

基于storm云平台的电力信息系统实时监理的研究

为解决电力信息工程系统监理质量控制过程中各要素不能实时处于受控状态的问题,利用蝙蝠算法对Elman神经网络的初始权值和阈值优化,并结合storm云平台,构建了电力信息工程系统实时分析预测模型。通过案例分析与标准的Elman神经网络模型对比,表明该模型具有更高的预测准确度和较快的收敛速度,能够有效地应用于电力信息工程系统监理质量控制中。

韩福霞¹ 储志高² 舒彬² 刘宏志¹ 尹璐² 丁仁山^{3/1} 1.北京工商大学计算机与信息工程学院 2.北京电力经济技术研究院 3.雅砻江水电集团



韩福霞/硕士研究生

0 引言

随信息工程监理^[1-2]过程中质量控制^[3-4]的不断研究,如:基于模糊理论、基于灰色系统理论和隐马尔可夫模型^[5]等,这些方法都能较好地适应于数据量小问题的解决,但随着大数据时代的到来,电力信息工程系统^[6]建设的数据量极具增加,需要对数据及时并准确的处理,另外,电力信息工程监理是动态监理,只有对信息工程监理各个阶段动态监理,实时对质量控制各要素进行分析,才能使得各要素一直处于实时受控状态,提高监理对电力信息工程的质量控制效果。

近年来,由于神经网络具有自学习能力、识别准确度高等优点^[7],已经被应

用到信息工程监理各目标控制过程中。Elman神经网络是J.L.Elmán^[8]于1990年提出的一种反馈动力学网络,不仅具有前馈神经网络的非线性映射,还具有动态反馈信息和记忆功能,能够实现动态系统的映射并直接反应系统的动态特性^[9]。将该算法应用于动态监理质量控制过程中,对各要素数据进行实时分析,可以有效地对质量进行控制。Elman神经网络主要是利用反向传播算法(BP算法)进行训练,通过学习训练,使得网络的实际的输出值和输出样本值之间的差值来不断地修改权值和阈值,从而使得网络输出层的误差平方和最小^[10]。但BP算法基于梯度信息来调整连接权值,不具有全局搜索能力,极易陷入局部极值点。

为保证电力信息工程监理质量控制过程中各要素能实时处于受控状态,本文将蝙蝠算法和Elman的神经网络相结合,优化Elman神经网络的初始权值和阈值。利用蝙蝠算法中蝙蝠个体之间的协作搜索最优解的全局搜索能力,能有效地避免局部极小值问题,并将该改进的Elman算法迁移到storm云平台上,建立了电力信息工程监理实时分析预测模型,能够及时得存储和处理信息工程监理数据,实现电力信息工程监理实时分析。

1 利用storm云服务平台处理电力信息系统数据

storm是一个分布式的、可伸缩的且容错的实时计算系统^[11],可以简单、可靠地处理大量的数据流,storm有很多

关键词/Keywords

电力信息工程监理 ·
storm云平台 ·
Elman神经网络算法 ·
实时分析 ·

北京市属高等学校科学技术与研究生教育创新工程建设项目(PXM2012_014213_000037)。

使用场景：如实时分析、在线机器学习、持续计算、分布式RPC和ETC等，本文利用了storm的实时分析功能来处理监理过程中各要素产生的数据。

storm云服务平台利用Nimbus对资源进行分配和任务调度，而Supervisor负责接受nimbus分配的任务，这两者通过zookeeper协调^[12]。worker是运行具体处理组件逻辑的进程，worker中每个spout/bolt的线程称为一个task，其中spout会从外部数据源中读取数据，bolt对数据进行过滤、函数操作和合并等操作，Storm实时处理模型如图1所示。

