

DOI:10.3969/j.issn.1001-4551.2017.04.017

# 基于 mRMR 的发电机 DCS 信号神经网络 趋势预测方法\*

杨家荣, 李 辉, 郭双全, 吕 伟  
(上海电气集团股份有限公司 中央研究院, 上海 200072)

**摘要:**针对神经网络方法在发电机信号趋势预测过程中如何合理选择网络输入节点的问题,提出了一种基于最大相关和最小冗余(mRMR)算法的神经网络输入信号选取准则。该方法研究了发电机分布式控制系统(DCS)监测数据的特点,采用了 mRMR 算法从原始特征集中选择了与被描述对象具有最大相关性,且特征集元素间冗余量最小的特征子集作为网络输入,进而有效地提高了网络模型对输入输出间非线性函数关系的拟合精度。研究结果表明,当对某电厂 DCS 信号进行分析时,与直接利用神经网络进行趋势预测的准确性相比,该方法预测准确性高、泛化能力好,具有良好的工程适用性。

**关键词:**分布式控制系统;最大相关最小冗余;特征选择;趋势预测;发电机

中图分类号:TM621.3

文献标志码:A

文章编号:1001-4551(2017)04-0408-04

## Trend prediction for generator DCS signal using mRMR based neural network

YANG Jia-rong, LI Hui, GUO Shuang-quan, LV Wei  
(Center Academe, Shanghai Electric, Shanghai 200072, China)

**Abstract:** Aiming at the problem that how to select the network input node reasonably in the process of the trend prediction of the generator signals, an input signal selection criteria method based on max-relevance & min-redundancy (mRMR) was proposed. The characteristic of generator distributed control system (DCS) monitoring data was researched, and the feature subset which had the maximum correlation with the described object and minimum redundancy between feature elements from original feature set was selected as the network input through mRMR, and then the fitting precision of the nonlinear function between input and output of network model could be improved effectively. The results indicate that compares with the accuracy of using neural network to trend prediction directly, the proposed method has high accuracy and good generalization ability, therefore has good engineering applicability.

**Key words:** distributed control system (DCS); max-relevance & min-redundancy; feature selection; trend prediction; generator

### 0 引 言

发电机的实时运行状态直接关乎电站的安全稳定运行、电力质量和电力生产成本等重要的经济效益指标。因此,对发电机组 DCS 信号进行趋势预测,实时掌握机组的运行状态,及时发觉设备的异常工况,具有

重要的理论及现实意义。

针对发电机组 DCS 信号的特点,近年来国内外已提出多种方法用于趋势预测,人工神经网络是其中较为常用的一种方法<sup>[1-2]</sup>。该方法把需要预测的信号作为网络输出信号,把所需相关信号做为网络输入信号,通过网络连接权重表示输入输出间隐含的非线性函数

收稿日期:2016-10-25

基金项目:上海市科委企业合作专项项目(15dz1180400)

作者简介:杨家荣(1981-),男,上海人,高级工程师,主要从事自动控制技术、智能系统等方面的研究。E-Mail: yangjr@shanghai-electric.com

关系,进而获得输出信号的变化趋势。Verma K 和 Niazi K R<sup>[3]</sup>通过径向基函数神经网络对电力系统进行暂态安全评估,刘红文等<sup>[4]</sup>采用 BP 神经网络对发电机定子线棒主绝缘击穿电压进行预测,徐小力等<sup>[5]</sup>提出基于均值函数新息加权的神经网络预测模型对旋转机械振动时间序列数据进行预测。在训练网络的过程中,针对已确定的输出信号,如何选择合适的输入信号是一个影响预测准确性的关键问题。所需输入信号期望与输出信号间相关性大,且输入信号间相关性小。此外,网络输入节点不宜过多,否则易出现信息冗余,过拟合以及计算速度慢等问题,进而影响训练网络的准确性和效率。

互信息是用于评价随机变量间相关性的重要指标之一。基于互信息的最大相关和最小冗余(mRMR)算法能够在较多输入特征中构造出最优特征空间,该空间能针对特定的输出特征将所有的输入特征进行重要性排序,进而为人工神经网络选择合适的输入信号提供依据。

本研究将采用基于 mRMR 的神经网络方法对发电机 DCS 信号进行趋势预测。

## 1 mRMR 简介

mRMR 是从原始特征集中选择与被描述对象具有最大相关性,且特征集元素间冗余量最小的特征子集的算法。即从一个具有  $n$  维样本的特征空间中选取一个  $m(m \leq n)$  维子空间  $S$ ,用于最优特征空间的构造。其中  $m$  是能够代表被描述对象的最小特征维数<sup>[6-7]</sup>。

