

基于案例推理的黄河流域水资源优化调配

彭少明^{1,2}, 张春生¹

(1. 黄河勘测规划设计有限公司, 河南 郑州 450003; 2. 中国水利水电科学研究院水资源研究所, 北京 100044)

摘要 利用水文要素的周期性、重现性、相似性等特征, 将嵌套案例推理技术引入到黄河流域水资源调配方案决策中, 建立 K-D 树定性搜索与相似度精确定位相结合的基于案例推理的黄河流域水资源调配决策模型。对黄河流域水资源调配中的水文要素进行定性识别, 以关键水文要素的相似度作为检索指标, 通过 K-D 树结构实现快速搜索定位与水文要素的相似性匹配, 可搜索出与决策问题高度相似的方案, 实现水资源的优化调配。

关键词 水资源调配; 案例推理; 相似度; K-D 树检索; 黄河流域

中图分类号: TV212.4 文献标识码: A 文章编号: 1006-7647(2012)02-0046-04

Case-based reasoning for water resource optimal allocation in the Yellow River Basin//PENG Shao-ming^{1,2}, ZHANG Chun-sheng¹(1. Yellow River Engineering Consult Co., Ltd., Zhengzhou 450003, China; 2. Department of Water Resources, China Institute of Water Resource and Hydropower Research, Beijing 100044, China)

Abstract: Using the characteristics of periodicity, repeatability and comparability of hydrology element, the method of case-based reasoning is introduced into the water resource allocation, and a model for water resource allocation decision in the Yellow River Basin is established based on case-based reasoning, combined K-D tree qualitative searching technology with analogical degree pinpoint technology. Taking the water resource allocation in the Yellow River Basin as an example, the hydrology elements are identified qualitatively and the analogical degree is regarded as a searching index, then the highly analogical case is searched out based on comparability matching of hydrology elements through K-D tree structure.

Key words: water resource allocation; case-based reasoning; analogical degree; K-D tree searching; Yellow River Basin

案例推理(case-based reasoning, CBR)是一种基于知识的问题求解和学习方法, 通过重用以前相似问题的方案信息来解决新问题。CBR 通过旧案例的检索、重用、匹配, 再现案例决策, 实现对新问题的求解, 是一种有别于传统人工智能的新技术。CBR 将定量分析与定性分析相结合, 具有动态知识库和增量学习的特点, 可以克服当前专家系统所面临的难题。

由于传统的人工智能面临知识获取的局限, 基于数据挖掘和分析的计算智能成为新的研究热点^[1]。Aamodt 等^[2]提出了 CBR 四阶段循环, 即案例的检索、复用、修正、保存; Gilboa 等^[3]将案例表示为问题描述、解描述、效果描述的三元组; Aha 等^[4]通过案例工程确定案例所必须包含的信息, 并从数据中提取信息; Anders^[5]提出了基于概率案例的 PEBM 模型, 采用概率传播技术评估和检索案例。进入 21 世纪, 随着计算机技术的快速升级, CBR 已被广泛应用

到故障诊断、商业管理、智能决策、知识提取等领域^[2-8]。当前 CBR 研究主要集中在案例的索引及检索技术、案例修正规则的获取方法、案例库维护技术及其性能评价、CBR 与其他方法的集成等方面^[9-11]。

1 基于 CBR 的水资源调配

1.1 调配原理

相似性技术是一种用于分析存在于自然界和社会科学中隐含信息的技术, 用来探求和解决真实世界中复杂问题的工具。与以往的人工智能不同, CBR 将过去处理过的问题描述成由问题特征集和解决方案组成的案例, 并存储在案例库中, 当出现新的问题时, 系统从案例库中检索出与新问题最相近的案例, 将旧案例的解决方案作为新案例的解决方案。基于 CBR 的流域水资源调配原理见图 1。

