

协同演化免疫算法提取水库调度规则研究*

王小林¹, 成金华¹, 尹正杰², 胡铁松³

(1. 中国地质大学(武汉)经济管理学院, 湖北武汉430074;

2. 长江科学院水资源研究所, 湖北武汉430010;

3. 武汉大学水资源与水电工程科学国家重点实验室, 湖北武汉430072)

摘要:从复杂适应系统角度探讨了水库调度的复杂性, 分析比较了水库调度规则提取过程与人体免疫反应中抗体种群协同演化的相似性, 提出了基于协同演化人工免疫系统(CoEISM)水库调度规则提取模型, 并应用于某水库供水调度规则提取, 所获规则对训练与检验样本识别率分别为92.5%和84.4%, 缺水指数为2.23(10¹⁴ m⁶), 与神经网络(RBF)提取结果相当。进一步研究表明, CoEISM用于水库调度规则提取是可行、有效的, 不仅所获规则易于理解, 能提供更为丰富的决策信息; 而且调度规则自适应调整有利于调度人员做出适应调度环境变化的水库蓄泄决策。

关键词: 水库供水; 调度规则; 协同演化; 人工免疫系统

中图分类号: TP393.8; TV697.1 **文献标志码:** A **文章编号:** 0529-6579(2010)06-0121-05

Reservoir Operating Rule Extraction based on Co-evolution Artificial Immune System

WANG Xiaolin¹, CHENG Jinhua¹, YIN Zhengjie², HU Tiesong³

(1. School of Economics and Management, China University of Geosciences, Wuhan 430074, China;

2. Water Resource Department, Yangtze River Scientific Research Institute, Wuhan 430010, China;

3. State Key Laboratory of Water Resources and Hydropower Engineering Science,
Wuhan University, Wuhan 430072, China)

Abstract: The background of reservoir operation is becoming complex. The release strategies has to be continuously adjusted to be adaptable for the change of reservoir operation environments. In this paper, the complexity of reservoir operation was discussed from the viewpoint of complex adaptive system and the similarity was analyzed between the process of extracting reservoir operating rule and self-adaptive learning of antibody and antigen in human immune system. A reservoir operating rule extraction model based on the co-evolution immune system (CoEISM) is developed and applied to extract the water-supply reservoir operating rule of a water-supply reservoir in North China. The obtained operating rules acquire the identification rate of 92.5% for training samples and 84.4% for testing samples, respectively, which is comparable to that by the RBF. The reseaches indicate that the co-evolution artificial immune system is feasible and effective for extracting reservoir operating rule. The obtained rules are easier to be interpreted and can provide more rich decision-making information. The self-adaptive adjustment of rules is helpful for decision-makers to make reseasonable release strategies to fit for environmental changes.

Key words: water-supply reservoir; operating rules; co-evolution; artificial immune system

* 收稿日期: 2009-12-21

基金项目: 国家自然科学基金重点资助项目(50639040); 中国地质大学(武汉)校优秀青年教师资助项目(CUGQNW0908)

作者简介: 王小林(1974年生), 男, 讲师, 博士; E-mail: wangxiaolin_cug@163.com

水资源系统是一个复杂适应系统 (Complex Sdaptive System, CAS), 构成系统的各主体 (Agent) 通过协作与竞争等方式协同演化实现系统整体最优^[1-4]。水库作为水资源系统一个重要组成部分, 面临的调度背景日趋复杂, 其调度运行往往受经济、社会与生态环境等因素影响, 需要不断调整自身性能与行为 (调度规则) 以适应环境变化。目前数据挖掘技术 (如 ANN、决策树等) 已逐渐应用于水库调度研究中, 弥补了传统方法在处理调度决策中的多目标、多变量与非线性等方面存在的不足^[5-7]。但其思路是寻求水库调度决策与调度影响因素之间的非线性映射关系, 或是对调度决策空间的静态划分, 还没有将调度规则提取过程上升到复杂系统演化的高度来看待, 因而无法洞察演化过程中调度规则自适应变化以及群体行为的涌现, 难以适应复杂调度环境下水库有效运行需要。

