

基于动态负载均衡策略的网格任务调度优化模型和算法

钟绍波

(重庆师范大学 初等教育学院, 重庆 400700)

(zshaob@126.com)

摘要:任务调度是一个 NP-hard 问题,而且是并行与分布式计算中一个必不可少的组成部分,特别是在网格计算环境中任务调度更加复杂。结合免疫克隆算法和模拟退火算法的优点,提出了一种网格任务调度优化模型和算法。仿真实验结果表明,这种调度算法有效地实现了资源的负载均衡,克服了遗传算法容易陷入局部最优的缺点,可以成功地应用于网格任务调度中。

关键词:网格计算;任务调度;免疫克隆算法;模拟退火算法;并行模拟退火克隆算法

中图分类号: TP313 **文献标志码:** A

Optimal resource model and task scheduling algorithm based on dynamic load balancing strategy in grid

ZHONG Shao-bo

(College of Elementary Education, Chongqing Normal University, Chongqing 400700, China)

Abstract: Task scheduling is a NP-hard problem and also an integral part of parallel and distributed computing. It becomes more complicated especially in the grid computing environment. An optimal task scheduling model and an algorithm were brought forward, which combined the advantages of immune clonal algorithm and simulated annealing. The simulation results show that this algorithm achieves resource load balancing, and it overcomes the shortcomings of genetic algorithm, and can be applied to the optimization of task scheduling successfully.

Key words: grid computing; task scheduling; immune clonal algorithm; simulated annealing algorithm; parallel simulated annealing clonal algorithm

0 引言

网格计算是当今互联网研究中的一个热点,也是并行和分布处理技术的一个发展方向。在网格计算中,任务调度是网格计算研究的核心问题。网格计算是一个由广域网连接大型设备形成的并行计算环境,计算节点的处理能力远大于传统的分布式计算环境中的计算节点,而节点间的通信代价比传统的基于局域网的分布式计算环境要高许多。网格计算环境可以看作是一个大的计算池,适合处理相互间几乎没有通信的粗粒度并行任务,所有被提交的任务放进这个计算池就会被根据资源管理方法分配适当的资源,并以一定的调度方法进行调度和执行。每个资源可根据自己的调度策略对分配给自己的任务进行进一步的调度^[1]。

由于网格计算任务调度,面临的是一个 NP 完全问题,它引起了众多学者的关注,成为目前网格计算研究领域的一个焦点。文献[2]提出了一种基于 Agent 的计算网格任务调度方案,并设计了实现网格资源管理的三个协议原型。文献[3]提出了一个可用于网格计算任务调度的高级时间 Petri 网模型。文献[4]在经济原则的基础上为调度问题引入了成本概念。文献[5]提出了利用不同站点多同时请求的分布调度算法。文献[6]设计了一个网格计算的蚂蚁(Ant Algorithm, AA)任务调度算法。文献[7-9]介绍了一种基于遗传算法(Genetic Algorithm, GA)的网格计算任务调度算法,其目的是为了尽可能地提高资源的使用率和吞吐量。文献[10]介绍

了模拟退火(Simulated Annealing Algorithm, SAA)等进化算法在网格计算任务调度中的应用,其目的是为了使总任务执行时间最少。文献[11]提出了一个具有模糊处理时间的网格任务调度免疫克隆算法(Clonal Algorithm, CA)。

在过去的研究过程中,人们主要关注的是任务执行时间,而没有充分考虑到网格的异构性,分布性,可扩展性等特点,因此当出现大批量的任务请求时,网格环境下经常会出现某个资源负载过重而导致无法继续提交任务或者负载过轻导致资源闲置的情况。在开发并行和分布计算应用时,负载均衡是一个关键问题^[12]。网格任务调度涉及到交叉域和大规模应用的调度,所以网格环境下的负载均衡问题显得更为突出。本文设计的新任务调度方法——并行模拟退火克隆算法(Parallel Simulated Annealing Clonal Algorithm, PSACA)是将模拟退火算法和克隆算法的优点充分结合起来,同时充分考虑到资源的负载均衡和任务的高效调度,这样有效地提高了算法的运行效率和求解质量。

