



基于 KPCA-GA-BP 神经网络的 POI 质量预测研究

刘璐¹, 杨丹², 陈睿杰², 李嘉², 周熹²

(1. 中国移动通信集团设计院有限公司重庆分公司, 重庆 401121;

2. 中国移动通信集团云南有限公司, 云南 昆明 650228)

摘要: 目前移动网络优化一般基于小区进行网络质量评估及预测, 遵循“升维研究, 降维实施”的研究思路, 提出了兴趣点 (point of interest, POI) 网络质量的柔性评价体系, 但其涉及较多网络关键绩效指标 (key performance indicator, KPI), 导致 POI 网络综合质量评价体系较为庞杂且预测精度不高, 为提高 POI 网络质量预测精准性, 采用核主成分分析 (kernel principal component analysis, KPCA) 算法对反向传播 (back propagation, BP) 神经网络的输入变量进行相关性压缩, 简化了 BP 神经网络结构, 然后通过遗传算法 (genetic algorithm, GA) 优化了 BP 神经网络连接权值及阈值参数。与传统 BP 神经网络预测结果进行对比, 在预测准确度方面提高了 10.90%, 均方误差性能显著降低, 对研究 POI 网络质量的预测可起到较好的支撑作用。

关键词: POI 柔性评价体系; 核主成分分析; 遗传算法; BP 神经网络; POI 质量预测

中图分类号: TN92

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2023004

Research on POI quality prediction based on KPCA-GA-BP neural network

LIU Lu¹, YANG Dan², CHEN Ruijie², LI Jia², ZHOU Xi²

1. Chongqing Branch of China Mobile Communications Group Design Institute Co., Ltd., Chongqing 401121, China

2. Yunnan Branch of China Mobile Communications Group Co., Ltd., Kunming 650228, China

Abstract: At present, in network optimization, network quality evaluation and prediction are generally based on communities, and a flexible evaluation system for POI network quality was proposed following the research idea of “research on dimensionality increase and implementation of dimensionality reduction”. However, it involves many network KPI, resulting in a relatively complex evaluation system for POI network comprehensive quality and low prediction accuracy. In order to improve the prediction accuracy of POI network quality, KPCA was used to compress the correlation of input variables of BP neural network, the structure of BP neural network was simplified, and then the connection weights and threshold parameters of BP neural network were optimized through GA. Compared with the prediction results of traditional BP neural network, the prediction accuracy is improved by 10.9%, and the mean square error performance is significantly reduced, it can play a better supporting role in the prediction of POI network quality.

Key words: POI flexibility evaluation system, kernel principal component analysis, genetic algorithm, BP neural network, POI quality prediction

0 引言

近年来,随着移动无线网络的大力建设,用户规模及业务量急剧增长,无线网络中覆盖、干扰、容量等质量问题日益凸显,传统面向小区级的优化策略过于片面。遵循“升维研究、降维实施”的研究思路,向上构建兴趣点(point of interest, POI)网络质量评价体系,基于该评价体系动态识别 POI 网络质量情况,并结合核主成分分析(kernel principal component analysis, KPCA)算法对评价体系中的多项关键绩效指标(key performance indicator, KPI)进行降维压缩处理,将此结果作为反向传播(back propagation, BP)神经网络的输入,并采用遗传算法(genetic algorithm, GA)对 BP 神经网络参数进行优化调整,最后对 POI 网络质量进行综合、全面的预测评估。

1 构建 POI 网络质量评价体系

1.1 POI 柔性识别主服小区

目前大数据及地理信息系统(geographic information system, GIS)快速发展,基于各场景功能具备的社会经济作用,POI^[1]被划分为最小无线网络的评价单元,POI 网络质量评价体系以最小化路测(minimization of drive tests, MDT)技术数据^[2]为主,结合测量报告(measure report, MR)数据辅助补充,动态识别 POI 内所辖主服小区,相较于传统簇群,人为静态锁定服务小区方式更加灵活,可随用户业务分布进行智能识别并归属 POI,形成柔性 POI 的“呼吸效应”。

基于 MDT 数据的 POI 动态划分小区示意图如图1所示,小区 MDT 采样点携带了详细的经度、纬度信息,且与 POI 场景是多对多关系,即一个小区的 MDT 采样点分布在多个 POI 场景中,而一个 POI 场景则纳入多个小区的 MDT 采样点,动态识别 POI 内所辖服务小区,而非人为定义其下挂小区数量,通过计算可获取 POI 场景内各小区采样点贡献度。

