

自适应遗传退火算法优化 BP 神经网络及其应用^①



裴 瑞, 白尚旺, 党伟超, 潘理虎

(太原科技大学 计算机科学与技术学院, 太原 030024)

通讯作者: 白尚旺, E-mail: swbai@tyust.edu.cn

摘 要: 以提高预测软件老化趋势为应用背景, 提出一种新型自适应遗传退火算法 (NAGSA) 优化 BP 神经网络模型, 该模型采用轮盘赌选择法与精英保留策略相结合的选择算子, 在迭代后期通过模拟退火算法对适应度函数进行拉伸, 相比传统的自适应遗传算法 (AGA) 在个体适应度较低时, 能够非线性地自适应调节交叉概率和变异概率, 从而对 BP 神经网络的权值和阈值优化并进行网络训练. 对在线售书网站注入内存泄漏的代码使之老化, 收集实验所需的老化数据进行仿真训练, 实验结果表明, NAGSA-BP 模型相比于传统遗传算法 (GA)、传统自适应遗传算法 (AGA)、传统自适应遗传退火算法 (NGSA) 优化的 BP 神经网络模型提高了预测精度和取得了优良的收敛效果, 在该应用领域验证了本文方法的有效性.

关键词: 优化; BP 神经网络; 遗传算法; 模拟退火算法

引用格式: 裴瑞, 白尚旺, 党伟超, 潘理虎. 自适应遗传退火算法优化 BP 神经网络及其应用. 计算机系统应用, 2019, 28(7): 109-113. <http://www.c-s-a.org.cn/1003-3254/6892.html>

Adaptive Genetic Annealing Algorithm for Optimizing BP Neural Network and its Application

PEI Rui, BAI Shang-Wang, DANG Wei-Chao, PAN Li-Hu

(School of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024, China)

Abstract: In order to improve the prediction accuracy of software aging, a New Adaptive Genetic Simulated Annealing algorithm (NAGSA) is proposed to optimize the BP neural network prediction model. The model's selection operator is combined with the elite retention strategy using the roulette selection method, stretching the fitness function by simulated annealing algorithm in the late iteration. Compared with the traditional Adaptive Genetic Algorithm (AGA), it can adaptively adjust the crossover probability and mutation probability nonlinearly when the individual fitness is low, thereby optimizing and weighting the BP neural network weights and thresholds, injecting a memory leak code into the online book-selling website to age it, collecting the aging data required for the experiment for simulation training. The experimental results show that the BP neural network model optimized by the NAGSA-BP model compared with the traditional Genetic Algorithm (GA), traditional AGA, and traditional Adaptive Genetic Simulated Annealing algorithm (NGSA) improves the prediction accuracy and achieves excellent results. The effectiveness of the proposed method is verified in this application field.

Key words: optimization; BP neural network; genetic algorithm; simulated annealing algorithm

① 基金项目: 山西省中科院科技合作项目 (20141101001); 山西省重点研发计划 (一般) 工业项目 (201703D121042-1); 山西省社会发展科技项目 (20140313020-1)

Foundation item: Science and Technology Collaborative Program Between Shanxi Province and Chinese Academy of Sciences (20141101001); Key Research and Development Program for General Industry of Shanxi Province (201703D121042-1); Science and Technology Program for Social Development of Shanxi Province (20140313020-1)

收稿时间: 2018-11-13; 修改时间: 2018-12-03, 2018-12-18; 采用时间: 2018-12-20; csa 在线出版时间: 2019-07-01

在一个长期运行的系统中,软件的退化和意外中断被称为软件老化现象,软件老化表现为系统的状态异常、性能退化、软件的挂起和失效^[1].已经证明了几种类型的系统遭受到软件老化,包括:Web服务器^[2],操作系统^[3],数据库管理系统^[4],云计算^[5]和虚拟化环境^[6]等.因此,对软件老化的趋势进行早期预测,并在系统崩溃之前采取有效的预防措施,有助于提高软件的可靠性和可用性,减少软件失效的发生和一些不必要的资源浪费及财产损失.本文通过对TPC-W基准测试标准的在线售书网站注入一段内存泄漏代码,收集JVM在120s内的平均内存变化数据作为实验的训练集和测试集.

由于BP神经网络独特的误差逆向传播算法不断的修正权值和阈值,不仅可以缩短训练时间还能提高拟合精度的特点,自适应遗传算法有着全局搜索能力和自适应优化方法,当种群规模较大时遗传算法可以在全局搜索范围内找到最优解,该过程简单易实现且求解快速的特点.