1 网格资源优化模型

为了便于分析问题,我们特建立以下网格资源优化模型。

定义 1 网格资源模型:假设用三元组 $G = (R, T, L)$ 表示网格环境。其中: $R = \{r_1, r_i, \dots, r_m\}$ 代表 m 个计算资源的集合, $T = \{t_1, t_2, \dots, t_n\}$ 表示 n 个相互独立的元任务, $L = \{l_1, l_2, \dots, l_m\}$ 表示 m 个资源节点的动态负载权值。

假设资源节点 r_j 当前的 CPU 利用率,内存利用率,当前网

络流量,磁盘 I/O 访问率,进程总数等负载参数分别用 $C_j\%$, $M_j\%$, N_j , $I_o_j\%$, P_j 表示,则资源节点 r_j 动态负载权值可以表示为:

$$l_j = \pi_1 \times C_j\% + \pi_2 \times M_j\% + \pi_3 \times N_j + \pi_4 \times I_o_j\% + \pi_5 \times P_j \quad (1)$$

其中 $\sum \pi_j = 1$, π_j 反映各个负载参数的重要程度, $j = 1, 2, \dots, m$ 。

定义 2 任务调度模型。假定用三元组 $S = (D, \Delta, F)$ 表示本文设计的一个网格任务调度系统。 $D = \{D_{ij} \mid D_{ij}$ 是任务 t_i 在资源 r_j 上的执行时间, $i = 1, 2, \dots, n$; $j = 1, 2, \dots, m\}$, Δ 是一个 $m \times n$ 矩阵,其元素 $\Delta_{ij} = 1$ 表示任务 t_i 分配到资源节点 r_j 上,否则, $\Delta_{ij} = 0$ 。 $F(x)$ 是任务调度目标函数:

$$F(x) = \sum_{i=1}^m \sum_{j=1}^n C \times \Delta_{ij} \times D_{ij} \times e^{\lambda \sum_{i=1}^m l_j} \quad (2)$$

$F(x)$ 是一个能够反映分配代价的函数,它表示在资源节点 r_j 上分配任务 q_i 的代价估计,负载权重越大的资源节点,在其上分配理想状态下执行时间越长的任务时,就会产生较高的分配代价。其中 C, λ 是常数。

定义 3 算法优化模型。 $PSACA = (\Omega, f, A, N, \Gamma, \Psi, \Xi)$ 是本文设计的 PSACA 算法。其中: Ω 是个体的编码方法, f 是抗体的亲和度评价函数, A 是初始抗体群, N 是群体规模, Φ 是克隆变异算子, Γ 是克隆交叉算子, Ψ 是克隆选择算子, Ξ 是终止条件。

2 PSACA 算法的结构与描述

2.1 PSACA 算法的思想

免疫克隆算法是模拟生物免疫系统的多克隆机理,不仅采用变异,交叉等操作实现抗体间的信息交换,而且还要充分利用抗体在变化过程中已经获得的对抗原反应的特性,进一步增加克隆的多样性。克隆算子是依据抗体与抗原的亲和度函数,将解空间中的一个点分裂成多个相同的点,并经过克隆变异,克隆交叉,克隆选择变换后获得新的抗体群^[13-15]。在克隆算法执行过程中,为了保持解的多样性而扩大空间搜索范围,采取对父代进行克隆复制的策略,其解空间变大是以计算时间增长为代价的。为了提高算法的运行效率,有效控制抗体规模,我们对免疫克隆算法融入模拟退火思想。模拟退火算法起源于统计物理学方法,并首次被 Kirkpatrick 等引入优化问题的求解,模拟退火算法具有很好的局部搜索能力^[16]。因此本文将二者相结合,提出了 PSACA 算法。PSACA 算法首先依据抗体与抗原的亲和度函数,将解空间中的一个点分裂成多个相同的点,并经过克隆变异、克隆交叉和克隆选择变换后获得新的抗体群,同时在变异和交叉过程中,按照模拟退火接受准则来决定是否接受操作后的新抗体,当退火温度接近 0 的时候,PSACA 算法自然结束。