人体免疫系统是一高度并行、分布式的复杂适应系统^[8-9]。文献 [10-11] 研究了免疫应答中抗体种群自适应学习行为与协同演化机制, 提出了协同演化人工免疫模型 (Coevolution Immune System Model, CoEISM)。水库调度规则提取过程与 CoEISM 相比有诸多相同之处, 例如: ①调度环境与调度决策之间的作用类似抗原提呈; ②调度规则提取可视为抗体自适应学习; ③最优规则可视为记忆细胞。由此可见, 抗体种群自适应学习以及协同演化方式应非常适用于水库调度规则提取和水库运行的适应性管理, 尽管目前 CoEISM 在水库调度研究中尚很少得到应用。基于此, 本文在文献 [7] 研究基础上, 以水库供水调度为例, 尝试利用 CoEISM 自适应提取供水调度规则, 为水库调度研究提供新的方法与思路。

1 水库调度规则提取的协同演化免疫算法

水库调度环境涉及政府部门、生态以及用水户等多个主体。不同主体对水库蓄泄要求不同, 它们与调度决策之间存在复杂非线性关系, 调度过程中需要综合考虑这些主体对调度决策的影响。水库调度是一个不断调整调度决策以适应环境主体变化的过程, 因而调度规则提取所面临的问题是, 如何从历史调度决策过程中挖掘出环境主体与水库调度决策之间相互作用的非线性关系或潜在模式, 以便更好地指导未来的调度决策。

对 CoEISM 而言, 调度环境为系统面临的抗原 (Antigen), 其基因编码由调度环境内不同主体的

基因模块 (编码) 组合而成, 而这些基因组合蕴含丰富的调度决策信息; 调度规则为 B 细胞。由此, 本文给出 CoEISM 自适应提取水库调度规则系统结构 (如图 1), 思路为: 利用协同演化机制对抗原进行学习, 从其基因模块组合中获取能适应调度环境变化的水库调度规则。其中, B 细胞由基因型与表现型构成 (如图 2); 表现型称为抗体 (Antibody), 条件部分含有通配符 “#”, 使抗体能匹配多个抗原。调度规则提取算法主要流程如下:

Step1: 抗原提呈。通过探测器将抗原转化为 B 细胞能识别的二进制编码, 用 “IF <Condition> THEN <Action>” 形式表示, 与 B 细胞的抗体类似但无通配符, 其中 <Condition> 为环境信息内各主体基因模块组合; <Action> 为水库蓄泄决策 (调度模式) 编码。对于供水水库而言, 这里考虑调度决策或称供水模式 (C) 主要受来水、用水及管理者知识 (或经验) 的影响^[6]。来水包括水库蓄水量 (S)、径流量 (I) 以及调度时段 (N) 三方面; 用水为不同用户的用水需求 (D); 管理者知识表现为对中长期来水情况的预测 - 即对水文年型 (P) 的判断。故这三个主体包含的 5 个影响因素构成了水库调度特征属性; 将这 5 个属性与调度决策 (C) 综合考虑便形成水库历史调度信息, 用 “IF <N, S, I, D, P> THEN <C>” 形式表示。

