Research on Water Supply Reservoir Operating Rules Extraction Based on Artificial immune Recognition System

Xianfeng Zhang[1,2] Xiaolin Wang[2] Zhengjie Yin[3] Huiqiang Li[1]

[1]School of Civil Engineering & Mechanics, Huazhong University of Science & Technology
Wuhan 430074, P. R. China
[2]School of Management, University of Geosciences, Wuhan 430073, P. R. China
[3] Water Resource Department, Yangtze River Scientific Research Institute, Wuhan 430010, P. R. China

Abstract

In this paper artificial immune recognition system (AIRS) is employed as an emerging technique of data mining to extract the reservoir operating rules with a case of water supply reservoir, and we mainly focus on the impacts of learning mechanisms of AIRS on the obtained operation rules, therefore the mechanisms are explored and different gene encodings, as knowledge representatives, and the uncertainties of annual hydrological conditions (AHC), one attribute of the operating data, are considered. In order to further illuminate the learning capabilities, the classification results of the rules through AIRS and RBF networks are compared, indicating AIRS can be better for mining the reservoir operating rules which are of more transparent and interpretive, and can be dynamically updated.

Keywords: Water supply, Data mining, Operating rules, Information entropy, AIRS

1. Introduction

Recently novel computational intelligent approaches related to artificial neural networks, genetic algorithms and artificial immune systems are gaining popularity and have been adopted progressively by the data mining as a new theory or technique in the areas of the water resources management[1], offering a novel approach of obtaining the operating rules. However in practice the techniques need to be improved further and expected to consideration and employment of the existing information to great degree and fusion of the experience or knowledge of the decision-maker. In addition, a mass of data and information are accumulated gradually in the long-term process of the reservoir operations and eventually form enormous information database which encourages the application of data mining for the reservoir operations.

Of several related definitions of data mining one that is most appropriate for real-world applications is given by Fayyad et al. (1996): Data mining is the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data [2]. In other words, data mining is capable of extracting the hidden patterns and discovering the valued information or knowledge from the observed data. Currently the decision tree [3]-[4] and artificial neural networks [5]- [7] have been wildly applied to obtain the reservoir operating rules but the two techniques are limitations. The data to be mined, expressed as points in the high dimension state apace, are composed of operating data attributes and the corresponding release decisions. Although obtaining the rules easy to be interpreted and deduced, the decision tree, a traditional data mining technique, can merely make divisions parallel to the axes of the state space which leads to the limitation of effective divisions in the space when facing the complex non-linear mappings between the attributes of operating data and the related decisions. On the other hand, compared with the decision tree, the artificial neural networks can make more effective divisions but the size of the network, especially the hidden node number, are difficult to make sure. Another point noted especially is that with changes of environments and demands, operating data are subject to continuous modification and supplements. Current and future operation decisions are more dependent on these new data. Unfortunately the above-mentioned data mining methods could not make dynamic calibration or updating of the acquired operating rules using available new operating data [8].

Artificial immune system (AIS), implementing a learning technique inspired by the human immune system is a remarkable natural defense mechanism that learns about foreign substances and protects the bodies from the invasions of the microorganism, bacteria or virus, and applied wildly to the domains such as machine learning, abnormal detection, computer security, optimization and scheduling [9]. By far many artificial immune systems or models have been proposed, and one of these models, a typical model for unsupervised learning, is resource limited artificial immune system (RLAIS) which has been proposed by Timmis (2001)[10], followed by the artificial immune recognition(AIRS)[11], which is advanced by Watkins(2004) based on the Timmis' research on the RLAIS. Up to now AIS, mostly applied to solving the optimal operating results in the water resource management domains[12]-[13], but for the purpose of classification or pattern recognition have never been reported. Therefore, we will attempt to obtain the water supply reservoir operating rules using AIRS as a novel data mining technique and aim to explore the impacts of the learning mechanisms of AIRS on extracting the classification operating rules.

The remainders of this paper are organized as follows. Section 2 primarily explored the learning mechanism of AIRS. Section 3 lists the rule derivation process via AIRS in detail while discussing the effect of the different encodings of operating data for the AIRS and analyzing the obtained operating rules. Section 4 introduces the information entropy theory to AIRS to establish a IEAIRS to decrease the influence of the subjectivities and uncertainties from the determination of AHC. In order to further describe the performance of AIRS, the classification results of AIRS, IEAIRS and RBF are compared in section 5 and then conclusions about the advantages and disadvantages of the AIRS for the derivation of the rules are drawn while presenting the future research work in section 6.