在此基础上提出了一种新型自适应遗传退火算法(NAGSA)优化BP神经网络模型,采用轮盘赌选择法和精英保留策略相结合的选择操作对种群中个体进行筛选,之后通过新型自适应遗传算法优化交叉概率和变异概率,由于遗传算法在运行后期个体的适应度逐渐趋于一致,优秀个体优势不足,这时需要通过模拟退火算法对适应度进行拉伸,之后将得到的种群中个体的解码赋值给BP神经网络的权值和阈值进行预测.实验表明本课题提出的NAGSA-BP神经网络模型,相比传统遗传算法改进的BP神经网络模型的预测精度更高、收敛性更好.

1 BP神经网络

BP(Back Propagation)神经网络是目前应用较广泛的一种多层前馈型神经网络,其网络训练采用误差逆传播算法^[7],它可以拟合任意的非线性映射从而降低预测误差.现有的基于BP神经网络模型虽在一定程度上取得了进展,但是依旧存在一些不足需加以改进.

2 自适应遗传退火算法优化BP神经网络

2.1 遗传算法优化BP神经网络模型

遗传算法的优点是全局式搜索,并且它可以利用历史信息来指导搜索进入到搜索空间内性能更好的区域,BP神经网络的优点是局部寻优,因此可结合两者

优势进行改进.遗传算法对BP神经网络有两方面的优化:权值优化和结构优化.权值优化:先通过遗传算法缩小搜索范围,再用BP算法使其快速收敛寻找最优解,以此来实现权值的优化.结构优化:通过形成不同长度的种群和不同结构的网络,各个种群经过进化之后选择一个最优个体,可确定各权值的初始值.但是传统的自适应遗传算法(AGA)随着进化的进行,一般先执行交叉操作后执行变异操作,这时种群的多样性不丰富并且进化速度很慢,最终导致收敛过慢或者不易收敛的后果.因此,本文采用闫春等人^[8]提出的新型自适应遗传算法,结合了模拟退火算法共同优化BP神经网络模型,该模型根据优化遗传算法的选择算子、交叉算子、变异算子克服了传统遗传算法易陷入“早熟收敛”的缺点,平衡了种群的多样性并且提高了收敛精度.

2.2 新型自适应遗传退火算法

2.2.1 编码

首先初始化种群,种群中的个体采用实数编码方式将神经网络中的权值和阈值进行优化,编码所得长度 H 为:

$$H = n \times l + l \times m + l + m$$

其中, n 为输入节点个数, l 为隐含层节点个数, m 为输出层节点个数, $n \times l$ 为输入层与隐含层之间的权重 ω_{ij} 的编码长度, $l \times m$ 为隐含层与输出层之间的权重 ω_{jk} 的编码长度, n 为隐含层阈值 α 的编码长度, m 为输出层阈值 β 的编码长度.

2.2.2 选择算子

基因选择遵从“优胜劣汰,适者生存”的原则进行基因选择.选择算子在整个种群中寻找适应度较高的个体去生成初始的交配池.本文选用轮盘赌法和精英保留策略相结合的方法计算个体的选择概率,将父代个体中适应度值最高的个体保留至下一代,剩下的个体按轮盘赌方法选择个体.

2.2.3 适应度

遗传算法通常采用评估函数——适应度函数来评估个体或者解的优劣,适应度越高的个体遗传给下一代的概率就越大,适应度越低的个体遗传给下一代的概率就越小.但是遗传算法容易在早期使个别优良的个体后代充斥整个种群,从而造成“早熟收敛”,而后期适应度函数基本趋于一致,使得优良个体后代优势不足造成种群进化停滞.本文引入模拟退火算法对适应度函数进行拉伸,在遗传算法前期温度较高,适应度相

似的个体产生后代的概率近似,到了遗传算法后期也就是温度逐渐下降之后,适应度相似的个体通过拉伸作用,放大了它们的适应度差值,使得优秀个体的优势更明显,可帮助算法跳出局部最优解.拉伸方法如下:

$$F_i = e^{f_i/T} / \sum_{i=1}^M e^{f_i/T}$$

$$T = T_0(0.99^{g-1})$$

其中, f_i 为第 i 个个体的适应度函数值, M 为种群个数, T 为退火温度, T_0 为初始退火温度, g 为遗传迭代次数.

