

# 等效分组级联 BP 网络模型及其应用

鲁智勇<sup>1,2</sup>, 张 权<sup>1</sup>, 张 希<sup>1</sup>, 唐朝京<sup>1</sup>

(1. 国防科技大学电子科学与工程学院, 湖南长沙 410073; 2. 六三八八零部队, 河南洛阳 471003)

**摘 要:** 松弛的和紧密的分组级联 BP 网络模型等概念的提出, 对于解决有限的小样本情况下高维 BP 网络的训练和预测问题有一定的参考价值. 定义了 BP 网络等效性和相关定理, 构建并证明了与 BP 网络等效的分组级联网络模型, 分析比较了两种网络模型所需训练样本的数量情况, 并将其应用于网络安全评估领域. 最后通过仿真试验结果证实了所提出分组级联 BP 网络模型的可行性和有效性.

**关键词:** 分组级联网络模型; BP 神经网络; 等效性; 小样本

**中图分类号:** TP183 **文献标识码:** A **文章编号:** 0372-2112 (2010) 06-1349-06

## Equivalent Grouping-Cascaded BP Network Model and Its Applications

LU Zhi-yong<sup>1,2</sup>, ZHANG Quan<sup>1</sup>, ZHANG Xi<sup>1</sup>, TANG Chao-jing<sup>1</sup>

(1. College of Electronic Science and Engineering, National University of Defense Technology, Changsha, Hunan 410073, China;

2. Unit 63880, Luoyang, Henan 471003, China)

**Abstract:** This paper puts forward to the concept of loosely and tightly grouping-cascaded BP network model to solve the training problem of high dimension BP neural network with limited small samples. After defining the equivalence definition and relative theorem of BP neural network, the grouping-cascaded model which is proved to be equivalent to BP network is constructed and the required number of training samples between them is compared. Then this method is applied in the evaluation of the network security. Finally, the feasibility and validity of the proposed grouping-cascaded BP network model is verified with simulation results.

**Key words:** grouping-cascaded network model; BP neural networks; equivalent; small samples

### 1 引言

BP 神经网络(back propagation, BP)是基于误差反向传播学习算法的多层前馈神经网络,其基本原理是采用梯度下降的方法在权矢量空间中求取误差函数的极小值,使误差函数极小化,直到误差低于一个预先设定的阈值.近年来,BP 神经网络已获得了广泛的实际应用:用于水下目标识别<sup>[1]</sup>;用于实时的故障诊断<sup>[2]</sup>;对工厂效益和业绩进行综合评估<sup>[3]</sup>;对雷达目标用于分类识别<sup>[4]</sup>;用于机器人控制和模糊控制领域<sup>[5]</sup>,如此等等.实践中、理论上已经证明,在不限制隐含层节点数的情况下,只有一个隐层的 BP 网络可以模拟任意线性与非线性函数,但当样本的维数很高时,为减少网络规模,BP 网络需要有两个隐含层<sup>[6,7]</sup>.我们知道,对于输入层有  $M$  个节点、第一隐含层有  $L_1$  个节点、第二隐含层有  $L_2$  个节点、输出层有  $d$  个节点的高维 BP 神经网络来说,记为 BP( $M, L_1, L_2, d$ )网络(如不加以说明,下文的高维 BP 神经网络都是含有两个隐含层),要具有良好的预测

能力和模式识别能力,实现一个好的泛化,训练样本集的大小  $N$  应满足条件<sup>[8,9]</sup>:

$$N = O(W/\epsilon) \quad (1)$$

其中,  $W$  是指网络中自由参数(即突触权值和偏值)的总数,  $\epsilon$  表示测试数据中允许分类误差.  $O(\cdot)$  表示所包含的量的阶数. 则高维 BP 神经网络总的自由参数  $W$  为:

$$W = M \cdot L_1 + L_1(L_2 + 1) + L_2(d + 1) + d \quad (2)$$

由式(1)和(2)可知,对于一个网络体系结构和允许分类误差固定的 BP 网络,要产生一个好的泛化,训练样本集的大小与输入节点数(训练样本维数)、隐含层节点数、输出层节点数是成线性关系的.