2.2 抗体编码及初始抗体群

在任务调度模型中,由矩阵 $\Delta = [\Delta_{ij}]_{m \times n}$ 的定义可以分析出矩阵 Δ 中,每行有且仅有一列的元素值为 1,其他的为 0。抗体是用来表示所求问题的候选解,每一个抗体对应着一种任务调度方案。我们对抗体采用自然数编码方式,把矩阵 Δ 每行中元素 1 所在列号作为抗体基因,抗体基因编号为 K_1, K_2, \dots, K_n ,其中 $K_j \in [1, m]$, $j \in [1, n]$,且为可重复相等的自然数。

采用随机分配的方法产生适量抗体组成的初始种群,为了能够取得全局最优解,群体要有一定的规模。具体方法:随机产生 N 个位数为 n 的抗体,然后用自然数编码的抗体串作为初始抗体种群。

2.3 亲和度

亲和度是用来表明抗体与抗原之间的匹配程度,亲和度越高,说明抗体越接近抗原,也就越接近所求问题的解。

我们对任务调度目标函数 $F(X)$ 进行适当变换,如下:

$$f(X) = \frac{1}{\min(F(X))} \quad (3)$$

其中 $f(X)$ 值为抗体 X 的亲和度, $X = (K_1, K_2, \dots, K_n)$ 。

2.4 抗体克隆

对当前代抗体进行克隆,生成一个临时的克隆群体 C ,每个抗体与抗原的亲和度越大,抗体的克隆规模也就越大。对抗体群 A 中每一个抗体 a_i 按规模 $\text{INT}(\alpha \times N/i)$ 克隆到新的抗体群 C 中,新抗体群规模为 N_c 。

$$N_c = \sum_{i=1}^N \text{INT}\left(\frac{\alpha \times N}{i}\right) \quad (4)$$

其中: α 为克隆系数, $\text{INT}(\ast)$ 为取整函数, N 为初始抗体群大小。

2.5 克隆变异算子

通过克隆扩大了群体的规模后,对克隆后的临时群体 C 中每个抗体进行变异,可以提高群体中抗体的多样性,扩大搜索范围,用来寻找更优秀的抗体,变异后临时的抗体群为 M 。克隆变异操作如下:

1) $v' = v + P_m \times \exp(-f) \times N(0, 1)$, v 和 v' 分别是父抗体和子抗体, $N(0, 1)$ 是均值为 0, 方差 $\sigma = 1$ 的高斯变量, P_m 是变异概率, $f(\ast)$ 是 v 的亲和度。

2) 计算 v' 和 v 的适应度 $f(V')$, $f(V)$, 若 $\min\{1, \exp(-(f(V') - f(V))/T_k)\} > \text{random}[0, 1]$, 则按照退火接受准则,接受新解;否则,放弃变异后的抗体 v' 。 T_k 为第 k 次进化的温度。

2.6 克隆交叉算子

交叉是指把两个父代抗体的部分结构加以替换重组而生成新抗体的操作,交叉后临时的抗体群为 C_r , 克隆交叉操作如下:

1) 在 $[0, 1]$ 内随机生成一个数 $r_c = \text{random}[0, 1]$, 若 $r_c < P_c$, 则按下式进行交叉操作: $V'_1 = r_c \times V_1 + (1 - r_c) \times V_2$, $V'_2 = r_c \times V_2 + (1 - r_c) \times V_1$ 。 V_1, V_2 分别是父抗体, V'_1, V'_2 分别为子抗体, P_c 是交叉概率。

