

光学学报

基于神经网络和多目标优化算法的
掺铋光纤放大器设计侯文强^{1,2}, 裴丽^{1,2*}, 王建帅^{1,2}, 郑晶晶^{1,2}, 徐文轩^{1,2}, 田梓辰^{1,2}, 王丁辰^{1,2}, 王丽红^{1,2}, 李晶^{1,2}, 宁提纲^{1,2}¹北京交通大学电子信息工程学院, 北京 100044;²北京交通大学光波技术研究所全光网络与现代通信网教育部重点实验室, 北京 100044

摘要 掺铋光纤放大器有助于将光纤通信系统拓展至新的传输波段。然而,其增益和噪声性能存在相互制约的关系,提升增益往往会导致噪声性能的恶化,反之亦然。因此,提出一种结合反向传播神经网络(BPNN)和带精英保留策略的快速非支配排序遗传算法(NSGA-II)的多目标优化方法,通过对两级掺铋光纤放大器结构进行设计,实现了增益和噪声性能的同时优化。使用经过训练的BPNN对增益和噪声系数预测的均方根误差分别为0.191和0.084,具有较高预测精度。以高增益和低噪声系数为目标,使用NSGA-II算法进行优化,得到包含500个解的Pareto最优解集。优化后,放大器所能实现的平均增益范围为15~37 dB,相应的平均噪声系数范围为4.95~5.31 dB。利用BPNN代替求解耦合微分方程来评价个体适应度,使得优化时间较传统方法由10⁶s左右降低为80s左右,大幅提升了优化效率。所提方法也为其他掺铋光纤放大器的高效、多目标结构优化设计提供了一种新的思路。

关键词 光纤通信; 掺铋光纤放大器; 反向传播神经网络; 多目标优化; 带精英保留策略的快速非支配排序遗传算法

中图分类号 O436

文献标志码 A

DOI: 10.3788/AOS231173

1 引言

多波段传输被认为是解决光纤通信系统容量日益紧缺的一种有效方法^[1]。G652.D单模光纤在除C、L波段外的O、E、S等波段的传输损耗均低于0.4 dB/km。然而,由于缺少成熟的光放大器,至今尚未在这些波段大规模部署密集波分复用技术进行远距离传输^[2-3]。近年来,随着掺铋光纤放大器(BDFA)的提出和发展^[4-7],研究人员发现不同组分的掺铋石英光纤在近红外区域具有广泛的荧光特性^[8-10],这一发现为解决上述难题带来了希望。就光纤放大器而言,通常以高增益、低噪声系数为目标,对结构参数进行优化设计。传统情况下,需要利用Runge-Kutta算法结合打靶法或松弛法求解耦合微分方程组以分析放大器的性能^[11-13]。若结合全局优化算法,例如遗传算法、粒子群算法等来优化光纤放大器结构,则需要进行成千上万次相关方程组的求解,过程复杂且耗时。而通过机器学习的方法建立光纤放大器结构参数和性能之间的非线性映射关系,可以代替传统求解微分方程组的过程,大幅降低计算时间,从而实现对放大器结构的高效优化设计。

2015年,Singh等^[14]使用遗传算法对EDFA拉曼

混合光放大器的结构参数进行优化。同年,陈静等^[15]使用最小二乘支持向量回归机建立多泵浦拉曼光纤放大器模型,并结合遗传算法实现在设计时间不超过19s的情况下,快速调节开关增益。2018年,Chen等^[16]使用极限学习机和差分进化算法相结合的混合优化算法对多泵浦拉曼光纤放大器进行优化设计,该算法缩短了计算时间,提高了整体优化效率。同年,Chen等^[17]还使用最小二乘支持向量机结合粒子群算法来提高拉曼光纤放大器的优化效率。2021年,巩稼民等^[18]采用神经网络建立泵浦波长和泵浦功率与拉曼净增益谱之间的映射关系,并结合人工蜂群算法来优化泵浦光参数。由此可见,上述方法主要对拉曼光纤放大器或者混合光放大器进行优化设计,少有益于对掺杂光纤放大器的结构优化设计。且上述方法大多使用单目标优化算法,每次优化只能得到一个最优解。而通常放大器的多个性能指标相互制约,不存在唯一的最优解,因此使用多目标优化算法进行优化设计是一个更好的选择。

本文采用带精英保留策略的快速非支配排序遗传算法(NSGA-II)优化设计BDFA结构。该多目标优化算法可将多个波长信号的增益和噪声系数作为目标,生成Pareto最优解集。在评价个体适应度方面,通

收稿日期: 2023-06-21; 修回日期: 2023-08-17; 录用日期: 2023-09-20; 网络首发日期: 2023-09-30

基金项目: 国家重点研发计划(2020YFB1805802)、国家自然科学基金项目(62235003,62221001)

通信作者: *lpei@bjtu.edu.cn

过建立反向传播神经网络(BPNN)模型代替求解耦合微分方程组,提升整体优化效率。最后,通过分析各个结构参数及对应放大器输出性能,得到两级分段泵浦式放大方案,为BDFAs的结构设计提供指导。

