Multiple Kernel Learning for Heterogeneous Anomaly Detection: Algorithm and Aviation Safety Case Study

Santanu Das, Bryan Matthews, Ashok Srivastava and Nikunj Oza

NASA Ames Research Center

16th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington DC, July 25-28, 2010
Motivation

Flight Data Monitoring

Automatic identification and causal analysis of hazards from data streams with mixed attributes
Fleet wide analysis

Sequences D and continuous data streams C interactions

How to integrate all information in a **concise** and **intuitive** manner?

Compression, Feature extraction, Fusion, Anomaly detection
Mining Framework

Preprocessing discrete data

Preprocessing continuous data

Sequence

Kernel on switching

Model

SAX representation

Kernel on measurand

Compression

Fusion

Detection

SAX was invented by Eamonn Keogh and Jessica Lin, 2002

Multivariate symbolic sequences.
Multivariate continuous sequences.
Pair wise Similarity Measure

For more information, please see B. Schölkopf, A. Smola, R. Williamson, and P. L. Bartlett. New support vector algorithms. Neural Computation, 12, 2000, 1207-1245.

- Solves a convex and quadratic optimization problem.
- Can appropriately introduce a mixture of kernels in the convex cost function.
- Enables using non-linear kernel functions to learn complex separating planes.
- Results a model that can be used to classify new examples.

Normalized Longest Common Subsequence

Detector

One Class nu-SVMs

\[
K_0(f_i, f_j) = \frac{L(h(s_i, s_j))}{\sqrt{L(s_i) \times L(s_j)}}
\]
Optimization problem

One class SVMs training algorithms require solving the quadratic problem

**Dual form**

\[
Q_{\text{min}} = \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j \left( \sum_{\lambda} \beta_{\lambda} K_{i,j}^{\lambda} \right)
\]

Subject to:

\[
\sum_i \alpha_i = 1
\]

\[
\nu \in [0,1],
\]

\[
0 \leq \alpha_i \leq \frac{1}{l \nu}, \forall i
\]

\(\alpha\) : Lagrange multipliers of the primal QP problem

**Linear equality constraint**

**Control parameter**

**Bounds on design variables**
Anomaly scores

Decision boundary is determined only by margin and non-margin support vectors obtained by solving the QP problem

\[ h(\alpha, \beta, f_z, \rho) = \sum \alpha_i \left( \sum \beta_{\lambda} K_{i,z}^\lambda \right) - \rho \]

Datapoints with \( \alpha_k > 0 \) will be the support vectors

Indicator

Sign of \( h \): if negative - outlier
if positive - normal

Value of \( h \): degree of anomalousness
Experiment

Simulation data

Type 1 - (Missing event) Flaps were not extended to normal full deployment at landing.

Type 2 - (Extra event) Landing gear was retracted after being deployed on final approach.

Type 3 - (Out of order event) Gear deployed before initial flaps below flaps limit.

Type 4 - (Continuous anomaly) High bank angles or rate of descent below 1,000 ft.
Case study: FOQA anomaly detection

- The traditional methods cannot detect and monitor these anomalous activities that may have occurred simultaneously and are heterogeneous in nature.
Conclusion

What can we summarize?

**Performs**

... anomaly detection on multivariate mixed attributes where sequences may influence the system dynamics which is reflected on the continuous data streams.

**Highlights**

.. High detection rate on most operationally significant anomalies in fleet wide analysis on large datasets

.. Discover some “unknown unknowns”

**Application**

1. Support flights safety experts
2. Schedule maintenance
• Contact and feedback:
  – Santanu Das
    Santanu.Das-1@nasa.gov

• More resources on Dashlink website:

• Acknowledgement to Robert Lawrence, Dr. Irv Statler and Dr. Kanishka Bhaduri for insightful discussions and suggestions during this research.