

# 容噪学习机制及其在 Robocup 中的应用研究<sup>\*</sup>

承文俊 沈建强 谢 琪 陈兆乾

(南京大学计算机软件新技术国家重点实验室 南京210093)

**摘 要** 噪声对机器学习影响极大, 不容忽视, 因此提出一种有效的容噪机制是当前研究的热点问题。本文研究了 BP 神经网络在不同噪声环境下的容噪能力, 利用容噪 BP 神经网络分析了机器人足球中的噪声问题及其解决方案, 并将这种 BP 网络应用到了我们的仿真球队 NDSocTeam 的截球动作上, 其效果良好。

**关键词** 容噪, 机器学习, RoboCup, Agent

## Research on Noise Tolerance Mechanism in RoboCup

CHENG Wen-Jun SHEN Jian-Qiang XIE Qi CHEN Zhao-Qian

(State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093)

**Abstract** Noise is an important issue in machine learning and it will bring great effect on machine learning, so to develop an effective noise tolerance mechanism is a popular topic in current research. This paper studies the noise tolerance capability of BP Neural Networks under different noisy circumstances, analyzes the noise problem in Robocup and solves the noise in Robocup with noise tolerance BP Neural Networks. It also describes the implementation of ball interception in NDSocTeam.

**Keywords** Noise, Machine learning, RoboCup, Agent

## 1. 引言

传统的机器学习机制研究前提侧重于理想环境, 即: 在学习过程前, 环境是精确的, 易于描述的; 在学习过程中, 训练数据是完整的, 所含噪声极少, 甚至没有。但现实世界并非如此, 往往专家知识是模糊的, 难以描述的; 训练数据缺少某些属性; 训练数据受噪声干扰, 出现畸变。这导致学习算法在实际应用中效率较为低下, 即学习的过程极为耗时, 而结果却不是非常可靠, 所以目前容噪学习机制是机器学习领域研究的热点问题。

早在1992年, Alan Mackworth 就提出机器人足球比赛是很好的机器人和 AI 研究的实验平台<sup>[1]</sup>。在一些学者的积极倡导下(如 CMU 的 Manuela M. Veloso 教授), Sony 公司的支持下, 成立了 RoboCup 联合会。经过几年的研究, 于1997年8月23~29日在日本举行了第一届 RoboCup 比赛及会议, 此后每年都举行一次 RoboCup 比赛。RoboCup 是典型的多 Agent 问题, 多 Agent 系统固有的复杂性, 如状态空间与行为空间十分巨大, 利用机器学习技术来解决这一复杂性有很多益处<sup>[1]</sup>。机器学习是多 Agent 系统研究的关键技术, 许多现有的机器学习技术都可以直接应用到多 Agent 系统中去, 有助于解决多 Agent 系统的固有复杂性。而 RoboCup 领域具有复杂、动态、实时、不确定和含有大量噪声的特点, 是容噪机器学习技术很好的实验平台。

本文主要以此为平台, 研究噪声条件下的机器学习, 设计和实现了一种容噪的 BP 网络, 并应用于我们开发的仿真足

球队系统 NDSocTeam (Nanjing University Robot World Cup) 中, 其效果很好。

## 2. 噪声条件下的机器学习

### 2.1 噪声模型

目前, 一般的机器学习算法和模型都假设数据是无误的, 较准确的, 而现实世界并非如此, 会产生各种噪声, 使其数据有所偏差。机器学习中噪声可以划分为如下四种:

1) 恶意噪声(MALV): 当获取数据时, 数据正确的概率为  $1-V$ ; 数据存在错误的概率为  $V$ , 而且是无法预知的错误<sup>[2]</sup>。该模型 MALV (malicious noise) 用来描述数据基本正确, 但其中有一部分( $V$ )被噪声干扰的情况。

2) 随机误分类噪声(RMCV): 获得正确数据( $X, S$ )的概率为  $1-V$ ; 存在概率  $V$ , 我们获得数据( $X, \hat{S}$ )<sup>[3]</sup>。该模型 RMCV (random misclassification noise) 用来描述数据基本正确, 但其中有部分( $V$ )的分类标记被逆反了。