2. Brief overview of learning mechanisms of AIRS

The human immune system is highly distributed, highly adaptive and self-organizing in nature, maintains a memory of foreign substances (called antigen) and has the ability to continually learn about new antigens[14]. Once the human body suffers from the attack of the antigen, the specific immune response of the immune system will be elicited and the immune cells of the body, called B lymphocyte cell which is produced by the bone marrow cell, are stimulated to secrete antibodies which undergo continuous evolution and can identify, match and eventually neutralize or eliminate the binding antigens. From an information-processing perspective, the patterns of the antigens are identified, learned and eventually remembered by the antibodies through the specific immune response, therefore this learning mechanisms of antibody against antigen are taken as an inspiration for the application of the machine learning, data mining and data analysis. Hunt et al. (1995, 1996) proposed an artificial immune network system for machine learning which succeeded in the recognition of promoters in DNA sequence [15]-[16], but the size of the network could not be controlled effectively. In order to solve this problem, the concepts of resource limited and the artificial recognition ball (ARB) were adopted by Timmis (2001) to advance a resource limited artificial immune system (RLAIS) for unsupervised clustering tasks[17]. Based on the Timmis' researches on the RLAIS, Watkin(2004) further puts forward to a new supervised learning classifier system, called artificial immune recognition system, for the classification and pattern recognition tasks. Although AIRS is inspired by the work on immune networks, the development of the system leads to the abandoning of the network principles in favor of a simple population-based model. Just as the RLAIS, the ARIS could be used to learn about the antigen not only one-shot, but also continuously. The main learning mechanisms of AIRS come from the thoughts of ARB, resource limited and memory pool.

In order to use AIRS as a data mining technique to effectively learn about and extract the patterns of the training data or antigen, the repertoire of the system should be complete, namely the there exists at least one antibody that can recognize any antigen, and the repertoire completeness of the system is achieved by the generalization and diversity of the antibodies. ARB of the AIRS, also known as B-cell, consists of an antibody, a count of the number of resources held by the cell, and the current stimulation value of the cell. The generalization of the cell is performed by the identifying any antigen entering the recognition radius of ε in the shape space, namely an ARB can acquire the common schema or feature from many different antigens(data), similar to the cross reaction in the human immune system. Once the antigen enters the AIRS, the ARB in the system becomes activated, undergoes clonal selection and somatic hyper-mutation and generates a great number of copies of the cell but not all of the copies could survive in the system. The idea of resource limitation is accepted by the AIRS and the total number of the system wide resource is set to a certain limit. Each ARB is allocated a number of resources based on its stimulation value and the clonal rate. In the AIRS, the stimulation of an ARB is based both on affinity to the antigen (training data) and the antigen class label, where highly stimulated ARBs are those of the same class as the antigen which are close to the pattern of antigen, or are of a different class and far from the pattern of the antigen. So the ARBs with high stimulation could compete for acquiring much more resources and survive, otherwise be removed from the system. This process illuminates the dynamic of the system, maintains the diversity of the ARB population and makes the patterns of antigens or training data to be identified.

Compared with the memory mechanisms of RLAIS that the patterns or characters of the antigen which have been remembered are maintained in the network, those of AIRS are different. The reason is that the ARBs against the antigens undergo the maturation of immune response and of which the highest stimulated one would be selected as a memory cell to enter the memory cell pool. On the other hand, another advantage is that AIRS can continuously learn about the new patterns of the antigens, namely once AIRS encounters new antigen or training data, the new pattern of antigens or data can be remembered and the old can never be forgotten.