而对于 BP 网络来说,隐层节点数的多少对网络性能的影响较大,当隐层节点数太多时,会导致网络学习时间过长,训练样本要很大,甚至不能收敛;而当隐层节点数过小时,网络的容错能力差,网络输出和希望输出之间的拟合度不高,不能充分发挥神经网络的高度拟合性特点.隐层神经元数目的选取是一个十分复杂的问

题,往往需要根据设计者的经验和多次实验来确定,因而,不存在一个理想的解析式来表示.隐单元的数目与问题的要求、输入/输出单元的数目都有着直接关系.隐含层的节点根据以下经验公式<sup>[6,7]</sup>进行设计:

$$L = \sqrt{M + d} + \lambda \quad (3)$$

式中,  $L$  为隐含层节点个数;  $M$  为输入节点个数;  $d$  为输出节点个数;  $\lambda$  为 1~10 的常数.

由式(3)可知,对于一个输入和输出节点数已知的 BP 网络,隐含层节点的个数  $L$  是可变的,但应在一个整数区间  $[\lceil \sqrt{M + d} + 1 \rceil, \lceil \sqrt{M + d} + 10 \rceil]$  范围内变化,其中输出节点个数  $d$  是已知确定的,那么隐含层节点个数主要由输入节点数确定.

综上所述,为获得一个具有很好泛化能力的 BP 网络,满足式(1)的训练样本集的大小主要由输入节点数(训练样本维数)确定.但在实际情况中,获得的训练样本集的大小往往是有限的,对于低维的 BP 网络,训练样本的个数还能够满足式(1)的要求,但对于高维 BP 网络,训练样本数量往往不能满足式(1)的要求,也就是不能使训练的高维 BP 网络具有很好的泛化能力.

那么,在有限的高维训练样本空间条件下,如何才能使高维 BP 网络具有很好的泛化能力,本文提出的等效分组级联 BP 网络模型,能很好的解决这方面的问题.

## 2 BP 网络的等效分组级联模型

为了说明问题的方便,先给出以下定义和定理:

**定义 1** 级联 BP 网络模型:前一级多个 BP 网络的输出作为下一级 BP 网络的输入,如此构建的 BP 网络的多级互联模型,称为级联 BP 网络模型(简称 CBP, Cascaded BP).

级联 BP 网络模型有二级级联 BP 网络模型、三级级联 BP 网络模型和多级级联 BP 网络模型.

二级级联 BP 网络模型如图 1 所示.

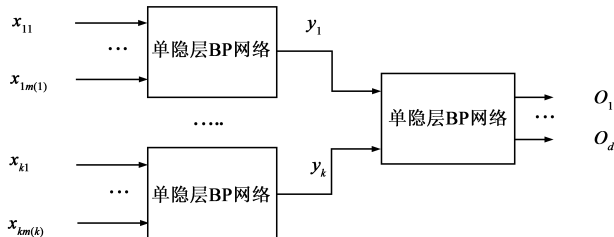


图1 二级级联BP网络模型

在图 1 所示的二级级联 BP 网络模型中,第一级共有  $k$  个单隐层 BP 网络,输入节点分别为  $(x_{i1}, x_{i2}, \dots, x_{im(i)}), \dots, (x_{k1}, \dots, x_{km(k)})$ , 输出为  $y_i$ , 其中  $1 \leq i \leq k, m(i)$  为第  $i$  个 BP 网络的输入维数;第二级是一个单隐层 BP 网络,是以第一级 BP 网络的输出作为输入,故其输入节点为  $(y_1, y_2, \dots, y_k)$ , 输出节点有  $d$  个,分别为  $o_1, o_2, \dots, o_d$ . 此二级级联

BP 网络模型可记为  $CBP(m(1), \dots, m(k), L_1, L_2, d)$ .

**定义 2** 松弛级联 BP 网络模型:在一个级联 BP 网络模型中,如果前一级多个 BP 网络的输出与下一级 BP 网络的输入在物理上并不相联,这种结构的级联 BP 网络模型称为松弛级联 BP 网络模型(简称 LCBP, Loosely Cascaded BP).

松弛级联 BP 网络模型实质上是由多个独立的 BP 网络组成,之所以称之为级联 BP 网络模型,这是由于在网络训练和预测时,前一级 BP 网络和后一级 BP 网络要遵循一定的规律:在级联 BP 网络模型的训练过程中,前一级 BP 网络的训练样本的输出作为下一级 BP 网络的输入;在级联 BP 网络模型的预测过程中,前一级 BP 网络的预测结果作为下一级 BP 网络的预测样本输入.

**定义 3** 紧密级联 BP 网络模型:在一个级联 BP 网络模型中,如果前一级多个 BP 网络的输出与下一级 BP 网络的输入在物理上直接相联,这种结构的级联 BP 网络模型称为紧密级联 BP 网络模型(简称 TCBP, Tightly Cascaded BP).