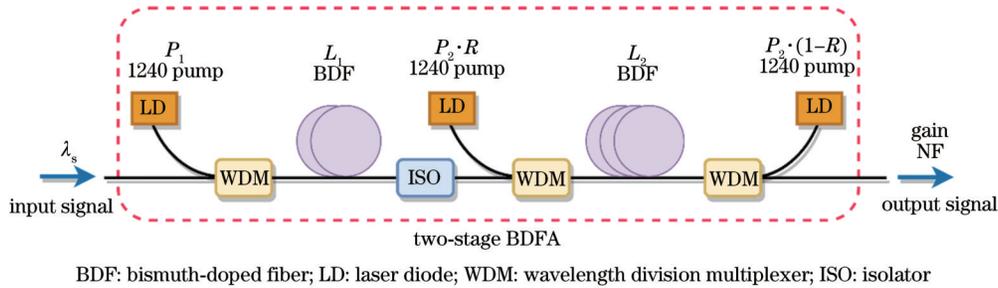
2 相关理论分析

2.1 掺铋光纤放大器

光纤放大器的主要性能参数包括:信号增益(Gain)、噪声系数(NF),以及增益带宽等。其中, Gain和NF是优化BDFAs所考虑的主要目标。为了获得这两个参数,通常需要利用Runge-Kutta算法结合打靶法或者松弛法对功率传播方程和粒子速率方程进行求解^[19]。当将光纤放大器用于线路放大时,需要同时满足高增益和低噪声系数这两个基本条件^[20]。然而,使用单级光纤进行放大很难同时满足这两个条件。为实现高增益,需要使用较长的掺杂光纤,但这会造成自发辐射噪声的累积,从而使噪声系数上升。另一方面,为实现低噪声系数,需要使用较短的光纤来保证较高的总体反转率,但这样又无法获得理想的增益。在以往的光纤放大器结构中,通常会采用多级放大的方式来改善线路放大器的性能。在多级放大

方式下,整个放大器的噪声系数很大程度上取决于第一级放大器的噪声系数。因此,若将第一级设计为低噪声放大器,以保证较好的噪声性能,将后续级设计为高效的功率放大器,以保证较好的增益性能,这样可以使得整个放大器同时具备高增益和低噪声系数^[21]。

本研究采用掺铋磷酸盐光纤作为增益介质,该光纤在1240 nm泵浦光的激发下,能够实现对O波段附近信号光的放大^[8]。掺铋光纤(BDF)中铋离子浓度较低,因此需要使用较长的光纤来实现所需的信号放大,长度通常为100~200 m。放大器采用两级分段泵浦式结构,如图1所示。通常情况下,正向泵浦具有最佳的噪声性能,双向泵浦具有最佳的增益性能^[22]。因此,本研究中第一级BDF采用正向泵浦,第二级BDF采用双向泵浦。为防止第二级的后向自发辐射光进入第一级,造成不必要的噪声累积,在两级之间加入隔离器。放大器的主要结构参数包括:第一级泵浦功率(P_1)、第二级总泵浦功率(P_2)、第一级BDF长度(L_1)、第二级BDF长度(L_2),以及第二级前后向泵浦功率比(R)。通过调整这些结构参数形成不同组合,可以实现不同的增益和噪声性能。



BDF: bismuth-doped fiber; LD: laser diode; WDM: wavelength division multiplexer; ISO: isolator

图1 两级掺铋光纤放大器结构示意图

Fig. 1 Schematic diagram of the two-stage BDFAs

对于两级BDFAs,每种结构参数配置下需要进行两次耦合微分方程组的求解操作,即以第一级光纤的末端值和第二级的正向泵浦功率作为第二级的初始值,然后进行第二次求解。整个过程相对复杂且非常耗时,平均每次计算需要50 s。在某些特殊情况下,还可能无法得到耦合微分方程组的收敛解。如果采用智能优化算法,例如遗传算法、粒子群算法等对放大器结构参数进行优化,每次迭代优化都需要进行成千上万次耦合微分方程组的求解操作,将花费更长的计算时间。

2.2 反向传播神经网络

BPNN是一种多层前馈神经网络模型^[23],具有较强的非线性映射能力和自适应能力。由于其原理和结构简单,对于小数据集模型,更容易避免过拟合问题。因此本研究利用其来映射两级BDFAs的结构参数和性能之间的非线性关系。BPNN主要包括输入层、隐藏层和输出层,如图2所示。输入层包含6个神经元,

分别对应输入信号波长和放大器的5个结构参数。输出层包含两个神经元,分别对应相应波长信号的增益和噪声系数。输入层和输出层之间是隐藏层,由于单层隐藏层神经网络训练的准确度不够高,而层数过多会导致过拟合现象,因此采用两层隐藏层。隐藏层神经元个数的选择没有准确的方法,通常根据经验公式来确定,具体如下:

$$H = \sqrt{I + O} + a, \quad (1)$$

式中: H 是隐藏层神经元数量; I 是输入层神经元数量; O 是输出层神经元数量; a 是介于1~10之间的常数。各层之间通过激活函数联系起来,激活函数的主要作用是引入非线性因素,以增强神经网络的表达能力。

BPNN的主要特点是信号的前向传递和误差的反向传播。它属于有监督学习方法,需要使用一组已知目标输出的学习样本集进行训练。训练过程中,首先使用随机值初始化网络的权值和阈值。然后,输入信号从输入层开始,经过隐藏层逐层处理并向前传递,直