3. Reservoir water-supply operating rules extraction procedures with AIRS

3.1. Description of operating rules

Water supply of the reservoirs needs to be comprehensively considered to meet the water demands of multi-objectives which have different requirements of priorities and guarantee rates. For instance, the domestic water demands of urban and rural inhabitants shall be satisfied firstly, while agricultural and industrial water demands as well as navigation requirements shall also be considered and taken care of. In this paper the water demands of three water supply objectives are considered[18], denoted as $D_{i,t}$ where i is water supply objective, equal to $1,2,3$, and t is water supply period of time, equal to $1,2,\ldots,T$, therefore $D_{1,t}$,$D_{2,t}$ and $D_{3,t}$ are listed in descending order of priorities of water supply. According to operation decision-making there are four possible operating patterns and only one of them shall to be selected to guide the release at any period of time. The four operating patterns are listed below: ①normal water supply for all the objectives. ②only limitation to $D_{1,t}$. ③limitation to both $D_{1,t}$ and $D_{2,t}$. ④limitation to the three objectives. Consequently, the decision issues of the water supply reservoir can be described briefly as that the water supply reservoir operating rules, obtained from the operation data, combined with current decision information, are served as the guidance to perform series of decisions for determining an operating pattern from the four patterns in order to minimize the shortage index of water.

3.2. Data mining procedures for the operating rules

Any sample of the operating dataset consists of five characteristic attributes of reservoir storage(S), operating period number(N), water demand(D), reservoir inflow (I) and annual hydrological condition (P), and the class label of operating pattern(C) related to the attributes. The operating patterns C, known as "true patterns"[7], obtained by the dynamic programming, are the optimal release decisions during the whole run-time of the reservoir. In AIRS, the gene representatives of the antigen and the antibody of the artificial recognition ball(ARB) are denoted as the rule formulation of "if (N, S, I, D, P), then C", and any rule is expressed as a six dimension vector in the shape space. The procedure of operating rule extraction is given in detail as follows.

Step1: Antigen presenting. This stage of the algorithm can primarily be viewed as a data pre-processing, where the training samples are regarded as antigens encoded in a proper manner and normalize all the antigen feature vectors so that the Euclidean distance between the feature vectors of any two items is in the range of [0, 1] while calculating the affinity threshold among the antigens.

Step2: Memory cell identification and ARB generation. Once the data pre-processing is completed, training will proceed as a one-shot incremental algorithm. Given a specific training antigen, find the memory cell of the same class as the antigen which is highly stimulated and selected to generate new ARBs through the clonal selection and hyper-mutation, and then those balls are placed into the population of pre-existing ARBs.

Step3: Competition for resources. ARBs compete for acquiring the resources based on the stimulation value to decide whether survival or not death and then the numbers of the cells in the system can be adjusted dynamically.

Step4: Operating rule extraction. The ARB of the same class and highest stimulation value as the antigen is selected as a candidate to enter the memory cell pool.

Step5: Once an antigen is dealt with by the AIRS, the procedure will go to the next antigen until all of antigens enter the system. Eventually the obtained memory set is the result of extraction of the water supply reservoir operating rules.

Data mining is the derivation of the potential useful information and knowledge hidden in the features of the operation data. Generally the feature attributes of the same data may be expressed as the binary, continuous, categorical, or hybrid representative, however different representative may lead to inductive bias of the AIRS [19], therefore it is necessary to explore the influences of different representatives of five feature attributes of the operation data on the operating rule extraction before AIRS is applied to the data mining of the rules.

3.3. Representatives choice of the operation data

Hamaker and Lois Boggess(2004) have performed series of experiences utilizing the standard data from the UCI machine learning repository to test the classification performances of AIRS, and the results demonstrated that different representatives related to the distance measures have great impact on the performances of AIRS[20]. Thus the continuous and hybrid representatives are accepted to denote attributes in this paper, where the distance of the former is calculated by the Euclidean distance measure whereas the latter by the heterogeneous value difference metric(HVDM). Five attributes of operation data or antigen (antibody), and the related representatives are listed in Table 1.

Ab(Ag)		continuous	hybrid
Condition	*N*	numerical	normal
	S	numerical	numerical
	I	numerical	numerical
	D	numerical	numerical
	P	numerical	normal
Action	*C*	normal	normal

Table 1: Different representatives of antibody (antigen).