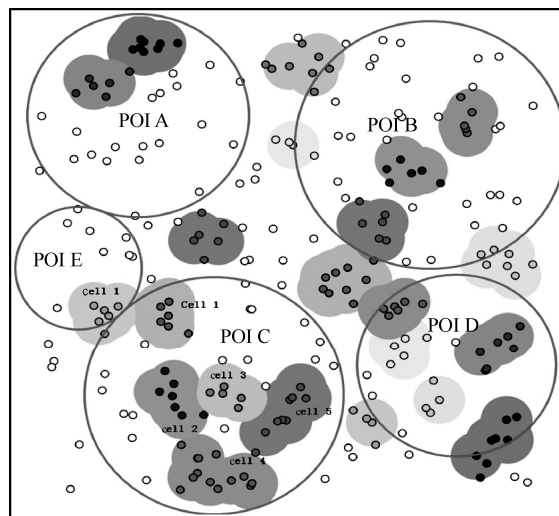


图1 基于 MDT 数据的 POI 动态划分小区示意图

(1) 以 POI C 与 POI E 为例,在 t 时刻,小区 1 (cell 1) 仅覆盖 POI C 及 POI E 两个场景,即:

cell 1 总采样点数 (N) = cell 1 在 POI C 中的采样点数 (n_1) + cell 1 在 POI E 中的采样点数 (n_2), 而 cell 1 在 POI C 中的贡献度 $p_1 = n_1 / N$, cell 1 在 POI E 中的贡献度 $p_2 = n_2 / N$, 若 $p_1 > p_2$, 则 cell 1 属于 POI C 所辖主服小区。

(2) 在 $t+t_0$ 时刻,如果用户在 POI C 内未占用 cell 1 的信号,则有:

cell 1 总采样点数 (N') = cell 1 在 POI E 中的采样点数 (n'_2), 而 cell 1 在 POI C 中的贡献度 = 0, 就可达到动态识别所辖主服小区的目的。

1.2 POI 网络质量评估

传统基于小区粒度的网络质量评估^[3],对于整体网络评估结果往往过于片面,优化分析方案可能会陷入局部最优问题,无法形成综合、全面的评估结果。为解决该问题,本文以用户业务分布为导向,进行 POI 场景级的网络质量评估分析,在网络分析层面及优化执行层面寻求平衡,达到“网随业动”的目的。

根据德尔菲(Delphi)专家调查方法^[4]选取影响无线网络的重要指标,17 项 KPI 样本见表 1。



表 1 17 项 KPI 样本

KPI	样本 ID									
	1	2	3	4	5	6	7	8	9	10
MR 覆盖率	95.95%	99.48%	99.74%	99.37%	99.46%	99.17%	99.50%	99.14%	96.40%	97.84%
无线接通率	99.64%	99.91%	100.00%	99.86%	99.94%	99.84%	99.96%	99.89%	99.98%	99.96%
掉线率	0.04%	0.03%	0.02%	0.05%	0.06%	0.09%	0.02%	0.04%	0.01%	0.03%
RRC 重建比例	4.88%	1.38%	1.60%	2.09%	4.18%	1.95%	2.10%	1.45%	4.07%	4.30%
切换成功率	99.94%	99.94%	99.95%	99.92%	99.85%	99.90%	99.92%	99.93%	99.83%	99.94%
上行 PRB 平均利用率_自忙时	9.10%	7.59%	17.25%	16.34%	17.23%	19.56%	26.97%	52.50%	34.14%	1.13%
下行 PRB 平均利用率_自忙时	7.63%	10.41%	6.81%	13.11%	9.67%	38.08%	16.97%	38.71%	25.77%	6.04%
重叠覆盖度	1.39%	1.22%	0.90%	0.26%	0.75%	2.68%	1.17%	2.57%	0.13%	2.17%
eSRVCC 切换成功率	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
VoLTE 接通率	99.76%	99.96%	100.00%	99.95%	99.96%	99.84%	99.96%	99.90%	99.99%	99.98%
VoLTE 掉话率	0.23%	0	0	0	0.15%	0	0	0	0.45%	0.44%
VoLTE 切换成功率	99.74%	100.00%	100.00%	100.00%	99.84%	100.00%	100.00%	100.00%	99.46%	99.52%
eNodeB 接收干扰功率 > -110 dBm 比例	19.89%	32.10%	0.59%	3.12%	0.69%	35.52%	4.32%	10.33%	0.28%	2.38%
VoLTE 上行平均丢包率	12.32%	10.22%	10.28%	10.36%	11.40%	13.91%	10.31%	12.40%	11.07%	9.57%
VoLTE 下行平均丢包率	11.25%	11.35%	10.48%	11.69%	10.74%	20.08%	14.60%	13.03%	13.24%	9.88%
下行丢包率	0.02%	0.10%	0.01%	0.03%	0.07%	0.16%	0.06%	0.05%	0.08%	0.01%
上行丢包率	0.01%	0.02%	0.01%	0.01%	0.09%	0.18%	0.02%	0.06%	0.06%	0.02%

注：无线电资源控制（radio resource control, RRC），物理资源块（physical resource block, PRB），增强的单一无线语音呼叫连续性（enhanced single radio voice call continuity, eSRVCC），长期演进语音承载（voice over long-term evolution, VoLTE），演进型 Node B（evolved Node B, eNodeB，特指 LTE 基站）