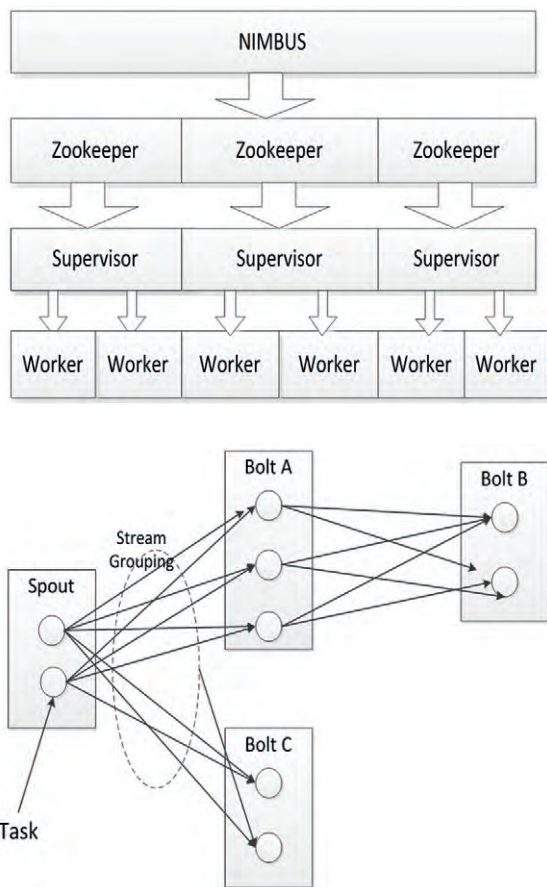


图1 Storm实时处理模型

2 利用改进的Elman神经网络分析数据

2.1 Elman神经网络模型

Elman神经网络结构如图2所示，从图2中可以看出，相比普通前馈神经网络在输入层多出了承接层并且其输入与隐藏层神经元的输出相关。

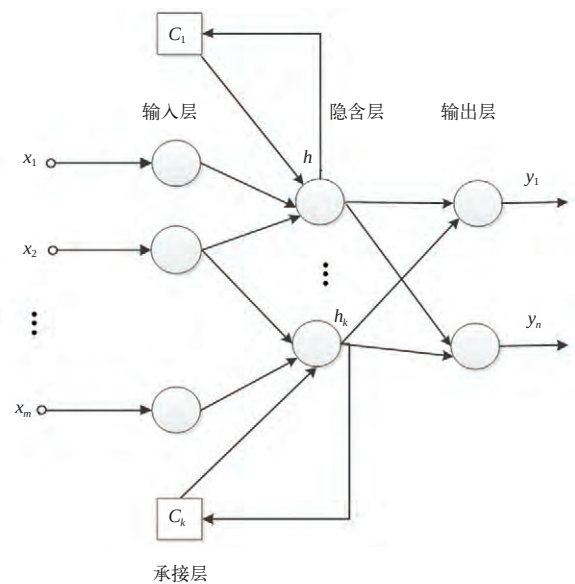


图2 Elman神经网络结构

Elman神经网络的运算规则由以下3个等式组成

$$h_o(k) = f[w_1 x(k) + w_2 C(k) + b_1] \quad (1)$$

$$y(k) = g[w_3 h_o(k) + b_2] \quad (2)$$

$$C(k) = h_o(k-1) \quad (3)$$

式中， $x(k)$ 为网络的输入； $y(k)$ 为网络的输出； $h_o(k)$ 为隐含层的输出； w_1 、 w_2 和 w_3 分别为输入层到隐含层、承接层到隐含层和隐含层到输出层的连接权值矩阵； $f(i)$ 为隐含层神经元的传递函数，采用tansig函数； $g(i)$ 为输出层传递函数，采用purelin函数； b_1 、 b_2 分别为输入层和隐含层的阈值。

2.2 利用蝙蝠算法训练Elman神经网络

蝙蝠算法^[13]是由Xin-She Yang于2010年提出的一种新型的启发式智能全局优化算法，通过模拟蝙蝠的回声定位能力，实现了群优化问题的求解。该算法具有模型简单、收敛速度快且可并行处理等特点。

改进过程如下：

蝙蝠算法中两个重要的元素就是位置和速度，其中位置代表问题的解，通过式(5)和式(6)更新每个个体的飞行速度和位置为

$$f_i = f_{\min} + (f_{\max} - f_{\min})\beta \quad (4)$$

$$v_i^{t+1} = v_i^t + (x_i^t - x_*) f_i \quad (5)$$

$$x_i^{t+1} = x_i^t + v_i^{t+1} \quad (6)$$

式中， x_i 、 v_i 分别为蝙蝠 i 在 t 时刻的位置和飞行速

度； x_* 为当前全局最优解； f_i 、 $f_{\max}$ 和 $f_{\min}$ 分别为在当前 t 时刻蝙蝠发出的脉冲频率、脉冲频率的最大值和最小值； $\beta \in [0,1]$ 为随机变量。

将所有需要优化的神经元间的连接权值和阈值变成个体，假设神经网络中包含 D 个优化参数，则每个个体将由一个 D 维相量为 $x_i=(x_{i1}, x_{i2}, \dots, x_{iD})$ 。

