

复杂系统可靠控制中的潜在问题互连神经网络分析方法

胡昌华¹ 刘丙杰¹

摘要 潜在问题是影响大型复杂系统安全性、可靠性的重要因素。神经网络是一种新的潜在问题分析方法,但其分析结果难以解释。本文提出了一种基于电路结构的神经网络模型 (Neural network model based on circuit architecture, CArNN), 将 CArNN 作为个体进行集成, 形成神经网络集成用于潜在问题分析。对 CArNN 模型的鲁棒性进行了分析, 提出了两个保证模型鲁棒性的约束条件。利用此方法对一个经典电路进行了分析, 结果显示, 此方法对潜在电路的正确识别率达到 94%。

关键词 潜在问题分析, 可靠控制, 神经网络集成, 克隆选择算法, 鲁棒性分析, 泛化性能
中图分类号 TP301

Sneak Circuit Analysis Based on Novel Coadjacent Neural Network Model for Reliability Control of Complex System

HU Chang-Hua¹ LIU Bing-Jie¹

Abstract Sneak circuits have important influence on systems' safety and reliability. Artificial neural network (ANN) is a new sneak circuit analysis (SCA) method. However, it is difficult to interpret why ANN gives the conclusion. A novel coadjacent neural network model based on circuit architecture, named CArNN, is proposed. Ensemble of the CArNN was employed to analysis the sneak circuits existing in systems. Two constraints guaranteeing the robustness of CArNN were given. A typical circuit containing sneak circuit was used to verify the method. The results showed 94% of the sneak circuits can be correctly recognized by the method.

Key words Sneak circuit analysis, reliability control, neural network ensemble, clonal selection algorithm, robustness analysis, generalization performance

随着自动化系统规模的不断扩大、复杂性的不断提高, 出现了一些大型复杂的电路系统。在这些大型复杂的电路系统中, 由于系统的复杂性、管道效应、人的因素、频繁的修改以及不同技术的交叉应用, 普遍存在一些潜在问题^[1-2]。所谓潜在问题是指系统在所有元件均正常工作的情况下, 导致系统出现非期望的功能或抑制期望功能的一些潜在的电路路径或潜在的电路状态^[3]。潜在问题分析 (Sneak circuit analysis, SCA) 技术是系统地分析和发现系统中存在这类问题的一类技术, 对保证系统可靠性与安全性具有非常重要的意义。

目前, 国内外 SCA 技术主要有三类: 全面的 SCA 技术、简化的 SCA 技术和智能 SCA 技术。1967 年美国国家宇航局在阿波罗登月计划中为提高系统的安全性和可靠性而提出了潜在问题分析技术, 该方法是波音公司的规范分析方法, 它试图揭示系统中的所有功能和行为, 因而称之为“全面的 SCA 技术”^[4]。其缺点是必须拥有适应系统规模的网络树

生成算法和拓扑识别算法, 并且这种技术只能在设计完成后、产品生产之前进行, 即使发现了潜在问题, 对整个系统设计的修改代价也是十分巨大的。正是这些因素限制了这种潜在电路分析方法的大范围推广。

简化的 SCA 分为两种^[4-5]: 潜在路径分析和设计缺陷分析。潜在路径分析是搜索出“源点”和“目标点”之间的所有路径, 并对其应用两类线索以识别潜在路径。设计缺陷分析则利用元件级线索识别设计缺陷。

智能 SCA 技术是指利用智能方法进行潜在问题分析的技术, 其中最主要的是基于神经网络的 SCA^[6]。这种方法的主要思想是利用系统的相关信息形成训练样本, 然后对神经网络进行训练, 最后利用训练过的神经网络预测所有开关组合下系统实现的功能, 通过与设计功能比较找到潜在问题。这种方法克服了传统 SCA 方法的一些缺陷, 如避免了线索表获取的瓶颈等。但是, 基于神经网络的潜在问题分析也存在一些难题, 如神经网络参数设置问题、解释问题、预测结果的可靠性问题等。

本文提出一种新的神经网络模型, 即基于电路结构的神经网络 (Neural network based on circuit architecture, CArNN)。在 CArNN 中, 神经元表示电路元件, 神经元作用函数为柔性 Sigmoid 函数 (F-

