

# 人工鱼群算法的改进

唐莉<sup>1</sup>, 张正军<sup>1</sup>, 王俐莉<sup>2</sup>

(1. 南京理工大学理学院, 江苏南京 210094;  
2. 海军指挥学院科研部, 江苏南京 210016)

**摘要:**人工鱼群算法(AFSA)是一种新型随机搜索优化算法。通过初步的研究表明,该算法具有许多优良的性质,但也有些不足之处。针对均匀随机行为和常数拥挤度因子而导致算法运行时间长或陷入局部最优的问题,引入对称正态随机行为,自适应调整该行为参数,减少了由于迂回搜索导致的无用计算,也使人工鱼可在解空间进行更为广泛的搜索,提高了搜索效率。采用了自适应拥挤度因子并提出新的适应度函数,加快了系统满意解的收敛速度,使数值解更加稳定。实验结果表明,与基本人工鱼群算法相比,该方法具有明显的优越性。

**关键词:**随机行为;拥挤度因子;适应度函数;人工鱼群算法;优化

**中图分类号:**TP301.6

**文献标识码:**A

**文章编号:**1673-629X(2016)11-0037-04

doi:10.3969/j.issn.1673-629X.2016.11.008

## Improvement of Artificial Fish Swarm Algorithm

TANG Li<sup>1</sup>, ZHANG Zheng-jun<sup>1</sup>, WANG Li-li<sup>2</sup>

(1. School of Science, Nanjing University of Science and Technology, Nanjing 210094, China;  
2. Research and Development Section, Nanjing Naval Command Academy, Nanjing 210016, China)

**Abstract:** Artificial Fish Swarm Algorithm (AFSA) is a new random search optimization algorithm. The preliminary study shows that it has many promising features, but also some disadvantages. Aiming at the problem of AFSA, such as long running time or being in local optimal, caused by uniformly random behavior and constant of congestion factor. Based on symmetric normality random behavior, self-adaption adjusts the parameter of this behavior, and a large number of unused circuitous searches are reduced, and a more complete search within solution space is obtained for artificial fishes so that a high search efficiency is arrived at. The self-adaption congestion factor is adopted and a new fitness function is proposed, increasing the convergence rate of satisfactory solution domain, making the result more stable. Results of experiments show that there is an obvious advantage for this improved method compared with the basic artificial fish-swarm algorithm.

**Key words:** random behavior; congestion factor; fitness function; artificial fish swarm algorithm; optimization

## 0 引言

随着动物的进化,根据优胜劣汰的自然法则,它们形成了各式各样的生存方式,这些方式为人类带来了许多灵感和启发。一个动物集群通常定义为一群自治体的集合,自治体是指在自然环境中具备自身活动能力的一个实体。将自治体的概念引入动物优化算法中,通过解决局部问题逐步达到解决最终问题的目的,从而形成人工鱼群算法。该算法是由李晓磊等长时间通过对鱼群生活习性的观察而模仿鱼类生存行动方式提出的,是一种基于动物行为的新型仿生优化随机搜索算法<sup>[1]</sup>,其归纳总结出符合鱼群行为并适用于鱼群

算法的几种典型行为:觅食行为、聚群行为、追尾行为和随机行为。该算法根据“富含营养物质最多的地方一般是水域中鱼生存数目最多的地方”这一特点来模拟各行为,从而实现全局优化,是集群智能思想的一个具体应用,具有并行性、简单性、快速性、较强的鲁棒性和较好的收敛性等特点。但也存在不足之处:

(1)算法在寻优过程中的随机行为,存在迂回搜索的问题,对于求解多极值函数,由于随机行为移动步长太小导致收敛速度慢甚至难以跳出局部最优解。

(2)拥挤度因子 $\delta$ 与人工鱼领域内伙伴的数目 $n_f$ 相结合决定人工鱼群是否执行追尾行为和聚群行为,

收稿日期:2016-01-27

修回日期:2016-05-05

网络出版时间:2016-10-24

基金项目:全国统计科学研究计划重点项目(2013LZ45)

