. 30, pp 18-31.

12. E. Borras, J. Ferre, R. Bogue, M. Mestres, L. Acena, A. Calvo, O. Busto, Prediction of olive oil sensory descriptors using instrumental data fusion and partial least squares (PLS) regression, 2016, Talanta, no. 155, pp 116-123.
13. E. Borras, J. Ferre, R. Boque, M. Mestres, L. Acena, A. Calvo, O. Busto, Olive oil sensory defects classification with data fusion of instrumental techniques and multivariate analysis (PLS-DA), 2016, Food chemistry, no. 203, pp 314-322.
14. A. Schoech, A. Salvadori, S. Carmignato, E. Savio, Enhancing multisensor data fusion on light sectioning coordinate measuring systems for the in-process inspection of freeform shaped parts, 2016, Precision engineering-journal of the international societies for precision engineering and nanotechnology, no. 45, pp 209-215.
15. D. Yang, H. Li, Y.G Hu, J. Zhao, H.W Xiao, Y.S Lan, Vibration condition monitoring system for wind turbine bearings based on noise suppression with multi-point data fusion, 2016, Renewable energy, no. 92, pp 104-116.
16. A. Schlicker, M. Michaut, R. Rahman, L.F.A. Wessels, OncoScape: Exploring the cancer aberration landscape by genomic data fusion, 2016, Scientific reports, no. 6, pp 28103.
17. B. Gomathi, P. Sakthivel, Enhancing information retrieval process using data fusion by ABC weighted based fuzzy retrieval in Health Care Analytic Software, 2016, Journal of medical imaging and health informatics, vol.6, no. 3, pp 863-868.
18. F.H Bijarbooneh, W. Du, E.C.H. Ngai, X.M Fu, J.C Liu, Cloud-assisted data fusion and sensor selection for internet of things, 2016, IEEE Internet of things journal, vol. 3, no. 3, pp 257-268.
19. Y.H Zhang, S. Prasad, Multisource geospatial data fusion via Local Joint sparse representation, 2016, IEEE Transactions on geoscience and remote sensing, vol. 54, no. 6, pp 3265-3276.
20. Q.F Zhou, H. Zhou, Q.Q Zhou, F. Yang, L.K Luo, T. Li, Structural damage detection based on posteriori probability support vector machine and dempster-shafer evidence theory, 2015, Applied soft computing, no. 36, pp 368-374.
21. Z.W Li, G.Q Wen, N.X Xie, An approach to fuzzy soft sets in decision making based on grey relational analysis and dempster-shafer theory of evidence: an application in medical diagnosis, 2015, Artificial intelligence in medicine, vol. 64, no. 3, pp 161-171.
22. M. Compare, E. Zio, Genetic algorithms in the framework of dempster-shafer theory of evidence for maintenance optimization problems, 2015, IEEE Transactions on reliability, vol. 64, no. 2, pp 645-660.

23. H.X Tang, A novel fuzzy soft set approach in decision making based on grey relational analysis and dempster-shafer theory of evidence, 2015, Applied soft computing, no. 31, pp 317-325.
24. K. Benmouiza, M. Tadj, A. Cheknane, Classification of hourly solar radiation using fuzzy c-means algorithm for optimal stand-alone PV system sizing, 2016, International journal of electrical power & energy systems, no. 82, pp 233-241.
25. O. Kesemen, O. Tezel, E. Ozkul, Fuzzy c-means clustering algorithm for directional data (FCM4DD), 2016, Expert systems with applications, no. 58, pp 76-82.
26. E. Esme, B. Karlik, Fuzzy c-means based support vector machines classifier for perfume recognition, 2016, Applied soft computing, no. 46, pp 452-458.