

Research Article

A Multiagent Dynamic Assessment Approach for Water Quality Based on Improved Q-Learning Algorithm

Jianjun Ni,¹ Li Ren,² Minghua Liu,¹ and Daqi Zhu³

¹ College of IOT Engineering, Hohai University, Changzhou 213022, China

² College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China

³ Laboratory of Underwater Vehicles and Intelligent Systems, Shanghai Maritime University, Shanghai 200135, China

Correspondence should be addressed to Jianjun Ni; njjhhuc@gmail.com

Received 30 March 2013; Revised 17 April 2013; Accepted 18 April 2013

Academic Editor: Guanghui Wen

Copyright © 2013 Jianjun Ni et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The dynamic water quality assessment is a challenging and critical issue in water resource management systems. To deal with this complex problem, a dynamic water assessment model based on multiagent technology is proposed, and an improved Q-learning algorithm is used in this paper. In the proposed Q-learning algorithm, a fuzzy membership function and a punishment mechanism are introduced to improve the learning speed of Q-learning algorithm. The dynamic water quality assessment for different regions and the prewarning of water pollution are achieved by using an interaction factor in the proposed approach. The proposed approach can deal with various situations, such as static and dynamic water quality assessment. The experimental results show that the water quality assessment based on the proposed approach is more accurate and efficient than the general methods.

1. Introduction

The assessment of water quality plays an essential role whether in engineering applications or in scientific research. However, due to the frequent occurrence of abrupt water pollution accident [1, 2], the general static assessment of water quality cannot meet the actual requirements any more. So, it is very important to assess the water quality of different regions accurately and dynamically, which is a hot field in water environment management system. The dynamic assessment of water quality can give out alarm timely before the pollutant reaches to some sensitive water regions. It is very helpful for these regions to make preparations and control water pollution effectively.

Various methods have been proposed to deal with the problem of water quality evaluation [3–5]. The main methods of the static water quality assessment include the comprehensive index method [6], fuzzy comprehensive evaluation method [7], BP neural network [8], and comprehensive water quality identification index method [9]. Although these methods have their own advantages, there are some shortcomings of these methods. For example, the calculation

of the comprehensive index assessment method is complex. The accuracy of the fuzzy comprehensive evaluation method is lower, which cannot give out the assessment for water in worse than Grade V level. The calculation model based on BP neural network is very complex, and the choice of training samples for BP neural network is difficult. The general comprehensive water quality identification index method cannot make specific analysis according to the characteristics of different water bodies because various indicators are considered to have the same effects in the water quality assessment.

The static evaluation methods can just assess water quality after the occurrence of water pollution. To solve this problem, more and more research has been focused on dynamic water quality assessment methods. For example, Yun et al. [10] evaluated the changes in river water quality during a period of time by using the probability transition matrix. Su et al. [11] studied the spatiotemporal patterns and source apportionment of pollution in Qiantang River (China) using neural-based modeling and multivariate statistical techniques. There is much research on dynamic water quality assessment methods, but few considered the problem of quick perception

for the abrupt water pollution. The methods to determine water quality of other regions according to the water quality change of a region in the same basin are few.

To control water pollution and improve water environment quality effectively, the trend of water pollution should be predicted accurately when water pollution accident occurs [12, 13]. Because it is a problem of complex system, the general method cannot deal with it efficiently. Recently, more and more focus has been put on the agent-based method, which is not only a feasible solution but also an efficient one [14, 15]. For example, Wen et al. [16] studied the problem of consensus in directed networks of multiple agents with intrinsic nonlinear dynamics and sampled-data information. Leon [17] proposed an interaction protocol for a task allocation system, which can reveal the emergent behaviors in social networks of adaptive agents. In the multiagent system, agent is defined as an entity, which has the capabilities of environment perception, problem solving, and communication with the outside world. Based on these features, the agent can be used to solve the complex problems in practice by sharing knowledge with each other [18]. To solve the problem of dynamic water quality assessment, a multiagent model of water environment is set up [19, 20], where different regions in the water environment are abstracted as various agents. An improved Q-learning algorithm is proposed to deal with the cooperation of multiagents and carry out the task of dynamic water quality assessment.