2) 按照 2.5 节中的变异步骤 2) 的退火接受准则决定是否接受交叉后的解。

2.7 克隆选择算子

对抗体群 $A(k), C, M, C_r$ 进行克隆选择操作,利用亲和度函数选出亲和度最高且互不相同的 N 个抗体组成新抗体群 $A(k+1)$ 。

2.8 算法描述

1) 随机在解空间产生初始抗体群 $A = \{a_1, a_2, \dots, a_N\}$, 并初始化种群规模 N , 交叉概率 P_c , 变异概率 P_m , 初始温度 T_0 ;

2) 将抗体群 A 中的抗体按照亲和度由大到小降序排列, 得到 $A(k) = \{b_1, b_2, \dots, b_N\}$, 且 $f(b_i) \geq f(b_{i+1})$, $f(\ast)$ 为亲和度函数, $k = 0$;

- 3) 计算抗体群中每个抗体 b_i 的亲密度 $f, i = 1, 2, \dots, N$;
- 4) 将第 i 个抗体克隆 $\text{INT}(\alpha \times N/i)$ 个, 得到规模为 N_c 的克隆后抗体群 $C, i = 1, 2, \dots, N$;
- 5) 对抗体群 C 进行克隆变异操作, 变异后临时的抗体群为 M ;
- 6) 对抗体群 M 进行克隆交叉操作, 交叉后临时的抗体群为 C_r ;
- 7) 对抗体群 C_r 进行克隆选择操作, 选择操作后产生新抗体群 $A(k+1)$;
- 8) 当 $T_k = 0$ 时, 退火过程自然结束, 输出最优解; 否则, $T_{k+1} = T_k \times (1 - k/M), k = k + 1$, 返回到 2)。

3 仿真实验

3.1 实验环境及参数选取

本文对任务调度算法进行了仿真实验, 以验证算法的正确性并对其性能进行评价。实验中, 模拟了由四个站点组成的网格计算环境下的网格任务调度情况, 每个站点的资源情况如表 1 所示。算法主要参数为: 种群大小为 100, 克隆系数为 150, 杂交概率为 0.85, 变异概率为 0.1, 初始温度为 1。

表 1 仿真环境

站点名称	cluster 配置	cluster 数量	管理软件
site1	64 节点	8	LSF 6.0
site2	128 节点	4	LSF 6.0
site3	128 节点	2	LSF 6.0
site4	256 节点	2	LSF 6.0

为更好地评估算法的性能, 下面定义资源平均利用率。

定义 4 资源平均利用率 v 。假设 v_j 为资源节点 r_j 的利用

率, 则 v_j 可以表示为: $v_j = \frac{\sum_{i=1}^n \Delta_{ij}}{l_j} \times 100\%$ 。

3.2 实验结果

图 1 为计算资源在 PSACA 算法部署前的负载统计情况。图 2 为该算法部署后, 计算资源的负载统计情况。图 3 是 PSACA 算法, GA 算法, CA 算法在 300 次进化中得到的最优解的完成时间。

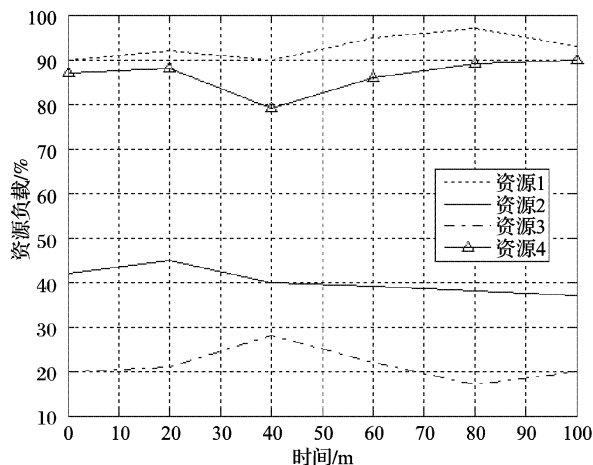


图 1 PSACA 算法部署前资源的负载统计

3.3 结果分析

从图 1 可以看出, 在 PSACA 算法部署前, 资源 1, 4 由于任务过多导致资源负载过重 (资源 1, 4 的平均利用率分别为 99.3%, 93.5%), 以至于部分网格用户无法继续提交任务, 而