以网络在训练集上产生的均方误差作为目标函数，构造如下的适应度函数，来计算个体的适应值，即

$$J = \frac{1}{NC} \sum_{i=1}^N \sum_{j=1}^C (y_{j,i}^d - y_{j,i})^2 \quad (7)$$

式中， y_{ij}^d 为神经网络计算输出； y_{ij} 为理想输出； N 为训练样本数； C 为输出神经元个数。

BA算法训练ELman神经网络的算法实现步骤及流程图如图3所示：

(1) 根据神经网络的输入、输出样本集，确定网络结构。

(2) 将权值和阈值做为蝙蝠算法中的个体。初始化蝙蝠种群的位置 $x(i)$ 和速度 $v(i)$ 为 $[-1, 1]$ 之间的随机数。

(3) 根据输入、输出样本，利用式 (7) 计算适应度值 J 。

(4) 利用式 (5) 和式 (6) 产生新解。

(5) 评估新解，若 $J > J_{\min}$ ，说明不能满足预设的误差，则更新位置和速度，其中， $J_{\min}$ 为预设的误差。

(6) 判断算法是否满足最大迭代次数，若满足，则退出，并输出最终结果，否则返回 (2)。

3 电力信息工程监理实时分析模型建立

3.1 训练样本选取

电力信息工程质量监理分为招标阶段、设计阶段、实施阶段和验收阶段四个阶段，各个阶段质量控制可以分为人员、软件元、开发方法、设备材料和工程环境五个要素，但不同阶段，相同要素所占的权重不同，而软件质量状态是依据各阶段五个要素的控制情况来判断，软件质量的状态可以分为优、良、中、合格和差五种状态。

为了能够使改进的神经网络较好的运用于信

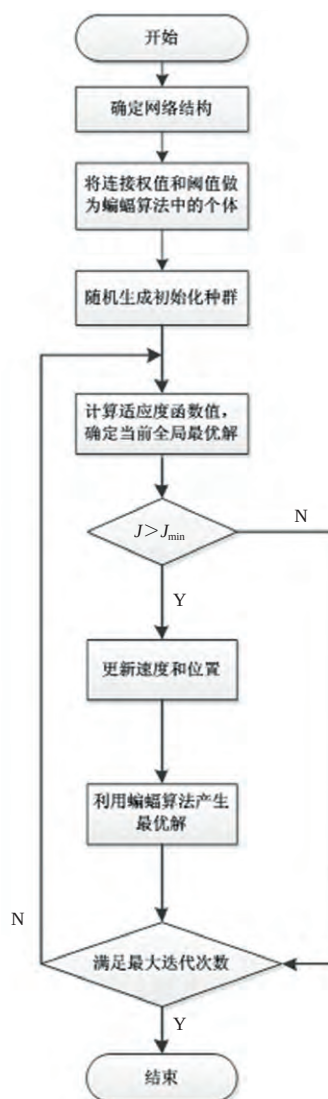


图3 蝙蝠算法训练神经网络流程图

息工程监理分析中，需要选定训练样本的输入和输出对神经网络进行有监督的训练，其中影响质量目标控制的各项因素作为训练样本的输入，软件质量状态作为训练样本的输出。

3.2 参数设定

做为神经网络训练样本输入，影响软件质量的要素有五个，用 $X=\{P,U,M,F,E\}$ 表示， $X_1=P$ (People)，表示样本输入人员； $X_2=U$ (Units)，表示样本输入软件元质量； $X_3=M$ (Methods)，表示样本输入开发方法； $X_4=F$ (Facilities)，表示样本输入设备材料； $X_5=E$ (Environment)，表示样本输入工程环境。设输入样本序列为 $X=\{X_1, X_2, X_3, X_4\}$ ，其中 x_i 取值范围为

[0, 1]。

神经网络训练样本输出, 定义监理软件质量的状态为 $S=\{A, B, C, D, E\}$, 即将软件质量分为A、B、C、D和E五个等级, $S_1=A$, 取值范围为[90, 100], 表示质量状态处于优; $S_2=B$, 取值范围为[80, 90], 表示质量状态处于良; $S_3=C$, 取值范围为[70, 80], 表示质量状态处于中; $S_4=D$, 取值范围为[60, 70], 表示质量状态处于合格; $S_5=E$, 取值范围为[0, 60], 表示质量状态处于差。通过对神经网络样本训练, 使得改进神经网络能够准确的对电力信息工程监理过程中的质量控制目标进行分析。

另外, Storm实时云平台是一个实时云计算系统, 能够保证每个消息都会得到处理。将改进的神经网络算法迁移到Storm云平台上, 可以实现实时的对电力信息工程监理数据进行分析。

