

Applying Machine Learning Algorithms to Support Root Cause Analysis – An Experimental Study in Automotive Engineering

Duc Tien, Vu – 22.05.2017, Munich

Advisor: Martin Kleehaus

Chair of Software Engineering for Business Information Systems (sebis)
Faculty of Informatics
Technische Universität München
www.matthes.in.tum.de

IAV GmbH
Ingenieurgesellschaft Auto und
Verkehr

iauv
automotive engineering

Type	Limited liability company: GmbH ^[1]
Industry	Automotive industry ^[1]
Founded	Berlin, Germany (1983) ^[1]
Founder	Prof. Dr. Hermann Appel
Headquarters	Berlin, Germany
Number of locations	16 operations across Germany, several Subsidiaries
Area served	Worldwide
Key people	Kurt Blumenröder, President, CEO; ^[2] Michael Schubert, President, CFO ^[2]
Services	Powertrain, Electronics and Vehicle Development
Revenue	595 million Euro (2013) ^[1]
Number of employees	5,700 worldwide (2013) ^[1]
Website	IAV.com



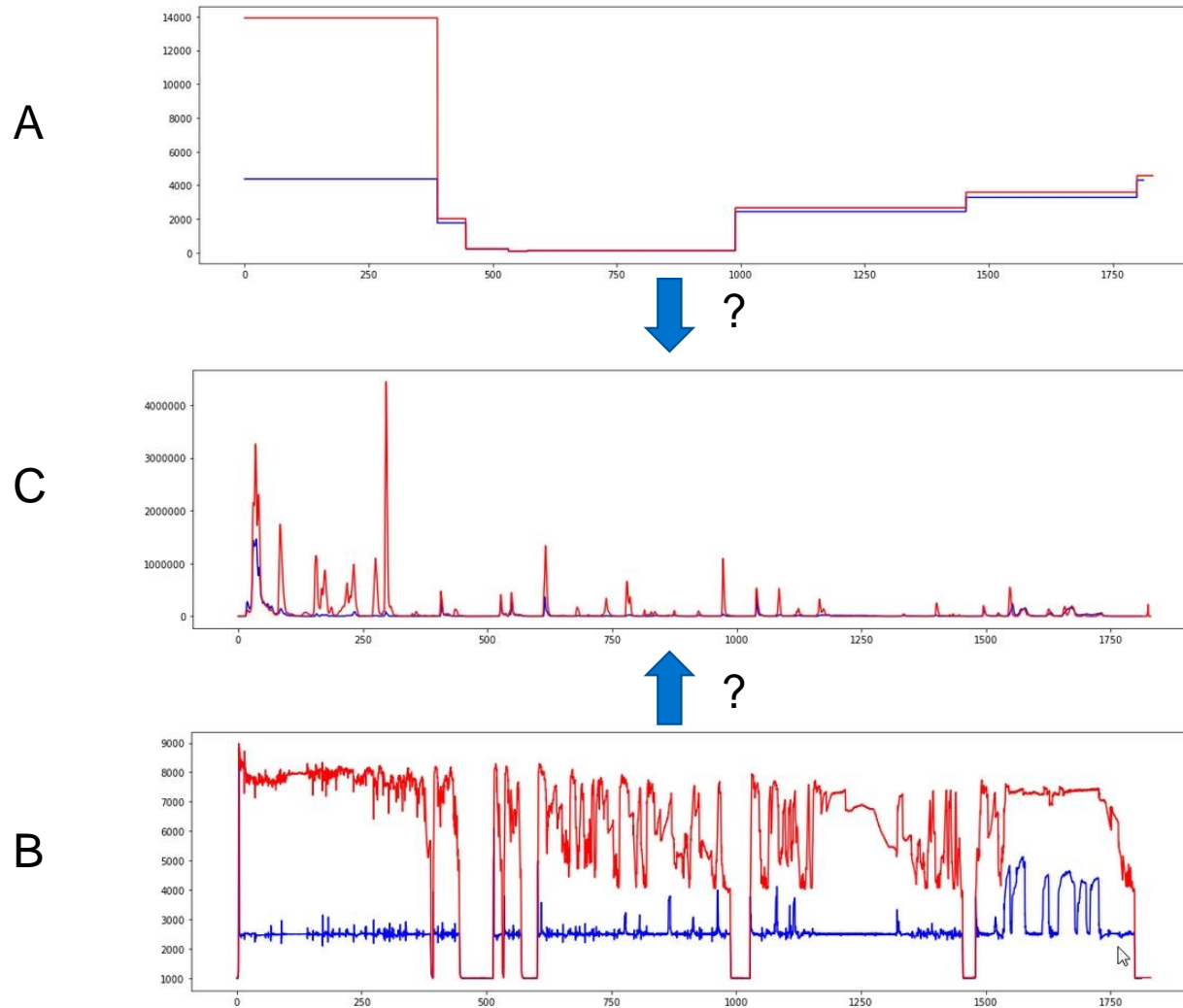
Agenda



1. A little start
2. Motivation
3. Problem Statement
4. Target
5. Research Questions
6. Timeline

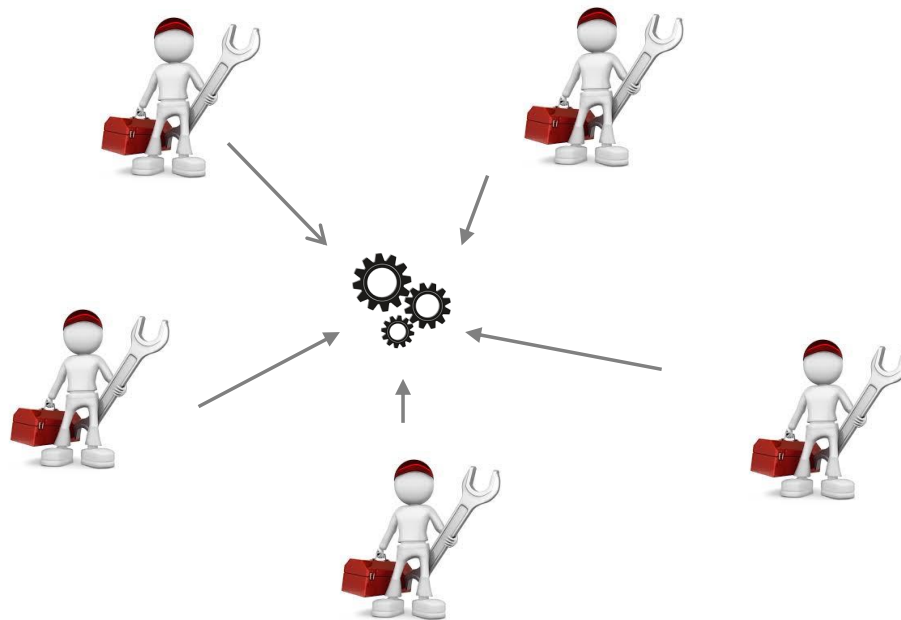
A little start

From A and B, which variable could be the cause for the progress of variable C?



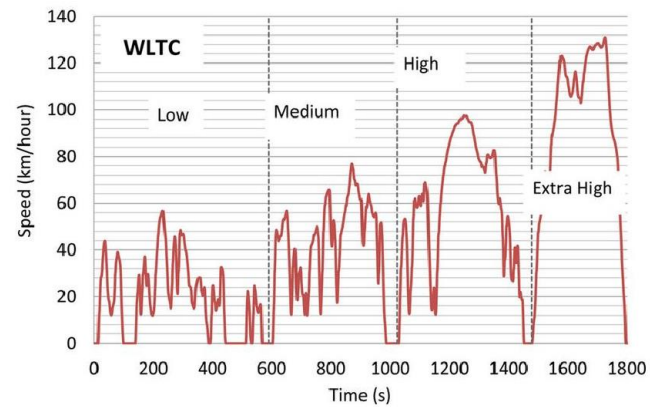
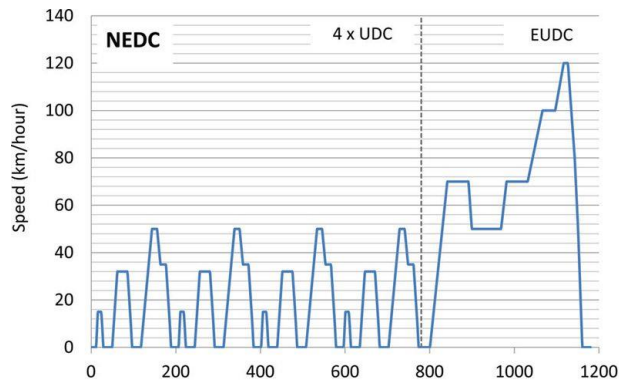
Motivation 1

- Engineer must manually check all the time series data (**800 Variables**) to find out which one is responsible for the anomaly
 - Requires a lot of expert knowledge and time consuming
- Multiple teams are responsible for different parts of the car
 - A coordination is hard



Motivation 2

- Currently: Fixed Driving Cycle (NEDC, WLTC, ...)