将小区质量指标升维至 POI 网络质量指标，需对表 1 中的各项 KPI 进行原始计数器（counter）的规则拆解，再通过第 1.1 节中 POI 动态识别小区的方法获取小区明细，构建 POI 网络质量评价模型，并依据正态分布原则划分高、中、低质量优先级的标识，最终输出统一标准化的数据源。

（1）小区 KPI 拆解及 POI 级指标汇聚

以无线接通率为例：小区无线接通率=(RRC 连接建立成功次数/RRC 连接建立请求次数)×(E-RAB 连接建立成功次数/E-RAB 连接建立请求次数)，其中 E-RAB 为演进的无线接入承载。

POI 无线接通率为 POI 内所辖所有服务小区 i 的 KPI counter 分别进行规则计算，即：

$$\text{POI 无线接通率} = \left(\sum_{\text{cell}=1}^i \text{RRC 连接建立成功次数} / \sum_{\text{cell}=1}^i \text{RRC 连接建立请求次数} \right) \times \left(\sum_{\text{cell}=1}^i \text{E-RAB 连接建立成功次数} / \sum_{\text{cell}=1}^i \text{E-RAB 连接建立请求次数} \right)$$

以上述方法进行各项指标最底层 counter 的拆解、合并后得到各项初始 POI 网络质量数据源，并对该数据进行标准化处理。

（2）数据标准化处理

利用式（1）对数据进行标准化处理，消除量纲差异。

$$\tilde{x}_{ij} = \frac{x_{ij} - \bar{X}_j}{\sigma_j} \quad (1)$$

对数据矩阵 $\mathbf{x}_{ij}$ 进行数据统一标准化处理, 其中 $\bar{X}_j$ 和 σ_j 分别是第 j 个观测指标的均值及标准差, $j=1,2,3,\dots,k$ 表示指标数量, $i=1,2,3,\dots,n$ 表示样本数量。

(3) POI 网络质量评价

对已标准化处理的矩阵 $\tilde{\mathbf{x}}_{ij}$ 进行线性评价, 得到每个 POI 网络质量样本的评价分数 Score。

$$\text{Score} = \sum_{j=1}^k \frac{\tilde{x}_{ij} - \tilde{x}_{j_{\min}}}{\tilde{x}_{j_{\max}} - \tilde{x}_{j_{\min}}} \times 100 / k \quad (2)$$

其中, $k=17$ 为 KPI 数量, $\tilde{x}_{j_{\min}}$ 为第 j 列指标标准化后的最小值, $\tilde{x}_{j_{\max}}$ 为第 j 列指标标准化后的最大值, 式 (2) 计算结果表示第 i 个 POI 网络质量样本的评价分数。

(4) POI 网络质量标识处理

按照标准正态分布原则对 POI 网络质量样本的评价分数进行标识划分处理, 则 POI 网络质量等级集合定义为 $M=\{\text{高质量, 中等质量, 低质量}\}$ 3 个等级, 映射类别向量 $M'=(1,2,3)$ 。

2 基于 KPCA 算法和 GA-BP 的预测模型

2.1 KPCA 算法

KPCA 的核心思想是对主成分分析 (principal component analysis, PCA) 算法的扩展^[5-6], 通过非线性核函数将原始数据从低维空间向高维特征空间映射, 再通过计算映射后数据的协方差矩阵, 得到相关特征值及特征向量, KPCA 算法可解决 PCA 算法不能处理非线性相关的问题, 降维输出结果可作为 BP 神经网络输入数据源, 再做进一步数据分析处理。具体实现步骤如下。

(1) 输入 POI 网络质量样本 $\tilde{\mathbf{x}}_{ij}$, 其中 $i=1,2,3,\dots,n$ 表示样本数量, $j=1,2,3,\dots,k$ 表示指标数量, 为方便演算, 令 $\mathbf{x}_i$ 为 1 个包含 k 维指标的 POI 网络质量样本向量, $\mathbf{x}_i \in \mathbf{R}^n$ 。

(2) 引入非线性映射函数 ϕ , 如果存在 $\mathbf{R}^n$ 到希尔伯特 (Hilbert) 空间的变换 $\phi: \mathbf{x} \rightarrow \phi(\mathbf{x})$,

使得:

$$\mathbf{K}(\mathbf{x}_1, \mathbf{x}_2) = (\phi(\mathbf{x}_1) \cdot \phi(\mathbf{x}_2)) \quad (3)$$

则该 $\mathbf{K}(\mathbf{x}_1, \mathbf{x}_2)$ 是定义在 $\mathbf{R}^n \times \mathbf{R}^n$ 上的核函数, 其中 $(\cdot)$ 表示内积。

本文采用高斯 (Gaussian) 核函数进行映射求解:

$$\mathbf{K}(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\sigma^2}\right) \quad (4)$$

式 (4) 定义为空间中任意一点 $\mathbf{x}$ 到某一中心点 $\mathbf{x}'$ 之间欧氏距离的单调函数, 随着两向量的距离增大, 高斯核函数单调递减, σ 控制高斯核函数的作用范围。