4 案例分析

该仿真实验是在Windows 07 系统下Matlab_R2012中运行的。实验数据来源于信息工程监理协同工作平台(IPSS)下的数据库。分别选取监理过程中的各阶段质量控制数据100组, 共400组作为训练样本, 并从这400组数据中均匀抽出100组作为测试数据。信息工程质量监理实时分析模型的输入神经元为5个, 输出神经元为1个, 隐藏节点数分别选择7、11、14和18, 根据多次仿真实验结果, 确定隐藏层为18个神经元, 如图4所示。

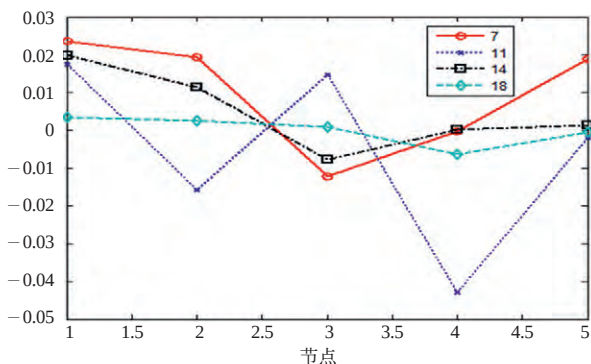


图4 Elman不同隐藏节点预测误差图

将隐含层和输出层的激励函数分别设置为tansig和logsig, 蝙蝠算法中的参数设置为: $\alpha=\beta=0.9$, 初始速度 $v=0$, 初始脉冲频率 $f=0$, 初始脉冲响度 $A \in [0, 1]$ 和初始脉冲发射速率 $r \in [0, 0.5]$ 均为随机生成函数rand生成。分别用基本的Elman神

经网络和本文改进的Elman神经网络训练, 结果如图5、图6所示。

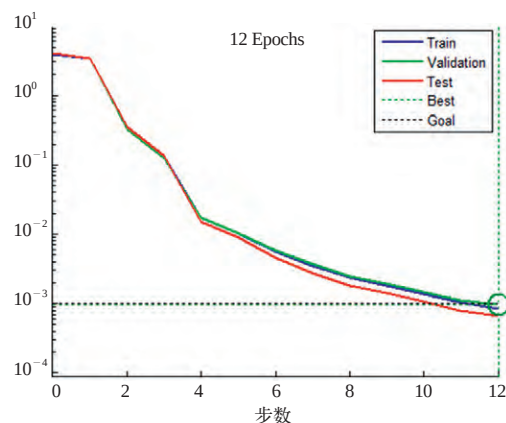


图5 标准Elman网络训练图

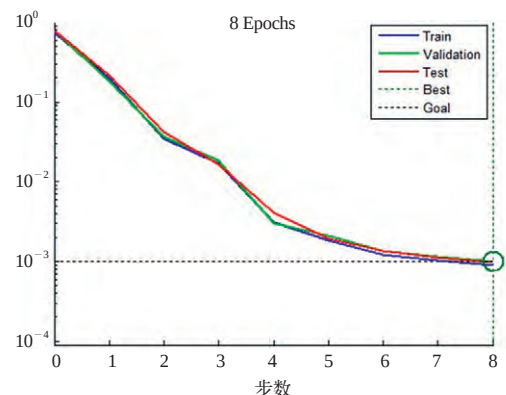


图6 改进Elman网络训练图

由图5和图6可知, 误差为0.001条件下, 改进Elman神经网络的收敛速度要比标准Elman神经网络的收敛速度快, 标准Elman神经网络训练时, 容易陷入局部最优。