作者简介:唐莉(1993-),女,硕士研究生,研究方向为数据挖掘与分析;张正军,博士,副教授,研究方向为数据挖掘与分析。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20161024.1114.060.html>

避免人工鱼过度拥挤而陷入局部极值,但在算法后期,固定数值的拥挤度因子会使得位于极值点附近的人工鱼之间相互排斥,不能进行觅食、聚群和追尾行为,从而无法精确地逼近极值点。

2009 年王联合国在基本人工鱼群基础上提出全局版人工鱼群<sup>[2]</sup>;2007 年范玉军等对人工鱼移动方式进行改进<sup>[3]</sup>;刘彦君等采用自适应视野和步长对人工鱼群进行改进<sup>[4-6]</sup>,采用跳跃行为<sup>[7]</sup>;黄光球等利用网格划分策略和禁忌搜索算法做了适当改进<sup>[8]</sup>,并证明了其全局性收敛<sup>[9-10]</sup>;2008 年王联合国等基于领域正交叉算子动态调整视野和步长<sup>[11]</sup>。

文中引入对称正态行为,自适应调整正态分布方差,并提出新的自适应拥挤度因子对基本人工鱼群算法进行改进。

# 1 基本人工鱼群算法及参数拥挤度因子 $\delta$ 的性能分析

## 1.1 相关定义

人工鱼状态定义为  $X = (x_1, x_2, \dots, x_n)$ , 其中  $x_i (i = 1, 2, \dots, n)$  为寻优变量;每条人工鱼当前食物浓度表示为  $Y = f(X)$ , 其中  $Y$  为目标函数值,是一个关于  $x_i$  的  $n$  元函数;任意两条人工鱼个体之间的距离表示为  $d_{ij} = \|X_i - X_j\|$ ;人工鱼所能感应到的视野范围表示为  $Visual$ ;  $Step$  表示人工鱼移动步长; $\delta$  表示拥挤度因子; $N_p$  表示鱼群个数; $try\ number$  表示尝试次数。

## 1.2 基本行为

### (1) 觅食行为。

设一条人工鱼  $i$  当前的状态为  $X_i$ , 在其视野范围  $Visual$  内随机选择另一个状态  $X_j$ ,  $X_j = X_i + Visual \cdot Rand()$ 。其中  $Rand()$  是一个介于 0 和 1 之间的随机数。根据  $Y = f(X)$  计算两条人工鱼的食物浓度并进行比较,若求解极大值问题,需  $Y_i < Y_j$  (求极小值为  $Y_i > Y_j$ , 由于极大值与极小值问题可以通过添加负号相互转换,以下考虑极大值问题),则该人工鱼  $X_i$  向该方向移动一步,到达位置  $X_i^{t+1}$ ,  $X_i^{t+1} = X_i^t + \frac{X_j - X_i^t}{\|X_j - X_i^t\|} \cdot Step \cdot Rand()$ 。若不满足  $Y_i < Y_j$ , 则要在视野范围内重新随机选择状态  $X_j$ , 继续判断是否满足前进条件  $Y_i < Y_j$ , 若反复尝试  $try\ number$  次后,仍不满足前进条件,则进行随机行为。

### (2) 聚群行为。

为了保证鱼群的生存和躲避危害,鱼在游动过程中会自然地聚集成群。在人工鱼群算法中对每条人工鱼做如下规定:一是避免每条人工鱼相互之间过分拥挤;二是人工鱼的移动尽量向临近伙伴的中心靠拢。

设人工鱼当前状态为  $X_i$ , 在当前领域 ( $d_{ij}$

$< Visual$ ) 内搜索,该领域内伙伴数目为  $n_f$ , 中心位置为  $X_c$ 。若满足  $Y_c/n_f > \delta Y_i$ , 则表示该领域不太拥挤且伙伴中心有较多食物,于是该人工鱼便朝伙伴中心位置  $X_c$  移动,到达位置  $X_i^{t+1}$ ,  $X_i^{t+1} = X_i^t + \frac{X_c - X_i^t}{\|X_c - X_i^t\|} \cdot$