3) 统一随机属性噪声(URAV): 设实例空间为  $\{0, 1\}^n$ , 数据形如  $(a_1, \dots, a_n, s)$ , 然后以概率  $V$  将每个  $a_i$  逆转为  $\bar{a}_i$ , 但其分类标记未改变<sup>[5]</sup>。显然, 该模型 URAV (uniform random attributes) 是一种较为良性的噪声模型。

4) 产生随机属性噪声(PRAV): 设实例空间为  $\{0, 1\}^n$ , 数据形如  $(a_1, \dots, a_n, s)$ , 然后, 对于每个  $1 \leq i \leq n$ , 存在概率  $V_i \leq V$  将  $a_i$  逆转为  $\bar{a}_i$ <sup>[10]</sup>。该模型 PRAV (product random attribute) 比 URAV 更常见, 且更加难以处理。

### 2.2 噪声对 PAC 学习的影响

<sup>\*</sup> 基金项目: 国家自然科学基金资助项目(60273033), 江苏省自然科学基金(DK2002081)。承文俊 硕士研究生, 主要研究方向为机器学习。沈建强 硕士研究生, 主要研究方向为机器学习。谢 琪 高工, 主要研究方向为智能应用。陈兆乾 教授, 博士生导师, 主要研究领域为专家系统、机器学习、神经网络。

2.2.1 PAC 模型 即可能近似正确模型是由 Valiant<sup>[2]</sup>首先提出来的,这是研究从实例中进行概念学习的精确模型,包括对学习效果如何定义。

在 PAC 模型中,学习器将目标概念的例子分类成正例和反例,在  $X^n$  上根据某个未知目标分类  $D$  随机提取例子。假设学习器有一个黑箱,称为 EX,每个 EX 返回一个分类实例,即例子  $(x, c(x))$ ,此处  $x$  是从  $D$  中随机产生的,根据  $c$  来分类。我们称收集的例子为样本。学习器给定输入为  $0 < \epsilon, \delta < 1$ ,那么我们定义可能近似正确模型如下:

可能近似正确模型(PAC):从  $D$  上随机产生的例子上,学习器的目标是输出一个假设  $h$ ,  $h$  与  $c$  的差异小于  $\epsilon$  的概率至少为  $1 - \delta$ ,即不一致的可能性最多为  $\epsilon$ (准确性至少有  $1 - \epsilon$ ,或者是  $\epsilon$ -good)。我们用  $\Delta$  表示两个概念间的区别,这样,以上可以定义为:学习器的目标是输出某个概念  $\hat{c}$ ,

$$P[D(c\Delta\hat{c}) > \epsilon] \leq \delta$$

如果学习算法存在(适合任何  $n \geq 1$ ,任何目标概念  $c \in C$ ,任何目标分类  $D$ ,任何  $\epsilon, \delta$  目标的算法),这个算法在多项式时间内完成,那么我们称该算法可 PAC 学习,称概念  $C$  是 PAC 可学习的。

2.2.2 噪声对 PAC 学习的影响 当前用于 PAC 学习最通用的方法是:选择一个概念,使这个概念对噪声数据样本具有最好(或者很好)的一致性,我们称之为最小化误差。

比较容易解决的情况是数据样本中,噪声的发生概率比较一致。在 Angluin<sup>[3]</sup>和 Laird<sup>[4]</sup>分别研究了离散和连续的实例空间中的噪声情况,证明当实例被随机误分类噪声修改时,最小化误差的方法符合 PAC 标准。Solon<sup>[5]</sup>将这些结论扩展到恶意分类噪声。Kearns 和 Li<sup>[6]</sup>证明这种最小化误差方法在离散的实例空间中能容忍少量恶意噪声。Blumer 等<sup>[7]</sup>提出的 VC 维方法能证明这种最小化误差的方法也能在连续实例空间中处理少量恶意噪声。在统一随机属性噪声情况下,如果使用最小化误差方法,那么获得的最小错误率将是噪声率的下界<sup>[5]</sup>。而任意的对抗性恶意噪声,是任何算法能容忍的最大噪声率<sup>[6]</sup>。