The first 30 years samples (from 1956 to 1986) are used to conduct 10-way cross validations for the continuous and normal representatives and the results are listed in the Table2,3 respectively. Comparative of results demonstrates that the average accuracy of the latter is 2.6% more than that of the former because the operation data is dealt with by the Euclidean distance metric and the HVDM respectively. In this paper operating period, divided into twelve months, are denoted by the integers from 1 to 12, while the AHCs, represented as six kinds of conditions of wet year, above average year, average year, below average year, dry year and extremely dry year, are denoted by the integers from zero to five. When the operating period and AHCs are measured by the Euclidean distance, the two features, regarded as real values which are arbitrarily decided , could not reflects the influence between the two attributes and the operating patterns. On the contrary, when measured by the HVDM, the two attributes, although denoted by the integer, are still viewed as categorical values because the posterior probability of Bayesian theory is accepted by the HVDM to exploit the inherent relationship between the two attributes and the related operating patterns. Consequently, the average accuracy of the latter is higher than that of the former. Therefore the hybrid representatives and the HVDM are absorbed by the AIRS to calculate the affinity between the antigens and the antibodies.

Data items	1	2	3	4	5
Accuracy	0.833	0.917	0.917	0.722	0.722
Data items	6	7	8	9	10
Accuracy	0.889	0.806	0.75	0.806	0.778
Average accuracy			0.8168	Var.	0.0709

Table 2: 10-way cross validations(Euclidean distance).

Data items	1	2	3	4	5
Accuracy	0.861	0.917	0.917	0.806	0.722
Data items	6	7	8	9	10
Accuracy	0.944	0.806	0.833	0.806	0.806
Average accuracy			0.8418	Var.	0.0681

Table 3: 10-way cross validations (HVDM).

Primary analysis of results

The first 30 years operating data are used as training samples by the AIRS to obtain the operating rule set and the second 15 years data are utilized as testing samples to validate the classification capabilities of the rules. The run process of AIRS is controlled by the key parameters which are set visa experiments as follows:TotalNumberResource=500,HyperClonalRate=2,ClonalRate=10,Pm=0.2,TS=0.09,StimulationThreshold=0.76, and k=3.

Once accomplishment of training, 210 ARBs are obtained in the memory set of the water supply operating rules(small parts of the rules are listed in Table 4), and the compression ratio of the system to the training samples is 58.5%. When the obtained rules are applied further to test the training samples and the classification accuracy of the rules arrives at 83.3%. We can draw some interesting conclusions from the analysis of the table 6 as follows:

(1) The classification accuracy of the pattern① is highest, followed by the pattern④, and the accuracies of the pattern②and③ are lowest, which illuminates that the memory dataset of the operating rules can be good for the high generalization of the training samples with the pattern①or④ and poor for those with the pattern②or③. Statistical analysis demonstrates that the operating decision-making is prone to conservative, and there are no abnormal events taking place in the classification results, namely the differences between the patterns, determined by the rules, and the “true pattern” are within the range of two levels. Another point we should note is that identification accuracy of

AIRS reflects the overall effectiveness of the obtained rules, e.g., the operating data in December 1989 is closest to the three rules of 9,10,11 in the shape space(seen in Table 4) and the classification pattern of the data is codetermined by the voting numbers of the three rules. Although the true pattern of the data belong to the "true pattern"④ is misclassified by the AIRS into the pattern③, the three rules could provide the reference information for the further decision-making.

(2) There are six operating data with the "true pattern"① which are misclassified into the pattern②, although the minimal storage in the six data, counting for 78 percent of the reservoir capacity, is much more than the reservoir inflow and the water demand at this time. These cases usually take place under the circumstances of the below average or dry years, which further makes clear the AHCs have affect on the operating decision-making.

(3)There are eight operating data with the "true pattern"② which are misclassified into the pattern① and five into the pattern③, generally occurring under the circumstances of large storage, low inflow, and large ware demand. In these cases the influence of the reservoir inflow on the decisions can be ignored while the satisfaction of the water demand depends mainly on the reservoir storage.

(4) Four operating data with the "true pattern"③ which are misclassified into the pattern② and five into the pattern④. These cases occurs under the conditions of low storage (around 60% of the reservoir capacity), low inflow, whereas quite large water demand (around 84.5% of the maximal water demand), which leads to the shortage of the water demand because the storage could not be supplemented by the inflow and be affected by the AHCs of above averages or below average years.

(5) There are two operating data with the "true pattern"④ which are misclassified into the pattern③, taking place under the circumstances of the dry or extremely dry years and the quite small storage (around 27% of the reservoir capacity) while the operation periods are within the low water seasons, thereby both the reservoir storage and the AHCs have great impacts on the operation decision-making.