紧密级联 BP 网络模型也是由多个 BP 网络组成,但并非每个 BP 网络之间是相互独立的,前后级的 BP 网络是直接相连的,级联 BP 网络模型的训练和预测要同时进行,可以看作是一个单独的神经网络,但已不是一个 BP 神经网络.

**定义 4** BP 网络的分组级联模型:级联 BP 网络模型第一级的多个 BP 网络的输入是由一个高维 BP 网络的输入样本经过分组得到的,这个级联 BP 网络模型称为此高维 BP 网络的分组级联模型(简称 GCBP, Grouping-Cascaded BP).

高维 BP 网络结构图如图 2 所示.

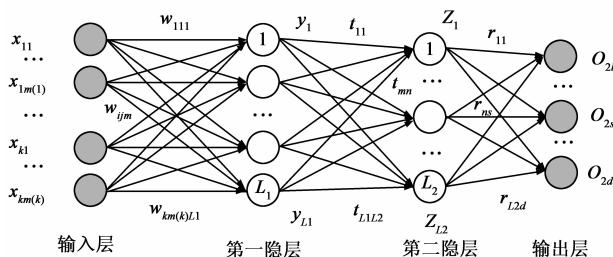


图2 高维BP网络结构图

在图 2 所示的高维 BP 网络中,输入节点共有

$\sum_{i=1}^k m(i)$  个,输入节点是  $(x_{i1}, \dots, x_{i, m(i)}), \dots, (x_{k1}, \dots, x_{k, m(k)})$ , 输出节点有  $d$  个,分别为  $o_{21}, \dots, o_{2s}, \dots, o_{2d}$ ;有两个隐含层,第一隐含层有  $L_1$  个节点,第二隐含层有  $L_2$  个节点;  $w_{ijm}$ 、 $t_{lmn}$  和  $r_{ns}$  分别表示输入节点与第一隐含层各节点、第二隐含层各节点和输出节点之间的权重关系.第一隐含层、第二隐含层和

输出层所用的激活函数分别为  $f_1(\cdot), f_2(\cdot), f_3(\cdot)$ , 激活函数可以是线性函数、logistic 函数或正切函数。其中,  $1 \leq i \leq k, 1 \leq j \leq m(i), 1 \leq m \leq L_1, 1 \leq n \leq L_2, 1 \leq s \leq d$  (如无声明, 下文相同变量的取值同此)。

如果图 1 中级联 BP 网络模型的第一级所有 BP 网络的输入是由图 2 所示高维 BP 网络的输入经分组得到的, 那么就定义图 1 中的级联 BP 网络模型是图 2 所示高维 BP 网络的分组级联模型。

**定义 5** BP 网络的等效性: 对于两个已训练好的理论上是无偏估计 (偏差和输出误差全为零) 的 BP 网络 (包括级联 BP 网络模型), 对任意一个有效输入样本, 如果两个 BP 网络的输出总是相同的, 则称这两个 BP 网络是等效的。

如果从实际应用角度出发, BP 网络的等效性可定义为: 对于两个已训练好的 BP 网络, 对任意一个有效测试样本集, 如果两个 BP 网络的输出都达到一定程度正确识别率, 则称这两个 BP 网络是等效的。

**定义 6** BP 网络的等效分组级联模型: 如果一个高维 BP 网络与其分组级联模型是等效的, 则称此 BP 网络的分组级联模型为此高维 BP 网络的等效分组级联模型。

**定理 1** 等效性定理: 对于一个高维 BP 网络, 总能找到一个与其等效的级联 BP 网络模型; 同样, 对于一个级联 BP 网络模型, 总能找到一个与其等效的高维 BP 网络。

定理 1 的证明见 4。

### 3 分组级联 BP 网络模型的构建

对于图 2 所示的高维 BP 网络, 其二级分组级联 BP 网络模型为: 第一级共有  $k$  个 BP 网络, 每个 BP 网络都是隐含层有  $L_1$  个节点的单隐层单输出网络, BP 网络的输入节点数分别是  $m(1), \dots, m(i), \dots, m(k)$ , 输入样本分别是  $(x_{11}, \dots, x_{1m(1)}), \dots, (x_{i1}, \dots, x_{im(i)}), \dots, (x_{k1}, \dots, x_{km(k)})$ ; 第二级 BP 网络是一个隐含层有  $L_2$  个节点的单隐层网络, 此 BP 网络的输入节点数为  $k$ , 输入数据为第一级 BP 网络的输出, 输出节点有  $d$  个, 分别为  $o'_{21}, \dots, o'_{2s}, \dots, o'_{2d}$ 。BP 网络的分组级联模型结构图见图 3。