The paper is organized as follows. In Section 2, the dynamic assessment model for water quality based on multiagent technology is introduced. Section 3 presents the proposed Q-learning algorithm for water quality assessment. Some experiments are conducted, and the results are discussed in Section 4. At last, the conclusions are given in Section 5.

2. The Multiagent Dynamic Assessment Model for Water Quality

In this paper, the dynamic assessment for water quality is studied. The dynamic water quality assessment has attracted much attention due to its complexity and significance. There are two main problems that need to be solved in the task of dynamic water quality assessment. The first one is how to assess the water quality of different regions efficiently, when the indicators of water quality in all these regions are obtained. The other one is how to assess the water quality of other regions, when the indicators of water quality in only one region are obtained.

To achieve the task of dynamic water quality assessment, an assessment model for water quality based on multiagent technology is proposed in this paper, where the water environment is divided into different regional agents based on the requirement of administration. By the information exchange among these regional agents, the task of dynamic water quality assessment can be accomplished efficiently. In each agent, there is a water quality assessment model, which is defined as follows in this study:

$$P = \omega_1 P_1 + \omega_2 P_2 + \cdots + \omega_n P_n, \quad (1)$$

where P is the level of water quality, P_1, P_2, \dots, P_n are various indicators used to assess the water quality and $\omega_1, \omega_2, \dots, \omega_n$ are the weights for these indicators.

To assess water quality, more accurately, the water quality level P is defined as the following form:

$$P = X_1 \cdot X_2, \quad (2)$$

where $X_1 \in \{1, 2, \dots, 6\}$ is the water quality level, which is determined by the Chinese national standard for water quality (see the Environmental Quality Standards for Surface Water in China (GB3838-2002)), X_2 is the relative position of the water quality level between two adjacent water quality grades, and the symbol \cdot is a separating character, which has the same function as plus. For example, when the water quality level $P = 2.5$, the mean is that the grade of this water quality is Grade II by the national standard, and the relative grade of the water quality is 0.5; namely, the water quality is at a relatively intermediate location between Grade II and Grade III. When the value of X_1 is 6, the water quality is worse than Grade V. To reduce the computation complexity, the value of X_2 is designed as a discrete value in this paper, namely, $X_2 \in \{0.0, 0.1, \dots, 0.9\}$.

In the water quality assessment model above, the weight for the i th indicator needs to be optimized based on the dynamic change of water environment. In this study, a Q-learning-based algorithm is proposed to deal with this problem, which will be introduced in detail at Section 3.

3. The Proposed Multiagent Q-Learning Algorithm

In the assessment model, the weights of various indicators need to be obtained. In general water quality assessment methods, these weights are always set by the experience. Recently, some artificial intelligent methods are introduced to optimize these weights, such as genetic algorithms and neural networks. However, those approaches cannot realize the information transmission and exchange among different regions. So, the weights obtained by those approaches are intrinsically static. To deal with this problem, the multiagent-based technology is introduced into the water quality assessment, and an improved Q-learning algorithm is proposed to realize the cooperation of multiagents. In general multiagent reinforcement learning, the Markov decision process is extended to realize the exchange learning for multiagent systems. In most of algorithms of multiagent reinforcement learning, it is required that each agent should know what action will be taken by other agents before it takes action. Thus, with the increase of the number of agents or the actions of each agent, it will cause that the state space of agent grows exponentially [21]. To solve these problems introduced above, an improved multiagent Q-learning algorithm is proposed in this study.

