

# 12th EUROSYS DOCTORAL WORKSHOP - 2018

## Multivariate temporal data analysis for vessels behavior anomaly detection

PhD Program in Computer Science and Engineering

Instituto Superior Técnico (Lisboa, Portugal)



Rui Maia (rui.maia@tecnico.ulisboa.pt)

### Temporal data and anomaly detection

**Temporal and time-series data analysis** is a broad research field with different relevant applications such as **cyber-security, health care but also for supporting civil and military operations.** **Anomaly detection** is a crucial specific field of data analysis since abnormal data behaviour typically represent critical situations that should be addressed Eg.:network traffic pattern change might indicate a cyber-attack, abnormal heart beating frequency can help anticipate and prevent heart-attacks, vessels behaviour might help detecting smuggling. **Generic Anomaly Detection Framework:** Figure 3

### Univariate vs Multivariate

Multivariate anomaly detection problem raise complex challenges due to the hidden data structure and semantics between time-series. **Observation, Sequence, Context and Collective anomalies** are still open and complex research challenges. Figure 1 illustrates a multivariate time-series.

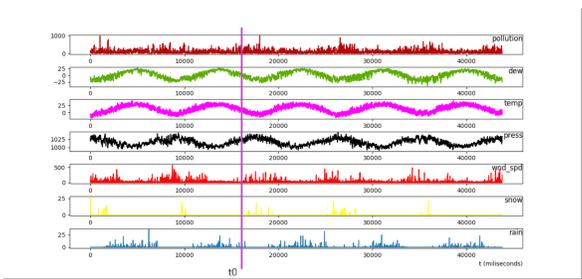


Figure 1: Air and weather conditions represented in a multivariate time-series measured by a network of time synchronized sensors.

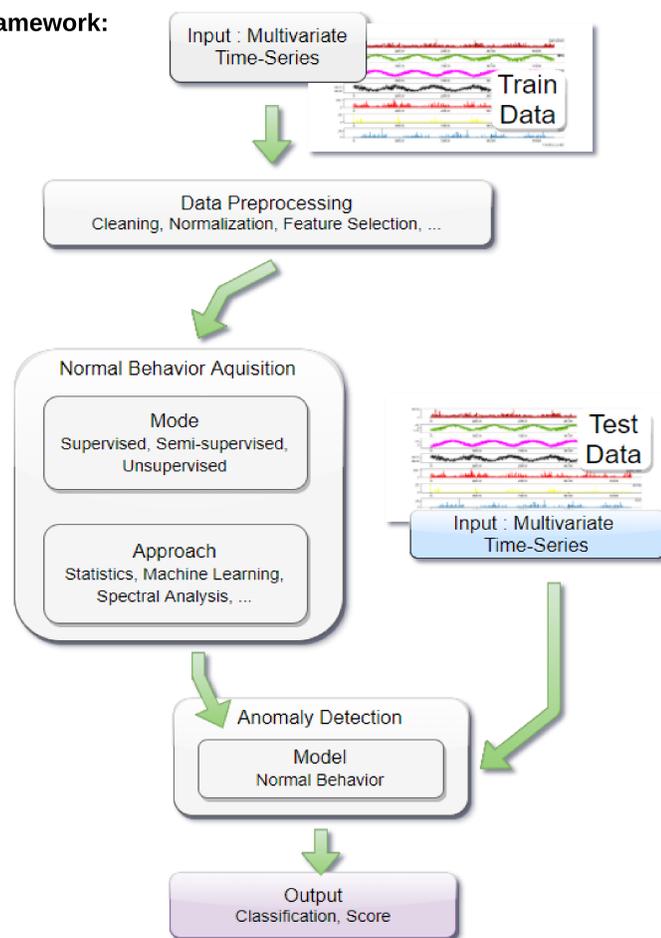


Figure 3: Multivariate temporal data generic anomaly detection framework

### Research Challenges & Goals

#### Research Challenges

- Categorical and real valued data parameters might have been manipulated by emitting entities.
- Complex relations of multiple dimensions affecting sensors data (Figure 4).
- Hidden semantic relations between different time-series or temporal data.
- Presence of different noise patterns and origins due to the use of complex networks of sensors.

#### Research Goals

New methods for - Observation, Sequence, Context and Collective - multivariate temporal data anomaly detection.

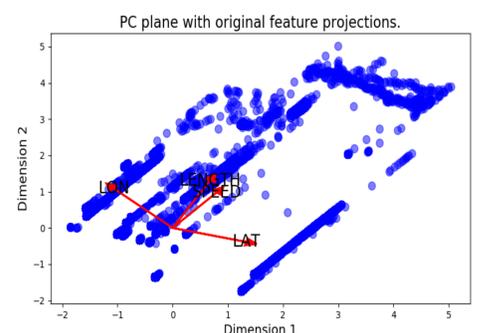


Figure 4: Observations representation using two dimensions. Multiple contributing dimensions vectors are also represented in the biplot.

### Marisa EU Project (\*)

This research work is developed in the scope of Maritime Integrated Surveillance Awareness (Marisa) H2020 European Project. We aim at the **development of new behavior analysis and anomaly detection methods and anticipate possible irregular activities, to support Search And Rescue operations and other civil and military operations.**

**Data sources:** weather and sea conditions, vessels characteristics, geographic positions, undersea informations or radar and satellite information.

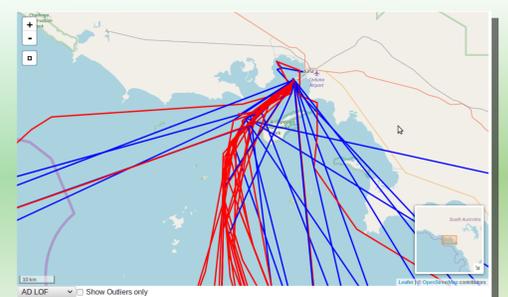


Figure 2: Maritime vessel tracks classification (Australia)

Supervisor: Professor Cláudia Martins Antunes (DEI-IST)

PhD Program in Computer Science and Engineering

\* Marisa - Maritime Integrated Surveillance Awareness is a project funded by the European Commission. This project has received funding from the European Union's Horizon 2020, research and innovation programme, under grant agreement N. 740698.