在图 3 所示的 BP 网络的分组级联模型结构图中: 第一级的第  $i$  个 BP 网络的各输入节点与隐含层各节点之间的权重关系用  $w'_{ijm}$  表示, 隐含层各节点与输出节点  $o_{1i}$  之间的权重关系用  $p_{im}$  表示; 第二级的 BP 网络的各输入节点与隐含层各节点之间的权重关系用  $q_{in}$  表示, 隐含层各节点与输出节点  $o'_{2s}$  之间的权重关系用  $r'_{ns}$  表示; 第一级 BP 网络的隐含层、第二级 BP 网络的隐含层

和输出层所用的激活函数分别为  $f_1(\cdot), f_2(\cdot), f_3(\cdot)$ , 激活函数可以是线性函数、logistic 函数或正切函数。图中的一条竖线物理上把前后级 BP 网络隔离开来, 表示模型中的级联关系是松弛的; 如果图中没有竖线, 则表示模型中的级联关系是紧密的。

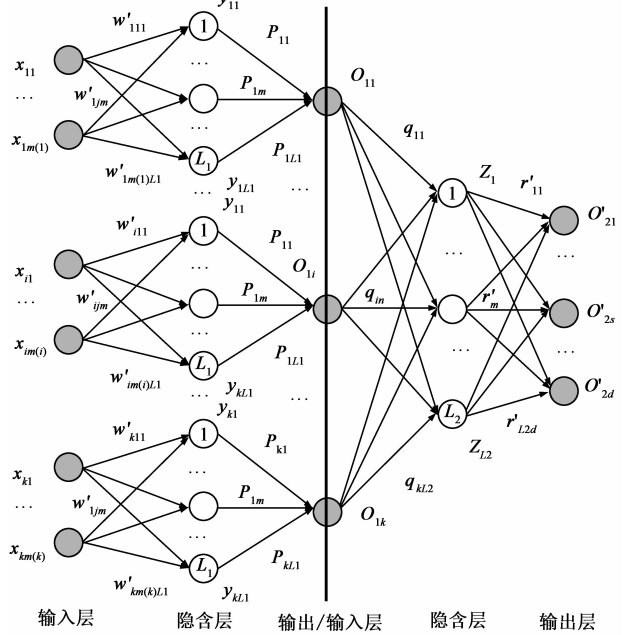


图3 分组级联BP网络模型结构图

### 4 BP 网络等效性的证明

根据 BP 网络等效性的定义, 定理 1 可以用仿真方法来验证: 用同一个训练样本集对 BP 网络和其分组级联网络模型进行训练, 直至均满足一定的误差要求, 然后选择任意一个有效测试样本集进行测试, 比对两个网络输出结果的正确识别率, 如果都能达到一定程度的正确识别率, 则说明这两个 BP 网络是等效的。

定理 1 也可用理论推导的方法进行证明, 具体如下:

对于已训练好的无偏估计的高维 BP 网络 (如图 2 所示), 如输入样本为  $(x_{11}, \dots, x_{1m(1)}), \dots, (x_{i1}, \dots, x_{im(i)}), \dots, (x_{k1}, \dots, x_{km(k)})$ , 则第一隐含层各节点的值为:

$$y_m = f_1\left(\sum_{i=1}^k \sum_{j=1}^{m(i)} (x_{ij} \cdot w_{ijm})\right) \quad (4)$$

第二隐含层各节点的值为:

$$\begin{aligned} z_n &= f_2\left(\sum_{m=1}^{L_1} (y_m \cdot t_{mn})\right) \\ &= f_2\left(\sum_{m=1}^{L_1} \left(f_1\left(\sum_{i=1}^k \sum_{j=1}^{m(i)} (x_{ij} \cdot w_{ijm})\right) \cdot t_{mn}\right)\right) \end{aligned} \quad (5)$$

BP 网络的第  $s$  个输出神经元的输出值为:

$$o_{2s} = f_3\left(\sum_{n=1}^{L_2} (z_n \cdot r_{ns})\right)$$

$$= f_3 \left( \sum_{n=1}^{L_2} \left( f_2 \left( \sum_{m=1}^{L_1} \left( f_1 \left( \sum_{i=1}^k \sum_{j=1}^{m(i)} (x_{ij} \cdot w_{ijm}) \right) \cdot t_{mn} \right) \right) \cdot r_{ns} \right) \right) \quad (6)$$

