# Predicting Mechanical Properties of Hot-Rolling Steel by Using RBF Network Method based on Complex Network Theory

## 摘要

近来，生产高精度，高质量的钢材成为大型钢铁企业的主要目标。由于产品的内部复用部件和生产过程中的复杂变化，在热轧生产过程中实现精确控制是非常困难的。本文采用径向基函数神经网络完成性能预测。它具有训练速度快，精度高的优点，克服了以前使用的BP神经网络的容易陷入局部最优等缺点。在确定径向基函数的中心时，我们利用复杂的网络可视化方法，根据节点的关系可以清楚地理解输入向量之间的关系，并接收中心和宽度。实验表明，社区发现算法与RBF相结合的方法具有稳定性高，训练时间短等优点，适用于大规模数据分析。更重要的是，它可以达到高精度。

## 介绍

我国热轧带钢生产处于相对落后的时期已经有很长一段时间了。但是，随着20世纪90年代中期各大钢铁企业引进国外先进技术和设备，热轧带钢生产进入了一个快速发展的时代。为了满足市场需求和提高竞争力，各大公司正致力于生产高精度，高质量的钢铁产品。研究钢材性能预测对提高生产率和追求高附加值产品具有重要的现实意义。根据之前对钢铁性能的研究和预测，人们可以通过调整钢铁部件和加工技术来达到预期的效果。该方法既可以降低生产成本，又可以开发针对性的客户群。公司可根据技术水平，资金存储量，市场地位等因素生产不同能力的热轧产品能力，以满足市场需求，力求实现自身发展最大化。

由于热轧产品内部结构的复杂性以及生产过程中出现的物理和相变，使得热轧钢材生产的质量控制更加困难。在以前的研究中，学者们用回归分析数据，建立了大量的数据模型[1] - [3]，然而，由于钢铁制造过程的复杂性，这些数据模型的效果并不是很满意，除了各种因素之间的非线性关系外，还有大量的影响因素。模拟人类神经系统信息传递的方式，提出了人工神经网络方法，该方法在钢铁生产质量控制方面显示出很大的优势。近年来，研究人员利用BP神经网络方法进行了大量实验，并在钢铁性能预测中取得了一定的成果[4]，[5]。同时，还提出了许多优化算法，如使用修剪算法改进输入数据[6]，采用贝叶斯算法防止网络“过训练”[7]，使用高维输入法提高网络的准确性[8]，采用遗传算法优化网络权值以避免局部最小[9]，采用两阶段混合算法寻找最佳逼近结果[10]。我们可以注意到，但是，BP神经网络也有其局限性。它还没有得到很好的解决，例如如何确定隐层节点的数量，使用何种训练方法以及训练时间长等问题。

在本文中，我们使用径向基函数神经网络来克服BP神经网络产生的问题，例如如何选择隐层和训练时间长。 在确定隐藏层时，我们使用从复杂网络可视化带来的社区发现算法。 根据输入数据之间的相似性，我们获得隐藏层节点和中心的数量，然后绘制它们。 该方法有效地解决了如何选择节点隐藏层的问题。

本文的结构如下：第二部分介绍相关工作。 第三部分提出了一种结合径向基函数和社区发现算法的高级RBF模型。 第四部分显示了基于钢铁实际生产数据的实验结果。 第五部分总结了这些工作并展示了我们未来的工作

## 相关工作

1985年，鲍威尔提出了一种多变量插值自由基函数（RBF）方法。径向基函数神经网络由Moody和Darken于1988年发展而来，已被证明是能以任何精度临近任何连续非线性函数的网络。 RBF网络是一个三层的前馈网络，包括输入层，隐藏层和输出层[11]。与以前的BP神经网络不同，隐含层的RBF网络是有意义的。我们可以根据不同的输入源来确定隐藏层的数量和值。隐层的传递函数为RBF，它是一个非负非线性函数，具有径向对称性和中心衰减特性。输入层的数据直接映射到隐藏层而不加权。当确定隐藏层的中心时立即确定映射关系。这个过程是非线性映射。从隐藏层到输出层是线性的，即隐藏层中的所有节点经加权后到达输出层。这个过程是线性映射。对于输出层，线性方程计算的方法可以用来计算权重，从而可以大大加快网络的学习速度。

已有不同的算法来改进RBF方法，例如何英生利用改进的K-means算法来确定RBF的中心[18]，王华立用熵聚类方法自动确定RBF神经中心和初始值[19]。 本文使用复杂网络中的社区发现算法来确定RBF中心和宽度。

最近，复杂的网络吸引了许多学者的关注[16]，[17]。 复杂的网络可以描述许多真实世界的复杂系统。 热轧产品性能预测也是一个复杂的问题。 复杂网络结合RBF模型作为研究复杂科学和复杂系统的有力工具，为热轧产品性能预测提供了新的视角。

