**工业工程类专业优秀课程设计展示选题**

**半导体晶圆生产过程中多元轮廓数据异常检测**

1. 题目背景

在半导体硅片的生产过程中，诸如晶圆生长、切割、光刻等工艺步骤均需高度精密的控制，任何微小偏差都可能导致产品缺陷。因此，实施异常监控对于保障产品质量和提升生产效率具有重要意义。在实际生产中，通常会部署多个传感器对关键过程变量（如温度、气流、电压、电流等）进行实时监测。这些传感器采集的数据随时间变化形成**轮廓数据（profile data）**，而多个传感器同时生成的数据集合即构成“多元轮廓数据”（multi-channel profile data）。由于不同传感器监测的变量之间可能存在物理或工艺相关性，如温度与气体流量之间的相互影响，因此这些轮廓数据不仅具有时间序列的内部结构（轮廓内相关性），还存在传感器间的相关性（轮廓间相关性）。若能对这类多维轮廓数据进行实时分析，将有望在产品仍处于加工过程阶段时及时识别潜在异常，显著降低由于离线检测延迟带来的质量风险和经济损失。

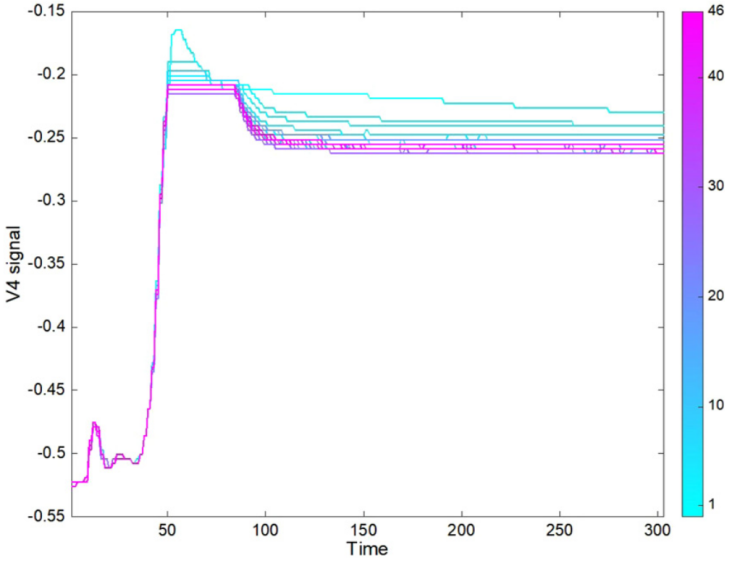
1. 数据介绍

数据文件*data.mat*提供了共 个样本（晶圆），其中前46个为正常样本（in-control），后6个为异常样本（out-of-control）。每个样本由一个 的数组表示，其中 表示共有129个传感器； 表示每个传感器采集的采样点数量（不同样本可能采样时间点不一致，故长度略有差异）,每个元素为一个传感器在对应时间点下的采集值。

1. 数据难点

这一批半导体晶圆生产数据分析主要有以下难点：

* 1. 多重相关性建模复杂：每条传感器数据具有时间序列结构,即存在轮廓内相关性；不同传感器之间可能存在冗余或协同变化，即存在轮廓间相关性，因此需联合建模。
  2. 冗余传感器存在：部分传感器输出为常数值或近似常值，缺乏信息；部分传感器与其他传感器高度冗余，可能造成“维度灾难”和模型过拟合；因此建模前需进行变量筛选或特征提取，以降低维度、提高效率。
  3. 长期漂移现象（long-term drift）：如图一所示，在46个正常样本中，即使系统无异常，传感器采集的轮廓数据也随生产批次推移而发生偏移。这种由于设备磨损或环境因素导致的慢性变化不应被误判为异常，模型需具备一定的漂移鲁棒性。



图一：长期漂移现象示意图

1. 题目要求

设计一套适用于实际晶圆生产线的在线异常监控系统，系统需满足以下核心要求：

* 1. 在线处理能力：每次仅使用当前及历史数据判断最新一个样本是否异常；
  2. 延迟最小化：期望首次检测异常的时间点 ，且越接近越好；
  3. 实时性能：每个样本处理时间应控制在5秒以内；
  4. 漂移鲁棒性：能有效识别系统异常而非漂移引起的波动；
  5. 可解释性与部署性：算法流程清晰、结果可视化、代码可实际部署。

1. 提交成果要求

（文件打包后统一命名：学校全称-轮廓数据异常检测-团队名称）

* 1. 用于展示项目的PPT（宽屏幕16\*9）；
  2. 项目代码，保证项目的可运行性和结果的真实性；
  3. 课程设计报告文件

报告内容包括但不限于：

1. 数据预处理流程（如对齐、归一化、冗余剔除等）；
2. 特征提取与建模方法（如主成分分析、函数型数据分析、深度学习模型等）；
3. 在线异常检测框架设计；
4. 漂移处理策略说明。

报告格式要求如下：

1. 摘要（200字以内）；
2. 问题描述；
3. 研究过程；
4. 结果分析；
5. 结论。

报告全文不超过30页。