Step2: B 细胞种群协同演化。设定系统中 B 细

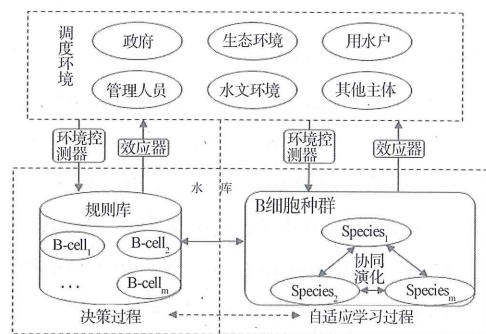


图 1 CoEISM 自适应提取水库调度规则系统结构

Fig. 1 CoEISM-based self-adaptive extraction of reservoir operation rules

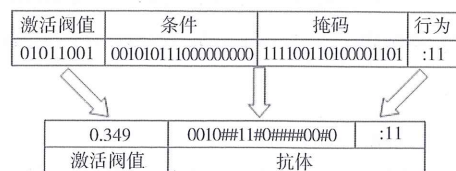


图 2 B 细胞结构及基因型与表现型关系

Fig. 2 Relationship between genotype and phenotype

胞种群数量为 m ；每一种群由 n 个拥有相同水库调度模式的规则构成的规则子集，其中适应度最大的规则为该种群代表 r_i ($1 \leq i \leq m$)。规则适应度通过 B 细胞种群协同（竞争或合作）机制实现^[8-9]，在本文中表为规则正确识别调度环境抗原数（条件与行动部分编码均匹配）与识别系统内规则数（条件部分匹配即可）之差。具体计算过程为：在 t 代，从种群 $S_i(t)$ 中选择一 $B_j^i(t)$ ($1 \leq j \leq n$)，与剩余种群代表 $r_l(t)$ ($1 \leq l \leq m, l \neq i$) 构成一个合作体 $(r_1(t), \dots, B_j^i(t), \dots, r_m(t))$ ，并将该合作体中所有 B 细胞适应度之和作为 $B_j^i(t)$ 适应度。在每一种群内所有调度规则适应度确定后，利用 GA 对该种群进行选择、交叉与变异，产生下一代种群。

Step3: B 细胞种群生灭。每隔固定演化代数 K ，利用式 (1) 检查系统演化是否停滞；当演化停滞时，则产生新调度规则子集；如果种群 $S_i(t)$ 的代表 r_l 的适应度小于消亡阈值 (Extinction threshold, ET)，则种群 $S_i(t)$ 被清除或该调度规则子集自动消亡。 $f(t)$ 为种群代表构成的合作体 $(r_1(t), \dots, r_i(t), \dots, r_m(t))$ 的适应度， C 为系统演化压力，为一设定值。

$$f(t) - f(t - K) < C \quad (1)$$

Step4: 调度规则生成。不断重复 Step2 - Step3，当演化一定代数后，从各 B 细胞种群中找出适应度最大的调度规则，形成水库调度规则库，用于未来调度决策。

2 实例研究

2.1 实例背景

以某供水水库为例，主要考虑 3 个供水目标：

农业灌溉用水 $D_{1,t}$ ；工业及市政用水 $D_{2,t}$ ；居民生活用水 $D_{3,t}$ 。3 个供水目标优先级依次从低到高，其中 t 为时段序号。依文献 [12]，水库在任一时段有 4 种可能调度模式分别为，模式①：正常供水；模式②：限制农业供水；模式③：限制农业和工业供水；模式④：限制所有用水户供水。根据水库长系列（1956 - 2000 年）来用水资料，以缺水指数 Z （供需缺水平方和）最小作为目标，建立动态规划模型，获得全运行期水库供水逐时段调度模式 (C) 的最优决策序列^[7]。将调度模式 (C) 与相应时段的 $\langle S、I、D、N、P \rangle$ 结合起来，则形成了该供水水库历史调度环境信息。

为减小水库特征属性 $S、I、D$ 的随机性以及水文年型判断 (P) 的主观不确定性对规则提取效果的影响，采用基于信息熵的 FUSINTER 算法^[13] 将其离散为几个区间（见表 1, 2）；调度决策是在年内不同时段做出的，故 N 直接采用二进制编码，按月分为 12 个时段（见表 2）。4 种供水调度模式编码为模式①：“00”；模式②：“01”；模式③：“10”；模式④：“11”。

