

文章编号: 1001-0920(2005)07-0764-05

双并联前向过程神经网络及其应用研究

钟诗胜, 丁刚

(哈尔滨工业大学 机电工程学院, 哈尔滨 150001)

摘要: 为克服多层前向过程神经网络收敛速度慢、精度低的问题, 提出了一种双并联前向过程神经网络模型。在输入空间中引入一组合适的函数正交基, 将输入函数和网络权函数表示为该组正交基的展开形式, 并利用基函数的正交性简化网络聚合运算过程, 给出了相应的学习算法, 并以飞机发动机状态监控中发动机排气温度的预测为例验证了模型和算法的有效性。

关键词: 双并联前向过程神经网络; 飞机发动机状态监控; 正交基函数; 学习算法

中图分类号: TP183

文献标识码: A

Research on Double Parallel Feedforward Process Neural Networks and Its Application

ZHONG Shi-sheng, DING Gang

(School of Mechanical and Electrical Engineering, Harbin Institute of Technology, Harbin 150001, China)

Correspondent: ZHONG Shi-sheng, E-mail: zhongss@hit.edu.cn

Abstract: To solve the problems of slow convergence speed and low accuracy of the multilayer feedforward process neural networks, a double parallel feedforward process neural networks model is proposed. By introducing a set of appropriate orthogonal basis functions into the input space, the input functions and the weight functions are expanded under the orthogonal basis functions, and the time aggregation operation of the process neurons is simplified by using the orthogonality of the basis functions. The corresponding learning algorithm is given and the effectiveness of this method is proved by the prediction of exhaust gas temperature in aircraft engine condition monitoring.

Key words: Double parallel feedforward process neural network; Aircraft engine condition monitoring; Orthogonal basis function; Learning algorithm

1 引言

现实世界中, 许多系统的输入往往是依赖于时间的连续函数或过程, 而其输出既依赖于输入的空间聚合, 又与输入的时间累积效应密切相关。为解决此问题, 文献[1, 2]提出了过程神经元和过程神经网络的概念。过程神经元与传统神经元在结构上类似, 其区别在于过程神经元的输入和连接权可以是时变函数或过程, 其聚合运算既体现了对空间的聚合, 又体现了对时间的累积。受输入同步瞬时限制, 传统神经网络难以表达连续输入信息的累积效应, 同时依

赖于时间的采样数据量较大, 且难于解决较大样本的学习和泛化问题, 因此, 传统神经网络在解决大容量非线性时变系统的信息处理问题时还存在不适应性^[3]。过程神经网络可以直接把时变函数或过程作为输入输出信号, 这是对传统神经网络在时间域上的扩展, 对于求解众多与过程有关的实际问题具有较好的适应性^[4, 5]。

过程神经网络是更一般化的人工神经网络模型, 研究过程神经网络的拓扑结构、函数逼近性质、学习算法等具有十分普遍的意义^[4]。从网络拓扑结

收稿日期: 2004-09-06; 修回日期: 2004-11-08

基金项目: 国家自然科学基金项目(60373102)。

作者简介: 钟诗胜(1964—), 男, 江西龙南人, 教授, 博士生导师, 从事人工智能、人工神经网络等研究; 丁刚(1976—), 男, 山东青岛人, 博士生, 从事人工神经网络及其应用的研究。

构来看, 人工神经网络主要可以分为前向型和反馈型网络, 目前应用最多的是多层前向神经网络结构。文献[6]在研究中发现, 多层前向神经网络结构中增加输入层与输出层的直接连接能够提高网络的训练速度和精度。

为克服多层前向过程神经网络收敛速度慢和精度低等问题, 本文在文献[1, 2, 6]的基础上提出了双并联前向过程神经网络的概念和模型。双并联前向过程神经网络在多层前向过程神经网络的基础上增加了输入层与输出层的直接全互连接。双并联前向过程神经网络的时间累积运算过程十分复杂, 本文选择合适的正交基函数, 并将网络的输入函数和连接权函数同时展开, 以达到简化运算的目的。文中给出了相应的学习算法, 并以飞机发动机状态监控中发动机排气温度(EGT)的预测为例验证了模型和算法的有效性。