收稿日期 2007-01-31 收修改稿日期 2007-06-21
Received January 31, 2007; in revised form June 21, 2007
国家自然科学基金 (60736026), 教育部新世纪优秀人才支持计划资助
Supported by National Natural Science Foundation of China (60736026), the Program for New Century Excellent Talents in University of China
1. 第二炮兵工程学院 302 教研室 西安 710025
1. Xi'an Research Institute of Hi-tech, Xi'an 710025
DOI: 10.3724/SP.J.1004.2008.00188

flexible sigmoid function, FSF), 神经元之间的连接与电路元件的连接关系相似. CArNN 的输入为电路开关状态组合, 输出为功能元件状态. 在一定的输入下, 只有一部分神经元被激活, 计算量较小, 比较符合实际情况. 由于 CArNN 是电路结构的抽象模型, 与其他网络模型相比, CArNN 的结构与电路结构相关联, 有一定的物理基础, 而且有较好的泛化能力. 由于 CArNN 是基于电路结构的网络模型, 不同的电路将产生不同的 CArNN 模型, 所以不能用传统的梯度下降等方法训练. 随机搜索优化算法对网络结构没有限制, 是一个较好的选择. 随机搜索优化算法主要包括: 遗传算法 (GA), 克隆选择算法 (CSA), 粒子群优化算法 (PSO), 蚁群算法 (ACA) 等. 这些算法及其改进算法的全局最优解求解能力和算法的收敛速度是很接近的. 克隆选择算法不仅能求解问题的全局最优解, 而且可以获得问题的局部最优解. 在神经网络集成中, 局部最优个体组成的集成模型有更好的泛化能力^[7-10]. 鉴于此, 本文采用克隆选择算法对网络训练, 训练后产生的抗体作为神经网络集成的个体, 通过平均加权预测电路实现的所有功能.

1 基于电路结构的神经网络及其集成

1.1 电路到 CArNN 模型的转换

电路系统表示为一个二元组: $G = (V, E)$, 其中 $V = (v_1, v_2, \dots, v_n)$ 为系统元件顶点集合, v_i 为第 i 个元件, n 为元件数; $E = (e_1, e_2, \dots, e_m)$ 为电路元件连接关系的集合, e_j 为第 j 个元件连接关系, m 为连接关系数. CArNN 模型可以表示为一个四元组: $C = (N, W, F, L)$, 其中 $N = (n_1, n_2, \dots, n_p)$ 为神经元节点集合, n_i 为第 i 个神经元节点, p 为神经元节点数; $W = (w_1, w_2, \dots, w_q)$ 为神经元之间的连接权值集合, w_j 为第 j 个连接权值, q 为连接权值数; $F = (f_1, f_2, \dots, f_p)$ 为神经元节点作用函数集合, f_i 为第 i 个神经元节点的作用函数; L 为神经网络的学习算法. 从 G 到 C 的转换规则为

1) $p = n, q = m$;

2) 对于任意 $v_i \in V, n_i \in N, f_i \in F, 1 \leq i \leq n$, 存在 $n_i = f_i(v_i), v_i = g_i(n_i)$.

上述转换规则的核心思想是电路元件的连接关系与神经网络的连接权值一一对应, 电路元件与神经元节点一一对应, 神经元节点的作用函数采用柔性 Sigmoid 函数

$$f(x, a, b) = \frac{1 - e^{-2ax}}{b(1 + e^{-2ax})} \quad (1)$$

不同的神经元节点具有不同的 FSF 函数.

1.2 CArNN 的数学模型

假设 CArNN 的输入为 $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$, 中间层神经元输入为

$$\mathbf{M}_I(1) = \begin{bmatrix} m_I^1(1) \\ m_I^2(1) \\ \dots \\ m_I^l(1) \end{bmatrix} = \begin{bmatrix} q_{11} & \dots & q_{1n} \\ q_{21} & \dots & q_{2n} \\ \dots & \dots & \dots \\ q_{l1} & \dots & q_{ln} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} + \begin{bmatrix} p_{11} & \dots & p_{1l} \\ p_{21} & \dots & p_{2l} \\ \dots & \dots & \dots \\ p_{l1} & \dots & p_{ll} \end{bmatrix} \begin{bmatrix} m_I^1(0) \\ m_I^2(0) \\ \dots \\ m_I^l(0) \end{bmatrix} \quad (2)$$