对于已训练好的无偏估计的 BP 网络的分组级联模型,输入样本为  $(x_{11}, \dots, x_{1m(1)}), \dots, (x_{i1}, \dots, x_{im(i)}), \dots, (x_{k1}, \dots, x_{km(k)}),$  则第一级第  $i$  个 BP 网络的隐含层各节点的值为:

$$y_{im} = f_1 \left( \sum_{j=1}^{m(i)} (x_{ij} \cdot w'_{ijm}) \right) \quad (7)$$

第一级第  $i$  个 BP 网络的输出值为:

$$\begin{aligned} o_i &= \sum_{m=1}^{L_1} (y_{im} \cdot p_{im}) \\ &= \sum_{m=1}^{L_1} \left( f_1 \left( \sum_{j=1}^{m(i)} (x_{ij} \cdot w'_{ijm}) \right) \cdot p_{im} \right) \end{aligned} \quad (8)$$

第二级 BP 网络的隐含层各节点的值为:

$$\begin{aligned} z_n &= f_2 \left( \sum_{i=1}^k (o_i \cdot q_{in}) \right) \\ &= f_2 \left( \sum_{i=1}^k \sum_{m=1}^{L_1} \left( f_1 \left( \sum_{j=1}^{m(i)} (x_{ij} \cdot w'_{ijm}) \right) \cdot p_{im} \cdot q_{in} \right) \right) \end{aligned} \quad (9)$$

第二级 BP 网络的第  $s$  个输出神经元的输出值为:

$$\begin{aligned} o'_{2s} &= f_3 \left( \sum_{n=1}^{L_2} (z_n \cdot r'_{ns}) \right) \\ &= f_3 \left( \sum_{n=1}^{L_2} \left( f_2 \left( \sum_{i=1}^k \sum_{m=1}^{L_1} \left( f_1 \left( \sum_{j=1}^{m(i)} (x_{ij} \cdot w'_{ijm}) \right) \cdot p_{im} \cdot q_{in} \right) \right) \cdot r'_{ns} \right) \right) \end{aligned} \quad (10)$$

根据定义 3, 如果图 2 和图 3 中的两个网络是等价的, 则对于任意有效的输入样本, 两个网络的输出值是相同的, 即式(6)和式(10)相等:

$$\begin{aligned} o_{2s} &= f_3 \left( \sum_{n=1}^{L_2} \left( f_2 \left( \sum_{m=1}^{L_1} \left( f_1 \left( \sum_{i=1}^k \sum_{j=1}^{m(i)} (x_{ij} \cdot w_{ijm}) \right) \cdot t_{mn} \right) \right) \cdot r_{ns} \right) \right) \\ &= o'_{2s} \\ &= f_3 \left( \sum_{n=1}^{L_2} \left( f_2 \left( \sum_{i=1}^k \sum_{m=1}^{L_1} \left( f_1 \left( \sum_{j=1}^{m(i)} (x_{ij} \cdot w'_{ijm}) \right) \cdot p_{im} \cdot q_{in} \right) \right) \cdot r'_{ns} \right) \right) \end{aligned} \quad (11)$$

对于已训练好的分组级联模型(如图 2 或图 3 所示), 令

$$\begin{cases} t_{mn} = \sum_{i=1}^k (p_{im} \cdot q_{in}) \\ r'_{ns} = r_{ns} \end{cases} \quad (12)$$

则由式(11)可推知:

当激活函数  $f_1(\cdot)$  是线型函数时, 有

$$w_{ijm} = w'_{ijm} \quad (13)$$

根据式(12)和式(13)计算出的权重值, 就可构建与分组级联模型等效的 BP 网络, 即对于一个级联 BP 网络模型, 总能找到一个与其等效的高维 BP 网络。

同样, 对于已训练好的高维 BP 网络(如图 2 所示),

令

$$\begin{cases} \sum_{i=1}^k (p_{im} \cdot q_{in}) = t_{mn} \\ r'_{ns} = r_{ns} \end{cases} \quad (14)$$

则由式(11)可推知:

当激活函数  $f_1(\cdot)$  是线型函数时, 有

$$w'_{ijm} = w_{ijm} \quad (15)$$

根据式(14)和式(15)计算出各权重值, 据此可构建与高维 BP 网络等效的分组级联模型, 即对于一个高维 BP 网络, 总能找到一个与其等效的级联 BP 网络模型。

