

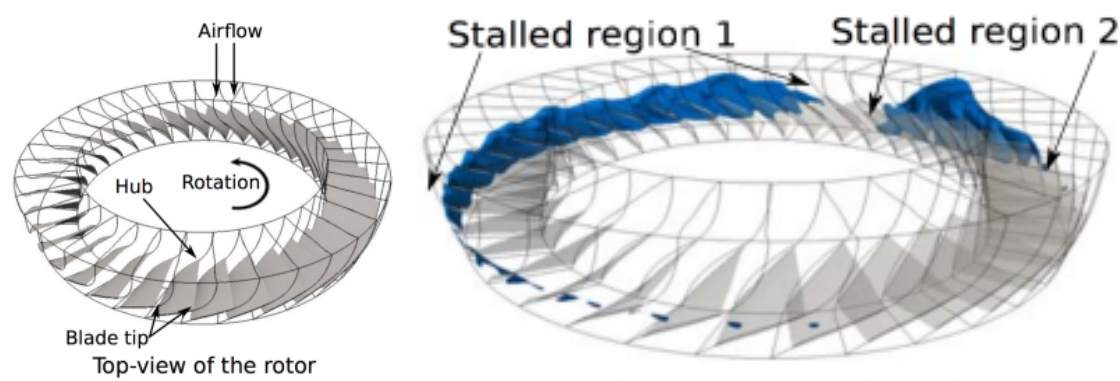
PKU Visualization Blog

北京大学可视化与可视分析博客

喷气机引擎模拟数据之原位预测驱动特征分析(In Situ Prediction Driven Feature Analysis In Jet Engine Simulations)

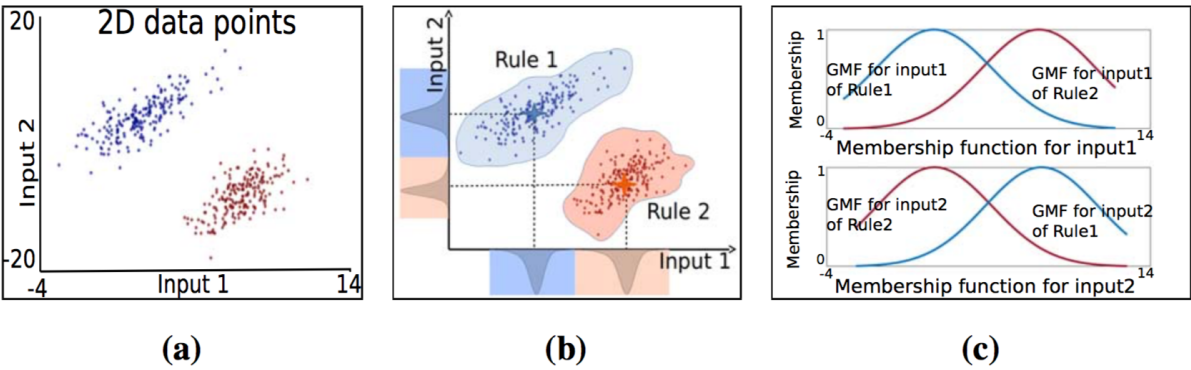
作者: 刘 灿 日期: 2018年4月30日

在喷气机发动机转子中，局部气流不稳可能会对发动机产生不可逆的损坏，这种局部气流不稳称为滞障（Stall）。发动机运行过程中，相关领域的专家希望观察气流滞障的产生过程以及实时观测探究不同参数设置对滞障产生过程的影响。本工作主要探究的参数设置为矫正块流率（CMF，corrected mass flow rate）。



滞障没有明确定义及解析表达，判断某区域是否处于滞障状态只能依靠专家的经验知识。而依靠人力进行判断在大规模数据多时间步实时探究的要求下不可行。因此本工作提出了预测驱动的特征分析。在此预测驱动是指通过机器学习方法，学习专家对滞障状态的解析表达的判断。

具体到本工作中，专家认为对滞障状态影响最大的三个因素为轴向速度、熵、温度，而专家基于此给出选定区域的滞障程度，滞障程度为从0到1连续的数值。故机器学习的核心便是拟合以轴向速度、熵、温度三个变量为输入、以滞障度为输出的函数。