令

$$Q = \begin{bmatrix} q_{11} & \dots & q_{1n} \\ q_{21} & \dots & q_{2n} \\ \dots & \dots & \dots \\ q_{l1} & \dots & q_{ln} \end{bmatrix}, \quad P = \begin{bmatrix} p_{11} & \dots & p_{1l} \\ p_{21} & \dots & p_{2l} \\ \dots & \dots & \dots \\ p_{l1} & \dots & p_{ll} \end{bmatrix} \quad (3)$$

$$\mathbf{M}_I(0) = Q\mathbf{x} \quad (4)$$

则有

$$\mathbf{M}_I(1) = Q\mathbf{x} + P\mathbf{M}_I(0) = (I + P)Q\mathbf{x} \quad (5)$$

由于 CArNN 中的中间层神经元之间可能存在连接关系, 所以中间层神经元的输入包括两部分: 来自于输入层的输入和来自于其他中间层神经元的输入. P 表示中间层神经元之间的连接权值, Q 表示输入层到中间层的连接权值, $\mathbf{M}_I(0)$ 表示输入层到中间层神经元的输入, $\mathbf{M}_I(1)$ 表示中间层神经元之间的二次输入.

中间层神经元的输出为

$$\mathbf{M}_O = f(\mathbf{M}_I(1)) \quad (6)$$

其中 $f(\cdot)$ 为 FSF.

令

$$H = \begin{bmatrix} h_{11} & \dots & h_{1l} \\ h_{21} & \dots & h_{2l} \\ \dots & \dots & \dots \\ h_{l1} & \dots & h_{ll} \end{bmatrix}, \quad W = \begin{bmatrix} w_{11} & \dots & w_{1m} \\ w_{21} & \dots & w_{2m} \\ \dots & \dots & \dots \\ w_{m1} & \dots & w_{mm} \end{bmatrix} \quad (7)$$

$$\mathbf{M}_O(0) = H\mathbf{M}_I \quad (8)$$

则有:

$$\mathbf{M}_O(1) = H\mathbf{M}_I + W\mathbf{M}_O(0) = (I + W)H\mathbf{M}_I \quad (9)$$

其中 H 表示中间层到输出层的连接权值, W 表示输出层神经元之间的连接权值, $\mathbf{M}_O(0)$ 表示中间层

到输出层神经元的输入, $\mathbf{M}_O(1)$ 表示输出层神经元之间的二次输入.

输出层神经元的输出为

$$\mathbf{M}_O = f(\mathbf{M}_O(1)) \quad (10)$$

其中 $f(\cdot)$ 为 FSF.

综上所述, 可得 CArNN 模型可表示为

$$\mathbf{M}_O = f((I + W)Hf((I + P)Q\mathbf{x})) \quad (11)$$

假设训练样本为 $\mathbf{S} = (\mathbf{X}, \mathbf{Y})$, $\mathbf{X} \rightarrow \mathbf{R}^n$, $\mathbf{Y} \rightarrow \mathbf{R}^m$, 则 CArNN 训练的目标函数为

$$J(W) = \min \|\mathbf{M}_O - \mathbf{Y}\| \quad (12)$$

CArNN 的训练要使式 (12) 最小. 从式 (11) 可以看出, CArNN 的可调参数矩阵随着电路的不同而不同, 而且可能存在闭环. 与规则结构的网络模型不同, ArNN 模型是一种互连神经网络.

1.3 CArNN 鲁棒性分析

定义 1^[11]. 系统鲁棒性是指系统参数或结构发生变化时, 仍能在一定的误差范围内给出正确输出的能力. 神经网络的鲁棒性分析是指网络权值或结构发生变化时, 仍能在一定的误差范围内给出正确输出的能力. 本文假设 CArNN 的结构不发生变化, 而仅仅考虑 CArNN 的权值发生变化时系统的鲁棒性. 衡量网络鲁棒性的指标就是网络在参数发生变化时, 网络的输出与期望输出的均方差.

定理 1. 不失一般性, 这里假设系统的矩阵均为一维实数, 存在常数 $\varepsilon_w > 0$, $\varepsilon_p > 0$, $\varepsilon_q > 0$, $\varepsilon_h > 0$, 神经网络的泛化误差为 e , CArNN 模型参数在满足以下条件时具有鲁棒性, 即 CArNN 可以满足潜在问题分析需求:

1) $w \neq -1, p \neq 1, q \neq 0$;

2) $\Delta w \leq (0.5 - e)/\varepsilon_w$, $\Delta p \leq (0.5 - e)/\varepsilon_p$,

$\Delta q \leq (0.5 - e)/\varepsilon_q$, $\Delta h \leq (0.5 - e)/\varepsilon_h$.