2.2 计算结果与分析

本文用 30a 数据（1956 - 1985 年，共 360 条）提取调度规则（如图 3 所示），后 15a 数据（共 180 条）用于检验。初始种群数为 4，对应 4 种供水模式，种群内个体 popsize = 100。模型中遗传算法参数采用经验取值：交叉率 $P_c = 0.4$ ，变异率 $P_m = 0.06$ 。经多次试验，种群动态变化控制参数 $ET = 2$ ， $K = 10$ ；而 C 的取值对提取效果影响不显著（试验过程不赘述）。规则提取过程如图 3 所示，识别率从初始的 68.6% 逐渐上升，经 500 代演化，最终接近 92.5%，获得 24 条调度规则（如表 3）。

表 1 蓄水量 (S)，来水量 (I)，需水量 (D) 离散区间与编码

Table 1 Discrete intervals of storage (S), inflow (I) and water demand (D), and corresponding binary code

区间数	区间 S	编码	区间 I	编码	区间 D	编码	区间数	区间 S	编码	区间 I	编码
1	2400 - 3814	0000	13 - 121	0000	416 - 797	000	9	8114 - 8150	1000	637 - 746	1000
2	3814 - 3975	0001	121 - 165	0001	797 - 864	001	10	8150 - 8299	1001	746 - 788	1001
3	3975 - 4644	0010	165 - 309	0010	864 - 967	010	11	8299 - 8302	1010	788 - 1098	1010
4	4644 - 4709	0011	309 - 342	0011	967 - 1212	011	12	8302 - 8495	1011	1098 - 1350	1011
5	4709 - 4856	0100	342 - 417	0100	1212 - 1229	100	13	8495 - 8512	1100	1350 - 1601	1100
6	4856 - 6401	0101	417 - 502	0101	1229 - 1255	101	14	8512 - 8645	1101	1601 - 3528	1101
7	6401 - 7525	0110	502 - 519	0110	1255 - 1480	110	15	8645 - 10167	1110	3528 - 5142	1110
8	7525 - 8114	0111	519 - 637	0111	1480 - 1480	111	16	10167 - 10800	1111	5142 - 13499	1111

表 2 水文年型 P 与调度时段 N 编码

Table 2 Discrete intervals of hydrological (P) and operation periods (N) and corresponding binary code

年径流量区间/(10 ⁴ m ³)	38921 ~ 17193	17193 ~ 12547	12547 ~ 8034	8034 ~ 5092	5092 ~ 4308	4308 ~ 0						
P	丰	偏丰	平	偏枯	枯	特枯						
编码	000	001	010	011	100	101						
N (月份)	1	2	3	4	5	6	7	8	9	10	11	12
编码	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100

表 3 水库供水调度规则

Table 3 Water-supply reservoir operating rules obtained by CoEISM

序号	阈值	条件					行动	适应度	序号	阈值	条件					行动	适应度
		S	I	D	N	P					C	S	I	D	N		
1	0.65	1###	####	###	####	###	00	139	13	0.47	01##	#0##	#1#	####	###	10	12
2	0.99	1#0#	####	###	#1#1	###	01	11	14	0.46	0#11	####	0#0	####	#0#	01	3
3	0.89	0###	####	###	####	###	10	41	15	0.99	1#0#	####	1#0	###1	###	01	16
4	0.66	#0##	####	###	####	###	01	12	16	0.38	0#11	####	#0#	###0	###	01	8
5	0.17	0##1	#1##	###	##1#	##1	10	7	17	0.82	0###	####	0##	#0##	##0	01	8
6	0.22	1#0#	###0	###	####	###	01	8	18	0.12	##0#	####	1##	##0#	###	11	8
7	0.50	##0#	#0##	#00	#1#0	##0	00	2	19	0.80	####	#0##	1##	0###	1#0	00	4
8	0.46	##1#	##10	000	1###	0##	00	26	20	0.49	010#	00##	##0	####	01#	10	8
9	0.28	##1#	####	1##	####	1##	01	5	21	0.19	0##0	####	##0	##0#	0#0	10	2
10	0.16	0#1#	####	0##	00##	##1	01	3	22	0.57	0##0	##11	1##	#01#	###	01	2
11	0.79	0#10	#0##	1#0	0###	##1	10	2	23	0.43	0#11	#1#0	#1#	#1##	0##	01	2
12	0.68	00##	####	###	0###	##1	11	3	24	0.35	0010	##11	#0#	####	0#0	11	1