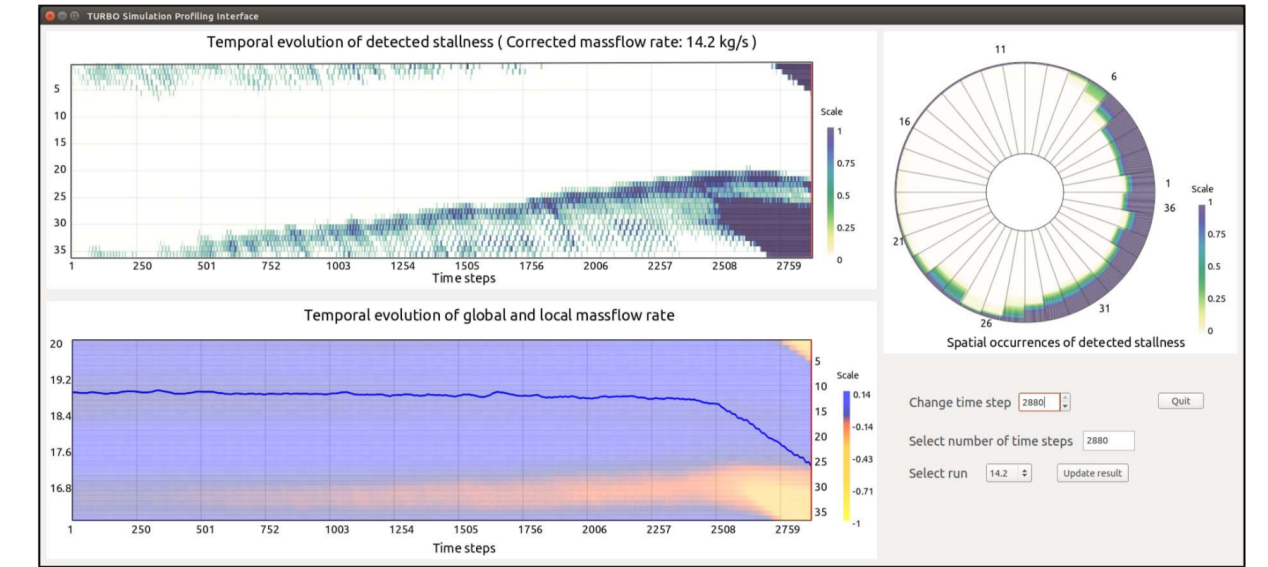
拟合的方式是采用片模糊聚类下的规则生成方法。以二维输入为例，如上图（a），输入数据有两个维度输入1和输入2，本工作中可以代入轴向速度、熵和温度，输出值是一个单一的变量，可以理解成滞障程度。在二维坐标图中，可以看到两个聚类、分别为蓝色和红色，其中蓝色部分代表滞障程度为0，红色部分代表滞障程度为1。直观理解，如果一个新的输入非常接近蓝色聚类的中心，那么它的输出将非常接近0，同理接近红色聚类的，输出将接近1。

对于如果一个点在两者中间，那么计算方法就是以接近两个聚类中心的程度来表示这个点属于两个聚类的程度。表示接近聚类中心的程度可以用高斯函数表示，在每个维度将每个聚类分布拟合成为高斯函数，判断某个点在这个聚类的接近程度就是所有维度上高斯函数的乘积。

每个聚类中对它给出一个预测值。预测值在聚类内部用线性函数逼近，这个线性函数的系数是机器拟合的关键。

有了每个聚类的预测值和接近程度，就可以求出各个聚类的加权输出。越接近的聚类的预测输出比重在总输出中越高。

引擎气流模拟数据具有极大的数据规模，传统的基于事后(post-hoc)式的探索模式在此数据规模下失效。因此本工作提出基于原位的分析，本工作最后仅仅展示在发动机转子中小部分采样数据，可以保证输入输出规模得到控制。支持原位分析。



工作的最后展示了可视化系统以显示机器学习效果，在正确率和对未知数据的拟合能力都有合理性。

References:

[1] In Situ Prediction Driven Feature Analysis In Jet Engine Simulations. Soumya Dutta, Han-Wei Shen, Jen-Ping Chen. In *Proceedings of IEEE Pacific Visualization Symposium (PacificVis 2018)*, Kobe, Japan, Apr. 10-13, 2018.