2 双并联前向过程神经网络

2.1 过程神经元

过程神经元由加权、聚合和激励 3 部分组成^[2]。过程神经元与传统神经元不同之处在于过程神经元的输入和权值可以是随时间变化的函数或过程, 其聚合运算既有对空间的多输入聚合, 也有对时间过程的累积聚合。过程神经元的结构如图 1 所示。

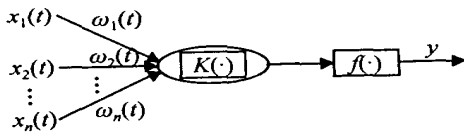


图 1 过程神经元结构

图 1 中, $x_i(t) (i = 1, 2, \dots, n)$ 为过程神经元的输入函数, $\omega(t)$ 为相应的连接权函数, $t \in [0, T]$, $[0, T]$ 为时间采样区间, $K(\bullet)$ 为过程神经元的时间聚合基函数, $f(\bullet)$ 为激励函数, 一般可取为 Sigmoid 函数。

过程神经元的输入与输出之间的关系为

$$y = f((W(t) \oplus X(t)) \otimes K(\bullet) - \theta). \quad (1)$$

其中: $X(t) = (x_1(t), x_2(t), \dots, x_n(t))$, $W(t) = (\omega_1(t), \omega_2(t), \dots, \omega_n(t))$; “ \oplus ”表示某种空间聚合运算, 可取为 $A(t) = W(t) \oplus X(t) = \sum_{i=1}^n x_i(t)\omega_i(t)$; “ \otimes ”表示某种时间聚合运算, 可取为 $A(t) \otimes K(\bullet) = \int_0^T A(t)K(\bullet)dt$; θ 为过程神经元的输出阈值。

2.2 双并联前向过程神经网络模型

2.2.1 网络拓扑结构

在多层前向过程神经网络的基础上, 增加从输入层神经元到输出层神经元的直接全互连接后形成

双并联前向过程神经网络。本文重点介绍 3 层双并联前向过程神经网络, 设该网络输入层有 n 个单元; 隐层由 m 个过程神经元组成; 为简便起见, 输出层只有 1 个单元。网络拓扑结构为 $n \times m - 1$, 如图 2 所示。

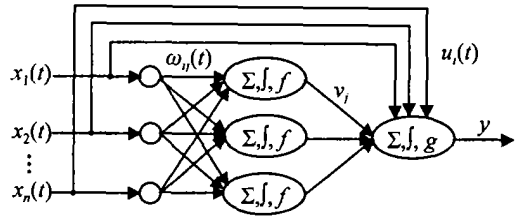


图 2 双并联前向过程神经网络

2.2.2 网络各层间的输入输出关系

系统输入为

$$X(t) = (x_1(t), x_2(t), \dots, x_n(t)).$$

隐层输入为

$$net_j = \int_0^T \sum_{i=1}^n \omega_j(t)x_i(t)dt \quad (2)$$

其中: $net_j (j = 1, 2, \dots, m)$ 为隐层第 j 个过程神经元的输入, $\omega_j(t)$ 为隐层第 j 个过程神经元与输入层第 i 个单元的连接权函数, $[0, T]$ 为系统输入过程区间。

隐层输出为

$$out_j = f(net_j - \theta^{(1)}). \quad (3)$$

其中: out_j 为隐层第 j 个过程神经元的输出, $\theta^{(1)}$ 为隐层第 j 个过程神经元的输出阈值, f 为隐层过程神经元的激励函数, 一般取为 Sigmoid 函数。

设输出层的激励函数 g 为线性函数, 则双并联前向过程神经网络的系统输出为

$$y = \sum_{j=1}^m out_j v_j + \int_0^T \sum_{i=1}^n x_i(t)u_i(t)dt - \theta^{(2)}. \quad (4)$$

其中: y 为双并联前向过程神经网络的系统输出, $u_i(t)$ 为输入层第 i 个单元与输出层神经元的连接权函数, $\theta^{(2)}$ 为输出层神经元的阈值。

综合式(2)~式(4), 双并联前向过程神经网络的输入输出关系为

$$y = \sum_{j=1}^m v_j f \left[\int_0^T \sum_{i=1}^n \omega_j(t)x_j(t)dt - \theta^{(1)} \right] + \int_0^T \sum_{i=1}^n x_i(t)u_i(t)dt - \theta^{(2)}. \quad (5)$$