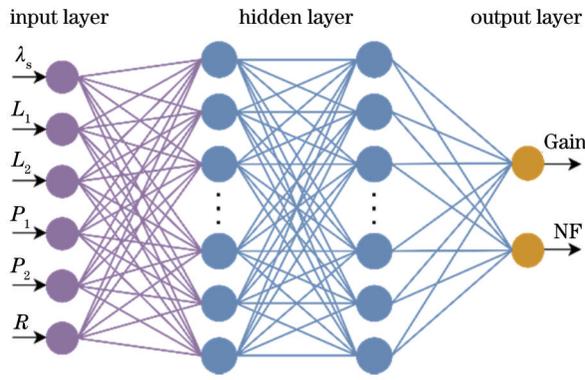


图 2 BPNN 结构图

Fig. 2 Structure diagram of BPNN

到输出层。如果输出层的预测值与期望输出不符,则进行误差的反向传播。利用梯度下降法,逐层调整神经网络的权值和阈值。这个调整的过程是反复进行的,直到网络的预测值逼近目标值且误差不再下降时,训练完成。

2.3 多目标优化算法

本研究涉及的多目标问题可由式(2)表示:

$$\text{minimize } [f_1(x), f_2(x)] \text{ s.t. } x_1 \leq x \leq x_u, \quad (2)$$

式中: $f_1(x)$ 和 $f_2(x)$ 分别与信号增益和噪声系数相关; $x = [L_1, L_2, P_1, P_2, R]^T$ 为放大器结构参数; x_1, x_u 分别为放大器结构参数取值范围的下限和上限。对于多目标问题,一种解决方法是将其转化为单目标问题进行处理。但这种方法需要预先知道权重因子,且每次计算只能获得一个最优解,即只能得到单个适应度函数的最大值或最小值。而通常情况下,多目标问题的目标值之间相互制约,不存在唯一的最优解。例如本文中, Gain 和 NF 相互制约,一者性能的提升会导致另一者的恶化,因此不存在一种结构参数配置能够使放大器同时实现最高的 Gain 和最低的 NF。相较之下,多目标优化算法能够获得一组相互独立的解,这组解中没有一个解在所有目标上都优于其他解,因此均属于最优解。工程设计人员可以根据实际需要进行选择。

NSGA-II 是在非支配排序遗传算法(NSGA)基础上改进的一种多目标优化算法^[24]。通过引入快速非支配排序方法、精英保留策略和拥挤度算子,降低了计算复杂度、提高了优化效率并能保证种群中个体的多样性。算法流程如图 3 所示。通过该算法找到的一系列最优解构成 Pareto 最优解集,也称 Pareto 前沿。非支配解及 Pareto 最优解集的定义如下:假设 x_a 和 x_b 是问题的两个可行解,若满足式(3)中的条件,则称 x_a 支配 x_b , 或称 x_b 被 x_a 支配。若解 x^* 不被其他任何可行解支配,则称此解为非支配解^[25],由所有非支配解构成的集合即为 Pareto 最优解集。

$$\forall i = 1, 2, f_i(x_a) \leq f_i(x_b) \cap \exists j = 1, 2, f_j(x_a) < f_j(x_b). \quad (3)$$

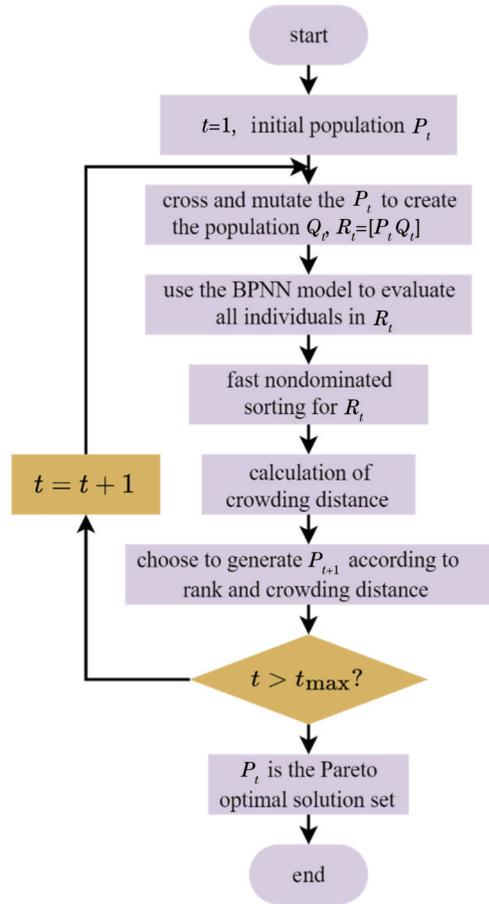


图 3 NSGA-II 算法流程图

Fig. 3 NSGA-II algorithm flow chart

NSGA-II 算法具体计算步骤如下^[26-28]:

- 1) 初始化迭代次数 $t = 1$, 设置最大迭代次数 t_{\max} 。根据需要优化的放大器结构参数,包括两级 BDF 的长度、两级泵浦功率和第二级前后向泵浦功率比,在其相应取值范围内随机生成一系列个体,构成大小为 N 的父代种群 P_t 。
- 2) 对种群 P_t 执行交叉和变异操作,生成大小为 N 的种群 Q_t 。合并种群 P_t 和 Q_t ,生成大小为 $2N$ 的种群 R_t 。
- 3) 使用经过训练的 BPNN 模型对种群 R_t 中每个个体进行评估,即获得每种参数配置下的两级 BDF 性能,包括信号增益和信号噪声系数。
- 4) 根据评估得到的适应度,对种群 R_t 中的个体执行快速非支配排序操作,确定每个个体的 Pareto 等级。从 Pareto 第一等级开始,将其个体加入下一代种群 P_{t+1} 中,直到 P_{t+1} 的大小为 N 。
- 5) 若 Pareto 第 i 等级加入后,种群 P_{t+1} 大小超过 N , 而第 $i-1$ 等级加入后, P_{t+1} 大小小于 N , 则计算第 i 等级中每个个体沿着目标两侧最近的两个相邻个体之间的平均距离。该距离称为拥挤距离,选择拥挤距离更大的个体进入下一代,直到种群 P_{t+1} 的大小为 N 。
- 6) 重复步骤 2)~5), 直到达到最大迭代次数时停