证明. CArNN 可表示为

$$m_o = f((1 + w)h \times f((1 + p)qx)) \quad (13)$$

假设系统的期望输出为 y , 则系统的均方差为

$$e = (y - m_o)^2/2 \quad (14)$$

系统的鲁棒性可以表示为均方差对参数的偏导数. 首先考虑 e 对 w 的偏导数的范数

$$\left| \frac{\partial e}{\partial w} \right| = |m_o - y| \cdot \left| \frac{4are^{-2ar}}{b(1+w)(1+e^{-2ar})^2} \right| \quad (15)$$

其中

$$r = (1 + w)h \frac{1 - e^{-2a(1+p)qx}}{b(1 + e^{-2a(1+p)qx})} \quad (16)$$

因为 $w \neq -1, p \neq 1, q \neq 0$, 且参数有界, 即存在常数 $\delta_w > 0, \delta_p > 0, \delta_q > 0, \delta_h > 0$, 使得 $|w| \leq \delta_w, |p| \leq \delta_p, |q| \leq \delta_q, |h| \leq \delta_h$, 由式 (16) 可知, 存在 $\delta_r > 0$ 使得

$$|r| < \delta_r \quad (17)$$

式 (14) 可变为

$$\left| \frac{\partial e}{\partial w} \right| \leq |m_o - y| \cdot \left| \frac{4a}{b(1 + \delta_w)} \right| \cdot \left| \frac{e^{-2a\delta_r}\delta_r}{(1 + e^{-2a\delta_r})^2} \right| \quad (18)$$

令 $D = \left| \frac{e^{-2a\delta_r}\delta_r}{(1 + e^{-2a\delta_r})^2} \right|$, 则存在 $\delta_D > 0$ 使得 $D \leq \delta_D$, 有

$$\lim_{\delta_r \rightarrow \infty} D = 0 \quad (19)$$

因为系统参数有界, 所以 $4a/[b(1 + \delta_w)]$ 有界, 对于给定的模型 (11) 以及训练样本, 其输出误差也是有界的, 即 $|m_o - y|$ 有界. 综上所述, 可知 $\partial e/\partial w$ 有界, 即存在 $\varepsilon_w > 0$, 使得

$$\frac{\partial e}{\partial w} < \varepsilon_w \quad (20)$$

假设其他参数不变, 上式可表示为

$$\frac{\partial e}{\partial w} = \frac{\Delta e}{\Delta w} = \frac{|e - e'|}{\Delta w} < \varepsilon_w \quad (21)$$

其中 e 为 w 不变时 CArNN 的均方误差, e' 为 w 变为 $w + \Delta w$ 时 CArNN 的均方误差. 上式通过整理可得

$$e - \varepsilon_w \Delta w \leq e' \leq e + \varepsilon_w \Delta w \quad (22)$$

由文献 [6] 可知, 神经网络用于 SCA 的一个必要条件就是其泛化误差小于 0.5, 即 $e' < 0.5$, 从而可得

$$\Delta w \leq \frac{0.5 - e}{\varepsilon_w} \quad (23)$$

上式给出了在神经网络的鲁棒性满足要求的条件下, 参数 w 的变化范围. 同理可得 e 对其他参数的偏导数的范数均有界, 即存在 $\varepsilon_p > 0, \varepsilon_q > 0, \varepsilon_h > 0$, 使得

$$\Delta p \leq \frac{0.5 - e}{\varepsilon_p}, \Delta q \leq \frac{0.5 - e}{\varepsilon_q}, \Delta h \leq \frac{0.5 - e}{\varepsilon_h} \quad (24)$$

同理, 可将系统推广到多维系统, 鲁棒性定理成立. \square

从定理 1 可以看出, $e = 0.5$ 是 CArNN 的一个鲁棒性临界点, 此时任意参数发生微小变化都会引起模型不稳定. 所以, 为了提高模型的稳定性就要尽可能减小泛化误差.