资源 2 和资源 3 的负载过低 (资源 2, 3 的平均利用率分别为 40.1%, 23.6%), 资源没有得到充分地利用。在图 2 中, PSACA 算法部署后, 虽然资源 1, 4 的负载有所降低 (资源 1, 4 的平均利用率分别为 85.1%, 80.5%), 但是资源 2 和资源 3 的负载有了很大的提高 (资源 2, 3 的平均利用率分别为 78.2%, 73.6%), 从整体上做到了对资源的负载均衡。在图 3 中, 如果以最小完成时间进行评估, PSACA 算法能有效地搜索到全局最优值, 同时得到的优化极值明显优于其他算法。如果从稳定性方面进行评估, PSACA 算法的任务完成时间最大值和最小值相差不大, 标准偏差小, 说明该算法具有较好的稳定性; 如果从早熟收敛方面进行评估, PSACA 算法收敛时迭代次数最多, 有效避免早熟收敛现象。在仿真实验中, PSACA 算法对第 i 个抗体克隆 $\text{INT}(\alpha \times N/i)$ 个, 抗体群克隆后的规模为 N_c , 则总的计算复杂度是 $O(n^\beta)$ ($1 \leq \beta \leq 2$) 数量级^[5], 远小于经典方法的 $O(2^k)$ 数量级。

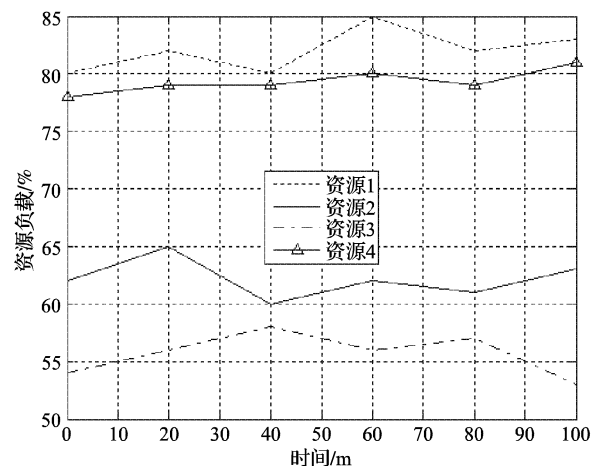


图 2 PSACA 算法部署后资源的负载统计

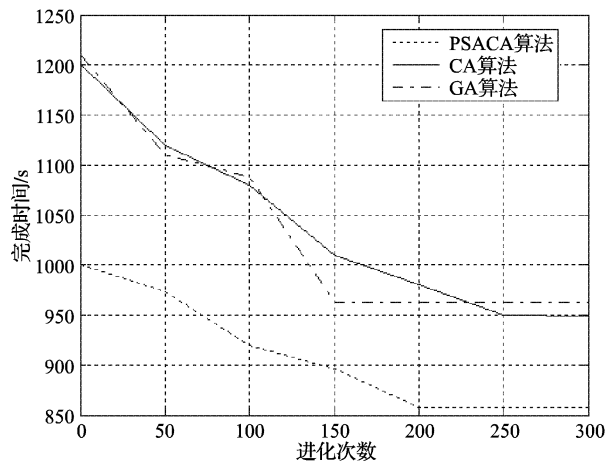


图 3 三种算法的任务完成时间比较

4 结语

针对网格计算中任务调度问题建立了优化调度模型, 并基于 PSACA 算法给出了问题的求解办法。PSACA 算法充分利用免疫克隆算法和模拟退火算法的优点, 有效地实现了网格环境下资源的负载均衡, 克服了遗传算法收敛速度慢, 易陷入局部最优的缺点。实验结果证明, PSACA 算法可以成功地应用于网格任务调度。

