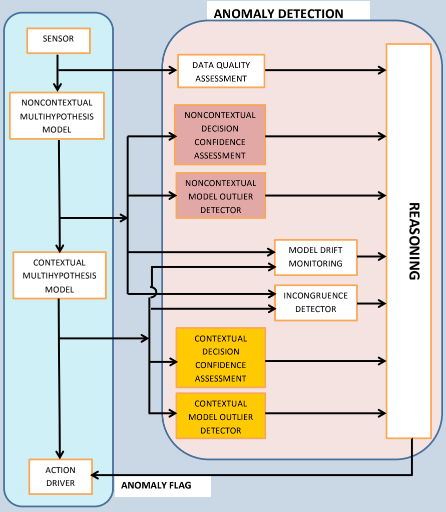
University Defence Research Collaboration (UDRC) Signal Processing in a Networked Battlespace

LSSC WP1: Automated Statistical Anomaly Detection and Classification in High Dimensions for the Networked Battlespace

WP Leaders: David Parish[1]*, Yulia* Hicks[2]*, Josef* Kittler[3]
Researchers: Francisco Aparicio Navarro[1]*, Ioannis* Kaloskampis[2]*, Cemre* Zor[3]
[1]*Loughborough Uni.(LU),* [2]*Cardiff Uni.(CU),* [3]*Uni. of Surrey(SU)*

Introduction:
This work package proposes the design of an automated statistical anomaly detection and classification system with advance methodology to be used in networked battlespace scenarios.


ANOMALY DETECTION
SENSOR
NONCONTEXTUAL MULTIHYPOTHESIS MODEL
CONTEXTUAL MULTIHYPOTHESIS MODEL
ACTION DRIVER
DATA QUALITY ASSESSMENT
NONCONTEXTUAL DECISION CONFIDENCE ASSESSMENT
NONCONTEXTUAL MODEL OUTLIER DETECTOR
MODEL DRIFT MONITORING
INCONGRUENCE DETECTOR
CONTEXTUAL DECISION CONFIDENCE ASSESSMENT
CONTEXTUAL MODEL OUTLIER DETECTOR
REASONING
ANOMALY FLAG

SU: Incongruence Detection for Statistical Anomaly Detection
LU: Statistical Anomaly Detection in Communication Networks
CU: Anomaly Detection in Video

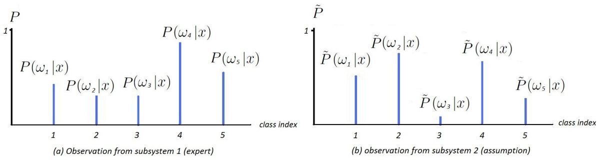
Objectives:

- Developing algorithms for automatic detection of anomalies from multidimensional, undersampled, non-complete datasets and unreliable sources
- Identifying the nature and statistical characteristics of these anomalies once they have been detected in a high dimensional complex network environment
- Determining the "normal" data characteristics and changes in "normal" behaviour to provide an acceptable balance between false positive/negative performance and computational cost
- Using data quality and ambiguity measures to ensure the models of normality are not corrupted by unreliable and ambiguous data

Incongruence Detection

Aims to aid the detection of anomaly in sensor data processed by a complex decision making system. Focuses on:

- Comparing the outputs of two classifiers with a view to detecting statistical anomaly in sensor data
  - The nature/nuance of anomaly should subsequently be identified based on a detailed analysis of the classifier outputs
- Analysing measures of surprise in Bayesian Analysis, Histogram Consistency / Similarity Tests, Bayesian Surprise
- Development of an alternative method which focuses on the dominant hypotheses flagged by the two experts: Max Difference(Δ_{max})



(a) Observation from subsystem 1 (expert)

(b) observation from subsystem 2 (assumption)

Statistical Anomaly Detection in Communication Networks

Aims to improve the 'Data Quality Assessment' module in the proposed anomaly detection system. Focuses on:

- Developing novel methodologies that automatically generate labeled network traffic datasets, using the outcome of an unsupervised IDS
  - In the current IDS system, the detection system provides levels of belief in three hypotheses: Normal, Attack, Uncertainty
  - The beliefs are fused using Dempster-Shafer Theory of Evidence
  - Labeling according to the final results of an unsupervised IDS
  - Considering only strong belief results (large difference between belief in Normal and Attack produces), new datasets fully composed of correctly labeled instances are created
- Feature selection techniques to automatically select the most appropriate set of metrics and include new metrics
  - A Genetic Algorithm (GA) based approach
- Incorporating information from different levels of the domain knowledge and contextual information

Anomaly Detection in Video

Aims to develop an accurate, data-driven anomaly detection method which is computationally efficient and which incorporates domain knowledge to detect anomalies in video. Focuses on:

- Searching and evaluating datasets: VIRAT, NGSIM Peachtree Street, Gun-Point, Unusual Crowd Activity, Technion, Thermal Imaging, In-house datasets
- Performing low-level feature extraction on the datasets
  - Using HOG features, colour space features
- Video analysis using Incremental Learning
  - Evolving statistical models: GMMs, HMMs
  - Combining low-level statistical models of video features with high-level event models for anomaly detection.

Future Work: Building a solid general framework for surprise measure thresholding including error sensitivity analysis by SU, including contextual/higher level information and considering non-attack anomalies by LU, and combining low-level statistical models of video features with high-level event models for anomaly detection by CU are planned as future work within this workpackage. Collaborations within and inbetween workpackages for advances and applications on feature selection, incremental learning and surprise measure thresholding are proposed.

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) and Dstl.