2.2.4 交叉、变异操作

传统 AGA 算法思想是在个体适应度值高于平均适应度值时自适应的调节交叉概率、变异概率,反之则采用固定值的交叉概率、变异概率,所带来的不足是一些较差的个体中携带的较为优良的个体会遭到破坏.因此,本文采用非线性自适应交叉概率 P_c 和变异概率 P_m :

$$P_c = \begin{cases} k_1 \left(1 - \frac{\arcsin(f_a/f_m)}{\pi/2} \right), & \arcsin(f_a/f_m) \geq \pi/6 \\ k_1 \frac{\arcsin(f_a/f_m)}{\pi/2}, & \arcsin(f_a/f_m) < \pi/6 \end{cases}$$

$$P_m = \begin{cases} k_2 \left(1 - \frac{\arcsin(f_a/f_m)}{\pi/2} \right), & \arcsin(f_a/f_m) < \pi/6 \\ k_2 \frac{\arcsin(f_a/f_m)}{\pi/2}, & \arcsin(f_a/f_m) \geq \pi/6 \end{cases}$$

其中 f_a 是种群个体平均适应度, f_m 是种群个体适应度最

大值,当 f_a 变化时, $\arcsin(f_a/f_m)$ 会随之快速变化,因为 $\sin(\pi/6) = 1/2$ 当 $\arcsin(f_a/f_m) \geq \pi/6$ 时自变量 $f_a/f_m \geq 1/2$ 经过反复多次实验,通过 NAGSA 算法和 NGSA 算法中的 P_c 、 P_m 取值对训练误差精度的影响,当满足设置条件 $\pi/12 \leq \arcsin(f_a/f_m) \leq \pi/3$ 时, NAGSA 算法优先执行交叉操作再执行变异操作,则精度高于 NGSA 算法 25%;当满足设置条件 $\arcsin(f_a/f_m) \leq \pi/12$ 且 $\arcsin(f_a/f_m) \geq \pi/3$ 时, NAGSA 算法优先执行变异操作再执行交叉操作,则精度高于 NGSA 算法 22%. 得出设置条件 $\pi/12 \leq \arcsin(f_a/f_m) \leq \pi/3$, 若满足则 NAGSA 算法优先执行交叉操作,反之则优先执行变异操作.

将 NAGSA 算法训练后的权值和阈值作为 BP 神经网络的初始权值和阈值,用实验获得的样本数据建立 BP 神经网络模型.

3 实验验证

3.1 数据获取

造成软件老化的一个重要因素是由于内存的泄漏,本课题在服务器端使用符合 TPC-W 基准测试标准的在线售书网站,为其注入一段内存泄漏代码以加速老化.实验平台模拟了一个电子商务 Web 服务系统,其中包括运行在同一台物理机的虚拟机中的一个 Web 服务器,一个数据库服务器以及一组模拟的客户端,如表 1 实验配置所示.

表 1 实验配置

	服务器	数据库	客户端
操作系统	Linux CentOS7 Kernel 3.10.0-693.17.1.el7	Linux CentOS7 Kernel 3.10.0-693.17.1.el7	Linux CentOS7 Kernel 3.10.0-693.17.1.el7
CPU	Intel® Xeon® CPU E5506@2.13 GHz	Intel® Xeon® CPU E5506@2.13 GHz	Intel® Core™2 Duo CPU E8400@3.00 GHz 3.00 GHz
JVM	JDK 1.6.0		
软件	Apache Tomcat 6.0.33	TPC-W 客户端	MySQL 5.6.35

在符合 TPC-W 规范的在线售书网站中,每一个网页页面都使用单独的 Servlet 实现,这些类都工作在同一个 Java 虚拟机 (JVM) 中. JVM 是 Java 技术体系的核心,操作系统为 JVM 的进程分配内存,并由 JVM 进行储存和管理.其中 Java 堆是运行时的数据区域并随着 JVM 的启动而创建,所有实例类型和数组的内存均从此处分配.程序运行时,对象的堆内存由称为垃圾回收器 (GC) 的自动内存管理系统回收.由于内存的泄漏大多发生在 Java 堆,所以 GC 主要针对 Java 堆进行^[9].

当 JVM 内存使用量持续增加并达到最大值时, GC 将自动启动以释放内存,但是如图 1 JVM 中的内存泄漏所示, GC 仅回收无用和未引用的对象,而无用且引用的对象无法从占用的内存区域释放,从而导致内存泄漏.同时,由于系统资源是有限的,垃圾收集器在内存释放期间将占用大量的 CPU 和系统资源.因此,连续出现的内存泄漏问题的累积效应将直接导致可用内存不足和与老化相关的系统故障.

本文采用加速退化测试^[10](Accelerated Degradation

Tests, ADT) 实验研究受软件老化影响的应用程序故障, 通过增加一个 HeapLeak 类, 让服务器的生命周期都保持对该类的 OOMObject 对象的引用, 当服务器创建的实例数达到最大堆容量时, 会产生 JVM 堆的溢出, 从而 OOMObject 对象不再被 GC 回收. 实验收集 JVM 的内存使用量, 共持续 14 400 s, 每隔 120 s 取一次平均值. JVM 在 120 s 内平均使用内存的变化趋势图如图 2.