止。此时的 P 为 Pareto 最优解集, 即具有最优增益性能和噪声性能的一系列放大器结构参数配置。

3 仿真与分析

3.1 BPNN 模型的建立

利用两级 BDF 的理论模型进行仿真实验, 通过求解功率传播方程以获得样本集。放大器的结构参数包括两级 BDF 的长度、两级泵浦功率和二级前后向泵浦功率比。除结构参数外, 还将输入信号波长作为仿真的变量参数, 而输入信号功率统一设置为 -20 dBm。相应参数的取值范围见表 1。通过设置这些参数的随机组合, 可以得到不同的 Gain 和 NF 数据, 这些数据将用于 BPNN 的训练和测试。本文共采集得到 3000 组样本集, 所花费的计算时间大约为 2500 min。

表 1 放大器仿真变量参数取值范围

Table 1 The ranges of parameters in the simulation

Parameter		Value range
Wavelength of signal /nm	λ_s	1320-1360
Length of first BDF /m	L_1	0-50
Length of second BDF /m	L_2	100-250
Pump power of stage 1 /S	P_1	0-0.1
Pump power of stage 2 /W	P_2	0.3-0.7
Pump power ratio of stage 2	R	0-1

通过试错法, 综合评估训练时间和预测精度后, 最终确定神经网络的结构为 6-12-8-2。其中, 输入层和隐藏层、隐藏层和隐藏层之间的激活函数选用 tansig 函数, 而隐藏层和输出层之间的激活函数选用 purelin 函数。采用 Levenberg-Marquardt (L-M) 方法训练 BPNN, 该方法的收敛速度较快并且均方误差较小, 适合训练中小规模的神经网络。具体的训练参数设置如下: 学习率为 0.0001, 最大迭代次数为 1000, 训练目标的最小误差为 0.001。模型的准确度由均方根误差 (E_{RMSE}) 和决定系数 (R^2) 来衡量:

$$E_{\text{RMSE}} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}, \quad (4)$$

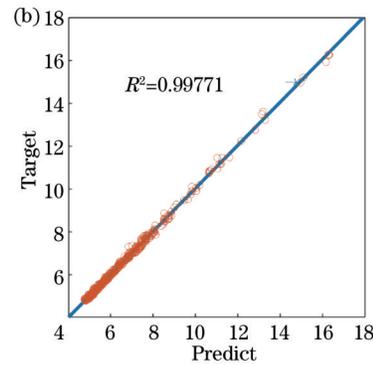
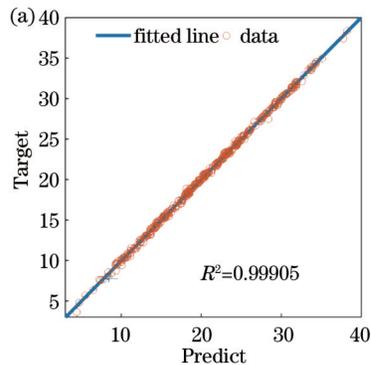


图 5 预测值与目标值的拟合情况。(a)(b)BPNN 预测的 Gain 和 NF

Fig. 5 Fitting of predicted value and target value. (a)(b) Gain and NF predicted by BPNN

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}, \quad (5)$$

式中: y_i, \hat{y}_i 分别为第 i 个测试样本的目标值和预测值; \bar{y} 为测试样本的平均目标值。在每次训练和测试之前, 需要对数据进行归一化处理, 网络输出则需要进行反归一化处理。这样的处理有助于加速网络收敛, 并避免出现数值问题。

分别在不同样本数量下, 训练 BPNN 并进行测试, 训练集和测试集大小的比值为 9:1。图 4 展示了 BPNN 模型在测试 Gain 和 NF 方面的性能。可以看到, 随着样本集的增大, E_{RMSE} 整体趋势逐渐变小, R^2 值逐渐增大, 在样本数量为 3000 的时候, BPNN 模型对 Gain 和 NF 测试的 E_{RMSE} 分别为 0.191 和 0.084, R^2 值分别为 0.999 和 0.998。这些结果表明, 随着样本数量的增加, BPNN 模型在预测 Gain 和 NF 方面的性能逐渐提高, 预测结果的准确度增加。

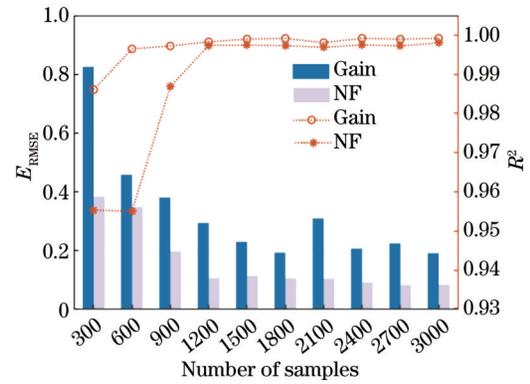


图 4 不同样本数量训练 BPNN 模型, 对 Gain 和 NF 的测试性能

Fig. 4 The test performance of Gain and NF after training BPNN model with different sample sizes