$Step \cdot Rand()$ ;若不满足前进条件,则进行觅食行为。

### (3) 追尾行为。

根据对自然习性的观察,若一条鱼或者几条鱼发现食物,邻近伙伴尾随其则可以达到快速到达食物点的目的,这是一种向着邻近且拥有最高目标函数值的人工鱼进行追逐的行为,在寻优算法中表示为向邻近最优伙伴前进的过程。设人工鱼  $i$  当前状态为  $X_i$ , 在当前领域 ( $d_{ij} < Visual$ ) 内搜索,令该领域内人工鱼目标函数值  $Y_j$  为最大值的伙伴所处位置为  $X_j$ 。若满足  $Y_c/n_f > \delta Y_i$ , 表明该伙伴  $X_j$  的状态具有较高的食物浓度且四周不太拥挤,则朝最优伙伴位置  $X_j$  移动,到达位置  $X_i^{t+1}$ ,  $X_i^{t+1} = X_i^t + \frac{X_j - X_i^t}{\|X_j - X_i^t\|} \cdot Step \cdot Rand()$ ;若不满足前进条件,则进行觅食行为。

### (4) 随机行为。

在需要移动的人工鱼视野范围内随机选择一个状态,然后向该方向移动,该行为事实上是觅食行为的缺省行为:

$$X_i^{t+1} = X_i^t + Visual \cdot Rand() \tag{1}$$

## 1.3 拥挤度因子 $\delta$ 对算法性能的影响

### 1.3.1 拥挤度因子定义

在求极大值问题中,一般设  $\delta = 1/(\alpha n_{max})$ ,  $\alpha \in (0, 1]$ 。其中,  $\alpha$  为极值接近水平;  $n_{max}$  为在该领域内聚集期望达到的最大人工鱼数目。在求极小值问题中,一般设  $\delta = \alpha n_{max}$ ,  $\alpha \in (0, 1]$ 。在基本人工鱼群算法中,  $\delta$  一般取固定值。

### 1.3.2 参数 $\delta$ 的性能分析

$\delta$  的引入避免了人工鱼过度拥挤而陷入局部极值,与  $n_f$  相结合,通过决定人工鱼是否执行追尾行为和聚群行为对寻优产生影响。实验证明<sup>[12]</sup>,  $\delta$  越大,人工鱼觅食或者随机行为比较突出,摆脱局部极值能力越强;  $\delta$  越小,聚群行为和追尾行为在寻优过程中的作用逐渐凸显,有利于提高极值的精确度。而由于每次迭代人工鱼只向鱼群中心位置或鱼群最优位置移动一步,因此影响了算法收敛速度。

## 2 改进的人工鱼群算法

### 2.1 引入对称正态随机行为

在基本人工鱼群算法中,觅食行为奠定了算法收敛的基础,聚群行为提供了算法收敛的稳定性,追尾行为增强了算法收敛的快速性和全局性,而随机行为可

以逃出局部极值域。然而对于一些多极值函数,特别是达到局部极值与全局最优值的人工鱼所处位置相差甚远时;第一,若该人工鱼  $X$  达到局部极值点,而随机行为的步长过小,那么迭代多次仍然可能跳不出局部极值;第二,当人工鱼  $X$  迭代至逼近局部极值的位置附近时,若根据该人工鱼的行为评价选择随机行为,那么随机移动步长的过小会使得人工鱼  $X$  在局部最优解的附近迂回,反之随机移动步长的过大容易使得人工鱼  $X$  跳过全局极值  $X_{\text{best}}$  附近,从而增加收敛时间。