The Q-learning algorithm is a kind of reinforcement learning method by the trial-and-error method. Compared with other machine learning methods, the Q-learning algorithm can initiatively find which action will produce the greatest reward, instead of being told which action should

be done [22–24]. To improve the learning efficiency of the Q-learning algorithm, the number of action-state pairs and the searching in the action-state pairs should be reduced. In a multiagent system, an agent needs to keep track of its environment as well as other agents, so the convergency and learning speed are the problems that need to be solved at first in the multiagent Q-learning algorithm [25, 26]. Some improvements have been done on the multiagent Q-learning to deal with the convergence problem [27, 28]. However, there are still some problems of those approaches in the literatures such that most of those approaches do not consider the interactions among the agents. In this study, a fuzzy membership function with distinguish weight [29] is used to reduce the size of the action-state set. And a punishment mechanism [30] is used in the proposed algorithm to reduce the searching frequency. The proposed algorithm has some better performances than the general Q-learning algorithm such as the high learning speed and good convergence rate. Furthermore, an interactive factor is introduced into the proposed Q-learning algorithm, to realize the information transmission and the interaction among the agents in the system. The flow chart of the proposed Q-learning is shown in Figure 1. The proposed approach is presented in detail as follows.

3.1. The State Reduction Based on Fuzzy Membership Function.

In the state preprocessing module of the proposed Q-learning algorithm, a fuzzy membership function with distinguish weight is used to reduce the size of state-action sets by removing the superfluous or unrelated information from the system. The membership function is defined as follows in this paper:

$$y = \begin{cases} 0, & x \leq x_b \text{ or } x \geq x_{b+1}, \\ \frac{x_{b+1} - x}{x_{b+1} - x_b}, & x_b < x < x_{b+1}, \end{cases} \quad (3)$$

where x is the original state (namely, the evaluation indicator for water quality in this study), y is the distinguish weight value of the indicator within its membership domain (namely, the parameter X_2 in (2)), x_b is the demarcation point of a grade, and x_{b+1} is the demarcation point of the next grade. When the grade of x_b is g (namely, the parameter X_1 in (2)), the membership value of the evaluation indicator is $v = g + y$. Because the size of the distinguish weight value y is 10 and the number of the grade g is 6 in this paper, the total number of states is $z = 10 * 6 = 60$. By this way, the state space can be reduced obviously.

3.2. The Information Transmission Based on Interaction Factor.

To transfer the information among these agents in the system, a concept of the interaction factor (denoted by σ) is proposed in this paper, which can transfer the information of the key state to other agents for water quality assessment. In practice, the value of the interaction factor σ should be learnt by experience. In this paper, it is calculated by

$$\sigma = \exp(-Kd), \quad (4)$$

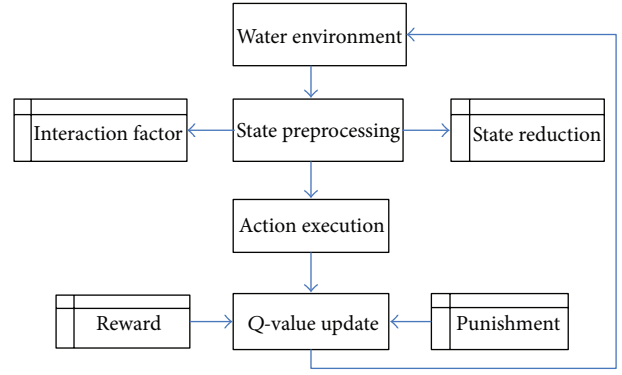


FIGURE 1: Flowchart of the proposed Q-learning algorithm for multiagent system.

where d is the distance between two agents, which can be an abstract concept or an actual physical distance and K is the attenuation coefficient, which can be calculated by the least square method:

$$\text{Min}_K F = \sqrt{\sum (s_c - s_r)^2}, \quad (5)$$

where s_c is the value of next state obtained by the proposed algorithm at the given K . Moreover s_r is the actual value of the next state. The value of K obtained is the attenuation coefficient, when the function F arrives to the minimum value. Based on (4) and (5), the time when the water quality grade of one region will be reached to the highest value (namely, the water quality will be worst) can be obtained. Then, we can make some preparations to prevent the water pollution for some sensitive areas.