综上所述, 定理 1 得以证明。

根据定理 1 可知: 对于任何高维的 BP 网络, 如果给出的训练样本数量不足以满足式(1)的要求, 可以对高维的输入样本进行分组, 依据分组结果构造 BP 网络的分组松弛级联模型, 然后分别对松弛级联模型中的各个 BP 网络进行训练, 在已训练好的分组级联 BP 网络模型上, 对于给出的任意输入样本, 级联模型最后一组 BP 网络的输出就是最终的预测结果; 也可以构建 BP 网络的分组紧密级联模型, 然后对紧密级联 BP 网络模型进行训练, 在已训练好的级联 BP 网络模型上, 对于给出的任意输入样本, 级联模型最后一组 BP 网络的输出就是最终的预测结果。

## 5 分组级联 BP 网络模型所需训练样本数量分析

根据式(1)可知: 在允许存在一定训练误差的情况下, 神经网络所需训练样本数由网络中自由参数(即突触权值和偏值)的总数  $W$  决定。而图 2 和图 3 中不同类型神经网络自由参数的总数分别如下:

(1) 对于图 2 所示的高维 BP 网络, 网络中自由参数的总数  $W_{BP}$  为:

$$W_{BP} = \left( \sum_{i=1}^k m(i) + 1 \right) L_1 + (L_1 + 1) L_2 + (L_2 + 1) d \quad (16)$$

(2) 对于图 3 所示的松弛级联 BP 网络模型, 网络中自由参数的总数  $W_L$  为:

$$W_L = \max \begin{cases} m(i) \cdot L_1 + 2L_1 + 1 \\ (k + d + 1) L_2 + d \\ 1 \leq i \leq k \end{cases} \quad (17)$$

这里选取最大值, 表示满足最大自由参数网络的训练样本数量能满足较小自由参数网络的训练要求。

(3) 对于图 3 所示的紧密级联 BP 网络模型, 网络中自由参数的总数  $W_T$  为:

$$\begin{aligned} W_T &= \left( \sum_{i=1}^k m(i) + k \right) L_1 + (k \cdot L_1 + k + k \cdot L_2 + L_2) \\ &\quad + (L_2 + 1) d \end{aligned} \quad (18)$$

比较式(16)、(17)和(18), 显然, 无论在何种情况下, 总有  $W_L < W_T$ ; 而在训练样本分组不是太多, 即  $k$  不

是很大的情况下,有  $W_L < W_{BP}$ . 因此有以下结论:

**结论 1** 在允许同样训练误差的情况下,松弛级联 BP 网络模型的训练样本数量总是小于紧密级联 BP 网络模型的训练样本数量,也小于 BP 网络的训练样本数量.

为比较  $W_{BP}$  和  $W_T$  大小,定义  $W_{BP-T} = W_{BP} - W_T$ ,则有

$$W_{BP-T} = L_1 \cdot L_2 + L_1 - 2k \cdot L_1 - k \cdot L_2 - k \quad (19)$$

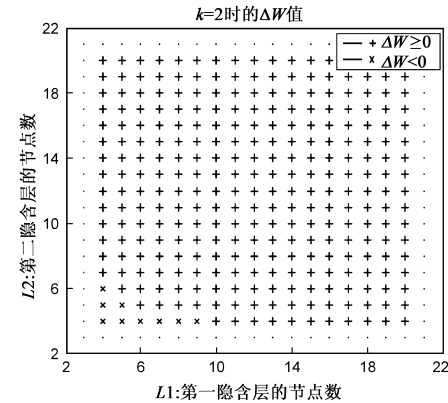


图4  $k=2$ 时 $\Delta W_{BP-T}$ 输出图

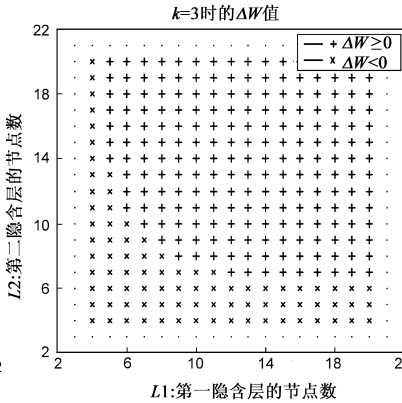


图5  $k=3$ 时 $\Delta W_{BP-T}$ 输出图

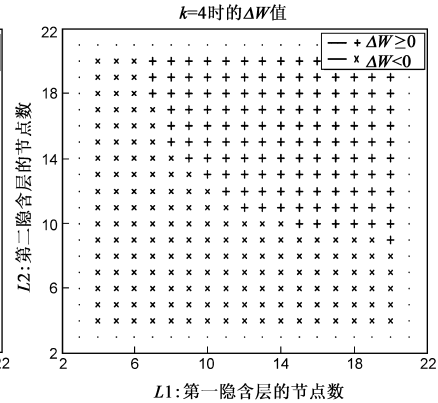


图6  $k=4$ 时 $\Delta W_{BP-T}$ 输出图

根据图 4~6 的统计数据可知:

- (1)当训练样本被分为 4 组时,二个隐含层节点数都不小于 11 的高维 BP 网络的自由参数总数比其等效分组紧密级联 BP 网络模型的自由参数总数要大;
- (2)当训练样本被分为 3 组时,二个隐含层节点数都不小于 8 的高维 BP 网络的自由参数总数比其等效分组紧密级联 BP 网络模型的自由参数总数要大;
- (3)当训练样本被分为 2 组时,二个隐含层节点数都不小于 5 的高维 BP 网络的自由参数总数比其等效分组紧密级联 BP 网络模型的自由参数总数要大.

因此有以下结论:

**结论 2** 在允许同样训练误差的情况下,若高维 BP 网络的二个隐含层节点数都大于 5,在训练样本被分为 2 组的情况下,那么 BP 网络训练样本数量要大于其等效的紧密级联 BP 网络模型的训练样本数量.

6 仿真试验比对

为了比较高维 BP 网络与其等效的分组紧密级联 BP 网络在允许同样训练误差情况下所需训练样本的数量,本文对这 2 种网络提供了相同的训练样本和相同的训练策略,然后比较网络的分类识别性能以比对所需训练样本的大小.

以一个包含有 8 台主机局域网的网络安全性评估为例,构建 BP(64,8,8,2)网络和 TCBP(2,32,32,8,8,2)网络分别对局域网的安全性进行评估.通过网络攻击分别获取每台主机的最高读权限、最高写权限和使主机拒绝服务的最大程度值,并以此作为神经网络的输入;(0,0)、(0,1)、(1,0)和(1,1)作为神经网络的输出

当训练样本被分为 2 组,即  $k=2$ ,并且  $L_1$  和  $L_2$  在 4~20 之间取值时, $W_{BP}$ 和  $W_T$  的差  $W_{BP-T}$ 的值见图 4.

当训练样本被分为 3 组,即  $k=3$ ,并且  $L_1$  和  $L_2$  在 4~20 之间取值时, $W_{BP}$ 和  $W_T$  的差  $W_{BP-T}$ 的值见图 5.

当训练样本被分为 4 组,即  $k=4$ ,并且  $L_1$  和  $L_2$  在 4~20 之间取值时, $W_{BP}$ 和  $W_T$  的差  $W_{BP-T}$ 的值见图 6.

值,分别表示安全、基本安全、不安全和很不安全.训练样本可以通过网络攻击测试、以往经验值或仿真等方法获得,获取 1500 个训练样本作为训练样本集,均分成 15 组.

神经网络的训练和测试方案为:前 10 组数据作为训练样本集对 2 种网络进行训练,然后用测试集对训练好的网络进行测试,分别记下 2 种网络的识别率;前 11 组数据作为训练样本集对 2 种网络进行训练,然后用测试集对训练好的网络进行测试,分别记下 2 种网络的识别率;依次进行训练,直到所有的组作为训练集对网络训练完为止.2 种网络对于网络安全性的分类评估正确识别性能比较见表 1.

表 1 BP 网络和 TCBP 网络正确识别结果的比较

训练集组数	BP 网络				TCBP 网络			
	(0,0)	(0,1)	(1,0)	(1,1)	(0,0)	(0,1)	(1,0)	(1,1)
11	23/30	25/30	20/25	11/15	25/30	26/30	22/25	13/15
12	23/29	25/31	21/25	12/15	26/29	28/31	22/25	13/15
13	24/29	27/31	22/26	11/14	26/29	29/31	24/26	13/14
14	25/30	26/30	23/25	12/15	27/30	28/30	23/25	14/15
15	26/30	27/30	23/25	12/15	28/30	29/30	24/25	15/15

从仿真试验结果可以看出,在同样的训练样本集情况下,TCBP(2,32,32,8,8,2)网络对于网络安全性评估的正确识别率稍大于 BP(64,8,8,2)网络的正确识别率,也就是说,在有限的、高维训练样本空间情况下,TCBP 网络的模式识别和预测能力要优于 BP 神经网络,在具体实际应用中可予优先考虑.