Based on analysis mentioned above, the AHCs have impact on all of the four operating patterns, so proper division for the hydrological year types helps to extract better rules form the operation data with the AIRS.

No.	*N*	*S*	*I*	*D*	*P*	*C*
1	8	0.983	0.097	0.982	1	①
2	3	0.873	0.073	0.800	3	②
3	6	0.649	0.076	0.445	2	③
4	6	0.843	0.020	0.479	2	①
5	7	0.731	0.325	0.854	1	②
6	9	0.881	0.033	0.554	3	①
7	4	0.694	0.112	0.833	3	②
8	4	0.815	0.086	0.818	4	①
9	12	0.336	0.029	0.001	5	④
10	10	0.336	0.014	0.068	5	③
11	10	0.402	0.077	0.157	5	③

Table 4: Small parts of operating rules in memory set.

Determination of the AHCs based on information entropy theory

The inflow of the reservoir is the principle source of the reservoir water supply, and changed greatly due to the variances of the AHCs. Generally speaking, the AHCs are determined only by the frequencies of the annual runoff, however the determination could not account for the influences of the AHCs on the operation decision-making while not directed specially at the reservoir operation due to the consideration of pure hydrology. On the other hand, it is uncertain and subjective for the operators to predict the AHCs next year in the run of the reservoirs. So it is necessary for us to seek for another new technique to exploit the internal relationships between the annual runoff and the operating patterns for determining the AHCs in favor of the operating decision-making from the operating decision data. Information entropy as the measures of the random signals can be applied to extract the information of operating decision-making hidden in the annual runoff as the hydrologic random variable. In this paper the supervised discrete method, FUSINTER proposed by Zighed (1998) [21], is adopted by AIRS to establish new model, information-entropy based AIRS (IEAIRS), to divide the 45 annual runoff into six intervals with the relation to the AHCs (illustrated in table 5), and the intervals by the FUSINTER are compared with those by the frequencies of the annual runoff (seeing the Fig. 1).

After the AHCs are determined by the FUSINTER, new operating data are obtained, thereby the new reservoir operating rule are be acquired by the IEAIRS and the classification accuracy of the new rule memory set is percent 86.1 which is 2.8% higher than that of the AIRS. The comparison analysis of the two results, seeing Fig. 1, show that the wet year interval of the former is wider than that of the latter but the above average year interval is narrower while the below average year interval is wider whereas the dry year interval is narrower. So the determination of the AHCs based on the information entropy theory can definitely divide the intervals of the 45-year annual runoff and further decrease subjectivity and uncertainty of the operators.

Interval	AHC	Symbol
[0.00～7945）	Extremely dry	0
[7945～8194）	Dry	1
[8195～8347）	Below average	2
[834～12279）	Average	3
[12279～12992）	Above average	4
[12992～+∝）	Wet	5

(Note：λ=1.0, α =0.95)

Table 5: The annual runoff Intervals and the AHC.

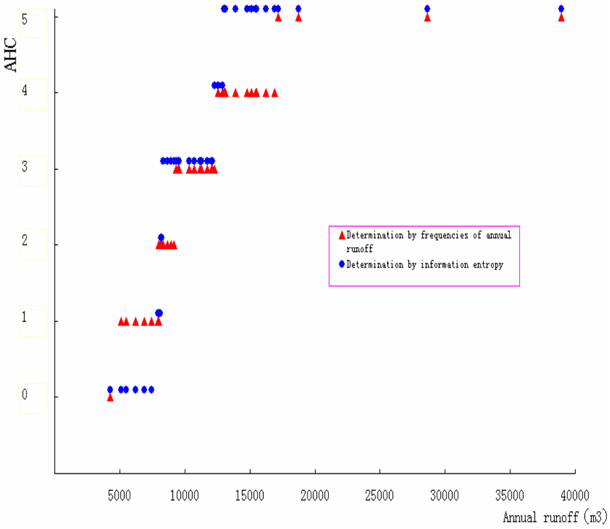

Fig.1: AHCs determined by information entropy and frequencies of annual runoff

4. Comparative analysis of the classification abilities of AIRS, IEAIRS and RBF for operating rules

Radius basis function network(RBF) is used as the technique of data mining to extract the water supply reservoir operating rules and the classification accuracy of the rules is percent 84(seeing the table 8), and the comparative analysis of the table 6, 7, and 8 illuminate the classification accuracy of the AIRS is near to that of the RBF and the accuracy of the IEAIRS is higher than that of the first two techniques while the AIRS accuracy of the operating data with the pattern ①and ④ are higher than that of RBF but lower than that of the data with the pattern ②and③. The explanations are given in detail as below:

(1) The information and knowledge contained in the operating data are acquired by the RBF through the training and eventually stored in the network nodes and the weights between the nodes. The operating data with the pattern ①and④ count for 48.3% and 3.9% of the overall data respectively, therefore the former data over-generalize the network but the latter can not carry out learning fully about the information of the decisions with the pattern④. So the pattern numbers of the training sample can affect the classification abilities of the RBF. On the contrary,, the learning abilities of the AIRS could not be subjective to the sample number during the training process and can perform the fine division in high dimension space due to depending mainly on the ARBs capturing the antigens (operating data) entering the radius of the ball and further extracting the common schema or features of the antigens.

(2) There are reasons about the low identification accuracies of the AIRS and IEAIRS for the operation data with the pattern②and③. One reason is that the data with pattern②and③ are easy to be interfered by its' neighbor data in the vector space because the identification regions of the neighbor ARBs usually overlap when these balls identifying the same data(antigen) Anther reason is that the matches between the rules in the memory set and the operating data, measured by the Euclidean distance or HVDM, ignore the interactions among the feature attributes of data to some degree (analyzed in Section 3), on the contrary the RBF, a learning method of connectionism，can implicitly expresses the information, knowledge or rules, so the nonlinear divisive capabilities of the RBF outperform those of the AIRS.

(3) Compared to the RBF, AIRS for the extraction of the operating rules has three advantages. The first one is that the size of the RBF is determined by trial and error but the parameters of the AIRS can be determined by the users in light of the problem-solving. The second one is that the operating information or rules acquired by the former are represented implicitly by the network and could not deliver more decision-making information to the operators. Especially, the operating rules in the memory set of the AIRS are expressed as definite formation of "IF-THEN" while an operating pattern of the operating data to be identified is codetermined by multiple rules in the memory set, so the operating rules acquired eventually by the AIRS are much more transparent and interpretative. The third one is that once the completeness of the training process of the RBF, the structure could not be changed, therefore if new reservoir operating data need to be trained by the network, the structure of the network will still be determined again by the trial and error, whereas the AIRS could continue learning the reservoir historical operating data including the present or future data due to the mechanisms of one-shoting, so the operating rules in the memory set could continue to be updated dynamically.

True pattern	Classification pattern				Classification accuracy
	①	②	③	④	
①	97	6			94.2%
②	8	26	5		66.7%
③		4	16	5	64%
④			2	11	84.6%

Table 6: Classification of AIRS.

True pattern	Classification pattern				Classification accuracy
	①	②	③	④	
①	98	5			95.1%
②	5	30	4		76.9%
③		4	16	5	64%
④			2	11	84.6%

Table 7: Classification of IEAIRS.

True pattern	Classification pattern				Classification accuracy
	①	②	③	④	
①	94	7	2		91.3%
②	2	32	4	1	82.1%
③	1	2	18	4	72%
④		2	4	17	63.6%

Table 8: Classification of RBF [7].

5. Conclusions and the future work

The AIRS is employed as a new technique of data mining to primarily explore the application of the extraction for the reservoir operating rules as a case of a water supply reservoir operation in this paper. We mainly focus on the influences of the different representations of the feature attributes of the operating rules and the different determinations of the AHCs on the derivation of operating rules. In order to deeply exploit the advantages of the AIRS learning performances for the extraction of the rules, Comparative analysis of the classification results of the AIRS and the RBF are carried out. The final analysis illuminates that the AIRS can be used as a new learning classifier to effectively extract the reservoir operating rules.

However there are some disadvantages about the AIRS for the extraction of operating rules, for examples, the redundancy of rules in the memory set with the increase of the memory cells, shortage of consideration of the weightiness of different features among the attributes of the rules, and lack of the associative representations between the features of the attributes and the operating patterns or among the features. Therefore, we will continue our research work as follows:(1) The co-stimulation between the T lymphocyte and the B lymphocyte will be introduced to the AIRS for regulating the weights of the features of the operating rules. (2) Self-adaptive mechanism will be adopted to adjust the radius of the ARBs to enhance the generalization capability of the AIRS.(3) In order to extract the more interpretative rules, the template instruction theory and the reinforcement learning strategy will be introduced.