在复杂水资源系统中, 一些水文现象通常具有

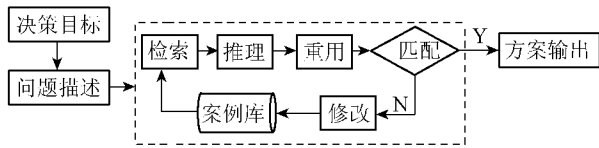


图1 基于 CBR 的流域水资源调配原理

周期性特征,即一些水文要素在一定的时段内会重复出现,基于这一特征将 CBR 引入水资源调配之中。利用新旧案例要素之间的相似性,从案例库中筛选出相似案例,重用决策信息,可避免多目标优化的“维数灾”问题,实现水资源的优化调配。

1.2 调配流程

基于 CBR 的流域水资源调配采用定性与定量相结合的方法,首先根据河流主要断面来水信息确定搜索方向,检索出潜在的案例组,然后根据相似度大小得出最匹配的方案,其调配过程主要分为以下 5 个步骤。

第 1 步 引入水文要素关系指示系数来归纳各个区间的水文特征,通过归纳水文要素的关系定性指示信息,开展定向检索。将流域划分为 K 个子区间,在搜索树的末端寻找各子区间水文要素关系指示系数完全一致的方案,即为潜在的案例组。

第 2 步 推理匹配。对决策问题的 n 个属性项计算潜在案例组与决策问题的相似度。

第 3 步 案例精选。系统按相似度降序排列潜在案例组,选择相似度最大的案例为精确案例,并保留这个案例。

第 4 步 参数调整。若决策者不满意这个案例,系统将启动交互决策界面修改决策参数,重复第 1~3 步。

第 5 步 系统存储第 3 步或第 4 步中选择的案例。

通过上述 5 个步骤归纳检索出潜在的案例组,再通过相似度确定精确案例,若该精确案例专家评价满意,即成为最合适的案例,并更新水资源调配案例库。

2 CBR 的检索与匹配结构设计

选择合适的检索和定位方法可以提高 CBR 系统的检索速度和匹配精度。考虑到黄河流域水资源调配具有多目标、多属性、高维度的复杂特征, CBR 采用 K-D 树^[12-17]检索算法和相似度匹配方法。检索过程中每选定一个检索方向可以看成是进行一次统计推断,首先判断较粗粒度空间中哪个部分最可能包含目标,然后在被选出的较细粒度空间上进一步判断被选出的较细粒度空间中哪个细粒度空间最有可能包含目标,这样逐层下降直到所有空间均包含搜索信息,确定出潜在案例组,然后计算相似度,

寻找出精确案例。

2.1 K-D 树检索技术

CBR 搜索结构设计采用 K-D 树建立多维索引, K-D 树是一种 K 维空间点二叉索引树^[14],具有数据结构简单、存取效率高等优点,是一种适用于高维数据结构的快速搜索结构。K-D 树内部节点与某一空间维 $x_i (1 \leq i \leq K)$ 对应,且每个内部节点都用 1 个正交于其空间维 x_i 的超平面 $x_i = c$ (c 为常数)将该节点所表示的 K 维空间分成两部分,这些超平面在 K 个方向上交替出现,并且每个超平面至少包括 1 个点数据。K-D 树将数据分散存储在每个节点上, K-D 树算法中,每个节点是 1 个二选结构, 2 个子节点代表将空间划分为 2 个子区。K-D 树的内部节点有 1 个相关联的属性 a 和一个值 V ,它将数据点分成 2 个部分: a 的取值小于 V 的部分和 a 的取值大于或等于 V 的部分。由于所有维的属性在层间循环,所以树的不同层上的属性是不同的。为获得块模式,对 K-D 树结构进行以下 2 处修改: ①内部节点只有 1 个属性,该属性的每个值确定地指向左、右分支; ②子节点是块,块空间中存放着尽可能多的记录。

2.2 K-D 树结构的定性推理

按照黄河流域主要断面(花园口、三门峡、河口镇、兰州、龙羊峡)将流域分为 6 层(级), 5 个断面分别对应于 K-D 树的 5 个节点,而 5 个流域区间则设计为 K-D 树的 5 级子系。图 2 为系统搜索的 K-D 树路径结构设计。