2 CArNN 集成 (Ensemble of CArNN, ECArNN)

2.1 ECArNN 算法

本文采用克隆选择算法 (CSA) 训练 CArNN, CArNN 作为克隆选择算法中的抗体, 通过训练调整 CArNN 的模型参数, 使得 CArNN 能够较好地拟合电路功能, 并具有一定的泛化能力, 具体步骤如下:

Step 1. 将 CArNN 编码为抗体, 每个抗体是一个 CArNN 模型;

Step 2. 初始化 CSA 参数, 如抗体规模、变异概率、克隆规模、亲和度阈值等;

Step 3. 计算抗体亲和度, 并归一化亲和度;

Step 4. 根据亲和度克隆并变异抗体;

Step 5. 从变异抗体群中选择最优抗体作为子代抗体;

Step 6. 如果迭代次数或最优抗体满足要求, 则训练结束, 否则转 Step 3.

抗体优化的结果是一个抗体群, 这些抗体能够保持多样性, 而且处于不同的极小值, 既包括全局极小值, 也包括局部极小值. 由于这些网络的差异性较大, 用这些网络集成可以提高泛化能力, 模型的集成方法采用简单平均集成, 集成输出由个体网络共同决定. 假设网络 i 在输入 \mathbf{x} 下的输出为 $\mathbf{y}_i(\mathbf{x})$, 则集成输出为

$$\sum_i w_i \mathbf{y}_i(\mathbf{x}), \quad \sum_i w_i = 1, w_i \geq 0 \quad (25)$$

由于神经网络的权值都是实数, 预测结果也是实数, 而不是 $(0, 1)$, 所以需要定义一个阈值来判断功能是否实现. 阈值 β_i ($i = 1, 2$) 和预测功能 f_i 之间有如下关系

$$f_i = \begin{cases} 1, & f_i > \beta_1 \\ 0, & f_i \leq \beta_2 \\ 0, & f_i < 0 \\ 1, & f_i > 1 \end{cases} \quad (26)$$

且 $0 \leq \beta_1, \beta_2 \leq 1, \beta_1 + \beta_2 = 1, \beta_i$ ($i = 1, 2$) 的取值与潜在问题的识别结果有密切关系, β_i 取值范围的不同可能会带来完全不同的结果. β_i 的选择还与神经网络的泛化误差有密切关系. 假设神经网络的泛化误差为 e , 则有下式成立

$$\|\mathbf{f} - \mathbf{y}\| < e \quad (27)$$

\mathbf{f} 为系统期望输出向量, \mathbf{y} 为网络实际输出向量, 则 β_i 应满足下式

$$\beta_1 > 1 - e, \quad \beta_2 < e \quad (28)$$

假设网络预测的功能为 f , 期望的输出为 1, 有

$$|1 - f| < e \quad (29)$$

根据式 (26) 可知

$$f > \beta_1 \quad (30)$$

由于 $f > 1 - e, f > \beta_1$, 所以需要满足

$$\beta_1 \geq 1 - e \quad (31)$$

假设网络预测的功能为 f , 期望的输出为 0, 有

$$|f| < e \quad (32)$$

根据式 (26) 可知

$$f < \beta_2 \quad (33)$$

由于 $f < e$, 所以需要满足

$$\beta_2 < e \quad (34)$$

综上所述, 只要满足

$$\beta_1 > 1 - e, \quad \beta_2 < e \quad (35)$$

则式 (26), (28) 同时成立.

值得说明的是, 式 (26) 确定了 f 与 β_i 有三种关系:

- 1) $\beta_1 \geq 0.5, \beta_2 < 0.5$;
- 2) $\beta_1 < 0.5, \beta_2 \geq 0.5$;
- 3) $\beta_1 = \beta_2 = 0.5$.

只有第三种关系可以识别所有功能.

从上述分析可知, 0.5 是 CArNN 泛化误差的一个临界点, 这就要求 CArNN 的泛化误差 $e < 0.5$. 如果泛化误差不满足要求, 则此方法失效, 需要更多的样本训练网络, 降低网络的泛化误差. 如果误差满足要求, 则选择 $\beta_1 = \beta_2 = 0.5$ 可以满足功能预测的要求.