Acknowledgement

We would like to thank professor Mo Hongwei, Harbin engineering university, for delivering some constructive advices to this paper and acknowledge the support of National Nature Science Foundation of (grant No.50479039)

References

[1] D.P. Solomatine, Applications of data-driven modeling and machine learning in control of water resources. *Computational intelligence in control*, pp.197-217, 2002.

[2] U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, *Advances in Knowledge Discovery and Data Mining*. AAAI Press and MIT Press, Chap.1:1-34, 1996.

[3] F.T. Bessler, D.A. Savic, G.A. Walters, Water reservoir control with data mining. *Journal of Water Resources Planning and Management*, *ASCE*, 129: 26-34, 2001.

[4] N.k. Ambujam, K. Venugopal, A data mining approach for deriving irrigation reservoir operatingrules.http://balwois.mpl.ird.fr/balwois/administration/full_paper/ ffp-643.pdf.

[5] T.R. Neelakantan, N.V. Pundarikanthan, Neural network-based simulation optimization model for reservoir operation. *Journal of Water Resources Planning and Management*, 126:57-64, 2000.

[6] V. Chandramouli, H. Raman, Multireservoir modeling with dynamic programming and neural networks. *Journal of Water Resources Planning and Management*, *ASCE*, 127:89-98, 2001.

[7] Z.J. Yin, X.L. Wang, et al. Water supply reservoir operating rules extraction based on data mining. *Systems Engineering-Theory & Practice*, 8:129-135, 2006.

[8] J.C. Xie, L.S. Suo, et al. Discussion on the calibration problem in dispatching rules of hydro-power reservoirs. *Journal of Xi'an University of Technology*, 12:226-231, 1996.

[9] J. Timmis, T. Knight, L.N. De Castro, E. Hart, *An Overview of Artificial Immune Systems. Computation in Cells and Tissues: Perspectives and Tools for Thought*. Springer, 2004.

[10] J. Timmis, M. Neal, A resources limited artificial

system for data analysis. *Knowledge-based Systems*, 14:121-130, 2001.

[11] A. Watkins, J. Timmis, L. Boddess, Artificial immune recognition system(AIRS): An immune-inspired supervised learning algorithm. *Genetic Programming and Evolvable Machines*, 5: 291-317, 2004.

[12] C.H. Song, C. Wang, et al. Optimal scheduling model of reservoirs system for irrigation district and its solution. *Water Resources and Power*, 24:80-83, 2006.

[13] X. Zuo, G.W. Ma, et al. The immune algorithm and its application in short-term scheduling of hydro plant. *Journal of Hydroelectric Engineering*, 25:21-25, 2006.

[14] H.W. Mo. *The theory and application of artificial immune system*. Harbin: Publishing house of Harbin Engineering University, 2003.

[15] D. Cooke, J. Hunt, Recognising promoter sequence using an artificial immune system. *Proceedings of Intelligent Systems in Molecular Biology*, 1995.

[16] J. Hunt, and D. Cooke, Learning using an artificial immune system. *Journal of Network and Computer Application*, 19:189-212,1996.

[17] J. Timmis, M. Neal, A resources limited artificial system for data analysis. *Knowledge-based Systems*, 14:121-130, 2001.

[18] Z.J. Yin, T.S. Hu, et al. Reservoir operation rule for multipurpose water supply. *Advances in Water Science*, 16:875 – 880, 2005.

[19] A. Freitas, J. Timmis, Revisiting the Foundations of Artificial Immune systems: A Problem-Oriented Perspective. *The Proceeding of Second International Conference on Artificial Immune Systems (ICARIS)*, Napier University, Edinburgh, UK, September 1–3, 2003.

[20] J. Hamaker, L. Boggess, Non-euclidean distance measures in AIRS, an artificial immune classification system. *Proceedings of the 2004 Congress on Evolutionary Computing*, 2004.

[21] Zighed, Rabaseda, Rakotomala. FUSINTER: A method for discretization of continuous attributes. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6:307-326, 1998.