接下来, 随机选取 2700 个样本作为训练集, 300 个样本作为测试集, 对 BPNN 模型进行训练和测试。图 5 展示了预测值和目标值拟合情况, 以及相应的 R^2 值,

从图中可以看到,使用BPNN模型对Gain和NF测试的线性情况均良好, R^2 值均接近于1。通过以上分析可知,所建立的BPNN模型具有较高的预测精度,能够很好地映射两级BDF A的结构参数和性能之间的非线性关系。

3.2 多目标优化

使用NSGA-II算法进行优化设计。将放大器的结构参数作为个体特征,初始化种群。根据表1在指定范围内生成500个随机分布的个体。这里选用平均增益和平均噪声系数作为目标函数,输入信号波长范围为1320~1360 nm,信号间隔为10 nm,共5个波长信号。最终的优化目标是使得这5个波长的信号平均增益尽可能大,并且噪声系数尽可能小。基于所建立的BPNN模型对目标函数进行评估,并根据得到的适应度值进行快速非支配排序。然后执行交叉、变异和选择操作,按照图3中的流程进行迭代计算。设置最大迭代次数为100,交叉和变异概率分别设置为0.9和0.5。最终得到的平均增益和平均噪声系数的Pareto最优解集如图6所示,共包含500个解。其中,每个数据点代表一种特定的放大器结构参数配置,且没有一种结构参数配置使得放大器的Gain和NF性能都优于其他各个解,因此这些解均属于最优解。所得到的平均增益范围为15~37 dB,相应的平均噪声系数范围为4.95~5.31 dB。可以看出,增益和噪声系数之间存在一定的制约关系。增益越大,噪声系数也会越高,这与之前的分析结果一致。

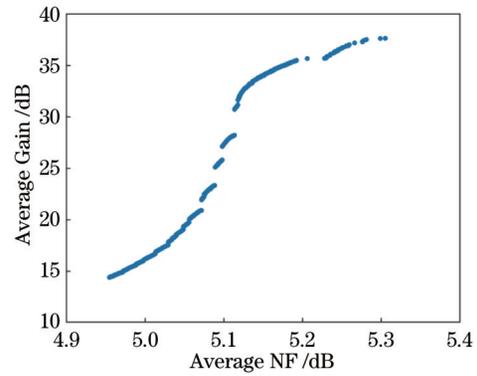


图6 Pareto最优解集

Fig. 6 Pareto optimal solution set

为了方便表示,从Pareto最优解集中选取5组具有代表性的解,如表2所示。可以看到,第一级泵浦功率和第二级总泵浦功率都接近最大值,而要实现不同的增益和噪声性能,主要取决于相应的光纤长度。增加光纤长度可以提高增益,但同时也会使噪声系数上升。由于所研究的BDF具有较小的吸收截面和发射截面,因此第一级BDF在0~50 m范围内变化,即光纤长度较短时,噪声性能的变化相对较小。第二级前后向泵浦功率比对放大器的性能有较大影响,在可实现的增益接近的情况下,第二级前向泵浦功率占比越高,相应的平均噪声系数越小。图7是这5组不同的放大器结构配置下,对不同波长的信号所能实现的增益和噪声性能情况。

表2 部分Pareto最优解集及相应的适应度值

Table 2 Partial Pareto optimal solution sets and corresponding fitness values

Group	Two-stage BDF A parameter					Average Gain /dB	Average NF /dB
	L_1 /m	L_2 /m	P_1 /W	P_2 /W	R		
1	48.82	246.43	0.10	0.69	0.28	37.67	5.31
2	18.65	249.94	0.10	0.67	0.58	35.01	5.18
3	5.33	249.93	0.10	0.69	0.83	32.71	5.13
4	6.11	203.15	0.10	0.70	0.80	28.21	5.11
5	6.10	185.59	0.10	0.69	0.93	25.81	5.10

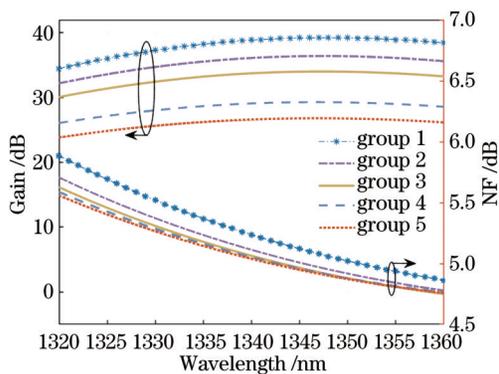


图7 不同结构配置下的Gain和NF

Fig. 7 Gain and NF in different configurations

3.3 性能比较

为进行比较,同时建立支持向量机(SVM)模型用于映射两级BDF A的结构参数和性能之间的非线性关系。选用径向基函数作为核函数,惩罚因子和gamma参数分别设置为20和0.9。在样本数量同样为3000、训练集和测试集按9:1的比例划分下,对SVM模型进行训练并测试。图8展示了分别使用训练后的BPNN和SVM两种模型进行测试时误差的概率密度函数(PDF)曲线,可以看到,使用BPNN模型对Gain和NF预测的误差更小,预测准确度更高。