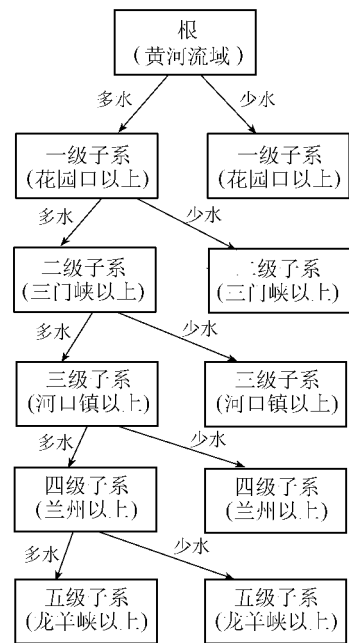


图2 基于 CBR 的黄河流域水资源调配系统 K-D 树路径结构设计

K-D 树定性搜索的方向是根据每一层级的来水特征进行判别,按照主要断面的水文信息(包括河道

来水量和降水量)与多年平均水量(由断面来水和区间降水量综合确定)的关系将区间的水文年份定性分为多水年和少水年。

K-D 树的案例检索规则是:首先采取归纳法抽取各节点来水量和降水量信息,按照水文要素的特征划分为不同水文年份,为决定检索分支方向提供依据,采用归纳法搜索,逐层递进,沿决策树越过节点的路径取各节点所包含信息的交集,在决策树的末端定位识别出与决策方案各区间水文年份均相同的潜在案例组。

2.3 基于水文要素相似度的案例匹配

基于 CBR 的黄河流域水资源调配方法从决策方案的降雨、径流预测入手,根据关键水文要素的相似程度,按照案例推理法则选择相似度最大(最接近)的案例为目标案例。案例相似度是水文要素(降雨量、径流量)的贴近程度。

a. 计算变量属性偏离度。对于数字型优化变量,采用偏离度作为属性差别的度量,设新案例的数字型优化变量 $W_j(x_1, x_2, \dots, x_n)$ 的属性值为 A_j , 与其相对应的旧案例的数字型优化变量为 $W_j^*(x_1^*, x_2^*, \dots, x_n^*)$, 其属性值为 A_j^* , 定义新案例中 $W_j(x_1, x_2, \dots, x_n)$ 与 $W_j^*(x_1^*, x_2^*, \dots, x_n^*)$ 的偏离度为

$$d_j = \left\| \frac{A_j^* - A_j}{A_j} \right\| \quad (1)$$

可以得到

$$d_j = \sqrt{\left(\frac{x_1^* - x_1}{x_1}\right)^2 + \left(\frac{x_2^* - x_2}{x_2}\right)^2 + \dots + \left(\frac{x_n^* - x_n}{x_n}\right)^2} \quad (2)$$

b. 计算案例的相似度。采用案例相似度判断 2 个案例之间综合水文要素的趋同或差异程度。设由 m 个属性组成的案例,其相似度可以定义为偏离度的集成,表达式如下:

$$\sin(i, m) = \sum_{j=1}^m \omega_j (1 - d_j) \quad (3)$$

式中 $\sin(i, m)$ 为旧案例 i 与目标案例 m 的相似度, ω_j 为各属性的权重。

为了利用相似度中包含的信息,将获取的相似度按从大到小的顺序排序,相似度最大的案例即为精确案例。

为了保证推理的质量,精确案例必须满足属性

表 1 基于 CBR 的黄河流域水资源调配定性搜索结果

决策项	V_1		V_2		V_3		V_4		V_5		V_6	
	$A_1/\text{亿 m}^3$	A_2/mm	$A_1/\text{亿 m}^3$	A_2/mm	$A_1/\text{亿 m}^3$	A_2/mm	$A_1/\text{亿 m}^3$	A_2/mm	$A_1/\text{亿 m}^3$	A_2/mm	$A_1/\text{亿 m}^3$	A_2/mm
目标信息	232.8	525.7	350.1	484.9	336.7	241.4	451.3	408.4	489.9	599.1	492.2	633.7
专家信息	212.0	485.9	329.9	478.9	331.8	261.8	482.7	493.5	532.8	658.2	534.8	647.8
定性识别	多水		多水		少水		少水		少水		少水	