(3) 求解核矩阵:

$$\mathbf{K} = (\phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j)) = \frac{1}{n} \sum_{i=1}^n \phi(\mathbf{x}_i) \phi(\mathbf{x}_i)^T \quad (5)$$

其中, $\phi(\mathbf{x}_i)$ 为 POI 网络质量样本 $\mathbf{x}_i$ 在高维特征空间中的映射, $i \in [1, n]$ 。

(4) 求解中心化后的核矩阵:

$$\bar{\mathbf{K}} = \mathbf{K} - \mathbf{I}_n \mathbf{K} - \mathbf{K} \mathbf{I}_n - \mathbf{I}_n \mathbf{K} \mathbf{I}_n \quad (6)$$

其中, $\mathbf{I}_n = \frac{1}{n} \mathbf{I}$, $\mathbf{I}$ 为 $n \times n$ 的单位矩阵。

(5) 求解核矩阵的特征值及特征向量:

$$\lambda \mathbf{V} = \bar{\mathbf{K}} \mathbf{V} \quad (7)$$

因为特征向量 $\mathbf{V}$ 位于 $\phi(\mathbf{x}_1), \phi(\mathbf{x}_2), \dots, \phi(\mathbf{x}_n)$ 张成空间, 所以在式 (7) 等式两边分别乘以 $\phi(\mathbf{x}_i)$ 构成新的等价式。

$$\lambda (\phi(\mathbf{x}_i) \mathbf{V}) = \phi(\mathbf{x}_i) \bar{\mathbf{K}} \mathbf{V} \quad (8)$$

存在系数 $\alpha_j (j=1, 2, \dots, n)$, 使 $\mathbf{V} = \sum_{j=1}^n \alpha_j \phi(\mathbf{x}_j)$,

并将其代入式 (8), 可得到:

$$\lambda \sum_{j=1}^n \alpha_j (\phi(\mathbf{x}_i) \phi(\mathbf{x}_j)) = \sum_{j=1}^n \alpha_j (\phi(\mathbf{x}_i) \bar{\mathbf{K}} \phi(\mathbf{x}_j)) \quad (9)$$

(6) 求解非线性主成分:

$$\theta_m = \mathbf{V} \phi(\mathbf{x}) = \sum_{i=1}^n \alpha_i^m (\phi(\mathbf{x}_i) \phi(\mathbf{x})) \quad (10)$$



式 (10) 表示原始低维空间任意样本 x 的映射数据 $\phi(x)$ 在特征向量 V 上的投影, θ_m 为第 m 维的非线性主成分。

(7) 对特征值进行从大到小排序, 并根据各主成分累积贡献率选取主成分, 一般累积贡献率 > 90% 所对应的主成分数量^[7]用来确定 BP 神经网络的输入变量。

2.2 GA 优化 BP 神经网络

BP 神经网络是一种按误差反向传播训练的多层前馈网络^[8-9], 其算法称为 BP 算法, 它的基本思想是梯度下降法, 利用梯度搜索技术, 以期使网络的实际输出值和期望输出值的误差均方差最小。基本 BP 算法包括信号的前向传播和误差的反向传播两个过程。即计算误差输出时按从输入到输出的方向进行, 而调整权值 W 和阈值 b 则从输出到输入的方向进行, 典型 BP 神经网络的 3 层结构如图 2 所示。

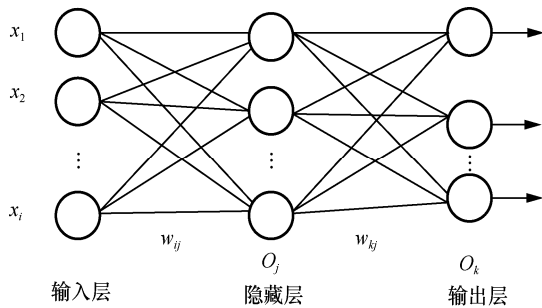


图 2 典型 BP 神经网络的 3 层结构

GA 优化 BP (GA-BP) 神经网络^[10-11]包括神经网络结构确定、遗传算法参数优化和 BP 神经网络预测 3 个部分。其中, 用 GA 赋值优化 BP 神经网络的初始权值 W 和阈值 b , 可有效解决 BP 神经网络对初始权值和阈值进行随机或人为赋值的不稳定缺陷, 使优化后的 BP 神经网络能够更好地预测函数输出, 其基本思想就是用个体代表网络的原始权值和阈值, 将初始化的 BP 神经网络的预测误差作为该个体的适应度值, 并通过选择、交叉、变异操作寻找最优个体, 即最优的 BP 神经网络

