一种新的前向神经网络部件冗余容错方法

许荔秦, 胡东成
(清华大学自动化系, 北京 100084)

摘 要: 多层前向神经网络(MLP)的容错性有两种主要的研究方法: 改进算法和部件冗余. 前一种方法需要耗用大量的学习时间, 对大型网络是不适用的. Phatak 曾提出了用后一种方法进行 MLP 的单故障容错的一种网络结构, 但是冗余部件数庞大, 尤其对于大型网络. 本文提出了一种新的冗余体系结构, 针对单隐层 MLP 的单故障容错问题. 这种体系结构充分考虑了不同权值的不同重要度, 解决了原体系结构的仅值瓶颈问题, 可以显著减少冗余部件数, 尤其对于大型网络, 更具有优越性.

关键词: 多层前向神经网络; 容错; 冗余; 单故障
中图分类号: TP18 文献标识码: A 文章编号: 0372-2112 (2000) 05-0099-03

A New Component Redundancy Method in the Fault Tolerance of MLP

XU Li-qin, HU Dong-cheng
(*Department of Automation, Tsinghua University, Beijing* 100084, *China*)

Abstract: There are 2 main methods in the research of fault tolerance of Multilayer Perceptrons (MLP): improvement in the learning algorithm and components redundancy. Phatak presented a network architecture of components redundancy to achieve the fault tolerance of the network, but the number of redundant components is too large, especially to the large network. A new architecture is presented in this paper, especially to the single fault tolerance of the single-hidden-layer MLP. In this architecture, the different importance of the weights is considered, thus the weight bottleneck problem is solved, and the number of redundant components is substantially reduced. So there's great superiority in this architecture, especially to the large network.

Key words: multilayer perceptrons; fault tolerance; redundancy; single fault

1 引言

多层前向神经网络(MLP)的容错性在近几年来已得到了很大的重视, 已有不少文献做过这方面的研究[1~6]. 从研究方法上来看, 提高 MLP 的容错性主要有两种方法: 对 MLP 进行结构冗余, 改进 MLP 的学习算法.

文献[1~3]都是对传统的学习算法进行改进来进行容错. 这种方法可以在不改变网络结构的条件下改善网络的容错性, 缺点是需要牺牲大量的学习时间. 对权值进行容错的效果并不好. 而通过对 MLP 结构进行冗余处理提高容错性则可以避免这样的限制. 文献[4]及[5]即用这种方法实现了对权值和神经元单元单故障的完全容错, 但这种方法也有其缺点, 所得网络结构非常庞大. 本文针对文献[4]的这一点不足之处提出了一种新的冗余结构, 这种冗余结构考虑了 BP 学习得到的各权值的不同重要度, 最终得到的冗余网络比文献[4]所得网络规模要小.

2 研究对象与故障模型

本文研究网络为模式识别中应用广泛的单隐层前向网络, 输出层神经元作用函数为硬限幅函数, 隐层神经元作用函数为 Sigmoid. 用 $W_{ij}^{(n)}$ 表示连接第 $n-1$ 层第 j 个神经元与第 n 层第 i 个神经元的权值, 用 $bias_i^{(n)}$ 表示第 n 层第 i 个神经元的偏置, 用 $O_i^{(n)}$ 表示第 n 层第 i 个神经元的输出, 用 I_i 表示第 i 个输出层神经元的净输入, 则有:

$$I_i = \sum_{j=1}^{N_1} (W_{ij}^{(2)} O_j^{(1)} - bias_i^{(2)}) \quad (1)$$

$$O_i^{(2)} = \mathrm{sgn}(I_i) \quad (2)$$

其中 N_1 为隐层神经元数目, $\mathrm{sgn}(\cdot)$ 为硬限幅函数.

MLP 中最常见的故障为权值连接断路与神经元损坏, 本文仅针对这两种故障进行讨论. 分别用 $W_{ij}^{(n)}$ stuck-at-0 及 $bias_i^{(n)}$ stuck-at-$B(-B)$ 来表示这两种故障, 其中 B 为一个很大的数, ($B > NW$, N 为神经元最大扇入系数, W 为权值 $W_{ij}^{(n)}$ 所能达到的最大值). 同时, 由于输入层与输出层单元故障为非常严重的故障, 会严重影响网络输出, 因此不在本文研究范围内. 本文所设定的故障模型为单故障模型, 即整个网络在运行中最多只发生一个故障, 包括权值故障与隐层神经元 bias 故障.