根据训练的结果, 利用标准的Elman神经网络和改进的Elman神经网络对样本进行预测, 将预测值与实际值进行比较, 预测结果分别如图7、图8所示。

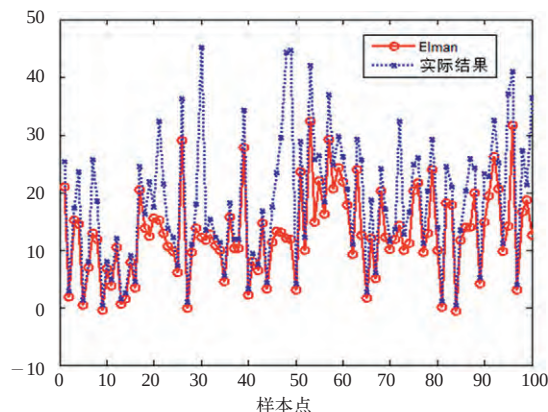


图7 标准Elman神经网络预测结果

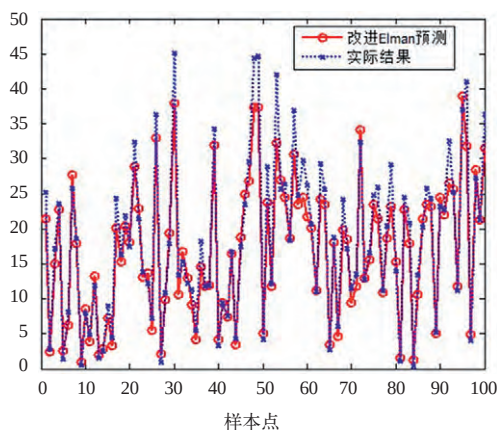


图8 改进Elman神经网络预测结果

由图7、图8对比可以看到,改进的Elman神经网络具有较高的预测准确度,能够较好的预测质量状态。

另外,随着大数据时代的到来,电力信息工程监理质量控制各要素的数据量极速增加,为了保证数据能够实时分析,使得各要素处于受控状态,将该改进Elman神经网络算法迁移到storm云平台上,将算法写到bolt中,而在storm中把对应的bolt部署到3台Supervisor中,整体性能提升2.4~2.6倍,如图9所示。

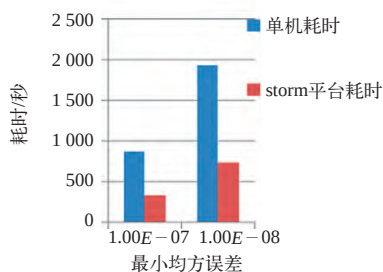


图9 单机和storm平台训练耗时对比

6 结束语

本文利用蝙蝠算法优化Elman神经网络的初始权值和阈值,一定程度上提高了算法的准确度,并结合storm云平台,建立了用于质量控制的电力信息工程监理实时分析模型,并尝试基于该模型对各电力信息工程监理质量进行控制,保证各要素处于实时受控状态。但本文只用了控制信息工程监理的五个基本要素作为训练样本,使得该模型稍显粗糙,需要进一步对该模型进行研究。另外,storm云技术具有实时计算的能力,但本文未充分的利用该技术,需要进一步对该技术应用进

行研究。

参考文献

- [1] 刘宏志,葛迺康. 信息化工程监理[M]. 北京: 中国电力出版社, 2009.
- [2] 韩福霞, 刘宏志. 基于蝙蝠算法的信息工程监理多目标优化研究[J]. 现代计算机, 2013(19): 3-6
- [3] 刘宏志, 朱方方. 基于云模型的信息工程监理服务质量评价的研究[J]. 电子技术应用, 2013, 39 (10): 133-135.
- [4] 张云霄, 刘宏志. 我国食品安全监理体系研究[J]. 食品科学技术学报, 2014, 32(1): 77-82.
- [5] 高丽, 刘宏志. 隐马尔科夫模型在信息工程质量监理中的应用[J]. 微型机与应用, 2010, 29(21): 91-93.
- [6] 姚楠. 电力信息系统工程监理的实施模型及其实践[J]. 电力信息化, 2009, 7(1):74-78.
- [7] 余伶俐, 周开军, 邱爱兵. 基于Elman神经网络的语音情感识别应用研究[J]. 计算机应用, 2012, 29(5): 1809-1814.
- [8] Raja S Toqeer, N Suha Bayindir. Speed estimation of an induction motor using Elman neural network [J]. Neuro Computing, 2003, 55(3-4): 727-730.
- [9] 党小超, 郝占军. 基于改进Elman神经网络的网络流量预测[J]. 计算机应用, 2010, 30(10): 2648-2652.
- [10] Shi XH, Liang YC, LeeHP, et al. Improved Elman networks and applications for controlling ultrasonic motors[J]. Applied Artificial Intelligence, 2004(18): 603-629.
- [11] 朱炳宇, 蔡莉. 基于Storm和Android的出租车乘客实时云平台调度系统[J]. 计算机科学, 2012(8A): 1-2.
- [12] 黄馥浩. 基于Storm的微博互动平台的设计与实现[D]. 广州: 中山大学, 2013.6
- [13] 朱炳宇. 基于Storm云平台的地图道路匹配算法研究[D]. 云南: 云南大学, 2013.

(收稿日期: 2015-05-15) EA

作者简介

刘宏志 (1964-), 男, 博士, 教授, 研究领域为信息工程监理、电子政务、软件工程。