3.3. The Action Execution Module. In the action execution module, the regional agents select their actions by the softmax strategy [31], which is defined as

$$P(a | s, Q) = \frac{e^{Q(s,a)/\tau}}{\sum_{a' \in A} e^{Q(s,a')/\tau}}, \quad (6)$$

where a is the action of agents, which is to increase or decrease the weights of indicators in this study, τ is the simulated annealing temperature parameter, which is used to control the searching rate, and $Q(s, a)$ is the Q-value function for the action-state pair. To reduce the searching times in the action-state set and accelerate the learning rate of the proposed algorithm, a punishment mechanism is introduced into the proposed algorithm. Then, the Q-value function $Q(s, a)$ is separated into a punishment Q-value and a reward Q-value function, respectively. The update algorithm of the punishment Q-value is

$$Q_c(s_t, a_t) = (1 - \alpha) Q_c(s_t, a_t) + \alpha [c + \gamma Q_{c \max}(s_{t+1}, b)]. \quad (7)$$

And the update algorithm for the cumulative reward Q -value is:

$$Q_r(s_t, a_t) = (1 - \alpha) Q_r(s_t, a_t) + \alpha [r + \gamma Q_{r \max}(s_{t+1}, b)], \quad (8)$$

where $\alpha \in (0, 1)$ is learning rate, $\gamma \in (0, 1)$ is discount factor, and r and c are the reward value and the punishment value, respectively.

3.4. The Work Flow of the Proposed Approach. The work flow of the proposed approach for dynamic water quality assessment is summarized as follows.

- (1) The initial state sets should be obtained from the water quality monitoring system, which is denoted as $S_1 = \{S_{11}, S_{12}, \dots, S_{1n}\}$, where S_{1i} is the actual concentration value of the i th indicator. The action set of agent is $A = \{\omega_i + 1, \omega_i - 1\}$, where ω_i is the weight of the i th indicator.
- (2) Initialize the value of Q_c and Q_r to 0, select a key state from the state sets, and initialize the interaction factor σ for this state to 0.
- (3) Reduce the initial state sets by (3), and a new state set can be obtained, which is denoted as $S_2 = \{S_{21}, S_{22}, \dots, S_{2n}\}$, where S_{2i} is the concentration value of the i th indicator after being processed.
- (4) Each agent calculates the real-time state P_{key} for the key state, based on the state sets after being processed. To easily compute and without losing generality, the state P_{key} is calculated by $P_{\text{key}} = S_{2\text{key}} + \sigma$, where $S_{2\text{key}}$ is the key state after being processed. The interaction factor σ can be obtained by (4) and (5).
- (5) Each agent selects the optimal action under current state s according to (6) and gets the next state s' after executing the action a . Then, a reward r and punishment c can be obtained from the environment feedback.
- (6) Calculate the cumulative punishment value Q_c by (7). If $Q_c \geq \theta$, then select a new action from the action sets (where θ is the upper limit for the punishment value and $\theta = 50$ in this paper). If $Q_c < \theta$, then set $Q_r(s_{t+1}, b)_{\max} = \max_{b \in A} Q_r(s_{t+1}, b)$. By (8), the value of Q_r can be obtained.
- (7) Repeat steps (5) and (6) to find out the weights of each group indicators, and calculate the average value of the weight for each indicator. Then, the water quality can be assessed by (1).

4. Experimental Studies

In order to test the performance of the proposed approach, some experiments are conducted. In these experiments, a water area of a lake is studied, which has six different regions (see Figure 2). The task of these experiments is to assess the water quality of the six regions. The pollution sources include the industrial pollution source, the agricultural pollution

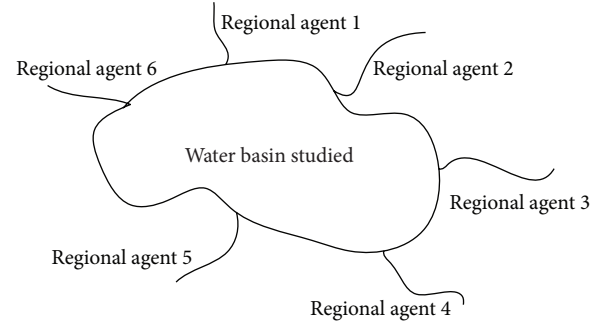


FIGURE 2: The schematic drawing of the water area studied.