根据上述分析,文中引入对称正态随机行为,采用正态分布随机数调整随机行为中的移动步长  $\text{step}$ ,定义该随机数的方差  $\sigma$  为:

$$\sigma = (\|\nabla f\| + \varepsilon)^{-1} \tag{2}$$

动态搜索增强了全局搜索能力,在一定程度上削弱了振荡现象对优化精度的影响并提高了收敛速度。

2.2 自适应拥挤度因子  $\delta$

随着  $\delta$  从大变小,人工鱼逐渐从觅食行为或者随机行为变为聚群行为或追尾行为,在进行全局搜索迭代过程后逼近全局极值点,再逐步进行精确搜索,使结果更加准确。根据上述分析,针对聚群行为和追尾行为中的拥挤度因子  $\delta$ ,定义  $X_i$  在当前位置的适应度函数  $f(X_i)$  [13-14] 为:

$$f(X_i) = \max\{0, \frac{1}{N(X_i)} \sum_{X_j \in N(X_i)} (1 - \frac{d_{i,j} \cdot d_{j,i}}{\alpha_{i,j}})\} \tag{3}$$

其中,  $N(X_i)$  表示当前人工鱼视野范围内感应到的人工鱼个数;  $d_{i,j}$  表示任意两条人工鱼之间的距离,通常用欧几里德距离;参数  $\alpha_{i,j}$  为:

$$\alpha_{i,j} = (\frac{1}{n} \sum_{k=1}^n d_{i,k}) \cdot (\frac{1}{n} \sum_{k=1}^n d_{j,k}) \tag{4}$$

采用指数式衰减变化策略[10]:  $\delta = \delta \cdot T$ 。其中,  $T = e^{-\beta \cdot f(X)}$ ,  $\beta$  为一固定值,那么  $T$  称为衰减因子。随着迭代次数的增加,  $\delta$  自适应减小,有利于提高搜索解的精度。

2.3 改进的人工鱼群算法步骤

(1) 初始参数设置,包括人工鱼群的总数  $N_p$ 、每条人工鱼的初始位置  $X$ 、移动步长  $\text{Step}$ 、视野范围  $\text{Visual}$ 、尝试次数  $\text{try number}$  以及拥挤度因子  $\delta$ 。

(2) 计算每条人工鱼当前的目标函数值  $Y = f(X)$ ,并在公告板中记录所有函数值中全局最优的人工鱼状态  $X$ 。

(3) 对每条人工鱼评价即到达全局最优的人工鱼保持不动,未到达的进行移动并选择所要执行的行为,包括觅食行为、聚群行为、追尾行为和对称正态随机行为。

(4) 将人工鱼移动后与公告板中的值进行比较,更新每条人工鱼数据前所处的位置信息。

(5) 更新全局最优人工鱼的状态。

(6) 根据循环条件,一般为判断连续所得最优目标函数值的均方差小于允许的误差,或判断聚在某个区域内的人工鱼数目达到某个值(某个比例),或判断连续多次所获取的值均不超过已寻找的极值的次数,或设置最大迭代次数等,则输出结果,否则返回步骤(2)。

3 仿真实现

在 Matlab7.0 环境下,选取两个经典测试函数进行实验。

实例 1:函数。

$$\max f(x,y) = \frac{\sin(x)}{x} \cdot \frac{\sin(y)}{y}, -10 \leq x,y \leq 10 \tag{5}$$

该函数在(0,0)处存在唯一极大值 1,原点附近散布着一些局部极值。

参数选取:人工鱼总数  $N_p = 20$ ,迭代次数  $\text{IT} = 50$ ,步长  $\text{Step} = 0.7$ ,尝试次数  $\text{try number} = 5$ ,视野范围  $\text{Visual} = 10$ 。下面分别在图 1 和图 2 中给出基本人工鱼群算法(AFSA)和改进人工鱼群算法(IAFSA)公告板中最优值的收敛曲线。其中,b\_value 曲线代表公告板最优值,afs\_value 表示每经过一次迭代后所有人工鱼中的最优值。并在表 1 中记录了两种算法的执行时间、最优人工鱼极值点(best\_x,best\_y)及公告板最优值 b\_value。