最后,记录使用松弛法求解一次两级BDF A耦合微分方程组花费的时间,若与NSGA-II算法结合进行优化设计,预测将花费 10^6 s。同时分别记录使用SVM

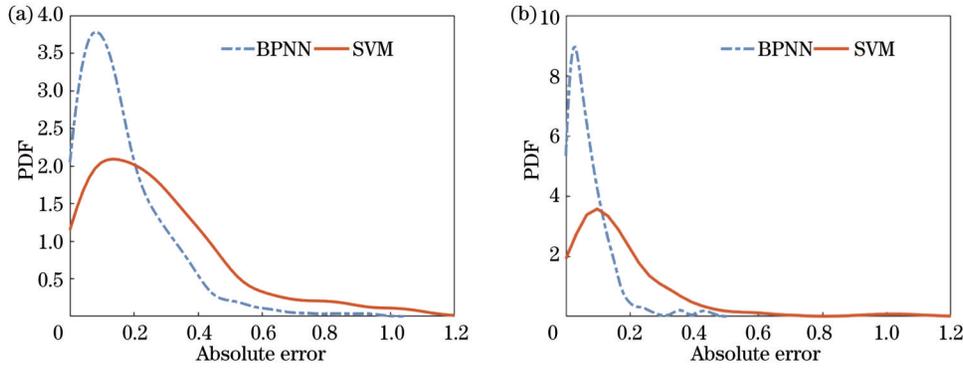


图 8 使用 BPNN 模型和 SVM 模型对 Gain 和 NF 测试的绝对误差 PDF 曲线。(a)Gain; (b)NF

Fig. 8 PDF curves of testing absolute errors for Gain and NF using BPNN and SVM models. (a) Gain; (b) NF

和 BPNN 与 NSGA-II 算法结合进行优化所需花费的时间,如图 9 所示。在每种方法中,保持种群大小、迭代次数等参数设置相同。可以看到,与使用松弛法结合 NSGA-II 相比,使用 BPNN-NSGA-II 进行优化设计所需的时间要低 5 个数量级,仅仅需要不超过 80 s 即可完成设计,相比使用 SVM-NSGA-II,时间要低一个数量级。需要注意的是,在采集样本时会花费一些时间,但一旦通过采集到的数据集训练好 BPNN 模型后,将不再需要额外的数据。因此,相比通过直接求解耦合微分方程组来评估个体适应度,使用 BPNN 模型能够更节省时间,优化设计的效率会更高。与以往的单目标优化方法相比,使用 BPNN-NSGA-II 所得到的 Pareto 最优解集中的各个解均属于最优解。能够为决策者提供更多选择,可以根据对放大器性能的不同需求来确定相应的放大器结构参数,增强了设计的灵活性。

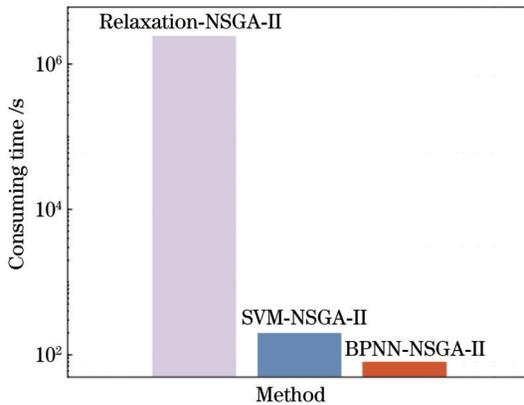


图 9 采用不同方法进行优化设计花费的时间

Fig. 9 Time consuming to optimize design with different methods

4 结 论

提出一种结合神经网络和 NSGA-II 算法的多目标优化方法,用于准确建模和高效优化设计两级掺铋光纤放大器。通过建立 BPNN 模型映射结构参数和

性能之间的非线性关系,避免了耦合微分方程组的大量重复求解过程。经过训练和测试,BPNN 模型表现出较低的 E_{RMSE} 和较高的 R^2 值。利用该 BPNN 模型结合 NSGA-II 算法,得到了一个包含 500 个解的 Pareto 最优解集,并给出 5 组不同放大器结构参数配置下的信号增益和噪声系数谱。与其他方法的比较结果表明,所提方法显著降低了优化设计时间,提升了优化效率,并能同时得到多个最优解,为决策者提供更多选择。综上所述,所提方法对掺铋光纤放大器以及其他掺杂光纤放大器的高效率、多目标结构优化设计具有重要意义。

参 考 文 献

- [1] Deng N, Zong L J, Jiang H Y, et al. Challenges and enabling technologies for multi-band WDM optical networks[J]. Journal of Lightwave Technology, 2022, 40(11): 3385-3394.
- [2] Renaudier J, Napoli A, Ionescu M, et al. Devices and fibers for ultrawideband optical communications[J]. Proceedings of the IEEE, 2022, 110(11): 1742-1759.
- [3] Rapp L, Eiselt M. Optical amplifiers for multi-band optical transmission systems[J]. Journal of Lightwave Technology, 2022, 40(6): 1579-1589.
- [4] Khagai A, Ososkov Y, Firstov S, et al. Gain clamped Bi-doped fiber amplifier with 150 nm bandwidth for O- and E-bands[J]. Journal of Lightwave Technology, 2022, 40(4): 1161-1166.
- [5] Dianov E M. Amplification in extended transmission bands using bismuth-doped optical fibers[J]. Journal of Lightwave Technology, 2013, 31(4): 681-688.
- [6] Bufetov I A, Melkumov M A, Firstov S V, et al. Bi-doped optical fibers and fiber lasers[J]. IEEE Journal of Selected Topics in Quantum Electronics, 2014, 20(5): 111-125.
- [7] Khagai A, Ososkov Y, Firstov S, et al. O+E band BDFAs with flat top 116 nm gain bandwidth pumped with 250 mW at 1256 nm[C]//Optical Fiber Communication Conference (OFC) 2021, June 6-10, 2021, Washington, DC. Washington, D.C.: Optica Publishing Group, 2021: Tu1E.4.
- [8] Firstov S V, Khagai A M, Kharakhordin A V, et al. Compact and efficient O-band bismuth-doped phosphosilicate fiber amplifier for fiber-optic communications[J]. Scientific Reports, 2020, 10: 11347.
- [9] Firstov S V, Khopin V F, Bufetov I A, et al. Combined excitation-emission spectroscopy of bismuth active centers in optical fibers[J]. Optics Express, 2011, 19(20): 19551-19561.
- [10] Firstova E G, Bufetov I A, Khopin V F, et al. Luminescence