3 基于正交基函数的学习算法

3.1 输入函数和网络权函数的正交基展开

对输入函数矢量 $X(t)$ 的分量函数 $x_i(t)$ 由 Weierstrass 第 1 逼近定理^[7]知, 对 $\forall \epsilon > 0$, 必存在多项式 $P(t)$, 使 $|P(t) - x_i(t)| < \epsilon$ 从而 $x_i(t)$ 可表示为 $x_i(t) = \sum_{p=1}^{c_i^{(p)}} c_i^{(p)} p_p(t), c_i^{(p)} \in R$ 。设多项式函数系

$\{p_1(t), p_2(t), \dots, p_p(t), \dots\}$ 是相互独立的, 则可以通过施密特正交化步骤^[8] 使之正交化, 于是可以得到 $C([0, T])$ 空间上的一组标准正交基函数(例如三角基函数、勒让德多项式函数等). 则 $x_i(t)$ 可表示为

$$x_i(t) = \sum_{p=1}^n a_i^{(p)} b_p(t), \quad a_i^{(p)} \in R,$$

其中 $b_1(t), b_2(t), \dots, b_p(t), \dots$ 为一组标准正交基函数. 同理, $\omega_j(t)$ 和 $u_i(t)$ 也可由同一组标准正交基函数分别展开为

$$\omega_j(t) = \sum_{p=1}^m \omega_j^{(p)} b_p(t), \quad \omega_j^{(p)} \in R;$$

$$u_i(t) = \sum_{p=1}^n u_i^{(p)} b_p(t), \quad u_i^{(p)} \in R.$$

对于 $\forall \epsilon > 0, \exists P_i$, 使 $\sup_{t \in [0, T]} |x_i(t) - \sum_{p=1}^{P_i} a_i^{(p)} b_p(t)| < \epsilon$, 并令 $P = \max\{P_1, P_2, \dots, P_n\}$, 则

$$x_i(t) = \sum_{p=1}^P a_i^{(p)} b_p(t),$$

$$\omega_j(t) = \sum_{l=1}^P \omega_j^{(l)} b_l(t),$$

$$u_i(t) = \sum_{l=1}^P u_i^{(l)} b_l(t).$$

根据正交函数的性质, 有

$$\int_0^T b_p(t) b_l(t) dt = \begin{cases} 1, & p = l; \\ 0, & p \neq l \end{cases}$$

因此, 式(5)可以转化为

$$y = \sum_{j=1}^m v_j f \left[\sum_{i=1}^n \sum_{p=1}^P a_i^{(p)} \omega_j^{(p)} - \theta^{(1)} \right] + \sum_{i=1}^n \sum_{p=1}^P a_i^{(p)} u_i^{(p)} - \theta^{(2)}. \quad (6)$$

3.2 学习过程

给定 S 个学习样本 $(x_{1s}(t), x_{2s}(t), \dots, x_{ns}(t), d_s), s = 1, 2, \dots, S$. 其中: $x_{ij}(t)$ 的第1个下标表示输入函数向量分量序号, 第2个下标表示学习样本序号; d_s 为对应于输入 $(x_{1s}(t), x_{2s}(t), \dots, x_{ns}(t))$ 的网络期望输出. 设 y_s 为对应于输入 $(x_{1s}(t), x_{2s}(t), \dots, x_{ns}(t))$ 的网络实际输出, 则双并联前向过程神经网络的误差函数可定义为

$$E = \frac{1}{2} \sum_{s=1}^S (y_s - d_s)^2 = \frac{1}{2} \sum_{s=1}^S \left[\sum_{j=1}^m v_j f \left[\sum_{i=1}^n \sum_{p=1}^P a_i^{(p)} \omega_j^{(p)} - \theta^{(1)} \right] + \sum_{i=1}^n \sum_{p=1}^P a_i^{(p)} u_i^{(p)} - \theta^{(2)} - d_s \right]^2, \quad (7)$$

其中 $a_{is}^{(p)} \in R$ 为函数 $x_{is}(t)$ 正交展开式中相对于基

函数 $b_p(t)$ 的系数. 为简便起见, 记 $z_{js} = \sum_{i=1}^n \sum_{p=1}^P a_{is}^{(p)} \omega_j^{(p)} - \theta^{(1)}$.