ECArNN 用于 SCA 的步骤如下:

Step 1. 模型转换, 将给定的电路转换成神经网络模型;

Step 2. 样本生成, 利用电路的有关数据, 生成以开关组合为输入, 功能元件状态为输出的学习样本;

Step 3. 利用 CSA 训练神经网络, 并形成集成模型;

Step 4. 估计 ECArNN 的泛化误差, 如果泛化误差小于 0.5, 则继续; 否则, 返回 Step 3, 并重新设置 CSA 参数, 重新学习;

Step 5. 利用 ECArNN 预测所有开关组合实现的功能;

Step 6. 比较预测功能与设计功能, 找到潜在问题;

Step 7. 生成分析报告.

2.2 应用实例

图 1 是美国红石火箭的发射电路，这是一个典型的潜在问题案例，从图 1 可以看出，当点火开关闭合后，如果出现尾部接地插头早于尾部脱落插头拔出，就会导致图中虚线所示的潜在问题，即导弹点火后又引起非期望的关机。

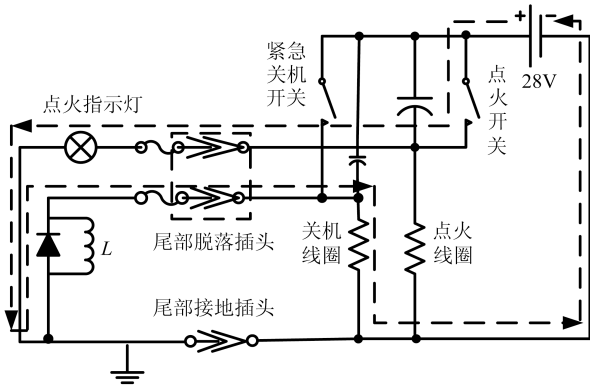


图 1 红石火箭发射电路简图

Fig. 1 Simplified diagram of Red-Stone rocket firing circuit

系统包括四个可控元件：点火开关、紧急关机开关、尾部脱落插头以及尾部接地插头。系统预期实现的功能有两个：点火和关机，对应的功能元件就是关机线圈和点火线圈。电路中的电源，点火指示灯是常态元件，电感线圈 L 以及电容是中间元件。点火开关和紧急关机开关与电源直接相连，关机线圈和点火线圈与地直接相连。转换后的神经网络模型如图 2。

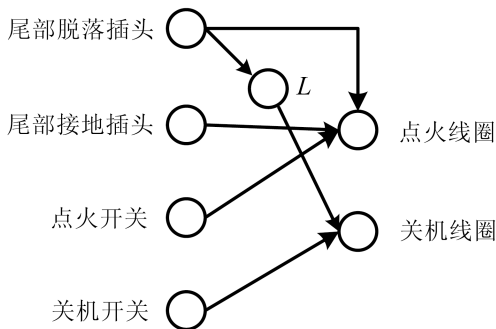


图 2 红石火箭发射的 CArNN 模型

Fig. 2 CArNN model of Red-Stone rocket firing circuit

设计需求：当点火开关关闭时，点火线圈工作，尾部脱落插头和尾部接地插头同时打开；关机开关闭合时，关机线圈工作。根据设计需求，由此可以得出系统的设计矩阵如表 1。

表中 1 表示开关闭合或元件工作，0 表示开关打开或元件不工作。

表 1 红石火箭发射电路设计矩阵

Table 1 Designed matrix of Red-Stone firing circuit

序号	点火开关	关机开关	尾部脱落插头	尾部接地插头	点火线圈	关机线圈
1	1	0	1	1	1	0
2	0	1	0	0	0	1
3	0	0	0	0	0	0

由于此电路中的功能元件状态单一，所以用 $(0, 1)$ 描述。由电路可以看出关机开关闭合时，关机线圈工作；点火开关闭合时，点火线圈工作，而且尾部脱落插头和尾部接地插头同时断开。但是，在实际操作时，由于工艺误差可能引起尾部脱落插头和尾部接地插头不是同时断开。由历史运行数据可知，在尾部脱落插头早于尾部接地插头断开时，系统仍然正常运行。另外，虽然关机开关和点火开关实际中禁止同时闭合，但是我们可以将此状态形成虚拟样本用于训练 CArNN。通过上述分析，可以得到表 2 所示的训练样本。