偏离度和案例相似度的临界阈值约束,即

$$\begin{cases} d_j < D_0 \\ \sin(i, m) \geq \lambda_0 \end{cases} \quad (4)$$

式中 D_0 和 λ_0 分别为目标案例属性偏离度和相似度阈值,其值由专家确定。

同时满足相似度最大和临界阈值约束的案例为优选案例。若不存在同时满足相似度最大和临界阈值约束的案例则修改参数,重复执行检索过程直至得到决策者满意的目标方案。

3 基于 CBR 的黄河流域水资源调配

3.1 案例的定向搜索

系统在接收到水文预报数据后,分析水文信息,对来水年份定性辨识归类,确定 K-D 树结构的搜索方向。基于 CBR 的黄河流域水资源调配系统 K-D 树设计为 6 个层级(含流域级),对 6 个层级水文要素进行辨识,确定搜索方向。根据 K-D 树结构,每个层级节点设有多水及少水 2 个方向,通过 6 层级的定性推理识别,确定 1 组与决策信息完全相同的案例。表 1 为基于 CBR 的黄河流域水资源调配定性搜索结果,表中 $V_1 \sim V_6$ 为流域 K-D 树自下而上的 6 个层级, A_1 和 A_2 分别为断面来水量及降水量。系统通过专家信息对决策方案进行定性辨识,例如对四级子系(兰州以上)进行水文要素辨识,若判断结果为少水年,则向右侧分支搜索(图 2)。决策方案的 6 个层级属性定性表达值为(多水,多水,少水,少水,少水,少水)。系统接收水文信息后通过 K-D 树搜索快速定位一组具有‘龙羊峡、兰州两断面多水,河口镇、三门峡、花园口及全流域少水’特征的潜在案例。

3.2 案例的定量匹配

系统完成搜索后,锁定了一组潜在案例,案例组的表达包括案例编号及 6 个层级的定性值。CBR 的目标是搜索与决策信息相似度最大的案例,即精确案例,根据专家信息得出目标案例属性偏离度和相似度阈值为 $D_0 < 5\%$ 和 $\lambda_0 > 0.95$ 。

采用式(2)~(4)计算潜在案例组所有案例与目标案例的相似度,并按照相似度大小进行排序。选择相似度最大的案例作为精确案例,其水文特征及各属性偏离度见表 2,案例编号为 BC99, $\sin(i, m) = 0.976 > 0.95$,各属性偏离度均小于 5%。流域水

表 2 精确案例的水文特征及各属性偏离度

层(级)	A ₁			A ₂			断面下泄水量/亿 m ³
	目标案例/亿 m ³	精确案例/亿 m ³	属性偏离度/%	目标案例/mm	精确案例/mm	属性偏离度/%	
V ₁	232.8	237.5	3.60	525.7	535.2	2.32	203.12
V ₂	350.1	366.4	4.45	484.9	497.2	2.47	296.41
V ₃	336.7	339.9	0.94	241.4	244.8	1.39	192.92
V ₄	451.3	453.2	1.07	408.4	409.7	2.23	272.51
V ₅	489.9	479.3	2.21	599.1	586.5	2.15	299.15
V ₆	492.2	489.8	2.37	633.7	629.1	4.04	206.85

资源调配成果见表 3。将水资源调配结果与国务院 1987 年颁布的黄河可供水量分配方案(简称“87 分水方案”);“87 分水方案”根据黄河来水量进行折扣分配)进行对比,可以看出农业灌溉用水量减少,发达地区用水量得到有效保障,体现了高用水的原则。

表 3 基于案例推理的黄河流域水资源调配结果 亿 m³

省(区)	需水量		供水量		缺水水量	“87 分水方案” 供水量
	流域内	流域外	地表	地下		
青海	24.5	0	18.5	3.3	2.7	20.80
四川	0.2	0	0.2	0	0	0.20
甘肃	54.4	2.0	43.8	5.7	6.9	47.50
宁夏	80.4	0	59.2	8.8	12.4	68.21
内蒙古	103.7	0	67.4	20.3	16.0	80.81
陕西	85.9	0	43.5	28.4	14.0	68.92
山西	59.3	0	35.4	21.1	2.8	62.68
河南	56.2	20.7	52.1	21.8	3.0	73.90
山东	23.7	60.0	67.4	11.6	4.7	77.80
河北、天津	0	5.0	4.9	0	0.1	18.40
合计	488.2	87.7	392.5	120.9	62.5	519.22