7 结论

分组级联 BP 网络模型在现实世界中也有很好的

物理意义:为解决一个大的、复杂的问题,通常把问题分解为递阶层次结构模型,小而简单的问题作为层次结构模型的底层,在解决完底层问题基础上,再逐层往上,直至解决目标问题.分组级联 BP 网络模型的第一级 BP 网络就如递阶层次结构模型的底层或大系统中的子系统,在完成对上一级 BP 网络训练或预测的基础上,再对下一级 BP 网络进行训练或预测,然后逐级完成对目标问题的求解.

对于高维的 BP 网络,把高维训练样本分割成多组低维训练样本,构建其分组级联 BP 网络模型,可以在有限小样本条件下较高的分类判别正确率,尽管理论上可以进行多级分组级联,但实验结果表明,BP 网络分类性能的改善在第二级最为显著,而在第三级已经大大减小,这是因为:对于松弛级联 BP 网络来说,在实际训练和预测过程中,松弛级联模型中的各个独立 BP 网络不可能达到完全的无偏估计,都存在一定的输出误差,而且这种误差随着级联层数的增加而逐渐被放大;对于紧密级联 BP 网络来说,随着级联级数的增加,网络中自由参数的总数也将有所增加,与 BP 网络相比已显示不出很大的优势.因此,分组级联 BP 网络模型一般应以级联二级或三级为限.

#### 参考文献:

- [1] Shen zhi-he, Liu fen. Applying improved BP neural network in underwater targets recognition[A]. International Joint Conference on Neural Networks[C]. Vancouver, BC, Canada: IEEE, 2006. 2588 – 2592.
- [2] Hua Zong, Yun Zhang. A real-time recognition of working patterns to fault diagnosis based on BP neural network research institute of electronic engineering [A]. Proceedings of the 6th World Congress on Intelligent Control and Automation[C]. Dalian, China: IEEE, 2006. 5769 – 5772.
- [3] Wang Jia-yang. Comprehensive evaluation method of the effect and the accomplishment of the enterprise according to BP neural network[J]. Journal of Chengdu Information Engineering College, 2004(1): 117 – 121.
- [4] 曾勇虎,王雪松,肖顺平,庄钊文.基于瞬态极化时频分布及奇异值特征提取的雷达目标识别[J].电子学报,2005, 33(3): 571 – 573.  
Zeng Yong-hu, Wang Xue-song, Xiao Shun-ping, Zhuang Zhao-wen. Radar target recognition based on instantaneous polarization time2 frequency distribution and singular value s feature extraction[J]. Acta Electronica Sinica, 2005, 33(3): 571 – 573. (in Chinese)
- [5] 高阳,胡景凯,王本年,王冬黎.基于 CMAC 网络强化学习的电梯群控调度[J].电子学报,2007, 35(2): 362 – 365.

- Gao Yang, Hu Jing-kai, Wang Ben-nian, Wang Dong-li. Elevator group control using reinforcement learning with CMAC[J]. Acta Electronica Sinica, 2007, 35(2): 362 – 365. (in Chinese)
- [6] 周开利,康耀红.神经网络模型及其 MATLAB 仿真程序设计[M].北京:清华大学出版社,2005.  
Zhou K L, Kang Y H. Neural Network Model and Its MTLAB Emulator Project[M]. Beijing: Tsinghua University Publishing House, 2005. (in Chinese)
  - [7] 葛哲学,孙志强.神经网络理论与 MATLAB 实现[M].北京:电子工业出版社,2007.
  - [8] Simon Haykin. 神经网络原理[M].叶世伟,史忠植,译.北京:机械工业出版社,2006.
  - [9] Simon Haykin. Neural Networks: A Comprehensive Foundation [M]. 2nd Edition. USA: Pearson Education, 1999.

#### 作者简介:



**鲁智勇** 男,1969 年生于河南,国防科技大学博士研究生,高级工程师,主要研究方向为电子靶场武器系统对抗、网络安全与评估.  
E-mail: lzy\_nudt@yahoo.com



**张 权** 男,1974 年生于上海,国防科技大学副教授,博士,硕士生导师,主要研究方向为信息安全、量子通信.



**张 希** 男,1984 年生于山东,国防科技大学博士研究生,主要研究方向为通信网络安全.



**唐朝京** 男,1962 年生于江苏,国防科技大学教授,博士生导师,电子科学与工程学院院长,“十一五”863-701 专题专家,中国通信学会理事,湖南省通信学会副理事长,主要研究方向为通信网信息安全与对抗.