初始权值及阈值, GA 优化 BP 神经网络算法流程如图 3 所示。

GA 优化 BP 神经网络步骤如下。

(1) 将第 2.1 节通过 KPCA 算法降维后的数据作为 BP 神经网络的预测模型输入变量, 根据式 (11) 遍历尝试隐藏层神经元个数, 直至选出最优解。

$$p = \sqrt{n+m} + c \quad (11)$$

其中, n 为神经网络输入变量个数, m 为输出变量个数, $c \in [1, 10]$ 任意整数。

(2) 设置神经网络训练参数学习率、最小误差目标值、训练迭代次数、训练函数等。

(3) 通过编码符号串可表示个体串长度 L :

$$L = \text{inputnum} \times \text{hiddennum} + \text{hiddennum} + \text{hiddennum} \times \text{outputnum} + \text{outputnum} \quad (12)$$

其中, $\text{inputnum} \times \text{hiddennum}$ 为神经网络输入层至隐藏层的权值 W_i 个数, $\text{inputnum} \times \text{outputnum}$ 为隐藏层至输出层权值 W_j 个数, hiddennum 为隐藏层阈值 b_1 个数, outputnum 为输出层阈值 b_2 个数。

(4) 遗传算法按与个体适应度成正比的概率决定当前种群中每个个体遗传至下一代种群的机会, 其个体适应度必须为正数或 0, 不能为负。期望输出结果 x_i 与预测输出结果 y_i 之间的误差绝对值为个体适应度 (fit) 值, k 为适应度函数系数。

$$\text{fit} = k \left(\sum_{i=1}^n \text{abs}(y_i - x_i) \right) \quad (13)$$

(5) 选择运算采用轮盘法进行选择, 即个体被选中的概率 p_i 取决于个体的相对适应度 f_i 。

$$p_i = f_i / \sum_{i=1}^m f_i \quad (14)$$

(6) 交叉运算对种群中的个体进行两两随机匹配, 若种群数量为 Q , 则有 $Q/2$ 对匹配的个体组, 每一对相互匹配的个体, 随机设置某一基因座的位置为交叉点。对每一对相互匹配的个体,

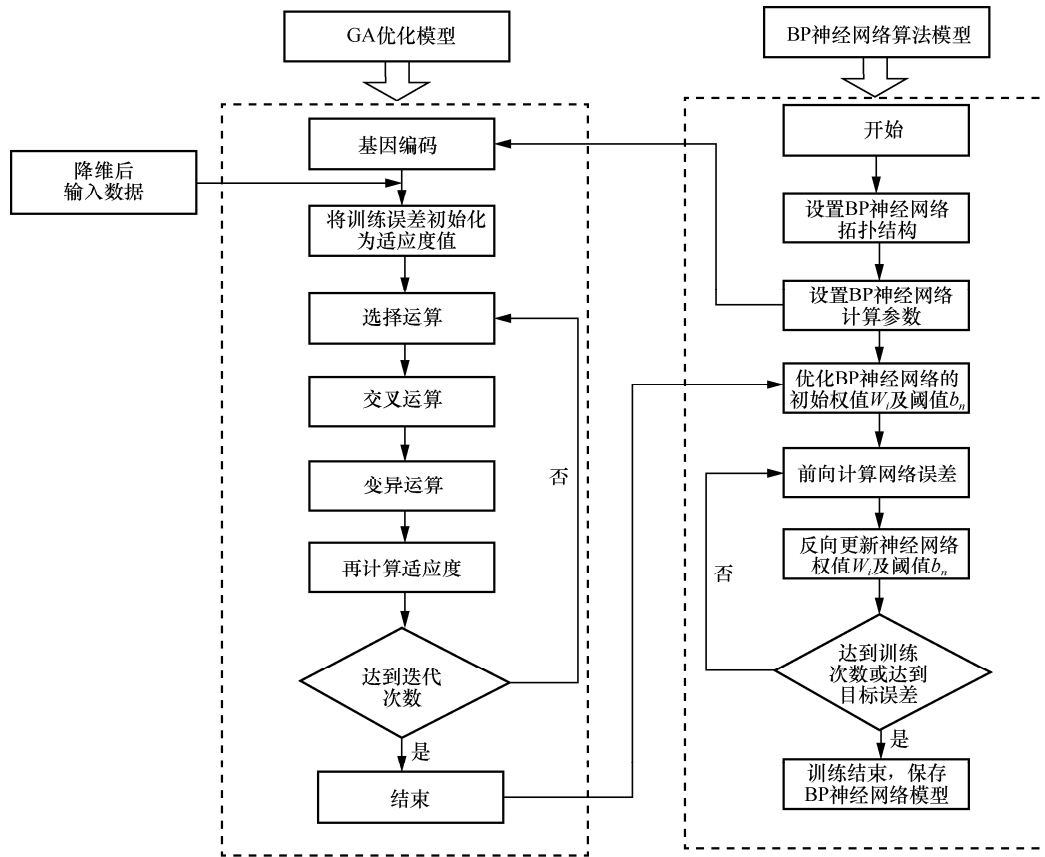


图3 GA优化BP神经网络算法流程

按预设的交叉概率 p_c 在其交叉点处交换彼此的部分染色体, 进而衍生两个新个体。

$$p_c = \frac{Q_c}{Q} \quad (15)$$

其中, Q_c 为群体中个体的数目。

(7) 变异运算针对个体的每一个基因座, 按照变异概率 p_m 指定其为变异点, 并针对每个指定的变异点, 对其基因值做取反运算或使用其他等位基因值进行替代, 进而产生新的个体, 则变异概率 p_m 为:

$$p_m = \frac{N}{QL} \quad (16)$$

其中, N 为每代中变异的基因数量, Q 为每代中群体拥有的个体数量, L 表示个体中基因串的长度。

(8) 如遗传算法模型达到设置最大遗传代数, 则输出优化 BP 神经网络的初始权值 W_i 及阈

值 b_n 至 BP 神经网络进行前向网络误差计算, 并反向更新其权值 W_i 及阈值 b_n , 直到达到设置的约束条件后结束模型参数优化, 输出最终的 BP 神经网络模型。

3 POI 网络质量预测算例

为验证该模型的有效性 & 准确性, 以某地市连续 3 个月的第 3 周无线网络 KPI 为原始数据源 (排除节假日、特殊事件影响), 通过第 1 节中 POI 网络质量数据的清洗及处理方法, 获取关联 POI 网络质量高、中、低标识的 KPI 指标数据 X_{ij} , 其中 i 表示样本数, $i=1,2,\dots,3\ 870$; j 表示 KPI 维度, $j=1,2,\dots,17$ 。

3.1 基于 KPCA 算法降维效果

根据 KPCA 算法进行标准化数据的特征提取, 可得到一组相互独立又包含大量原始数据信



息的主成分,一般将累计贡献率大于 90%作为降维标准,将原先 17 维输入变量通过 KPCA 算法降维后得到 4 维变量数据, KPCA 算法降维效果如图 4 所示。

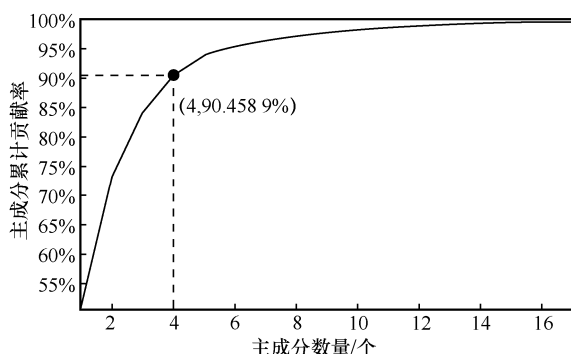


图 4 KPCA 算法降维效果

将 KPCA 算法与 PCA 算法在同一条件下,即主成分累计贡献率大于 90%,进行各项性能比较分析, KPCA 算法与 PCA 算法降维效果对比见表 2。

表 2 KPCA 算法与 PCA 算法降维效果对比

降维算法	累计贡献率>90%的主成分个数/个	降维幅度	算法耗时/s
PCA	12	29.41%	4.58
KPCA	4	76.47%	40.78

KPCA 算法在降维效果上比 PCA 算法提升约 47.06%,但耗时却增加 36.2 s,这是因为 PCA 算法协方差矩阵大小与输入变量样本维数相关,而 KPCA 算法中核矩阵大小与输入变量的样本点相关,本例中样本相对较小,核矩阵运算耗时较短,其在实际运用中是可接受的。

3.2 GA-BP 神经网络的模型预测

采用 KPCA 算法进行数据压缩降维后,获得一个 4 维变量数据,将其作为 GA-BP 模型的输入变量, POI 网络质量等级(高、中、低)作为输出结果,将 3 870 组样本随机划分为两个数据集,即 3 770 组样本的训练集和 100 组样本的测试集。通过多次逐步求精尝试,设定遗传算法的种群规模 Popsiz 为 40,最大遗传代数 Maxgen 为 100,

交叉概率 p_c 为 0.80,变异概率 p_m 为 0.01, BP 神经网络结构则采用典型的 3 层结构,即输入节点数 inputnum=4,隐藏层神经元数 hiddennum=10,输出节点数 outputnum=1,训练迭代次数为 100,训练函数则选择收敛速度较快的 trainlm 函数,训练学习率为 0.01,训练误差目标为 0.000 1。

经过上述 GA-BP 神经网络模型对输入变量进行训练,种群适应度与遗传代数关系曲线(终止代数=100)如图 5 所示,种群最佳适应度值在 0.759 54 及遗传代数为 34 时趋于收敛稳定。

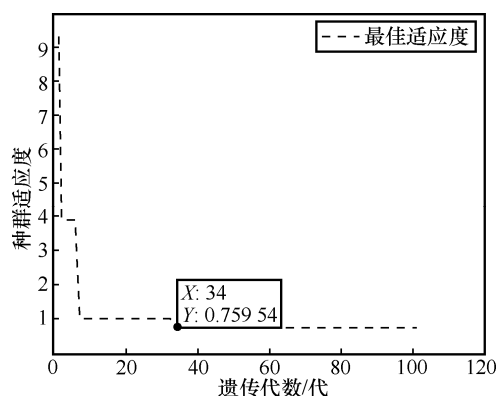


图 5 种群适应度与遗传代数关系曲线(终止代数=100)