参考文献:

- [1] BUYYA R, ABRAMSON D, GIDDY J. Grid resource management, scheduling, and computational economy [C]// Proceedings of the

- 1st International Workshop on Grid and Cooperative Computing. Washington, DC: IEEE Computer Society, 2000: 1734 – 1739.
- [2] 李春林, 卢正鼎, 李腊元. 基于 Agent 的计算网格资源管理[J]. 武汉理工大学学报: 交通科学与工程版, 2003, 29(15): 110 – 115.
- [3] HAN YAO-JUN, JIANG CHANG-JUN, YOU FU. Resource scheduling algorithm for grid computing and its modeling and analysis using petri net[C]// Proceedings of the 2nd International Workshop on Grid and Cooperative Computing. Washington, DC: IEEE Computer Society, 2003: 814 – 820.
- [4] WENG CHU-LIANG, LU XIN-DA. A cost-based on-line scheduling algorithm for job assignment on computational grids[C]// Proceedings of the 5th International Workshop on Advanced Parallel Processing Technologies: APPT 2003, LNCS 2834. Berlin: Springer-Verlag, 2003: 343 – 351.
- [5] SUBRAMMANI V. Distributed job scheduling on computational grids using multiple simultaneous requests[C]// Proceedings of the 11th IEEE International Symposium on High Performance Distributed Computing. New York: American Scholar Press, 2002: 359 – 367.
- [6] XU ZHI-HONG, HOU XIANG-DAN, SUN JI-ZHOU. Ant algorithm-based task scheduling in grid computing[C]// Proceedings of the Canadian Conference on Electrical and Computer Engineering. Washington, DC: IEEE Computer Society, 2003: 1107 – 1110.
- [7] DI MARTINO V, MILIOTTI M. Scheduling in a grid computing environment using genetic algorithm[C]// Proceedings of the 16th International Parallel and Distributed Processing Symposium. Washington, DC: IEEE Computer Society, 2002: 678 – 686.
- [8] DI MARTINO V, MILIOTTI M. Sub-optimal scheduling in a grid using genetic algorithm[J]. Parallel Computing, 2004, 30(5/6): 553 – 565.
- [9] YAO WEN-SHENG, LI BAI-YAN, YOU JIN-YUAN. Genetic scheduling on minimal processing elements in the grid[C]// Proceedings of the 15th Australian Joint Conference on Artificial Intelligence: Advances in Artificial Intelligence, LNCS 2557. Berlin: Springer-Verlag, 2002: 465 – 476.
- [10] ABRAHAM A, BUYYA R, NATH B. Nature's heuristics for scheduling jobs on computational grids[C]// Proceedings of the 8th IEEE International Conference on Advanced Computing and Communications: ADCOM 2000. Washington, DC: IEEE Computer Society, 2000: 27 – 38.
- [11] 李季, 钟将, 吴中福. 具有模糊处理时间的网格任务调度免疫算法[J]. 计算机科学, 2006, 33(2): 35 – 37.
- [12] 刘红, 白栋, 丁炜. 应用于 MPLS 网络负载均衡的启发式自适应遗传算法研究[J]. 通信学报, 2003, 24(10): 39 – 45.
- [13] 刘芳, 李阳阳. 量子克隆进化算法[J]. 电子学报, 2003, 31(12A): 2066 – 2070.
- [14] 李阳阳, 焦李成. 求解 SAT 问题的量子免疫克隆算法[J]. 计算机学报, 2007, 30(2): 176 – 183.
- [15] 焦李成, 杜海峰. 人工免疫系统进展与展望[J]. 电子学报, 2003, 31(10): 1540 – 1548.
- [16] YU HONG-MEI, YAO PING-JING. Combined genetic algorithm/simulated annealing algorithm for large-scale system energy integration[J]. Computers and Chemical Engineering, 2000, 8(24): 2023 – 2035.