source, and the domestic pollution source. According to the characteristics of the water area, it is assumed that the main pollution indicators are Permanganate Index (COD_{Mn}), Total Nitrogen (TN), and Total Phosphorus (TP); namely, the initial state set is $S_1 = \{S_{1C}, S_{1N}, S_{1P}\}$. Then, the reduced state sets can be obtained based on the membership function (3); that is, $S_2 = \{S_{2C}, S_{2N}, S_{2P}\}$. The assessment model for water quality is $P = \omega_1 P_C + \omega_2 P_N + \omega_3 P_P$. In this paper, two experiments were conducted, where the interaction factor σ is set as $\sigma = 0$ and $\sigma \neq 0$ to test the performance of the proposed approach in the static and dynamic assessment, respectively.

4.1. Static Water Quality Assessment ($\sigma = 0$). In this experiment, the interaction factor σ is set as 0, which means that there is no information transmission among these regional agents in the water area. Each regional agent assesses its own water quality based on the monitoring data of various indicators. The training data set for the Q-learning algorithm is shown in Table 1, which is used to learn the weights ω_1 , ω_2 , and ω_3 for COD_{Mn} , TN, and TP, respectively.

The training data of S_{1C} , S_{1N} , and S_{1P} in Table 1 are collected from the monitoring points for each regional agent. The value of the S_{2C} , S_{2N} , and S_{2P} is the corresponding value of the three indicators reduced by the membership function. P is the water quality assessed by water quality experts. From these training data, the optimal weights for the three indicators can be obtained, which are $\omega_1 = 0.5$, $\omega_2 = 0.3$, and $\omega_3 = 0.2$. Based on these optimal weights above, the water quality of different regional agents can be assessed. To show the advantages of the proposed Q-learning approach (QL), it is compared with the approach based on the fuzzy comprehensive evaluation method (FC) and the comprehensive identification index evaluation method (CI). The test data and the water quality assessment results are shown in Table 2, where the test data are the indicator data collected in each monitoring point.

The results in Table 2 show that the assessment results of the water quality are almost the same by the three methods (see the water quality of the regional agent 1, agent 2, agent 3, and agent 4). The water quality assessment result for the regional agent 5 shows that the water quality assessment by the proposed method is more accurate than the method based on the fuzzy comprehensive evaluation method (FC). The proposed approach can not only give out the water quality

TABLE 1: The training data set for the Q-learning algorithm.

S_{1C}	S_{1N}	S_{1P}	$P_C = S_{2C}$	$P_N = S_{2N}$	$P_P = S_{2P}$	P
1.2	0.1	0.01	1.4	1.5	1.5	1.5
2.4	0.4	0.08	2.4	2.7	2.5	2.5
4.6	0.6	0.12	3.6	3.4	3.4	3.6
6.8	1.3	0.23	4.4	4.8	4.6	4.5
10.5	1.6	0.34	5.2	5.4	5.8	5.4
16.9	2.8	0.68	6.3	6.8	6.6	6.4

TABLE 2: The test data and results of water quality assessment.

Regional agent	COD _{Mn} (mg/L)	TN (mg/L)	TP (mg/L)	The assessment results		
				QL	FC	CI
Agent 1	2.605	0.356	0.056	2.5	II	2.4
Agent 2	2.597	0.481	0.123	2.6	II	2.8
Agent 3	2.536	0.224	0.124	2.6	II	2.5
Agent 4	6.625	1.434	0.321	4.4	IV	4.7
Agent 5	17.603	2.741	0.646	6.5	V	6.5
Agent 6	10.618	1.914	0.788	5.4	V	6.0

grade but also evaluate the pollution degree of the water in this grade. In addition, the proposed approach can assess the water quality which is worse than Grade V (see the assessment results for agent 5 in Table 2). The assessment results for agent 6 show that the assessment based on the comprehensive water quality identification index method will become incorrect, when some indicators exceed the range in the national standard. Because the weight of each indicator is considered in the assessment model, the results based on the proposed approach are more accurate.