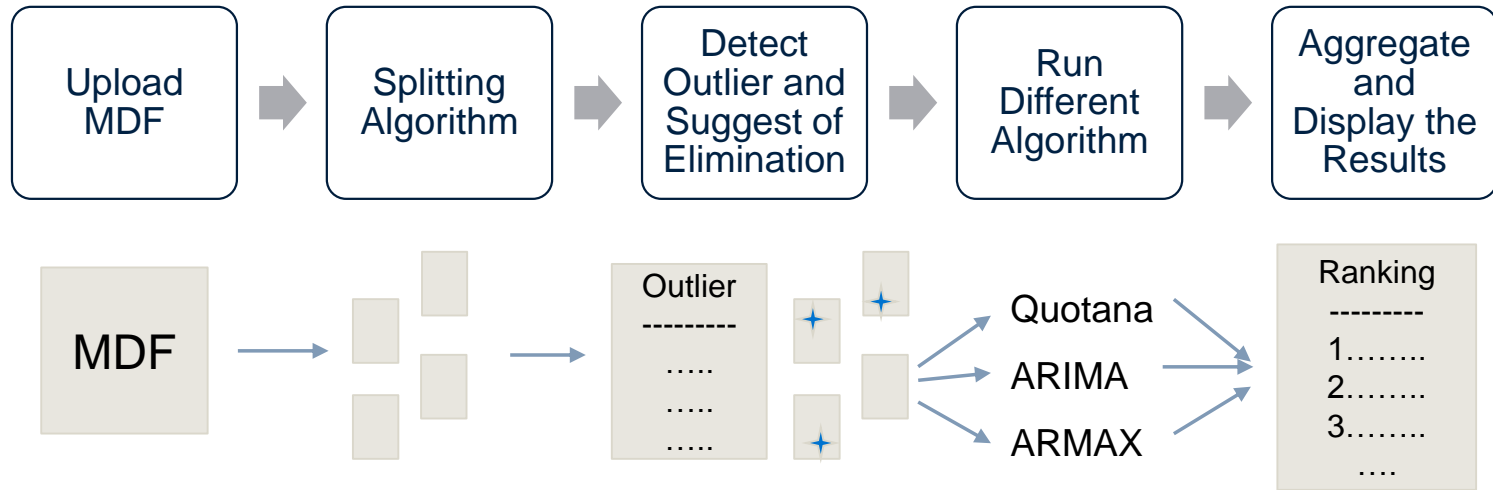


- Future: No Fixed Driving Cycle more, that means the car will be driven and measured in real life situation

→ Support from machine learning techniques is required

Target

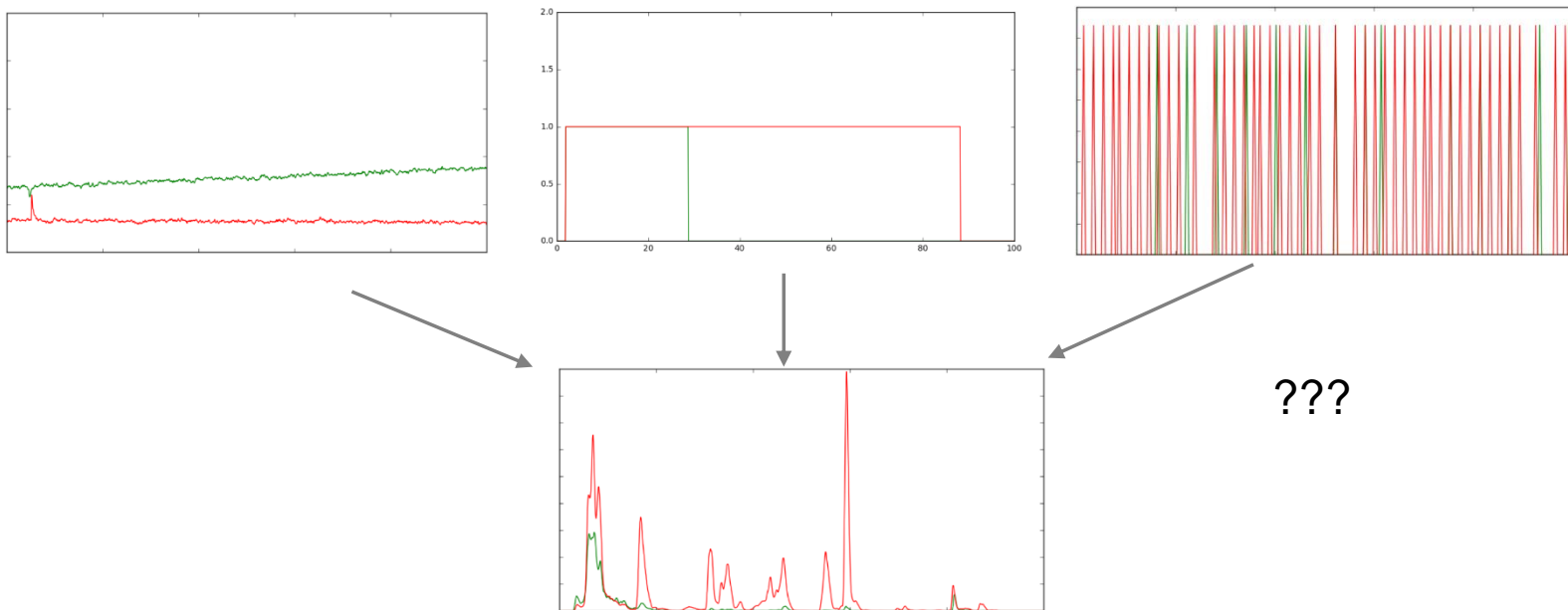
1. A process to analyze the time series



2. Evaluation of different algorithm

Problems

- Limited number from data sets because each emission test is very costly
- Traditional time series analysis deals is only used for curve fitting problem
- Different types of time series require different algorithm
- How to correctly aggregate results from different algorithm



Research questions

- 1. What are the related types of time series?**
- 2. Which algorithm can be used to detect anomaly in the time series**
- 3. In which way can the root-cause analysis be supported?**

Timeline