表 2 神经网络训练样本

Table 2 Training samples of CArNN

序号	输入 1	输入 2	输入 3	输入 4	输出 1	输出 2
1	0	0	0	0	0	0
2	0	0	0	1	0	0
3	0	0	1	0	0	0
4	0	0	1	1	0	0
5	0	1	0	0	0	1
6	0	1	0	1	0	1
7	0	1	1	0	0	1
8	0	1	1	1	0	1
9	1	0	0	0	1	1
10	1	0	1	1	1	0
11	0	0	1	0	0	0
12	1	1	0	0	1	1
13	1	1	0	1	1	1
14	1	1	1	0	1	1
15	1	1	1	1	1	1

克隆选择算法的群体规模为 10，克隆总数目为 500，学习代数为 200，终止条件为误差 $e < 0.1$ ，抗体间亲合度阈值为 0.1，抗体抗原阈值为 0.1，取输出

表 3 神经网络训练样本
Table 3 Training samples of CArNN

	CSANNE	CArNN	BP-NNE	BP(15)	BP(10)	BP(8)	BP(5)	BP(3)
正确率 (%)	94	92	70	50	30	30	25	45
泛化误差	0.3445	0.2907	0.3024	0.4632	0.4241	0.4024	0.4947	0.4048
鲁棒性	0.3539	0.3143	0.3415	0.5329	0.4313	0.4262	0.5194	0.4352

阈值. 我们将第 4, 14 组样本作为测试样本, 其余 13 组样本为训练样本. 采用 5 倍交叉验证实验方法训练网络并用网络在验证集上的平均误差作为网络的泛化误差估计. 为了验证 CArNN 的泛化能力, 实验同时利用 5 个结构不同的 BP 网络进行训练, 结构分别为: 4-15-2, 4-10-2, 4-8-2, 4-5-2, 4-3-2. 实验还对集成的泛化能力作了比较, 集成规模均为 10. 为了说明模型的鲁棒性, 对各个训练后的网络参数进行人为调整, 然后再进行测试, 将测试误差作为鲁棒性指标. 调整的方法是随机选择模型中一定比例的参数, 假设选择的参数为 w , 调整后的参数为: $w' = \tau \times w \times rand$, w' 为变化后参数, τ 为变化幅度参数, $rand$ 为随机数. 实际电路中, 在 (1, 0, 0, 1) 组合下的功能为 (1, 1), 即点火和关机同时工作. 这个功能与设计需求不一致, 是一个潜在状态. 我们分别利用 CArNN, CSANNE, BP, BP-NNE 进行分析, 各实验分别运行 50 次, 统计结果如表 3 所示.

2.3 仿真分析

从实验可以看出, CArNN 分析的正确率达到了 92%, 远远高于几种 BP 网络. 从泛化能力来看, CArNN 的泛化能力要强于 BP 网络, 其主要原因是 CArNN 模型的结构比 BP 网络简单, 即 CArNN 的中间层神经元最少, 而且神经元之间不是全连接的, 而是稀疏连接的. 从实验的时间复杂度看, 由于 CArNN 的结构简单, CArNN 的计算时间要远远小于 BP 网络.

实验通过变化 CArNN 模型的参数来测试模型的鲁棒性. 结果显示, 模型的鲁棒性较强, 参数变化后的模型的泛化能力与原模型相比相差不大, 泛化能力变化平均为 4%. 参数变化后, 预测结果发生变化率为 1%. 值得注意的是, 在某些情况下, 变化后的模型的泛化能力要优于原模型. 主要原因是在训练过程中原模型的泛化能力没有达到最优, 当模型发生变化时, 正好朝着优化的方向变化, 在这种情况下, 泛化能力就会提高. 实验中使模型的某些参数发生较大幅度的变化, 实验结果表明, 模型仍有较强的泛化能力, 模型具有较好的鲁棒性.

实验结果显示, 当网络泛化误差 $e > 0.5$ 时, 预测正确率仅为 2%, 当 $e < 0.5$ 时, 预测正确率达到了 85%, 当 $e < 0.4$ 时, 预测正确率达到了 93%. 这个结果和文中的理论分析是一致的. 实验显示泛化误差越小, 预测的正确率越高. 为了保证分析的正确性, 一般要求模型的泛化误差小于 0.4.