收稿日期: 1999-01-22; 修订日期: 1999-08-05
基金项目: 国家自然科学基金资助课题

本文只对单故障进行考虑的原因在于: 对于一个较为可靠的系统(系统中任一部件发生故障的概率 P_f 较小, 可认为 < 0.2), 单故障在故障发生中占绝大多数. 为简单计, 不妨设各部件发生故障的随机变量独立同分布, 概率都为 P, 总部件数为 N, 则发生故障的概率 $P_f = 1-(1-P)^N$, 而发生单故障的概率为 $P_{single} = NP(1-P)^{N-1}$, 我们对 P_{single}/P_f 进行计算, 即可得单故障在故障总数中所占比例, 对 P 取值 0.1~0.00001, 对 N 取值 10~10000, 所得结果如图 1 所示. 由图可知, $P_f < 0.2$ 时, 单故障在总故障中所占比例> 90%, 传统上的神经网络是比较可靠的, 因此本文对于单故障的分析还是有较大可信度的.

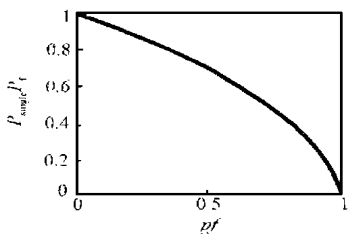

图 1 不同故障发生率下单故障点总故障中的比例

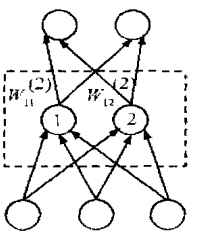

图 2 原网络

3 冗余容错方法

文献[4]的冗余容错体系结构如下:

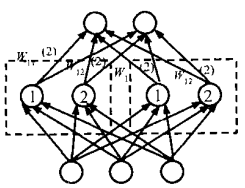

图 3 对隐层单元整组冗余的网络($m=1$)

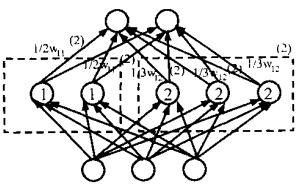

图 4 改进的冗余结构, 对不同隐层单元做不同的冗余处理

对原网络隐层单元整组冗余, 冗余 m 组. 各权值连接 $W_{ij}^{(n)}$ 的值不变, 隐层单元 $bias_i^{(1)}$ 不变, 输出层单元 $bias_i^{(2)}$ 变为原来的 $m+1$ 倍.

完全容错的 m 值按如下步骤确定: 对所有输入输出模式注入所有的单故障, 分别对每个输出节点按下式计算 m 值(为书写简便起见, 将第 j 个输出节点净输入 I_i 记为 I):

$$m = (I_d - I_f)/I_s \tag{3}$$

其中 I_d 为对该节点冗余后期望的净输入, I_f 为该节点发生该单故障时原网络的净输入, I_s 为该节点未发生故障时原网络的净输入. 对所有 m 值取其中最大的一个作为最终 m 值. 若要对网络进行完全容错, I_d 应选取与 I_s 同号的数.

上述方法能对所有的单故障容错, 但实际实现中所需冗余数非常大, 尤其对于一些大的网络来说. 主要原因在于, 传统的 MLP 学习算法得到的权值不是均衡分布的, 而上述方法对所有的权值与节点一并考虑, 导致了很大的浪费.

对于上述缺点, 一个解决办法就是对不同的隐层节点做不同的冗余处理(不同冗余数), 这样就可以避免冗余处理中个别权值的瓶颈效应. 其冗余结构如图 4 所示.

对原网络各隐层单元分别冗余, 第 i 个陷层单元冗余 m_i 组. 隐层单元组包括与某个隐层神经元相连的所有权值(其中包括与输出层相连的权值以及与输入层相连的权值)以及该神经元的 bias, 这样, 在冗余之后此陷层单元组中即包含 m_i+1 个隐层神经元. 该隐层单元组内与输出层相连的各权值连接为原值 $W_{ij}^{(2)}$ 的 $\frac{1}{m_i+1}$, 与输入层相连的各权值连接不变, 隐层单元 $bias_i^{(1)}$ 不变, 输出层单元 $bias_i^{(2)}$ 不变.

若要对一个隐层单元组 i 内的单故障完全冗余, 则对所有输入输出模式注入所有该隐层单元组内的单故障, 对每一个单故障 f 都分别对每个输出节点 j 按下式计算 $m_i(f,j)$, 取其中最大的一个作为最终的 m_i:

$$m_i(f,j) = (I_d - I_f)/(I_s - I_d)$$
$$m_i = \max_{f,j}\{m_i(f,j)\} \tag{4}$$

式(4)与式(3)中 I_d 与 I_s 同号, 而与 I_f 符号相反. 对照式(4)与式(3)可以发现, 在 $I_d = 0$ 时, 式(4)与式(3)同时变为

$$m_i = -I_f/I_s \tag{5}$$

即式(3)中 m 的取值为式(4)所得各值中的最大值, 而对输出层为硬限幅函数的神经元来说, 取 $I_d = 0$ 可以达到对单故障的完全冗余, 因此是完全可行的. 这表明, 经改进后的冗余结构可以减小冗余部件, 至少冗余部件数不会大于原冗余结构, 而对于某些权值很不均衡的网络来说, 则可以大幅度减少冗余部件.