### 1.1 相关性与冗余性

根据 Shannon 理论,给定两个随机变量  $X$  和  $Y$ ,其概率密度分别为  $p(x)$  和  $p(y)$ ,联合概率密度为  $p(x,y)$ ,则  $X$  与  $Y$  的互信息为:

$$I(X,Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \quad (1)$$

若  $X,Y$  为连续随机变量,则  $X$  与  $Y$  的互信息为:

$$I(X,Y) = \iint p(x,y) \log \frac{p(x,y)}{p(x)p(y)} dx dy \quad (2)$$

两个随机变量的互信息  $I(X,Y)$  等于 0,则表示两个随机变量之间不相关, $I(X,Y)$  越大表明两个变量之

间的相关性越强。

若所选特征和输出类别之间的互信息越大,说明所选特征包含的分类信息越多,与输出之间的相关性越大,能够较好地描述输出变量。若两个特征之间的互信息较大,则说明这两个特征对于输出类别的判定有较多重复信息,冗余量过多,容易出现过拟合现象。

### 1.2 算法原理

最大相关原则要求子空间  $S$  中的自变量与相应的目标变量有最大的相关性,即:

$$\max D(S,C), D = \frac{1}{|S|} \sum_{x_i \in S} I(x_i, c) \quad (3)$$

最小冗余原则要求子空间  $S$  中的  $m$  维样本之间的相关性最小,即:

$$\min R(S), R = \frac{1}{|S|^2} \sum_{x_i, x_k \in S} I(x_i, x_k) \quad (4)$$

优化准则可归纳为:

$$\max_{x_j \in F_{m-1}} [I(x_j, c) - \frac{1}{m-1} \sum_{x_i \in S_{m-1}} I(x_j, x_i)] \quad (5)$$

式中: $S$ —特征空间维数; $I(x_i, c)$ —子空间  $S$  中特征  $x_i$  和目标变量  $c$  之间的互信息; $I(x_i, x_k)$ —子空间  $S$  中特征  $x_i$  和特征  $x_k$  之间的互信息<sup>[8]</sup>。

## 2 最优特征空间选择

在使用神经网络描述输出时,最优特征空间是能够较好描述输出信号的最小维数空间,其中的特征符合 mRMR 准则。通过选择最优特征空间可以在保证描述输出信号准确性的同时减少特征维数。

本研究采用 BP 神经网络方法对某电厂 DCS 信号做趋势预测,信号的采样间隔为 1 min。以发电机定子线圈层间温度、定子上层线圈出水温度和定子下层线圈出水温度作为神经网络输出信号。首先选取 27 组与输出信号具有一定相关性的信号作为预选输入信号<sup>[9-10]</sup>,本例中共使用了某电厂连续 33 天的 DCS 监测数据进行分析,其中前 30 天的数据构成训练样本,共 43 200 个数据点,后 3 天的数据构成测试样本,共 4 320 个数据点。预选信号如表 1 所示。

表 1 预选信号

序号	信号	序号	信号
1	发电机目标负荷	2	发电机定子电流
3	发电机定子冷却器温度调节阀	4	发电机定子线圈冷却水进水流量
5	发电机氢气流量	6	发电机定子线圈冷却器出口冷却水温
7	发电机定子线圈进水温度	8	选择后发电机冷氢温度
9	发电机空侧密封油温度 1	10	发电机氢侧密封油温度 1
11	发电机空侧密封油温度 2	12	发电机氢侧密封油温度 2
13	发电机空侧密封油回油温度(汽端)	14	发电机空侧密封油回油温度(励端)
15	发电机氢侧密封油回油温度 01	16	发电机氢侧密封油回油温度 02
17	发电机定子冷却器出口冷却水温度	18	发电机定子铁心中部近汽端轭部温度
19	发电机定子铁心中部近汽端轭部温度	20	发电机定子铁心中部近励端轭部温度
21	发电机定子铁心中部近励端轭部温度	22	发电机定子铁心励端边端齿部温度
23	发电机定子铁心励端磁屏蔽温度	24	发电机定子铁心励端齿压板温度
25	发电机定子铁心汽端齿部温度	26	发电机定子铁心汽端磁屏蔽温度
27	发电机定子铁心汽端齿压板温度		

本研究通过 mRMR 算法对 27 组预选输入信号进行重要性排序。首先需确定神经网络的输入节点数。输入节点数过少会丢失有效训练信息,节点数过多使网络运算负担较重,易出现计算时间过长、过拟合等问题。