尽管最小化误差的方法对产生随机属性噪声无效,但已

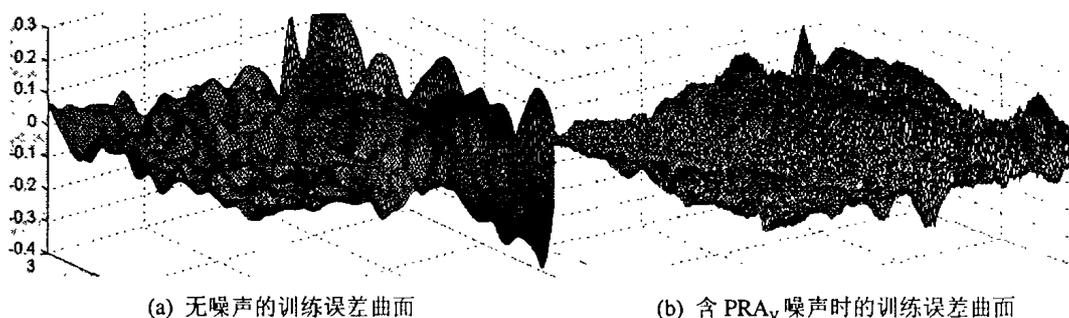


图1 有无噪声的训练误差对比

2)PRA<sub>v</sub> 模型要比 URA<sub>v</sub> 模型难处理,BP 神经网络在 PRA<sub>v</sub> 模型下性能有所恶化,较小的噪声率都会引起网络性能的大幅下降。隐层节点为10个,部分实验数据如表1,设  $X$  分量噪声率为  $V_x$ ,  $Y$  分量噪声率为  $V_y$ ,令  $V_{mn} = \min(V_x, V_y)$ 。在网络的承受范围内,噪声率  $V_x \neq V_y$  时,网络的学习效果比  $X, Y$  噪声率都为  $V_{mn}$  要差。

我们是这样理解有噪声的训练集能增强 BP 网络的泛化能力的:如果网络接受的总是单一训练集中的数据,会导致网络最终记住给定的模式;给训练集中的数据注入噪声能避免

有技术处理统一随机属性噪声。Shackelford 和 Volper<sup>[8]</sup>提出了能容忍大量随机属性噪声的算法,可用于学习  $k$ -DNF 公式。这个算法需要给出精确的噪声率(至少有很好的噪声估计率值)。Littlestone<sup>[9]</sup>提出一种算法,能解决几种不同模型的属性噪声。

到目前为止,还没有非常有效的算法处理恶意噪声和产生随机属性噪声,这也是当前的国际研究热点。

### 2.3 噪声对 BP 神经网络的影响

BP 网络算法是使用较广泛的 PAC 学习算法,1992年, J. I. Minnix<sup>[11]</sup>发现用噪声数据对 BP 网络进行训练可以增强容错能力。1994年, A. F. Murray 和 P. J. Edwards<sup>[12]</sup>分析了在多层网络训练过程中给连接权加入模拟噪音对容错性产生的影响。上述工作在当时也被认为是对容错性的研究,但由于它们主要解决的是神经网络在噪声环境下的学习,实际上是对网络的容噪性(noise-tolerance)进行研究。他们证明了神经网络的一个重要特性,即网络有较好的容噪性能。

上述研究主要针对的是输入向量中,各个分量噪声率一致的情况,即主要是研究统一随机属性噪声(URA<sub>v</sub>)模型下的容噪性能,对产生随机属性噪声(PRA<sub>v</sub>)基本没有考虑。本文使用函数模拟的方法测试了 PRA<sub>v</sub> 模型下 BP 网的容噪性。考虑到实际应用中,各个输入分量间的重要性不同,本文使用非线性、非对称函数  $\sin(X^2 + Y)$  作为被模拟对象,变量范围为  $0 \sim 3$ ,变量步长设为  $0.02$ 。我们设计的实验用 BP 网络为两输入、单输出的三层结构,各神经元间的传递函数分别为 tansig 函数和 purelin 函数,网络的反馈函数为 trainlm 函数,采用学习率自适应调整策略,训练终止条件为  $0.01$ 。噪声信号采用加入高斯噪声的  $\sin$  函数,噪声率由  $\sin$  的频率控制。我们作了大量实验,总结出以下结论:

1) 适量的 PRA<sub>v</sub> 噪声能改善 BP 神经网络的性能,降低其误差,使网络更好地模拟真实情况。使用10个隐层节点,无噪声的训练误差如图1(a),  $X$  的噪声率均为  $\pi/2$ ,  $Y$  的噪声率为  $\pi$  时的训练误差如图1(b),很明显,在有噪声的情况下,网络的误差较小、误差曲面平滑,网络的性能得到了改善。