根据最速梯度下降算法, 双并联前向过程神经网络的待训练参数调整规则为

$$\omega_j^{(p)}(k+1) = \omega_j^{(p)}(k) + \alpha \Delta \omega_j^{(p)},$$

$$v_j(k+1) = v_j(k) + \beta \Delta v_j,$$

$$u_i^{(p)}(k+1) = u_i^{(p)}(k) + \gamma \Delta u_i^{(p)},$$

$$\theta^{(1)}(k+1) = \theta^{(1)}(k) + \eta \Delta \theta^{(1)},$$

$$\theta^{(2)}(k+1) = \theta^{(2)}(k) + \lambda \Delta \theta^{(2)}. \quad (8)$$

其中: $\alpha, \beta, \gamma, \eta, \lambda$ 为学习速率, k 为迭代次数. $\Delta \omega_j^{(p)}, \Delta v_j, \Delta u_i^{(p)}, \Delta \theta^{(1)}, \Delta \theta^{(2)}$ 可分别表示为

$$\Delta \omega_j^{(p)} = - \frac{\partial E}{\partial \omega_j^{(p)}} = - \sum_{s=1}^S (y_s - d_s) f'(z_{js}) a_{is}^{(p)},$$

$$\Delta v_j = - \frac{\partial E}{\partial v_j} = - \sum_{s=1}^S (y_s - d_s) f(z_{js}),$$

$$\Delta u_i^{(p)} = - \frac{\partial E}{\partial u_i^{(p)}} = - \sum_{s=1}^S (y_s - d_s) a_{is}^{(p)},$$

$$\Delta \theta^{(1)} = - \frac{\partial E}{\partial \theta^{(1)}} = - \sum_{s=1}^S (y_s - d_s) f(z_{js}) (-1),$$

$$\Delta \theta^{(2)} = - \frac{\partial E}{\partial \theta^{(2)}} = - \sum_{s=1}^S (y_s - d_s) (-1). \quad (9)$$

若取激励函数 f 为 S 函数, 则

$$f(z) = \frac{1}{1 + e^{-z}},$$

$$f'(z) = f(z)(1 - f(z)).$$

3.3 学习算法描述

双并联前向过程神经网络基于正交基函数展开的学习算法可完整描述如下:

Step 1: 确定双并联前向过程神经网络结构参数;

Step 2: 选取合适的正交基函数 $b_p(t), p = 1, 2, \dots, P$, 将双并联前向过程神经网络的输入函数矢量 $X(t)$ 的分量函数 $x_i(t) (i = 1, 2, \dots, n)$ 和输入层与隐层之间的连接权函数 $\omega_j(t) (j = 1, 2, \dots, m)$ 以及输入层与输出层之间的连接权函数 $u_i(t)$ 同时展开;

Step 3: 给定网络学习误差精度 ϵ , 学习迭代次数 $k = 0$, 最大学习迭代次数 M 及学习速率;

Step 4: 初始化双并联前向过程神经网络待训练参数 $\omega_j^{(p)}, v_j, u_i^{(p)}, \theta^{(1)}$ 和 $\theta^{(2)}$;

Step 5: 根据式(7)计算双并联前向过程神经网络的误差函数 E , 如果 $E < \epsilon$ 或 $k > M$ 则转 Step 7;

Step 6: 根据式(8)和(9)的规则调整双并联前向过程神经网络待训练参数 $\omega_j^{(p)}, v_j, u_i^{(p)}, \theta^{(1)}$ 和 $\theta^{(2)}$, $k+1 \leftarrow k$, 转 Step 5;

Step 7: 输出学习结果, 结束

4 仿真试验

随着航空运输业的不断发展,对飞行安全提出了更高的要求,对飞机发动机实行实时状态监控是保证飞行安全的可靠措施。同时,为降低维修成本,世界各大航空公司均采取了“视情维修”的维修方式,即根据发动机状态来确定发动机是否需要返厂修理,这样可在保障安全的前提下,大大提高发动机的装机使用时间,降低运营成本。飞机发动机排气温度(即 EGT),是飞机发动机的重要参数之一^[9]。因为受到发动机热部件材料温度的限制,必须控制 EGT 值,通过监视 EGT 值能够判断飞机发动机的性能及衰退情况。在飞机巡航期间,EGT 是一个随时间变化而变化的过程。由于 EGT 受许多复杂因素的影响,无法以确定的数学模型预测其变化规律,本文尝试以双并联前向过程神经网络方法解决此问题。