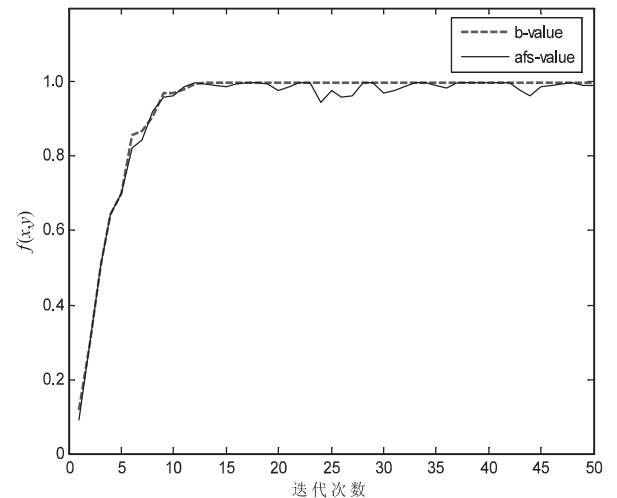


图 1 AFSA 实例 1 的收敛曲线

表 1 算法性能对比(1)

指标	AFSA	IAFSA
执行时间/s	1.591 000	1.576 000
best_x	3.322 899e-001	-1.723 285e-001
best_y	-2.316 029e-001	-3.416 143e-002
b_value	9.999 037e-001	9.999 134e-001

从表 1 可以看出,在时间上 IAFSA 算法比 AFSA 加快 1%,距极大值点(0,0)近 0.229 356 9,且更加收敛于极大值 1。

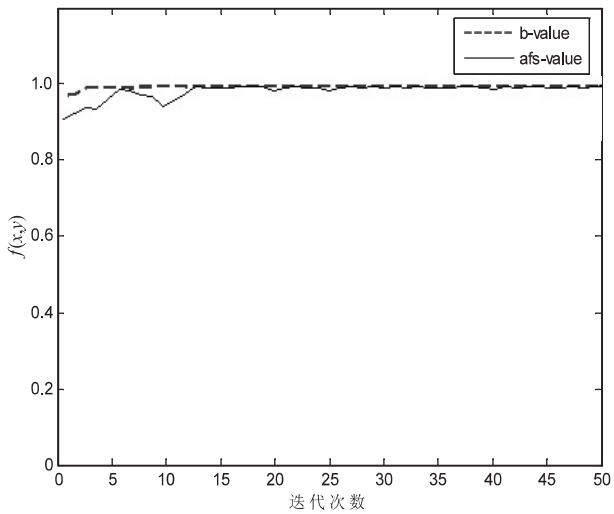


图 2 IAFSA 实例 1 的收敛曲线

实例 2;函数( Needle-in-a-haystack )。  
$$\max f(x,y) = \left( \frac{3}{0.05 + (x^2 + y^2)} \right)^2 + (x^2 + y^2)^2, \\ -5.12 \leq x,y \leq 5.12 \quad (6)$$

该函数为典型的多极值问题,其中(0,0)点为全局极值点,且全局最优值为 3 600,(-5.12,5.12)、(-5.12,-5.12)、(5.12,-5.12)、(5.12,5.12)为函数的四个局部极值点,且局部极值为 2 748.782 3。

参数选择:人工鱼总数  $N_p = 60$ ,迭代次数  $IT = 200$ ,步长  $Step = 0.2$ ,尝试次数  $try\ number = 1$ ,视野范围  $Visual = 2$ 。

图 3、图 4 分别为 AFSA 和 IAFSA 的收敛曲线,表 2 为对应的执行时间、最优人工鱼位置(best\_x,best\_y)及公告板最优值 b\_value。