3.3 案例的评价与输出

系统锁定的精确案例将由专家从经济效益、社会效益和生态环境效益 3 个方面进行综合评价,评价满意的方案方可作为最终推荐方案。首先分析生态效益,6 个断面的下泄水量(203.12 亿 m³,296.41 亿 m³,192.92 亿 m³,272.51 亿 m³,299.15 亿 m³,206.85 亿 m³)满足生态环境需水量(140 亿 m³,160 亿 m³,197 亿 m³,205 亿 m³,200 亿 m³,200 亿 m³)要求,流域生态环境需水可得到满足;其次分析经济效益,与黄河流域长期实行的“87 分水方案”比较,宁夏、陕西、内蒙古及山东等新兴工业区用水适度增加,而青海、甘肃及山西由于长期用水量不足,分配水量适度减少,这样既可保证能源化工产业区的用水需求,又可刺激农业节水发展,支持经济社会的快速发展;最后分析社会效益,青海、甘肃、宁夏、内蒙古的民生用水得到了保障。因此,综合评价该精确案例为满意方案,推荐实施。

4 结 语

将专家智慧和人工智能技术相结合,引入水文要素的相似度作为推理检索指标,针对水文过程的

周期性和相似特征,建立基于 CBR 的黄河流域水资源调配决策模型,通过水文要素的检索与匹配,再现案例决策,避免了人为因素的干扰。与以往传统的水资源调配决策相比,具备以下优点:①水量调度快捷、高效,据检测基于 CBR 完成一个案例调度决策的时间在 3 min 以内,而传统的调配、会商决策均在数天以上;②保障决策方案有效,方案库中存储的方案均为已实施或专家推荐实施的方案,可行性已在实践中得到检验;③推荐的案例为优化方案,通过搜索与匹配环节的筛选,使目标案例与推荐案例的水文要素高度接近(相似度 95% 以上),且经过了专家的综合评价,因此可保证推荐案例为优化方案。

参考文献:

- [1] 姜丽红. 智能化预测支持系统(IFSS)的理论与方法的研究[D]. 天津: 天津大学, 1996.
- [2] AAMODT A, PLAZA E. Case-based reasoning: foundational issue, methodological variation, and system approaches[J]. AI Communications, 1994, 7(1): 39-59.
- [3] GILBOA I, SCHMEIDLER D. Case-based decision theory[J]. Quarterly Journal of Economics, 1995, 110: 605-639.
- [4] AHA D, BRESLOW L. Refining conversational case libraries [C]//Proc of Second International Conference on Case-Based Reasoning. Berlin: Springer Verlag, 1997: 267-278.
- [5] ANDERS F M R. A probabilistic exemplar based mode[D]. Salford: University of Salford, 1998.
- [6] 田喜平, 赵红丹. 基于案例推理的重复谈判研究[J]. 新乡学院学报: 自然科学版, 2010, 27(1): 9079-9080.
- [7] 郑国宁, 马力. 基于案例推理技术在故障案例中的应用[J]. 电脑知识与技术, 2010, 6(32): 50.
- [8] 梁昌勇, 顾东晓. 面向不确定多属性决策问题的范例检索算法研究[J]. 中国管理科学, 2009, 17(1): 131-137.
- [9] 李锋刚. 基于案例推理的优化型智能决策技术的研究[J]. 淮北煤炭师范学院学报, 2006, 27(2): 58-61.
- [10] 杨瑾, 尤建新, 蔡依平. 基于案例推理的供应商选择决策支持系统研究[J]. 计算机工程与应用, 2006, 3(6): 17-28.
- [11] 蔡芝蔚. 基于知识密集型的 CBR 构件检索方法研究与实现[J]. 牡丹江教育学院学报, 2009, 26(9): 37-39.

(下转第 94 页)