本文的 EGT 数据取自北京飞机维修工程公司所维修的中国国际航空公司的飞机,其发动机所属机群编号为 767-ER,发动机型号为 JT9D-7R4E,发动机编号为 716928,飞机号为 B2553。采样时间为 2000-01-04~2000-12-18,采样时间间隔一般为 1 周左右,这里近似看作等间距采样,共采得 44 个离散数据。将连续的 8 个离散数据进行拟和,构成一个时序函数作为双并联前向过程神经网络的输入函数,以第 9 个数据作为输出。双并联前向过程神经网络的拓扑结构为 1-80-1。将双并联前向过程神经网络的输入函数及连接权函数由三角基函数展开,基函数个数为 8 个,共得到 36 组样本,前 30 组样本作为双并联前向过程神经网络的训练样本。网络误差精度设定为 0.01,学习速率为 0.05,最大迭代次数为 10 000 次。网络经 4 699 次学习迭代后收敛,学习误差曲线如图 3 所示。

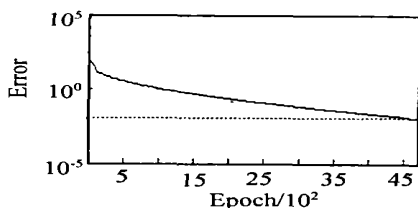


图 3 双并联前向过程神经网络学习误差曲线

为测试学习完成后双并联前向过程神经网络的泛化能力,用后 6 组非训练样本进行测试。测试结果如表 1 所示。平均相对误差为 2.14%,最大相对误差不超过 4.00%,满足实际工程要求,能够达到对发动机健康状况进行实时监控的目的。

为进行对比分析,在同样条件下,采用相同规模

表 1 双并联前向过程神经网络预测 EGT 测试结果

样本号	实际值/	预测值/	绝对误差/	相对误差/%
1	35.400 0	36.122 9	0.722 9	2.04
2	35.900 0	36.006 1	0.106 1	0.30
3	35.700 0	35.056 1	0.643 9	1.80
4	37.100 0	35.696 6	1.403 4	3.78
5	36.600 0	37.527 3	0.927 3	2.53
6	38.700 0	37.783 4	0.916 6	2.37

表 2 3 层前向过程神经网络预测 EGT 测试结果

样本号	实际值/	预测值/	绝对误差/	相对误差/%
1	35.400 0	36.880 5	1.480 5	4.18
2	35.900 0	37.263 7	1.363 7	3.80
3	35.700 0	35.604 8	0.095 2	0.27
4	37.100 0	35.879 2	1.220 8	3.29
5	36.600 0	37.709 0	1.109 0	3.03
6	38.700 0	37.797 3	0.902 7	2.33

的 3 层前向过程神经网络进行 EGT 预测,网络经 6 711 次学习迭代后收敛。用后 6 组非训练样本进行测试,测试结果如表 2 所示。

由表 1 和表 2 可以看出,虽然 3 号样本的预测结果中 3 层前向过程神经网络优于双并联前向过程神经网络,但在其他样本点中 3 层前向过程神经网络的预测误差却迅速上升,有较大起伏。3 层前向过程神经网络预测 EGT 的测试结果的平均相对误差为 2.82%,双并联前向过程神经网络预测 EGT 的测试结果的平均相对误差仅为 2.14%。这说明 3 层前向过程神经网络虽然能实现对 EGT 的预测,但其性能并不稳定,且收敛速度慢。双并联前向过程神经网络由于增加了输入层到输出层的全互直接连接,实际上实现了一个单层前向过程神经网络和一个多层前向过程神经网络的并联。单层前向过程神经网络能够快速给出线性解,多层前向过程神经网络利用其隐节点修改线性解,使网络性能趋于稳定,不仅收敛速度变快,而且精度有所提高。

5 结 语

本文在多层前向过程神经网络的基础上,提出了双并联前向过程神经网络模型,并给出了基于函数正交基展开的学习算法。基于函数正交基展开的学习算法,较好地解决了双并联前向过程神经网络在训练过程中对空间和时间开销较大的问题,可提高网络学习效率和对实际问题求解的适应性。以飞机发动机状态监控中 EGT 预测为例,验证了网络模型及学习算法的有效性,同时也验证了双并联前向过程神经网络比多层前向过程神经网络具有较快的收敛速度和较好的精度。