本研究通过枚举法,分别选取 1 组至 27 组信号作为神经网络输入信号用于模型训练,记录训练时间并计算该模型相对于训练样本的均方误差,所得结果如图 1 所示。

由图 1 可知,随着输入信号的增多,均方误差总体呈现逐渐减小的趋势。但当输入信号数目大于 8 个时,均方误差的变化已不太明显,甚至出现波动。这说明前面几组主要的输入信号已基本包含所有的有效信息,选择过多的输入信号反而容易出现过拟合现象。

表 2 神经网络输入信号

序号	信号	序号	信号
1	发电机定子铁心汽端齿压板温度	2	发电机目标负荷
3	发电机定子线圈冷却器出口冷却水温	4	发电机定子铁心中部近汽端轭部温度
5	发电机定子铁心励端磁屏蔽温度	6	发电机定子铁心中部近励端轭部温度
7	发电机定子铁心中部近励端齿部温度	8	发电机定子冷却器出口冷却水温度

### 3 实例分析

本研究选择表 2 所示的 8 组信号作为神经网络输入信号,选择发电机定子上层线圈出水温度作为神经网络输出信号进行趋势预测,得到的结果如图 2 所示。

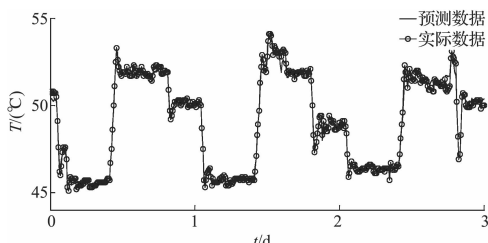


图 2 基于 mRMR 的发电机定子上层线圈出水温度预测

同时通过经验选择 8 组信号作为神经网络输入信号进行对比分析,得到的结果如图 3 所示。

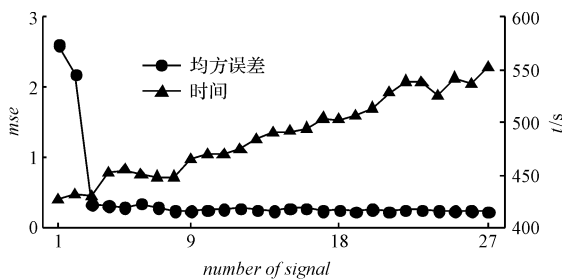


图 1 训练时间与均方误差随输入信号个数变化情况

与此同时,随着输入信号的增多训练网络所需的时间明显增长。本研究综合考虑均方误差和训练时间,最终选取经 mRMR 算法处理后的前 8 组信号作为神经网络的输入信号用于神经网络模型的训练。这 8 组信号如表 2 所示。

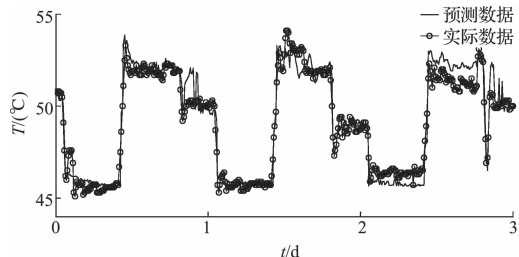


图 3 基于经验选择的发电机定子上层线圈出水温度预测

从图 2 和图 3 可以看出,当本研究采用不同的方法对神经网络输入信号进行选择时,趋势预测的准确性有明显的差异。基于 mRMR 的发电机定子线圈各测点温度的预测值和实际值较为接近,而基于经验选择的发电机定子线圈各测点温度的预测值和实际值偏差较大。为了进一步验证本研究所提出方法的优越

性,本研究利用准则来判断基于 mRMR 输入节点选择和基于经验输入节点选择的神经网络趋势预测的准确性<sup>[11-12]</sup>。首先,本研究计算测试数据实际值与预测值偏差的标准差,然后以为阈值来判断偏差是否超限,从而判断预测数据的准确性。计算结果如图4、图5所示。