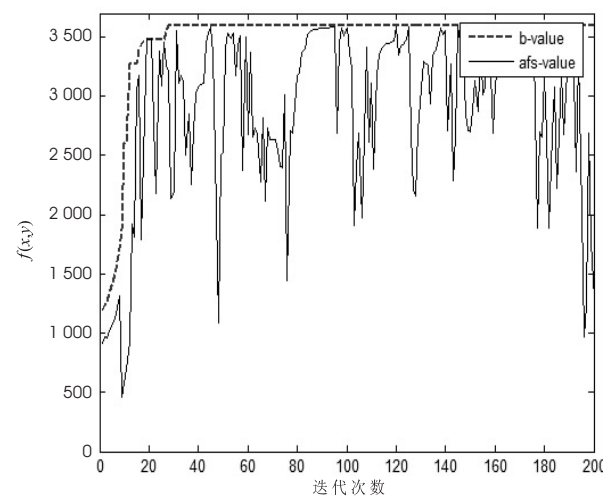


图 3 AFSA 实例 2 的收敛曲线

从表 2 可见,IAFSA 算法比 AFSA 时间加快了 1%,距极大值点(0,0)近 0.008 938 7,且同样更收敛

于极大值 1。

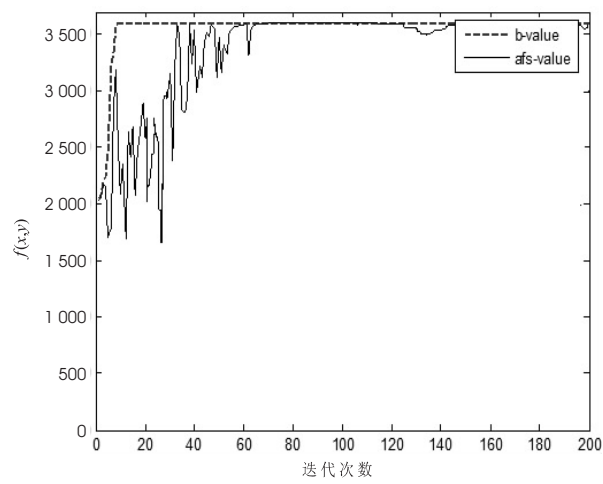


图 4 IAFSA 实例 2 的收敛曲线

表 2 算法性能对比(2)

指标	AFSA	IAFSA
执行时间/s	33.290 000	32.963 000
best_x	-3.070 736e-002	2.420 133e-002
best_y	2.586 932e-002	1.971 042e-002
b_value	3.597 964e+003	3.599 593e+003

综上所述,基本人工鱼群算法获取的是近似最优解,改进的人工鱼群算法不仅缩短了算法运行时间,而且具有更好的稳定性和更高的优化精度。

4 结束语

人工鱼群算法是一种新型随机搜索优化算法,具有许多优良的性质。文中引入对称正态随机行为,对基本人工鱼群算法中的随机行为进行改进,减少了迂回搜索的无用计算,同时采用自适应拥挤度因子并提出新的适应度函数,加快了系统满意解的确定。结果表明,与基本人工鱼群算法相比,该方法有效可行。

参考文献:

[1] 李晓磊,邵之江,钱积新. 一种基于动物自治体的寻优模式:鱼群算法[J]. 系统工程理论与实践,2002,22(11):32-38.

[2] 王联国,洪毅,施秋红. 全局版人工鱼群算法[J]. 系统仿真学报,2009,21(23):7483-7486.

[3] 范玉军,王冬冬,孙明明. 改进的人工鱼群算法[J]. 重庆师范大学学报:自然科学版,2007,24(3):23-26.

[4] 刘彦君,江铭炎. 自适应视野和步长的改进人工鱼群算法[J]. 计算机工程与应用,2009,45(25):35-37.

[5] Jiang Mingyan, Mastorakis N E, Yuan Dongfeng. Multi-thresh-old image segmentation with improved artificial fish swarm algorithm[C]//Proc of European computing conference. Athens Greece; Springer-Verlag,2007.