参考文献(References)

- [1] He X G, Liang J Z. Process Neural Network [A]. *World Computer Congress 2000, Proc of Conf on Intelligent Information Processing* [C]. Beijing: Publishing House of Electronics Industry, 2000: 143-146
- [2] 何新贵, 梁久祯. 过程神经网络的若干理论问题[J]. *中国工程科学*, 2000, 2(12): 40-44
(He X G, Liang J Z. Some Theoretical Issues on Procedure Neural Networks[J]. *Engineering Science*, 2000, 2(12): 40-44)
- [3] 许少华, 何新贵. 基于函数正交基展开的过程神经网络学习算法[J]. *计算机学报*, 2004, 27(5): 645-650
(Xu S H, He X G. Learning Algorithms of Process Neural Networks Based on Orthogonal Function Basis Expansion[J]. *Chinese J of Computers*, 2004, 27(5): 645-650)
- [4] 何新贵, 梁久祯, 许少华. 过程神经网络的训练及其应用[J]. *中国工程科学*, 2001, 3(4): 31-35
(He X G, Liang J Z, Xu S H. Learning and Applications of Procedure Neural Networks [J]. *Engineering Science*, 2001, 3(4): 31-35)
- [5] 钟诗胜, 丁刚. 基于过程神经网络的飞机发动机状态监控[A]. *第14届中国神经网络学术会议论文集*[C]. 合肥: 合肥工业大学出版社, 2004: 614-618
(Zhong S S, Ding G. Monitoring Aircraft Engine Condition Using Process Neural Networks[A]. *Proc of The 14th National Conference on Neural Networks*[C]. Hefei: Hefei University of Technology Publishing House, 2004: 614-618)
- [6] 何明一. 双并联前向神经网络及其在飞行故障检测仿真研究中的应用[J]. *航空学报*, 1994, 15(7): 877-881.
(He M Y. Double Parallel Feedforward Neural Network with Application to Simulation Study of Flight Fault Inspection [J]. *Acta Aeronautica et Astronautica Sinica*, 1994, 15(7): 877-881)
- [7] 徐利治, 王仁宏, 周蕴时. *函数逼近的理论与方法*[M]. 上海: 上海科学技术出版社, 1983: 33-40
(Xu L Z, Wang R H, Zhou Y S. *Theory and Method of Function Approximation* [M]. Shanghai: Shanghai Scientific and Technical Publishers, 1983: 33-40)
- [8] 柳重堪. *正交函数及其应用*[M]. 北京: 国防工业出版社, 1982: 1-6
(Liu Z K. *Orthogonal Function and Its Application* [M]. Beijing: National Defence Industry Press, 1982: 1-6)
- [9] 北京飞机维修工程有限公司. *ΔEGT与EGT裕度在发动机状态监控中的作用及运用*[R]. 民航发动机可靠性研究课题总结, 2000, 7: 33-43
(Beijing Aircraft Maintenance and Engineering Corporation. *The Function and Handling of ΔEGT and EGT Margin in Aircraft Engine Condition Monitoring* [R]. *The Reliability Research Program Summary of Civil Aviation Engine*, 2000, 7: 33-43)

第六届全国智能控制与自动化大会征文通知

全球智能控制与自动化大会(WCICA)是每两年一次在中国召开的重要国际会议。会议为全球从事智能控制和自动化的专家、学者和工程技术人员提供一个交流、研讨和报告他们最新研究成果的良机。录用的论文将由正式出版社出版论文集, 版权属IEEE。论文同时被国际重要检索机构EI收录。

会议主题: 智能控制和自动化系统的理论、方法和应用

征文范围: 控制理论和方法; 智能控制; 建模、检测与故障诊断; 控制技术与综合自动化系统; 流程工业自动化与智能化; 制造业自动化与智能化; 其他应用系统

时间地点: 2006年6月21—23日 大连

论文要求: 请投稿者登录大会网址 <http://wcica06.dlut.edu.cn>, 网上提交用中文或英文书写的论文全文(pdf File), 不超过5页。论文格式请参照IEEE标准版式。

截稿日期: 2005年11月1日

联系地址: 116024 大连理工大学电信学院

联系人: 庄严博士 wcica06@dlut.edu.cn 潘学军副教授 panxj@dlut.edu.cn

电话: 86-411-84707576; **传真:** 86-411-84707579