The results of this experiment show that the proposed Q-learning approach can assess the water quality accurately and can deal with some abnormal conditions such that some indicators become abnormal. Furthermore, the proposed approach can assess the water quality of worse Grade V.

4.2. Dynamic Water Quality Assessment ($\sigma \neq 0$). To test the performance of the proposed approach in the dynamic water quality assessment task, this experiment is conducted. In this experiment, an abrupt water pollution occurs in the regional agent 1, which is an industrial pollution, and the main contamination in the waste water is COD_{Mn}. So, the interaction factor σ is used to transfer the concentration information of COD_{Mn} among these regional agents. In the dynamic water quality assessment model, $P_C = S_{2C} + \sigma$, $P_N = S_{2N}$, and $P_P = S_{2P}$, respectively.

In order to have an easy analysis, the assumptions in this experiment are as follows. (1) The value of σ is only related to the physical distance among the regional agents. (2) The change step of K in (4) is assumed as 0.1. (3) The water speed is set as 0.02 km/h and assumed as fixed. (4) The water quality of each regional agent is known before the occurrence of the abrupt water pollution accident, which is set as the same data in the first experiment (see the water quality assessed by the proposed approach in Table 2).

TABLE 3: The actual monitoring concentration of COD_{Mn} for each agent.

Regional agent	The distance d To agent 1 (km)	The actual concentration of COD _{Mn} (mg/L)
Agent 1	0	6.00
Agent 2	2	3.08
Agent 3	4	1.84
Agent 4	6	0.48
Agent 5	5	1.10
Agent 6	3	2.33

In this experiment, the actual concentration of COD_{Mn} in the six regional agents before this abrupt water pollution accident and the physical distance between other agents to agent 1 are listed in Table 3. The value of K can be calculated by (5) based on the information of Table 3, where K is 0.4. With this K and d , the interaction factor σ of each agent can be obtained by (4). After the water pollution accident occurs, the concentration of COD_{Mn} in the regional agent 1 increases by 1.8 mg/L. Based on the proposed approach, the change of the COD_{Mn} concentration in other regional agents and the time when the COD_{Mn} concentration reaches to the highest value are shown in Figure 3. The results of the dynamic water quality assessment for each regional agent are shown in Figure 4.

The results in Figure 4 show that the water quality becomes worse too, when the concentration of COD_{Mn} in the regional agent 1 increased. This experimental results show that the proposed approach can give out the water quality assessment for different regions in the same water area, when there is just some information about the concentration of indicator in one region. Furthermore, the proposed approach can calculate the time when the concentration of the indicator will reach to the highest value. This performance is very

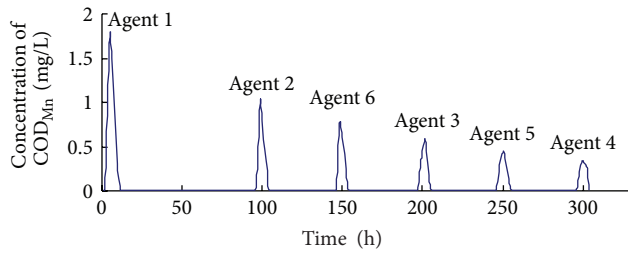


FIGURE 3: Changes of water quality and the diffusion time of pollutant in each agent.



FIGURE 4: The dynamic assessment results of water quality for different agents.

important for the sensitive regions to prepare for the water pollution control.

5. Conclusions

The dynamic water quality assessment for a whole water basin has been investigated. A water assessment model based on multiagent technology is set up, and an improved multiagent Q-learning algorithm is proposed. The proposed approach can deal with various situations. It can deal with the water quality assessment at the static situations, and the assessment results are more accurately than the general methods. In addition, it can deal with the dynamic water quality assessment, which is very important for the water pollution prewarning and control. The feasibility and efficiency of the proposed approach have been discussed and illustrated through experimental studies. The results show that the proposed approach can assess the water quality efficiently, without any complex mathematical model nor any prior knowledge about the water environment. The proposed approach is applicable to other real-time cooperative tasks of multiagent systems, such as the fire disaster response for wide tracts of forest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (61203365), the Jiangsu Province Natural Science Foundation (BK2012149), the Open Fund of Changzhou Key Laboratory of Sensor Networks and Environmental Sensing (CZSN201102), and the Yangtze River

Delta Region scientific and technological project (1259-5810200).