对所获规则进行检验，得到识别率为 84.4%，缺水指数为 $2.23 \times 10^{14} m^6$ ，与 RBF 提取结果相当 (识别率为 84%，缺水指数为 $2.36 \times 10^{14} m^6$)^[7]。

对复杂调度环境；因而，在实际调度中调度人员可以通过观察规则动态调整与演变轨迹，做出合理决策以适应调度环境的变化。

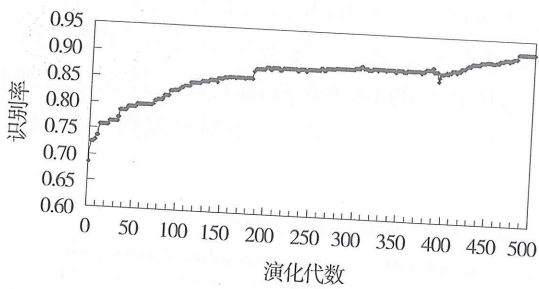


图 3 调度规则提取自适应学习过程

Fig. 3 Process of extracting reservoir operating rules by self-adaptive learning of CoEISM

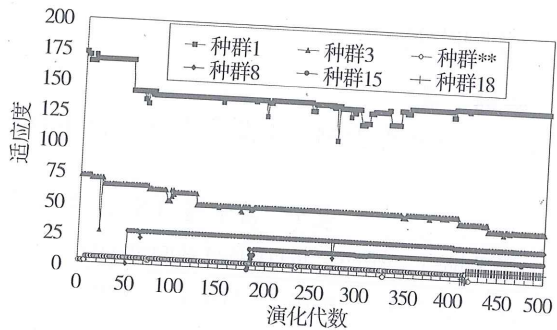


图 4 部分供水调度规则动态产生与消亡演化过程

Fig. 4 Creation and extinction of operating rules

演化过程中大约每隔 10 代识别率会产生跳跃，其因源于调度环境压力使得规则在演化过程中发生分化 (如图 4)，即新规则不断产生 (如种群 8、15、18)，原有规则因适应度较小而消亡。例如种群 (**) 在 410 ~ 420 代因适应度小于 $ET = 2$ ，故该种群被清除。这一过程清晰表明，调度规则通过协同演化的方式自适应调整，产生多样性，以应

所获 24 条调度规则编码含有大量通配符“#”，如将其与表 1、2 对应，则每一规则具有明确物理意义，能涵盖不同基因模块组合而形成的调度环境。例如，规则 12 包含 2^{13} 种对 3 个用水户限制性供水的决策信息，可解释为：

$$IF: S \in (4644, 2400] \cap I \in [13, 121] \cup$$

$(165, 309] \cup (342, 417] \cup (502, 637] \cap D \in [416, 1480] \cap N \in \{1 \text{月}, 2 \text{月}, \dots, 7 \text{月}\} \cap P \in \{\text{丰}, \text{偏丰}, \text{平}, \text{偏枯}\} \text{ THEN } : C \in \{\text{调度模式④}\}。$

仅就该规则给出的解释来看似乎不合理, 因为如果遇到连续丰水年份, 由于水库蓄水较大, 可正常供水。但调度规则对调度信息识别, 不仅取决于自身基因结构(模式), 而且更取决于同其他规则竞争和协作。例如, 1990年6月样本编码为“0000 0010 101 0110 101: 11”, 能同时被多条规则匹配(如表3), 但规则3, 4, 12的匹配值最大(值均为1.0), 然而规则12所含通配符数最少, 其物理意义更明确, 故调度模式由规则12给出, 即为供水模式④。由此可见, 所获调度规则不仅合理, 而且具有可解释性, 能为水库调度管理提供更为丰富的决策信息。