这个问题,这样输入空间的单个点就成为围绕这个点的一个区域了,它的大小与噪声大小有关,加入的噪声越大这个区域就越大。

## 3. 容噪 BP 网络的设计及应用

### 3.1 Robocup 中的噪声

RoboCup 比赛是以 client/server 方式进行的,每个队员 client 都是独立的进程,通过给定的端口和 server 连接。球员发送他们下一步要做的动作请求给 server, server 接收到这

些消息后,执行请求,并相应地更新环境。另外,server 向所有的球员提供感知信息。

表1 输入分量噪声率不同时部分实验对比数据

X分量 噪声率	Y分量 噪声率	最大绝对 对误差	最小绝对 对误差	标准差	相关系数
Pi	Pi/2.	0.5791	2.6577e-007	0.0872	0.9920
	Pi	0.3611	9.0877e-007	0.0615	0.9956
	3Pi/2	0.5788	1.2370e-005	0.0637	0.9938
	2Pi	0.2413	5.7857e-006	0.0686	0.9946
3Pi/2	Pi/2	0.3452	5.4393e-006	0.0806	0.9943
	Pi	0.9944	3.0952e-006	0.0849	0.9924
	3Pi/2	0.3415	1.1054e-006	0.0633	0.9960
	2Pi	0.4042	9.7019e-007	0.0867	0.9921

为了反映出实际比赛中球及球员运动的不确定性,server 在球及球员的运动过程中以及一些命令参数中加入了一定的干扰因素,因此每个 Agent 不能准确地感知环境,同时也不能精确地改变环境。运动中,干扰是以如下方式加入的:

$$(u_x^{t+1}, u_y^{t+1}) = (v_x^t, v_y^t) + (a_x^t, a_y^t) + (\bar{r}_{r_{max}}, \bar{r}_{r_{max}})$$

其中,  $\bar{r}_{r_{max}}$  为属于  $[-max, max]$  的随机数,  $r_{max}$  的定义为:  $r_{max} = rand \cdot |(v_x^t, v_y^t)|$ ,  $rand$  是  $player-rand$  和  $ball-rand$  的参数。

命令中的 Moment 和 Power 参数的干扰为(用 argument 表示):

$$argument = (1 + \bar{r}_{rand}) \cdot argument$$

可以看出,RoboCup 中参数的噪声率会随时间出现变化,属于 PRA<sub>v</sub> 模型。有效的控球是 RoboCup 获胜的关键,为了达到控球的目的,必须有高效的截球技术,即将队友或对手传出的球控制住。由于训练数据有很大的噪声,而 BP 神经网络有较好的容噪性能,因此本文采用 BP 神经网络来进行截球训练。

### 3.2 容噪的 BP 网络

在 RoboCup 环境下,服务器随机地对环境内运动的物体产生一定的干扰,运动物体的运动路线会产生一定的偏差,而且每个球员的视野都只有一定的范围,SoccerServer 环境中定义的球员的视野正常状态下是 90°。所以当一队员准备截球的时候,直接计算最佳截球点并保证在截球点准确截球是一个很困难的过程。BP 神经网络具有很好的容噪性,我们采用基于多层神经网络的 BP 学习算法来训练截球,并将这个技术应用到我们的仿真球队 NDSocTeam 上。

我们对截球速度进行离散化处理,将其分为 5 个类别:低速、慢速、中速、快速、高速,每个速度档分别用不同的神经网络表示,分别训练、分别使用。使用的神经网络为 3 层单输入单输出网络,输入为踢球角度,输出为截球角度。隐含层神经元数的选取是一个十分复杂的问题,与问题域的具体要求、输入层神经元个数和输出层神经元个数都有直接的关系。一般地,隐含层神经元数目太少,网络可能无法训练出结果,或网络不“健壮”,不能识别以前没有看过的样本,容错性差;但数目过多又会使学习时间过长,而且效果也不一定最佳。因此,应该选择一个合适的隐层神经元数,使其尽可能接近于最佳值。在本系统中,我们参考了式(1)选择隐层神经元数目:

$$n_B = \sqrt{n_A + n_C + \alpha} \quad (1)$$

其中,  $n_A$  为输入神经元数,  $n_C$  为输出神经元数,  $\alpha$  为 1~10 之间的常数。

在本系统中,根据式(1)可得:

$$n_B = \sqrt{1+1+\alpha} = \sqrt{2+\alpha} \approx 3 \sim 12$$

考虑到对测试样本的容错性,经过反复的实验后,取  $n_B = 7$ 。以后的实践证明,这一隐层神经元数目的选择是合适有效的。

各神经元间的传递函数分别为 tansig 函数和 purelin 函数,网络的反馈函数为 trainlm 函数。经过神经网络在噪声环境下的训练以后,神经网络可以准确表示踢球角度与截球角度之间的映射关系;对于不同的初始速度使用不同的网络。

### 3.3 实验结果

训练后,神经网络的平均误差率不超过 10%,基本满足要求。较好地解决了环境干扰问题,球员的截球动作基本有效(目测成功率在 70%左右)。

我们选择了 2000 年第三届机器人足球 RoboCup 的冠军队,卡内基梅隆大学的 ATTCMUnited2000 作为对手,和我们自行设计并实现的 NDSocTeam 进行比赛。在实验比赛中,应用 BP 网络来进行截球训练的 NDSocTeam 误差明显减小,并且能够取得一些场次的胜利。考虑到 NDSocTeam 是一支尚处于研制阶段的球队,在对手是前世界冠军的情况下,取得这样的成绩还是令人满意的。实验显示,我们使用 BP 网络来进行截球训练,球员在球场上能够取得较好的表现。

**总结** 机器人足球领域,具有复杂、动态、实时、不确定的特点,是机器学习技术很好的实验平台,近些年来,研究者在这一领域也取得了许多成绩,但大多是在无噪声的情况下开展的研究工作。本文研究了不同噪声模型下 BP 网络的容噪性能,并分析了 RoboCup 中的噪声问题,将 BP 网络应用到我们的仿真球队 NDSocTeam 的截球动作上,取得了良好的效果,为 NDSocTeam 高层动作的实现打下了基础。

### 参考文献

- 1 Wei B G, Sen S. Adaptation and Learning in MultiAgent Systems. Springer Verlag, Berlin, 1996
- 2 Valiant L G. A theory of the learnable. Communications of the ACM, 1984, 27(11): 1134~1142
- 3 Angluin D, Laird P. Learning from noisy examples. Machine Learning, 1988, 2(4): 343~370
- 4 Laird P D. Learning from Good and Bad Data. Kluwer international series in engineering and computer science. Kluwer Academic Publishers, Boston, 1988
- 5 Sloan R H. Types of noise in data for concept learning. In: First Workshop on Computational Learning Theory. Morgan Kaufmann, 1988. 91~96
- 6 Kearns M, Li M. Learning in the presence of malicious errors. SIAM Journal on Computing, 1993, 22(4): 807~837
- 7 Blumer A, et al. Learnability and the Vapnik-Chervonenkis dimension. Journal of the ACM, 1989, 36(4): 929~965
- 8 Shackelford G, Volper D. Learning k-DNF with noise in the attributes. In: First Workshop on Computational Learning Theory, Cambridge, Mas. Morgan Kaufmann. 1988. 97~103
- 9 Littlestone N. Redundant noisy attributes, attribute errors, and linear-threshold learning using winnow. In: Fourth Workshop on Computational Learning Theory, 1991. 147~156
- 10 Goldman S A. Can PAC Learning Algorithms Tolerate Random Attribute Noise? Algorithmica, 1995, 14: 70~84
- 11 Minnix J I. Fault Tolerance of the Backpropagation Neural Network Trained on Noisy Inputs. IEEE transaction, 1992, 1: 847~852
- 12 Murray A F, Edwards P J. Enhanced MLP Performance and Fault Tolerance Resulting from Synaptic Weight Noise During Training. IEEE Transactions on Neural Networks, 1994, 5(5): 792~802
- 13 Boyan J A, Moore A W. Generalization in reinforcement learning: Safely approximating the value functions. In: Advances in Neural Information Processing Systems: Proc. of the 1994 Conf. San Mateo, CA. Morgan Kaufmann. 1995. 369~376