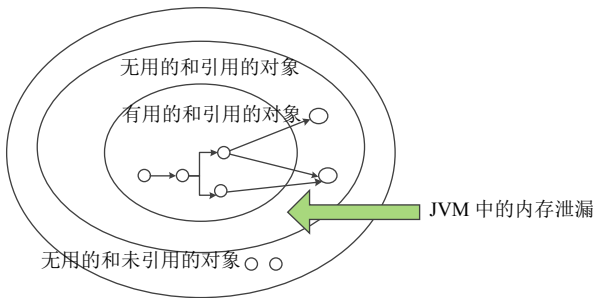


图 1 JVM 中的内存泄漏

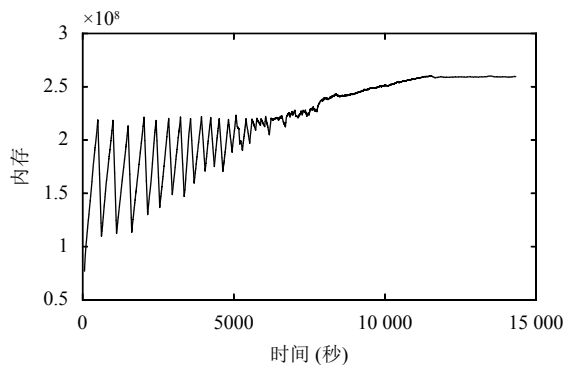


图 2 JVM 在 120 s 内平均使用内存变化趋势图

3.2 仿真实验

3.2.1 数据预处理

本次实验共获取到 3 个输入参数, 1 个输出参数, 神经网络结构为 3-10-1, 每组实验都选择工作负载为 100 个客户端, 共收集到 14 281 个数据, 取前 7000 个作为训练集, 后 7281 个作为验证集. 原始数据由于数量级差别较大, 会对 BP 神经网络的收敛性产生影响, 可能会加大训练难度, 耗费训练的时间, 因此为了提高网络的训练效率, 用 Matlab 中的归一化函数 mapminmax, 将输入层和输出层数据归一化到[-1, 1]之间, BP 神经网络训练后得到的数据最后进行反归一化处理, 得到正常值.

3.2.2 建立训练模型

本次实验的输入节点数为 3, 隐含层节点数为 10,

输出层节点数为 1, 种群规模为 60, 进化次数为 100 次, 交叉和变异概率分别选择 0.6 和 0.02. 为了验证本文提出的 NAGSA-BP 算法有着更好的收敛效果和验证软件老化预测上的有效性, 分别与 NGSA-BP、AGA-BP、GA-BP 算法进行比较, 选取 MAE、MSE 作为评价标准, 结果如表 2 所示.

$$MAE = \frac{1}{n} \sum_{h=1}^n |M_h - P_h| \quad MSE = \frac{1}{n} \sum_{h=1}^n (M_h - P_h)^2$$

表 2 改进后的遗传算法优化 BP 神经网络模型误差对比

类别	NAGSA-BP	NGSA-BP	AGA-BP	GA-BP
MSE	4.1105e+12	6.3589e+12	4.3326e+13	1.4711e+14
MAE	2.0274e+06	2.5216e+06	6.5822e+06	1.2128e+07

由表 2 可以看出, 本文提出的 NAGSA-BP 神经网络模型, 通过对交叉概率、变异概率的非线性优化, 相比传统的自适应遗传退火算法 (NGSA) 优化的 BP 神经网络模型在预测精度上提高了 35%, 相比传统的自适应遗传算法 (AGA) 优化的 BP 神经网络模型在预测精度上提高了 81%, 相比传统的遗传算法 (GA) 优化的 BP 神经网络模型在预测精度上提高了 97%.

由图 3、图 4 可以看出, NAGSA-BP 神经网络模型在加入了非线性自适应调整的交叉概率、变异概率之后, 其算法的稳定性和收敛性都优于未改进的 NGSA-BP 神经网络模型; NAGSA-BP、NGSA-BP 神经网络模型通过后期模拟退火算法对适应度函数的拉伸作用, 相比未改进的 AGA-BP、GA-BP 神经网络模型都在预测精度和拟合效果上有了显著的提高. 由此可得出, 本文提出的新型自适应遗传退火算法优化 BP 神经网络模型, 无论是在预测精度上还是拟合效果上都存在明显的优势.