- properties of IR-emitting bismuth centres in SiO₂-based glasses in the UV to near-IR spectral region[J]. *Quantum Electronics*, 2015, 45(1): 59-65.
- [11] Shang L, Qi L M, Liao Y L, et al. A combined algorithm for simulating fiber lasers based on the shooting and relaxation methods[J]. *Optical Fiber Technology*, 2012, 18(6): 502-508.
- [12] Lali-Dastjerdi Z, Kroushavi F, Rahmani M H. An efficient shooting method for fiber amplifiers and lasers[J]. *Optics & Laser Technology*, 2008, 40(8): 1041-1046.
- [13] Liu X M. Fast methods for solving nonlinearly coupled equations describing fiber amplifiers[J]. *Applied Mathematics and Computation*, 2006, 172(2): 1018-1035.
- [14] Singh S, Kaler R S. Performance optimization of EDFA-Raman hybrid optical amplifier using genetic algorithm[J]. *Optics & Laser Technology*, 2015, 68: 89-95.
- [15] 陈静, 江灏, 刘瞰东, 等. 基于最小二乘支持向量回归模型的拉曼光纤放大器优化设计[J]. *光学学报*, 2015, 35(11): 1123004.
Chen J, Jiang H, Liu T, et al. Optimization for Raman fiber amplifiers based on least squares support vector regression model[J]. *Acta Optica Sinica*, 2015, 35(11): 1123004.
- [16] Chen J, Jiang H. Optimal design of gain-flattened Raman fiber amplifiers using a hybrid approach combining randomized neural networks and differential evolution algorithm[J]. *IEEE Photonics Journal*, 2018, 10(2): 7101915.
- [17] Chen J, Qiu X J, Yin C Y, et al. Efficient design of gain-flattened multi-pump Raman fiber amplifiers using least squares support vector regression[J]. *Journal of Optics*, 2018, 20(2): 025702.
- [18] 巩稼民, 刘芳, 吴艺杰, 等. 基于神经网络和人工蜂群算法的拉曼光纤放大器设计方案[J]. *光学学报*, 2021, 41(20): 2006002.
Gong J M, Liu F, Wu Y J, et al. Design of Raman fiber amplifier based on neural network and artificial bee colony algorithm[J]. *Acta Optica Sinica*, 2021, 41(20): 2006002.
- [19] Giles C R, Desurvire E. Modeling erbium-doped fiber amplifiers[J]. *Journal of Lightwave Technology*, 1991, 9(2): 271-283.
- [20] Delavaux J M P, Nagel J A. Multi-stage erbium-doped fiber amplifier designs[J]. *Journal of Lightwave Technology*, 1995, 13(5): 703-720.
- [21] Mishra A R, Kakade R N, Kakade P D. Dual-stage EDFA for improving the performance of long-haul optical systems[J]. *IEEE Access*, 2022, 10: 13496-13514.
- [22] Wang D C, Pei L, Zheng J J, et al. Analysis of gain and noise characteristics of O-band Bi-doped fiber amplifier under different pumping schemes[J]. *Optik*, 2022, 251: 168491.
- [23] Hannan M A, Lipu M S H, Hussain A, et al. Neural network approach for estimating state of charge of lithium-ion battery using backtracking search algorithm[J]. *IEEE Access*, 2018, 6: 10069-10079.
- [24] Deb K, Pratap A, Agarwal S, et al. A fast and elitist multiobjective genetic algorithm: NSGA-II[J]. *IEEE Transactions on Evolutionary Computation*, 2002, 6(2): 182-197.
- [25] Zhang Y P, Liu D, Shen X E, et al. Design of iodine absorption cell for high-spectral-resolution lidar[J]. *Optics Express*, 2017, 25(14): 15913-15926.
- [26] Verma S, Pant M, Snasel V. A comprehensive review on NSGA-II for multi-objective combinatorial optimization problems[J]. *IEEE Access*, 2021, 9: 57757-57791.
- [27] 江灏, 林朝晖, 陈静, 等. 考虑非对称重叠光谱的FBG传感网络带宽资源多目标优化配置[J]. *光学学报*, 2021, 41(10): 1006002.
Jiang H, Lin C H, Chen J, et al. Multi-objective optimal configuration of bandwidth resources in FBG sensor network for asymmetric overlapping spectra[J]. *Acta Optica Sinica*, 2021, 41(10): 1006002.
- [28] 王燕燕, 李家豪, 舒林森, 等. 基于RSM与NSGA-II算法的激光熔覆参数多目标优化[J]. *激光与光电子学进展*, 2022, 59(7): 0714004.
Wang Y Y, Li J H, Shu L S, et al. Multi-objective optimization of laser cladding parameters based on RSM and NSGA-II algorithm[J]. *Laser & Optoelectronics Progress*, 2022, 59(7): 0714004.