(上接第 2849 页)

BP 神经网络都有较好的分类能力,可以在一定程度弥补其缺陷。本文所提出的 RGATAFSVM 模型相对其他传统方法而言,分类精度高,对水质数据具有较好的分类能力,同时方差较低,说明该方法对数据有较好的稳定性。

4 结语

TAFSVM 能够解决线性和非线性问题,具备标准支持向量机的所有优点,还具有处理过拟合问题、自适应性、提高泛化误差边界等优点。运用实值遗传算法优选 TAFSVM 参数,速度快,效率高,是进行 TAFSVM 参数调节的一种有效方法。RGATAFSVM 模型对水质分类具有较高精度和较高稳定性,是进行水质分类的一种有效方法。

参考文献:

- [1] 李如忠. 水质预测模式理论研究进展与趋势分析[J]. 合肥工业大学学报, 2006, 29(1): 26 – 30.
- [2] 胡惠彬. GM(1, 1) 模型在地表水浓度预测中的应用[J]. 中国环境监测, 1993, 9(4): 45 – 46.
- [3] 高学民, 陈静生, 王立新. BP 网络应用于长江水质研究[J]. 环境科学研究, 2001, 14(1): 49 – 52.
- [4] MOATAR F, FESSANT F, POIREL A. PH modeling by neural networks[J]. Application of control and validation data series in the Middle Loire river, 1999, 120(2/3): 141 – 156.
- [5] LOUCKS D P, LYNN W R. Probabilistic models for predicting stream quality[J]. Water Resources Research, 1996, (3): 593 – 605.
- [6] SIVAKUMAR B, LIONG EY. Evidence of chaotic behavior in Singapore rainfall[J]. Water Resources Research, 1998, 34(2): 301 – 310.
- [7] CRISTIANINI N, SHAW-TAYLOR J. 支持向量机导论: 英文版[M]. 北京: 电子工业出版社, 2005.
- [8] CHEN XIAO-LING, LI Y S, LIU ZHI-GANG, *et al.* Integration of multi-source data for water quality classification in the Pearl River estuary and its adjacent coastal waters of Hong Kong[J]. Continental Shelf Research, 2004, 24(16): 1827 – 1843.
- [9] 徐红敏, 杨天行. 基于支持向量机分类算法的湖泊水质评价研究[J]. 吉林大学学报: 地球科学版, 2006, 36(4): 570 – 573.
- [10] 徐劲力. 支持向量机在水质评价中的应用[J]. 中国农村水利水电, 2007(3): 7 – 9.
- [11] LIU Y H, CHEN Y T. Face recognition using total margin-based adaptive fuzzy support vector machines[J]. IEEE Transactions on Neural Networks, 2007, 18(1): 178 – 192.
- [12] KUMAR S, NARESH R. Efficient real coded genetic algorithm to solve the non-convex hydrothermal scheduling problem[J]. International Journal of Electrical Power and Energy Systems, 2007, 29(10): 738 – 747.
- [13] Campbell C. Kernel methods: A survey of current techniques[J]. Neurocomputing, 2002, 48(1/4): 63 – 84.
- [14] DUAN K, KEERTHI S S, POO A N. Evaluation of simple performance measures for tuning SVM hyperparameters, CD-01-11[R]. Singapore: National University of Singapore, 2001.
- [15] 谢金星. 全国大学生数学建模竞赛[EB/OL]. (2005 – 09 – 15). [2008 – 01 – 02]. <http://mcm.edu.cn/DEFAULTc.HTM>.
- [16] 孙涛, 潘世兵, 李永军. 人工神经网络模型在地下水水质评价分类中的应用[J]. 水文地质工程地质, 2004(3): 58 – 61.
- [17] 王洪礼, 王长江, 李胜朋. 基于支持向量机理论的海水水质富营养化评价研究[J]. 海洋技术, 2006, 24(1): 48 – 51.