将相对误差、均方误差(mean square error, MSE)以及预测准确度作为模型性能的验证标准,并与传统 BP 神经网络进行对比分析,结果如下。

基于 BP 算法、GA-BP 算法预测结果与实际值对比如图 6 所示,基于 BP 算法及 GA-BP 算法对 100 个测试样本组进行预测,依据相对误差表达式对比预测值 $\hat{y}_i$ 与实际值 y_i 如式(17)所示。

$$\varepsilon = \frac{|\hat{y}_i - y_i|}{y_i} \times 100\% \quad (17)$$

基于 BP 算法的相对误差范围为 $7.722 7 \times 10^{-5} \sim 4.335 \times 10^{-1}$,基于 GA-BP 算法的相对误差范围为 $5.413 2 \times 10^{-5} \sim 1.919 4 \times 10^{-1}$ 。从结果数据上分析,GA-BP 比 BP 算法更为平稳收敛,对误差“削峰”更有效。

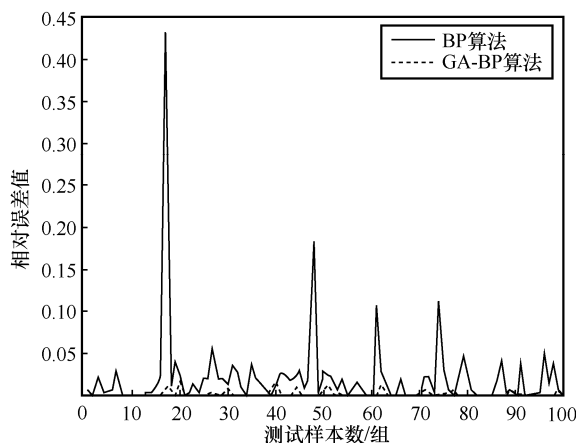


图6 基于BP算法与GA-BP算法预测结果与实际值对比

基于BP与GA-BP算法的预测准确度对比见表3, 基于BP与GA-BP算法对POI网络质量的预测准确度进行对比, 在同一标准条件下分别执行10组验证测试, BP预测准确度均值为86.40%, GA-BP较BP提升10.90%, 平均预测准确度可达到97.30%, 其预测性能整体明显优于BP算法结果。

表3 基于BP与GA-BP算法的预测准确度对比

实验组数	算法	
	BP	GA-BP
1	90.00%	98.00%
2	86.00%	97.00%
3	84.00%	99.00%
4	85.00%	94.00%
5	87.00%	99.00%
6	88.00%	98.00%
7	82.00%	99.00%
8	84.00%	97.00%
9	91.00%	98.00%
10	87.00%	94.00%

基于BP算法与GA-BP算法预测结果均方差性能对比如图7所示, 基于BP算法预测结果与实际结果的MSE收敛梯度较为平缓, 在BP神经网络训练100轮时, MSE性能指标仍未达到设置的目标值 1×10^{-4} , 而基于GA-BP算法的MSE下降梯度明显较快, 在模型训练至43轮时MSE达到 3.2883×10^{-5} , 已优于目标值, 模型性能提升显著。

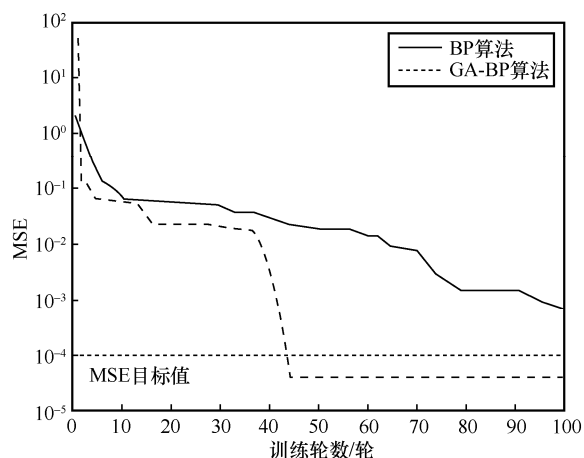


图7 基于BP算法与GA-BP算法预测结果均方差性能对比

4 结束语

本文依据“升维研究、降维实施”的研究思路, 首先构建POI网络质量评价体系, 从传统小区级转变为POI级优化分析的策略, 结合MDT定位数据柔性识别POI所辖主服小区, 实现网随业动的动态感知能力。并基于KPCA对原始数据进行压缩降维, 将此输出变量作为BP神经网络的输入数据源, 再采用GA对BP神经网络的权值、阈值进行初始化调优, 根据训练及测试的验证效果, 该基于KPCA算法降维的GA-BP神经网络模型具备较好的稳定性及准确性。