尽管所获24条规则对检验样本识别率与RBF的识别率相当, 但分别从4种调度模式识别结果来看(见文献[7]), 前者对模式①②④的识别率高于RBF(如表4, 5), 模式③则相反。可见, CoEISM对样本的非线性划分能力总体上高于RBF, 所获调度规则同样能表示水库调度特征属性之间的高维非线性关系。不仅如此, 前者在实际调度中可输出与调度环境信息匹配的所有规则, 并结合已有调度知识与经验做出决策; 而后者提取的规则隐含于网络中, 其结果不易为调度人员理解。

表4 CoEISM 识别结果

Table 4 Identification results by CoEISM

真实模式	分类模式				分类精度 %
	①	②	③	④	
①	95	8			92.2
②	4	33		2	84.6
③	3	4	15	3	60.0
④		2	2	9	69.2

表5 RBF 识别结果^[7]Table 5 Identification results by RBF^[7]

真实模式	分类模式				分类精度 %
	①	②	③	④	
①	94	7	2		91.3
②	2	32	4	1	82.1
③	1	2	18	4	72.0
④		2	4	7	53.8

3 结论

1) CoEISM 模型用于用于水库调度规则提取,

可微观观察水库调度规则自适应演变轨迹与动态调整, 有利于调度决策者做出合理蓄泄决策, 以适应调度环境变化;

2) 所获调度规则不仅具有高维非线性识别能力, 而且具有明确的物理意义, 易于理解, 为水库调度管理提供更为丰富的决策信息。

参考文献:

- [1] JEFFERY R., SEATON R. Evaluation methods for the design of adaptive water supply systems in urban environments [J]. *Water Science and Technology*, 1997, 35(9):45-51.
- [2] 赵建世, 王忠静, 翁文斌. 水资源复杂适应配置系统的理论与模型[J]. *地理学报*, 2002, 7(6):639-647.
- [3] 王慧敏, 伶金萍, 马小平, 等. 基于CAS范式的流域水资源配置与管理及建模仿真[J]. *系统工程理论与实践*, 2005, 23(5):34-36.
- [4] 陈洋波, 曾碧球. 基于水库调蓄的流域可供水量优化模型研究与应用[J]. *中山大学学报:自然科学版*, 2005, 44(2):94-97.
- [5] BESSLER T, FLORIAN S, DRAGAN A & GODFREY A. Water reservoir control with data mining [J]. *Journal of Water Resources Planning and Management*, 2003, 129(1):26-34.
- [6] NEELAKANTAN T. R., PUNDARIKANTHAN N. Neural network-based simulation optimization model for reservoir operation [J]. *Journal of Water Resources Planning and Management*, 2000, 126(2):57-64.
- [7] 尹正杰, 王小林, 胡铁松, 等. 基于数据挖掘的水库供水调度规则提取[J]. *系统工程理论与实践*, 2006, 26(8):129-135.
- [8] DASGUPTA D. An artificial immune system as a multi-agent decision support sytem [C] // *Proceeding of the SMC98, IEEE International Conference on Systems, Man, and Cybernetics*, Sandiego California, 1998, 4:3816-3820.
- [9] 张灵, 陈晓宏, 翁毅, 等. 人工免疫算法在洪水分类中的应用[J]. *中山大学学报:自然科学版*, 2008, 47(5):110-114.
- [10] POTTER M A, DE JONG K A. The coevolution of antibodies for concept learning [C] // *Proceedings of the Fifth International Conference on Parallel Problem Solving from Nature*, Springer-Verlag, 1998, 530-539.
- [11] POTTER M A, DE JONG K A. Cooperative coevolution: An architecture for evolving coadapted subcomponents [J]. *Evolutionary Computation*, 2000, 8(1):1-29.
- [12] 尹正杰, 胡铁松, 崔远来, 等. 水库多目标供水调度规则研究[J]. *水科学进展*, 2005, 16(6):875-880.
- [13] ZIGHED D A, RABASEDA S, RAKOTOMALA R. FUSINTER: A method for discretization of continuous attributes [J]. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 1998, 6(3):307-326.