3 总结

本文提出了一种基于电路结构的神经网络模型, 并将其作为集成中的个体, 然后利用集成模型预测电路功能. 针对不同的电路, 会产生不同的神经网络模型, 为了适应这种结构可变的模型的优化学习, 本文利用 CSA 训练网络, 训练后的抗体作为 NNE 的个体. 理论和实验证实, 由这种方法构造的 CArNN 模型不仅具有较高的潜在问题识别能力, 而且对模型的参数变化有较强的鲁棒性. 值得进一步研究的问题包括以下几个方面:

- 1) 目前的分析方法还不能直接完成潜在定时、潜在标志等潜在问题的分析;
- 2) 当系统比较复杂时, 会出现组合爆炸问题, 一个可行的办法就是将系统按功能分解为若干个功能子网络, 然后再利用此方法进行分析.

References

- 1 Vogas J L. The complementary roles of simulation and sneak-analysis. In: Proceedings of IEEE Annual Reliability and Maintainability Symposium. Anaheim, USA: IEEE, 1994. 396-401
- 2 Ren Li-Ming. Sneak circuit analysis. *Quality and Reliability*, 1998, **12**(2): 31-34
(任立明. 潜在电路及潜在电路分析技术. 质量与可靠性, 1998, **12**(2): 31-34)
- 3 Walkers F E. Sneak circuit analysis automation. In: Proceedings of IEEE Annual Reliability and Maintainability Symposium. Atlanta, USA: IEEE, 1989. 502-506
- 4 Li Xue-Feng, Qian Ling. Sneak circuit analysis and its application in astronautic systems. *Aerospace Control*, 2000, **18**(1): 26-30
(李学锋, 钱玲. 航天控制系统潜在分析技术应用研究. 航天控制, 2000, **18**(1): 26-30)

- 5 Dore B, Taylor J R. New developments on automation of sneak analysis. In: Proceedings of IEEE Annual Reliability and Maintainability Symposium. Anaheim, USA: IEEE, 1994. 409-414
- 6 Liu Bing-Jie, Hu Chang-Hua. Sneak circuit analysis based on neural network. *Journal of Astronautics*, 2006, **27**(3): 474-477
(刘丙杰, 胡昌华. 基于神经网络的潜在通路分析. 宇航学报, 2006, **27**(3): 474-477)
- 7 Perrone M P, Cooper L N. When networks disagree: ensemble method for neural networks. In: Proceedings of Artificial Neural Networks for Speech and Vision. New York, USA: Chapman Hall, 1993. 126-142
- 8 Zhou Z H, Wu J X, Tang W. Ensembling neural networks: many could be better than all. *Artificial Intelligence*, 2002, **137**(1): 239-263
- 9 West D, Dellana S, Qian J X. Neural network ensemble strategies for financial decision applications. *Computers and Operations Research*, 2005, **32**(10): 2543-2559
- 10 Granitto P M, Verdes P F, Ceccatto H A. Neural network ensembles: evaluation of aggregation algorithms. *Artificial Intelligence*, 2005, **163**(2): 139-162
- 11 Wu Huai-Ning, Li Yong, Cai Kai-Yuan. Fuzzy robustness analysis based on importance sampling and neural network. *Control Theory and Applications*, 2005, **22**(2): 335-340
(吴淮南, 李勇, 蔡开元. 基于重要抽样法和神经网络的模糊鲁棒性分析. 控制理论与应用, 2005, **22**(2): 335-340)



胡昌华 第二炮兵工程学院导航制导与控制国家重点学科教授. 1996 年于西北工业大学获工学博士学位. 主要研究方向为控制系统自动检测与故障诊断、容错控制与可靠性工程. 本文通信作者.
E-mail: hch6603@263.net

(**HU Chang-Hua** Professor at Xi'an Research Institute of Hi-tech. He received his Ph.D. degree from Northwestern Polytechnical University in 1996. His research interest covers automatic detection and fault diagnosis of system, fault-tolerant control, and reliability engineering. Corresponding author of this paper.)



刘丙杰 第二炮兵工程学院博士研究生. 2001 年获得海军航空工程学院电子工程系学士学位, 2004 年获得海军潜艇学院军事装备学硕士学位. 主要研究方向为故障诊断, 潜在通路分析, 系统可靠性、安全性分析, 智能算法.

E-mail: liubingjie@tom.com

(**LIU Bing-Jie** Ph.D. candidate at Xi'an Research Institute of Hi-tech. He received his bachelor degree from Navy Aviation Engineering College in 2001, and master degree from Navy Submarine Academy in 2004. His research interest covers fault diagnosis, sneak circuit analysis, reliability and safety of systems, and intelligent algorithm.)