参考文献:

- [1] 丁彦文, 许捍卫, 汪成昊. 融合OSM路网与POI数据的城市功能区识别研究[J]. 地理与地理信息科学, 2020, 36(4): 57-63.
DING Y W, XU H W, WANG C H. Research on urban functional area recognition integrating OSM road network and POI data[J]. Geography and Geo-Information Science, 2020, 36(4): 57-63.
- [2] 刘璐, 陈睿杰, 李嘉. 基于MDT重叠覆盖度数据的KNN-DBSCAN参数自适应调优研究[J]. 电信科学, 2022, 38(2): 119-129.
LIU L, CHEN R J, LI J. Research on adaptive optimization of KNN-DBSCAN parameters based on MDT overlapping coverage data[J]. Telecommunications Science, 2022, 38(2): 119-129.
- [3] 苏灿. 移动通信网络优化实证研究[J]. 电信技术, 2019(S1): 19-21.



- SU C. An empirical study on mobile communication network optimization[J]. Telecommunications Technology, 2019(S1): 19-21.
- [4] 刘伟涛, 顾鸿, 李春洪. 基于德尔菲法的专家评估方法[J]. 计算机工程, 2011, 37(S1): 189-191, 204.
- LIU W T, GU H, LI C H. Expert evaluation method based on Delphi method[J]. Computer Engineering, 2011, 37(S1): 189-191, 204.
- [5] 彭令, 牛瑞卿, 赵艳南, 等. 基于核主成分分析和粒子群优化支持向量机的滑坡位移预测[J]. 武汉大学学报·信息科学版, 2013, 38(2): 148-152, 161.
- PENG L, NIU R Q, ZHAO Y N, et al. Prediction of landslide displacement based on KPCA and PSO-SVR[J]. Geomatics and Information Science of Wuhan University, 2013, 38(2): 148-152, 161.
- [6] 李新春, 房赓赓, 张春华. 基于 KPCA 和改进 GBRT 的室内定位算法[J]. 传感技术学报, 2019, 32(3): 430-437.
- LI X C, FANG Z X, ZHANG C H. Indoor positioning algorithm based on KPCA and improved GBRT[J]. Chinese Journal of Sensors and Actuators, 2019, 32(3): 430-437.
- [7] 董鸽, 闵建中, 陈立范, 等. 基于 MATLAB 的中药主成分分析数学实验教学案例[J]. 现代商贸工业, 2021, 42(34): 142-144.
- DONG G, MIN J Z, CHEN L F, et al. Math experiment teaching case of principal component analysis of traditional Chinese medicine based on MATLAB[J]. Modern Business Trade Industry, 2021, 42(34): 142-144.
- [8] 冯玉芳, 卢厚清, 殷宏, 等. 基于 BP 神经网络的故障诊断模型研究[J]. 计算机工程与应用, 2019, 55(6): 24-30.
- FENG Y F, LU H Q, YIN H, et al. Study on fault diagnosis model based on BP neural network[J]. Computer Engineering and Applications, 2019, 55(6): 24-30.
- [9] 温廷新, 孔祥博. 基于 KPCA-GA-BP 的煤矿瓦斯爆炸风险模式识别[J]. 安全与环境学报, 2021, 21(1): 19-26.
- WEN T X, KONG X B. Risk pattern recognition of the coal mine gas explosion based on the KPCA-GA-BP[J]. Journal of Safety and Environment, 2021, 21(1): 19-26.
- [10] 宋超, 宋娟. 基于遗传算法优化和 BP 神经网络的短期天然气负荷预测[J]. 工业控制计算机, 2012, 25(10): 82-84.
- SONG C, SONG J. Parameter optimization for BP neural network with GA on short-term gas load prediction[J]. Industrial

Control Computer, 2012, 25(10): 82-84.

- [11] 庞尔军, 于虹, 唐贵基, 等. 基于遗传算法优化的 BP 网络变压器可靠度计算[J]. 机电工程, 2013, 30(8): 997-1000, 1028.
- PANG E J, YU H, TANG G J, et al. Reliability calculation of transformer based on GA optimized BP[J]. Journal of Mechanical & Electrical Engineering, 2013, 30(8): 997-1000, 1028.

[作者简介]



刘璐 (1986-), 男, 中国移动通信集团设计院有限公司重庆分公司工程师、高级咨询设计师, 主要从事无线网络智能优化业务及相关咨询设计工作。



杨丹 (1986-), 男, 中国移动通信集团云南有限公司工程师, 主要从事无线优化大数据分析及相关管理工作。



陈睿杰 (1987-), 男, 中国移动通信集团云南有限公司工程师, 主要从事无线优化大数据分析、智能运维工作。

李嘉 (1989-), 男, 现就职于中国移动通信集团云南有限公司, 主要从事无线优化大数据分析、智能运维工作。

周熹 (1993-), 女, 现就职于中国移动通信集团云南有限公司, 主要从事网络优化研究工作。