References

- [1] Q.-G. Wang, X.-H. Zhao, W.-J. Wu, M.-S. Yang, Q. Ma, and K. Liu, "Advection-diffusion models establishment of water-pollution accident in middle and lower reaches of Hanjiang river," *Advances in Water Science*, vol. 19, no. 4, pp. 500–504, 2008.
- [2] P. Tang, L. Zhao, L. Ren, Z. Zhao, and Y. Yao, "Real time monitoring of surface water pollution using microwave system," *Journal of Electromagnetic Waves and Applications*, vol. 22, no. 5-6, pp. 767–774, 2008.
- [3] N. Pochai, "A numerical treatment of nondimensional form of water quality model in a nonuniform flow stream using Saul'yev scheme," *Mathematical Problems in Engineering*, vol. 2011, Article ID 491317, 15 pages, 2011.
- [4] S. Rodgher, H. de Azevedo, C. R. Ferrari et al., "Evaluation of surface water quality in aquatic bodies under the influence of uranium mining (MG, Brazil)," *Environmental Monitoring and Assessment*, vol. 185, no. 3, pp. 2395–2406, 2013.
- [5] L. KeGang and H. KePeng, "Dynamic extension evaluation of soil and water environmental quality in metal mine and its improvement measures," *Research Journal of Chemistry and Environment*, vol. 16, no. 2, pp. 97–101, 2012.
- [6] Q. Wu, C. Zhao, and Y. Zhang, "Landscape river water quality assessment by nemerow pollution index," in *Proceedings of the International Conference on Mechanic Automation and Control Engineering (MACE '10)*, pp. 2117–2120, Wuhan, China, June 2010.
- [7] J. Shu, M. Hong, L. Liu, and Y. Chen, "A water quality monitoring method based on fuzzy comprehensive evaluation in wireless sensor networks," *Journal of Networks*, vol. 7, no. 1, pp. 195–202, 2012.
- [8] S. Ni and Y. Bai, "Application of BP neural network model in groundwater quality evaluation," *System Engineering Theory and Practice*, vol. 20, no. 8, pp. 124–127, 2000.
- [9] Z.-X. Xu, "Comprehensive water quality identification index for environmental quality assessment of surface water," *Journal of Tongji University*, vol. 33, no. 4, pp. 482–488, 2005.
- [10] Y. Yun, Z. Zou, W. Feng, and M. Ru, "Quantificational analysis on progress of river water quality in China," *Journal of Environmental Sciences*, vol. 21, no. 6, pp. 770–773, 2009.
- [11] S. Su, J. Zhi, L. Lou, F. Huang, X. Chen, and J. Wu, "Spatio-temporal patterns and source apportionment of pollution in Qiantang River (China) using neural-based modeling and multivariate statistical techniques," *Physics and Chemistry of the Earth*, vol. 36, no. 9–11, pp. 379–386, 2011.
- [12] W. Sun and Z. Zeng, "City optimal allocation of water resources research based on sustainable development," *Advanced Materials Research*, vol. 446–449, pp. 2703–2707, 2012.
- [13] G.-H. Wei, F. Liu, and L. Ma, "Fuzzy optimization of water resources project scheme based on improved grey relation analysis," in *Proceedings of the 3rd International Conference on Computer Research and Development*, vol. 4, pp. 333–336, Shanghai, China, 2011.
- [14] Y. Cao, W. Yu, W. Ren, and G. Chen, "An overview of recent progress in the study of distributed multi-agent coordination," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 1, pp. 427–438, 2013.