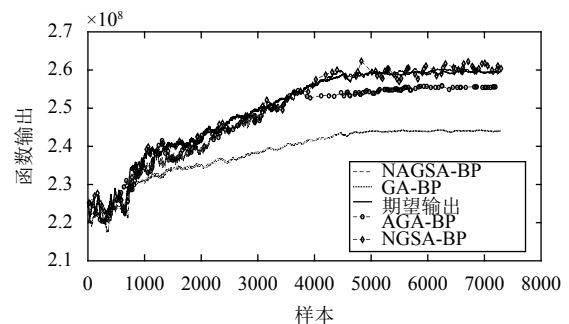


图 3 预测输出与期望输出对比图

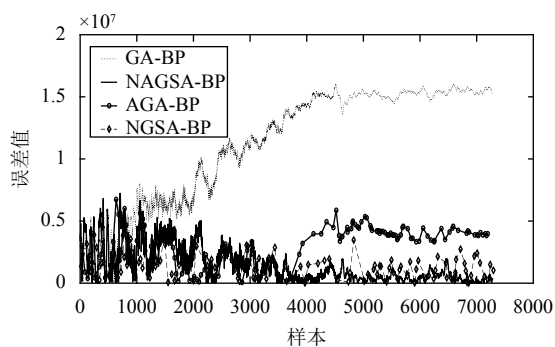


图4 误差对比图

4 结束语

本文通过建立一个软件老化测试平台,通过内存泄漏代码的注入,使得原本正常工作的在线书店网站出现老化现象,收集反应网站老化的系统参数,选取JVM在120s内平均使用的内存作为预测的数据集.实验表明,NAGSA-BP神经网络模型相比于几个传统遗传算法改进的BP神经网络模型,不仅提高了预测精度,还在预测结果的收敛性上有显著成效,验证了本文方法的有效性.

参考文献

- 1 Yan YQ, Guo P, Cheng B, *et al.* An experimental case study on the relationship between workload and resource consumption in a commercial web server. *Journal of Computational Science*, 2018, 25: 183–192. [doi: [10.1016/j.jocs.2017.05.019](https://doi.org/10.1016/j.jocs.2017.05.019)]
- 2 Zhao J, Trivedi KS, Grottke M, *et al.* Ensuring the performance of apache HTTP server affected by aging. *IEEE Transactions on Dependable and Secure Computing*, 2014,

- 11(2): 130–141. [doi: [10.1109/TDSC.2013.38](https://doi.org/10.1109/TDSC.2013.38)]
- 3 Cotroneo D, Natella R, Pietrantuono R, *et al.* Software aging analysis of the linux operating system. *Proceedings of the IEEE 21st International Symposium on Software Reliability Engineering*. San Jose, CA, USA. 2010. 71–80.
- 4 Bovenzi A, Cotroneo D, Pietrantuono R, *et al.* On the aging effects due to concurrency bugs: A case study on MySQL. *Proceedings of the IEEE 23rd International Symposium on Software Reliability Engineering*. Dallas, TX, USA. 2012. 211–220.
- 5 Araujo J, Matos R, Maciel P, *et al.* Experimental evaluation of software aging effects on the eucalyptus cloud computing infrastructure. *Proceedings of the Middleware 2011 Industry Track Workshop*. Lisbon, Portugal. 2011. 4.
- 6 Machida F, Kim DS, Trivedi KS. Modeling and analysis of software rejuvenation in a server virtualized system. *Proceedings of 2010 IEEE Second International Workshop on Software Aging and Rejuvenation*. San Jose, CA, USA. 2011. 1–6.
- 7 Huang DM, He SQ, He XH, *et al.* Prediction of wind loads on high-rise building using a BP neural network combined with POD. *Journal of Wind Engineering and Industrial Aerodynamics*, 2017, 170: 1–17. [doi: [10.1016/j.jweia.2017.07.021](https://doi.org/10.1016/j.jweia.2017.07.021)]
- 8 闫春, 厉美璇, 周潇. 基于改进的遗传算法在函数优化中的应用. *计算机应用研究*, 2019, 36(10): 1–6.
- 9 刘飞. Java虚拟机内存管理与内存泄漏. *信息与电脑*, 2017, (6): 69–71. [doi: [10.3969/j.issn.1003-9767.2017.06.025](https://doi.org/10.3969/j.issn.1003-9767.2017.06.025)]
- 10 Matias R, Barbetta PA, Trivedi KS, *et al.* Accelerated degradation tests applied to software aging experiments. *IEEE Transactions on Reliability*, 2010, 59(1): 102–114. [doi: [10.1109/TR.2009.2034292](https://doi.org/10.1109/TR.2009.2034292)]