块大小对于内存消耗的影响,但 PS 合并算法较之单文件合并算法,NameNode 节点内存消耗节省了 11.5%。由此可见,通过文件体积分布和利用文件关联关系,PS 文件合并算法有更好的效率和空间使用率。

4 结束语

文中提出了一种基于 HDFS 的小文件合并算法-PS 合并算法,能够优化 HDFS 在小文件存储方面表现出的不足,通过对文件关联关系的挖掘和文件体积的合理分布实现对小文件的合并重组,大幅降低了 Hadoop 集群中 NameNode 节点的内存消耗,减少了上传大量小文件所需时间。接下来的工作中可以在文件类型方面进行合并工作的研究,探讨文件类型对于数据处理以及文件合并的影响,以便能达到更好的效果。

参考文献:

[1] 李国琦. 中国产业互联网峰会[EB/OL]. 2015. <http://www.sootoo.com/content/548884.shtml>.

[2] Finta I, Farkas L, Szenasi S, et al. Buffering strategies in HDFS environment with STORM framework[C]//16th IEEE international symposium on computational intelligence and informatics. [s. l.]: IEEE, 2015: 297-302.

[3] 杜忠晖,何 慧,王 星. 一种 Hadoop 小文件存储优化策略研究[J]. 智能计算机与应用, 2015, 5(3): 28-32.

[4] Zhang Q F, Zhang W D, Li W J, et al. Cloud storage system for

small file based on P2P[J]. Journal of Zhejiang University (Engineering Science), 2013, 47(1): 7-8.

[5] Mackey G, Sehrish S, Wang J. Improving metadata management for small files in HDFS[C]//IEEE international conference on cluster computing and workshops. [s. l.]: IEEE, 2009: 1-4.

[6] Apache. The homepage of Hadoop[EB/OL]. 2012. <http://Hadoop.apache.org/>.

[7] Liu X, Han J, Zhong Y, et al. Implementing WebGIS on Hadoop: a case study of improving small file I/O performance on HDFS[C]//IEEE international conference on cluster computing and workshops. [s. l.]: IEEE, 2009: 1-8.

[8] 赵晓永,杨 扬,孙莉莉,等. 基于 Hadoop 的海量 MP3 文件存储架构[J]. 计算机应用, 2012, 32(6): 1724-1726.

[9] 游小容,曹 晟. 海量教育资源中小文件的存储研究[J]. 计算机科学, 2015, 42(10): 76-80.

[10] 江 柳. HDFS 下小文件存储优化相关技术研究[D]. 北京: 北京邮电大学, 2011.

[11] 李 铁,燕彩蓉,黄永锋,等. 面向 Hadoop 分布式文件系统的小文件存取优化方法[J]. 计算机应用, 2014, 34(11): 3091-3095.

[12] 张 丹. HDFS 中文件存储优化的相关技术研究[D]. 南京: 南京师范大学, 2013.

[13] 胡海峰,贾玉辰. 一种 Hadoop 存取海量小文件的优化方法: CN, CN104536959A[P]. 2015-04-22.

[14] 张春明,芮建武,何婷婷. 一种 Hadoop 小文件存储和读取的方法[J]. 计算机应用与软件, 2012, 29(11): 95-100.

(上接第 40 页)

[J]. 计算机工程, 2012, 38(2): 204-206.

[10] Iisufescu M. Finite Markov processes and their application[M]. Chichester, UK: Wiley Press, 1980.

[11] 王联国,施秋红. 人工鱼群算法的参数分析[J]. 计算机工程, 2010, 36(24): 169-171.

[12] 徐晓华,陈 峻. 一种自适应的蚂蚁聚类算法[J]. 软件学报, 2006, 17(9): 1884-1889.

[13] 朱 峰,陈 莉. 一种改进的蚁群聚类算法[J]. 计算机工程与应用, 2010, 46(6): 133-135.

[14] 陈广洲,汪家权,李传军,等. 一种改进的人工鱼群算法及其应用[J]. 系统工程, 2009, 27(12): 105-110.