- [15] C. Li, F. Wang, X. Wei, and Z. Ma, "Solution method of optimal scheme set for water resources scheduling group decision-making based on multi-agent computation," *Intelligent Automation and Soft Computing*, vol. 17, no. 7, supplement 1, pp. 871–883, 2011.
- [16] G. Wen, Z. Duan, W. Yu, and G. Chen, "Consensus of multi-agent systems with nonlinear dynamics and sampled-data information: a delayed-input approach," *International Journal of Robust and Nonlinear Control*, vol. 23, no. 6, pp. 602–619, 2013.
- [17] F. Leon, "Emergent behaviors in social networks of adaptive agents," *Mathematical Problems in Engineering*, vol. 2012, Article ID 857512, 19 pages, 2012.
- [18] J. Wang, K. Gwebu, M. Shanker, and M. D. Troutt, "An application of agent-based simulation to knowledge sharing," *Decision Support Systems*, vol. 46, no. 2, pp. 532–541, 2009.
- [19] J. Ni, C. Zhang, and L. Ren, "An intelligent decision support system of lake water pollution control based on multi-agent model," in *Proceedings of the International Conference on Computational Intelligence and Security (CIS '09)*, vol. 1, pp. 217–221, Beijing, China, December 2009.
- [20] J. Ni, M. Liu, J. Fei, and H. Ma, "Reinforcement learning based multi-agent cooperation for water price forecasting decision support system," *Information-An International Interdisciplinary Journal*, vol. 15, no. 5, pp. 1889–1899, 2012.
- [21] M.-L. Xu and W.-B. Xu, "Fuzzy Q-learning in continuous state and action space," *Journal of China Universities of Posts and Telecommunications*, vol. 17, no. 4, pp. 100–109, 2010.
- [22] S. Zheng, J. Han, X. Luo, and J. Jiang, "Research on cooperation and reinforcement learning algorithm in multi-agent systems," *Pattern Recognition and Artificial Intelligence*, vol. 15, no. 4, pp. 453–457, 2002.
- [23] A. Bonarini, A. Lazaric, F. Montrone, and M. Restelli, "Reinforcement distribution in fuzzy Q-learning," *Fuzzy Sets and Systems*, vol. 160, no. 10, pp. 1420–1443, 2009.
- [24] X. Xu, C. Liu, S. X. Yang, and D. Hu, "Hierarchical approximate policy iteration with binarytree state space decomposition," *IEEE Transactions on Neural Networks*, vol. 22, no. 12, part 1, pp. 1863–1877, 2011.
- [25] M. L. Littman, "Value-function reinforcement learning in Markov games," *Cognitive Systems Research*, vol. 2, no. 1, pp. 55–66, 2001.
- [26] X. Xu, D. Hu, and X. Lu, "Kernel-based least squares policy iteration for reinforcement learning," *IEEE Transactions on Neural Networks*, vol. 18, no. 4, pp. 973–992, 2007.
- [27] J. Wu, X. Xu, P. Zhang, and C. Liu, "A novel multi-agent reinforcement learning approach for job scheduling in Grid computing," *Future Generation Computer Systems*, vol. 27, no. 5, pp. 430–439, 2011.
- [28] K. Fujita and H. Matsuo, "Multiagent reinforcement learning with the partly high-dimensional state space," *Systems and Computers in Japan*, vol. 37, no. 9, pp. 22–31, 2006.
- [29] K.-D. Liu, Y.-J. Pang, and W.-G. Li, "Membership transforming algorithm in multi-index decision and its application," *Acta Automatica Sinica*, vol. 35, no. 3, pp. 315–319, 2009.
- [30] X.-H. Zhao, K.-K. Zhao, Q.-Q. Wang, and F.-Q. Ma, "Research and application of reinforcement learning based on constraint MDP in coal mine," in *Proceedings of the WRI World Congress on Computer Science and Information Engineering (CSIE '09)*, vol. 4, pp. 687–691, Los Angeles, Calif, USA, March 2009.
- [31] V. Derhami, V. J. Majd, and M. N. Ahmadabadi, "Exploration and exploitation balance management in fuzzy reinforcement learning," *Fuzzy Sets and Systems*, vol. 161, no. 4, pp. 578–595, 2010.