4 实例验证

为验证以上结论, 用一个实例来作验证.

下面的例子是一个实际应用的例子, 采用 MLP 进行数码识别,'0',...,'9' 十个数字采用七段数码管表示, 每段用三点表示, 因此, 输入向量为 21 维, 输出 10 维, 为识别结果. 采用带惯性量的 BP 算法, 分别采用 7, 14, 21 个隐层神经元进行学习, 对所得网络分别用文献[4]中的冗余结构和本文中改进的

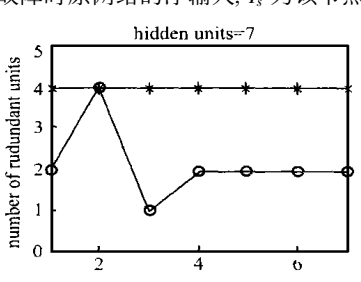

图 5 改进后, 冗余单元数为原来的 61%

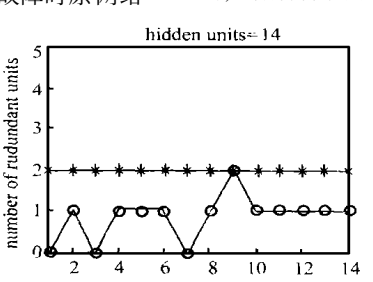

图 6 改进后, 冗余单元数为原来的 46%

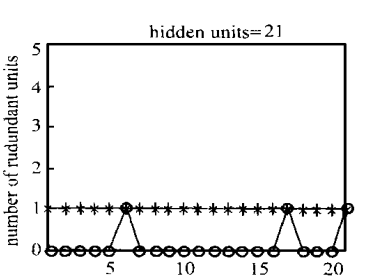

图 7 改进后, 冗余单元数为原来的 24%.

冗余结构做单故障的完全容错冗余处理,结果如下(以下图中横坐标表示隐层神经元号,纵坐标表示所需冗余单元数,图中“*”表示文献[4]中冗余方法,“°”表示本文中改进的冗余方法).

由上图可看出,当网络规模增大时,BP 算法所得权值会很不均衡,因此,采用改进的冗余容错方法有较为明显的效果,最多时,改进的冗余结构只需要原冗余方法 24% 的冗余单元即可达到同样的冗余容错效果.

5　结论

从上文中的分析可以看出,本文提出的改进的冗余结构针对不同的隐层单元做不同的冗余处理,减小了冗余后网络的规模.

通过两个实例分析,证明了这一点.在网络规模较大时,改进的冗余体系结构具有更大的优点.

参考文献

[1] Alan F. Murray and Peter J. Edwards. Enhanced MLP performance and fault tolerance resulting from synaptic weight noise during training. IEEE Trans. on Neural Networks, 1994, 5(5): 792~802

[2] Philippe Kerlinzin and Philippe Refregier. Theoretical investigation of the robustness of multilayer perceptrons: Analysis of the linear case and extension to nonlinear networks. IEEE Trans. on Neural Networks, 1995, 6(3): 560~571

[3] Behnam S. Arad and Ahmed El-Amway. On Fault Tolerant Training of Feedforward Neural Networks. Neural Networks, 1997, 10(3): 539~557

[4] D. S. Phatak and I. Koren. Complete and partial fault tolerance of feedforward neural nets. IEEE Trans. on Neural Networks, 1995, 6(2): 446~456

[5] D. S. Phatak and I. Koren. Fault Tolerance of Feedforward Neural Nets for Classification Tasks. IJCNN, 1992: II-386~391

[6] 张涛,胡东成. 一种神经网络控制器的容错性设计与可靠性评估方法. 电子学报, 1999, 27(2): 4~7

许荔秦　1973 年生,1996 年获清华大学自动化系自动控制专业学士学位,现为自动化系检测技术与装置专业直读博士研究生.现主要研究方向为人工神经网络的鲁棒性分析与容错性设计.

胡东成　1946 年生,1970 年毕业于清华大学电机工程系,留校于自动化系任教.1983 年至 1985 年由国家公派赴西德学习,后又数次出国访问或合作研究.现为清华大学副校长、教授、博士生导师,中国自动化学会理事兼教育工作委员会主任委员,中国电子学会高级会员,国家教委电子课程教学指导委员.长期从事电子与自动化方面的教学与科研工作,主要研究方向为自动测试、故障诊断与可靠性.