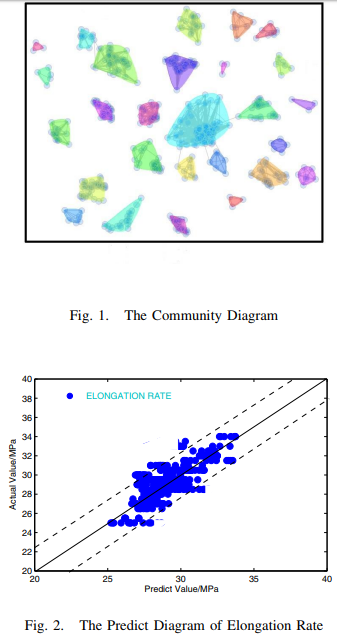
复杂网络是指连接关系复杂的网络，由系统内各元素提取的节点和元素关系下的边连接组成。 这个网络反映了系统的拓扑结构[15]。 复杂网络的社区结构是复杂网络研究的重要方向，其定义为：同一社区中的节点处于紧密连接状态，而不同社区中的节点之间的连接松散，这种结构称为集群，内聚性 组或模块。 从图论的角度看，社区发现只是聚类问题，但它解决了一些传统的聚类问题，如聚类数量，初始中心节点需要预先定义。 同时，由于考虑到复杂网络节点之间的完整拓扑特征，聚类精度大大提高.

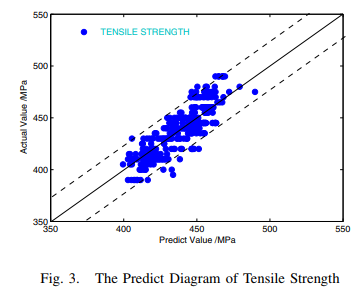
## 4.实验和结果

我们的实验使用钢铁企业的实际产品数据。 共有2000套数据，分为1650个训练和350个测试。 首先，在训练集上使用社区发现算法确定RBF神经网络的隐层节点中心和宽度。 在培训数据上绘制边缘，并将它们放在社区中，因为它们有相似之处。 图2显示了社区地图。

根据公式（3），我们可以计算中心节点，同时利用欧几里得距离来计算每个社区中节点和中心之间最小路径的最大值。 之后，隐藏图层中心和宽度已确定。 通过公式（1）（2）计算网络权重，并通过梯度下降算法调整权重，以降低网络错误率，最终实现期望的网络精度

如图（2）（3）（4）所示，采用我们提出的方法，三种力学性能的预测误差率的拉伸强度分别为±25Mpa，屈服强度±20Mpa和伸长率±3％ 都在错误率限制内。 而且，训练总时间和频率远远小于BP神经网络算法以及使用K-means算法确定中心和宽度的方式。





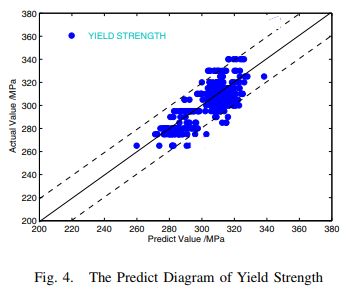
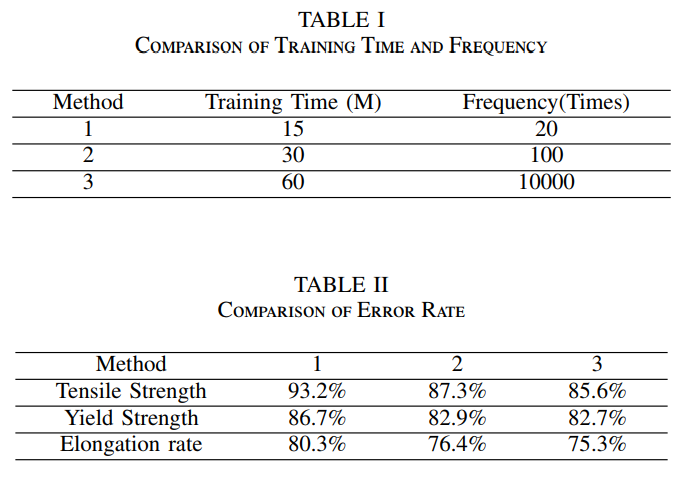


表1显示了三种方法之间训练时间和频率的比较。 方法1表示社区发现算法和RBF的组合。 方法2表示K均值和RBF的组合。 方法3代表BP神经网络。那么他们在表2中



从表1我们知道我们的方法需要较少的训练时间，因为社区发现算法帮助我们正确地找到隐藏层。 与K-means算法不同，我们不需要考虑隐藏层节点的初始数量。 所以这种方法既可以提高速度，又可以保证高精度。 如表2所示，误差率是指测试数据集中真值与预测值误差率**小于0.05的数据率**。

## 5.结论

进行建模和分析，可以获得较高的精度，较高的稳定性和较短的训练时间。 同时，值得关注的是，皮尔逊相关系数忽略了输入值的不同影响。 因此，一些输入矢量具有很大的相关性，但是输出有很大不同。 这是预测错误的最重要原因。 如何根据不同的影响来确定相关性将成为下一个重点。 另外，考虑到社区发现算法的聚类速度可以是线性的，**这种方法适用于大规模数据。 我们将在大型数据集上进行实验，指导未来的实际热轧生产过程**。