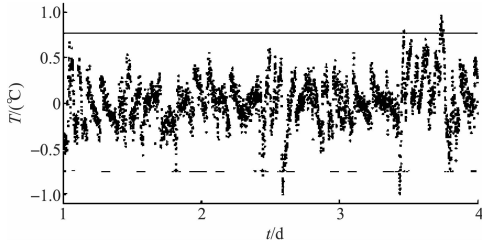


图4 基于 mRMR 的发电机定子上层线圈出水温度  
预测结果偏差值

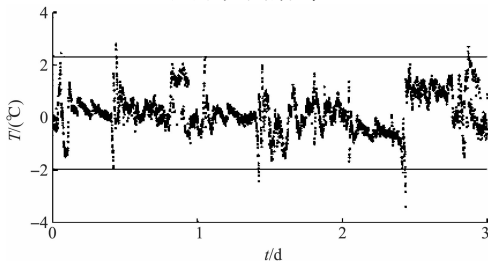


图5 基于经验选择的发电机定子上层线圈出水温度  
预测结果偏差值

通过对比可知,基于 mRMR 输入节点选择的神经网络预测结果偏差范围较小,超阈值数据点少,而基于经验输入节点选择的神经网络预测结果则偏差范围较大。实验结果说明基于 mRMR 输入节点选择的神经网络趋势预测方法的准确性高,同时具有良好的学习泛化能力,能够有效预测信号中的非平稳突变成分。

## 4 结束语

本研究提出一种基于 mRMR 的发电机 DCS 信号神经网络趋势预测方法。该方法首先利用 mRMR 算法对多组 BP 神经网络预选输入信号进行重要性排序,综合考虑训练精度和训练时间,选取能够较好描述输出信号的几组预选信号构成最优特征空间作为神经网络的输入。与直接利用神经网络进行趋势预测的结果相比,mRMR 算法能够剔除输入节点中的信息冗余,选取与期望输出量间相关性最大的输入量输入网络,进而有效地提高网络模型对输入输出间非线性函

数关系的拟合精度。本研究利用该方法对某电厂 DCS 信号做趋势预测,数据分析结果表明,该方法预测准确性高、泛化能力好,为实时掌握机组运行状态,及时发现设备异常工况提供参考,具有良好的工程适用性。

## 参考文献 (References):

- [1] CARTA J A, CABRRA P, MATIAS J M, et al. Comparison of feature selection methods using ANNs in MCP-wind speed methods, a case study [J]. *APPLIED ENERGY*, 2015 (158):490-507.
- [2] ZHANG Z, CHENG H, XU H, et al. Neural network based combining prediction model and its application in condition based maintenance of turbo-generator set [J]. *Proceedings of the Chinese Society of Electrical Engineering*, 2003, 23 (9):204-206, 211.
- [3] VERMA K, NIAZI K R. Rotor trajectory index for transient security assessment using radial basis function neural network [C]. *IEEE Power and Energy Society General Meeting*. New York: IEEE, 2014.
- [4] LIU H, WANG K, MA Y, et al. Aging rule and breakdown voltage prediction for insulation of generator stator winding [J]. *High Voltage Apparatus*, 2015, 51 (3):105-110, 116.
- [5] 徐小力,韩秋实,许宝杰,等.旋转机械的新息加权神经网络工作状态趋势预测研究[J]. *机械强度*, 2006, 28 (S1):37-39.
- [6] YAO M, WANG N, QI M, et al. Study on feature selection method of modified maximal relevance minimal redundancy [J]. *Computer Engineering and Application*, 2014 (50):116-122.
- [7] LI Y, GU X. Feature selection for transient stability assessment based on improved maximal relevance and minimal redundancy criterion [J]. *Proceedings of the Chinese Society of Electrical Engineering*, 2013, 33 (34):179-186.
- [8] 曹静.基于最大相关最小冗余的特征选择算法研究 [D]. 秦皇岛:燕山大学机械工程学院, 2010.
- [9] 张锋,李和明,李永刚,等.大型汽轮发电机定子绕组温度水力模型 [J]. *华北电力大学学报*, 2003, 30 (3):17-19.
- [10] 蒲莹.汽轮发电机故障在线综合诊断系统的研究 [D]. 北京:华北电力(北京)大学能源动力与机械工程学院, 2001.
- [11] 何世彪,杨士中.  $3\sigma$  准则在小波消噪中的应用 [J]. *重庆大学学报:自然科学版*, 2002, 25 (12):58-61.
- [12] 李辉,郭双全,张梦航.基于时间序列的发电机温升趋势分析方法 [J]. *上海电气技术*, 2016, 9 (1):49-52.

[编辑:周昱晨]

## 本文引用格式:

杨家荣,李辉,郭双全,等.基于 mRMR 的发电机 DCS 信号神经网络趋势预测方法 [J]. *机电工程*, 2017, 34 (4):408-411.

YANG Jia-rong, LI Hui, GUO Shuang-quan, et al. Trend prediction for generator DCS signal using mRMR based neural network [J]. *Journal of Mechanical & Electrical Engineering*, 2017, 34 (4):408-411.

《机电工程》杂志: <http://www.meem.com.cn>