Design of Bismuth-Doped Fiber Amplifier Based on Neural Network and Multi-Objective Optimization Algorithm

Hou Wenqiang^{1,2}, Pei Li^{1,2*}, Wang Jianshuai^{1,2}, Zheng Jingjing^{1,2}, Xu Wenxuan^{1,2},
Tian Zichen^{1,2}, Wang Dingchen^{1,2}, Wang Lihong^{1,2}, Li Jing^{1,2}, Ning Tigang^{1,2}

¹School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China;

²Key Laboratory of All Optical Network & Advanced Telecommunication Network of EMC, Institute of Lightwave Technology, Beijing Jiaotong University, Beijing 100044, China

Abstract

Objective Multi-band transmission is considered an effective solution to address the increasing capacity constraints in fiber optic communication systems. However, due to the lack of mature optical amplifiers, the large-scale deployment of dense wavelength division multiplexing (DWDM) technology for long-distance transmission in bands such as O, E, and S has not yet been achieved. In recent years, researchers have discovered that different dopants in bismuth-doped silica fibers exhibit broad fluorescence characteristics in the near-infrared region. This finding brings hope for addressing the aforementioned challenges. In traditional approaches, the performance analysis of amplifiers often requires solving a set of coupled differential equations using methods such as the Runge-Kutta algorithm combined with the Shooting method or Relaxation method. When incorporating global optimization algorithms, it becomes necessary to solve thousands of related

equations, resulting in a complex and time-consuming process. Previous research methods have mainly focused on the optimization design of Raman fiber amplifiers or hybrid optical amplifiers, with fewer studies specifically targeting the structural optimization design of doped fiber amplifiers, particularly bismuth-doped fiber amplifier (BDFA). Moreover, most of these studies have employed single-objective optimization algorithms, resulting in obtaining only one optimal solution at a time. In general, there is a trade-off relationship between the gain and noise performance of amplifiers. Increasing the gain often leads to the deterioration of the noise performance, and vice versa. As a result, there is no unique optimal solution. Therefore, it is necessary to design a method that can accurately model the amplifier and efficiently optimize multiple performance metrics simultaneously.

Methods The backpropagation neural network (BPNN) is a type of multilayer feedforward neural network consisting of input layer, hidden layers, and output layer. The input layer contains six neurons corresponding to the input signal wavelength and five structural parameters of the amplifier. The output layer contains two neurons corresponding to the Gain and noise figure (NF) of the respective wavelength signals. The main characteristic of BPNN is the forward propagation of signals and the backward propagation of errors. It belongs to the supervised learning methods. For multi-objective problems, the objective values are typically mutually constrained, and there is no unique optimal solution. Using multi-objective optimization algorithms can provide a set of independent optimal solutions, allowing engineering designers to choose based on their actual needs. NSGA-II is a multi-objective optimization algorithm that improves upon the non-dominated sorting genetic algorithm (NSGA). By introducing fast non-dominated sorting, elite preservation strategy, and crowding distance operator, NSGA-II reduces computational complexity, improves optimization efficiency, and ensures the diversity of individuals in the population.

Results and Discussions Simulation experiments were conducted using a theoretical model of a two-stage BDFA to obtain a sample set. The BPNN model was trained and tested with different sample sizes, with a training-to-testing set ratio of 9 : 1. It was observed that as the sample set size increased, the overall trend of RMSE decreased while the R^2 value increased (Fig. 4). When the sample size reached 3000, the BPNN model achieved an RMSE of 0.191 for Gain and 0.084 for NF in the testing phase, with R^2 values of 0.999 and 0.998, respectively. The established BPNN model exhibits high prediction accuracy and can effectively capture the nonlinear relationship between the structural parameters and performance of the two-stage BDFA. Based on the established BPNN model, the objective function is evaluated, and after 100 iterations, a Pareto optimal solution set containing 500 solutions is obtained (Fig. 6). Furthermore, a comparison is made between the performance of using SVM and BPNN for predicting Gain and NF. The results show that the BPNN model has smaller prediction errors and higher accuracy in predicting Gain and NF. Additionally, the time required for optimization design using BPNN-NSGA-II is five orders of magnitude lower than using Relaxation method combined with NSGA-II, taking less than 80 seconds to complete the design. Compared to SVM-NSGA-II, the time is reduced by one order of magnitude (Fig. 9).

Conclusions This paper proposes a multi-objective optimization method that combines BPNN and NSGA-II algorithms for accurate modeling and efficient design optimization of two-stage BDFA. By establishing a BPNN model to map the nonlinear relationship between structural parameters and performance, it avoids the need for repetitive solving of coupled differential equations. After training and testing, the BPNN model exhibits low RMSE and high R^2 values. Using this BPNN model in conjunction with the NSGA-II algorithm, a Pareto optimal solution set containing 500 solutions is obtained. The paper also provides the Gain and NF spectra for five different amplifier configurations. Compared to other methods, the proposed approach significantly reduces the optimization design time, improves optimization efficiency, and enables the simultaneous attainment of multiple optimal solutions, providing decision-makers with more choices.

Key words fiber optic communication; bismuth-doped fiber amplifier; backpropagation neural network; multi-objective optimization; fast and elitist nondominated sorting genetic algorithm