University of Glasgow

School *of* Computing Science

Essence: Pervasive & Distributed Computing

# Adaptive Novelty Detection over Contextual Data Streams at the Edge using One-class Classification


Olga Jodelka, Christos Anagnostopoulos, Kostas Kolomvatsos

School of Computing Science, University of Glasgow, UK


WORLD CHANGING GLASGOW


ICICS 2021 @ Valencia, Spain (May 24-26, 2021)

# Motivation

- **Concept drift** – fundamental problem in Statistical Learning related to data streams
- Improvement of Sequential Learning due to recognition of **novelties** and elimination of **anomalies**
- Maintaining data quality in **Edge Computing** environments

# Contributions

- **Adaptive** mechanism for novelty detection
- Methodology to investigate impact of time and frequency on **model re-training**
- Technique for **automated** creation of labeled dataset
- Technique to cope with **adaptability changes** imposed by concept drift
- Comprehensive experimental evaluation

# OCSVM: One-Class SVM

School *of* Computing Science
Essence: Pervasive & Distributed Computing

- Multivariate Greenhouse datasets:
  - Collected streams of humidity, air temperature and soil temperature every 3 minutes
  - Generate two 2D datasets with features with strongest correlation
  - Soil & Air Temperature datasets (D1)
  - Humidity & Air Temperature dataset (D2)
- OCSVM
  - Learns a **decision boundary** that separates <u>inliers</u> from <u>novelties</u>
  - Parameters ν and γ control the shape of the boundary and number of Support Vectors

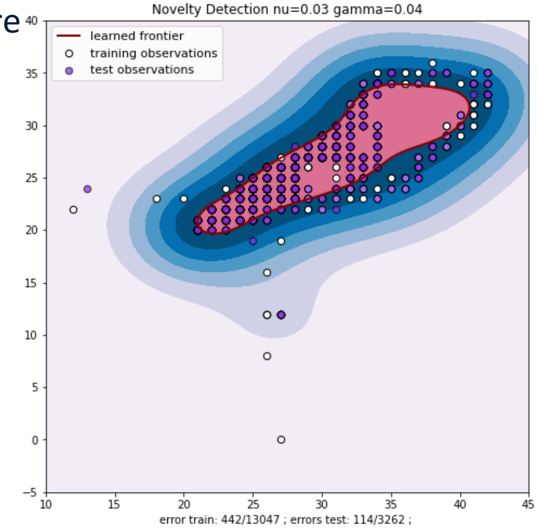

Fig.2 Learnt frontier D1

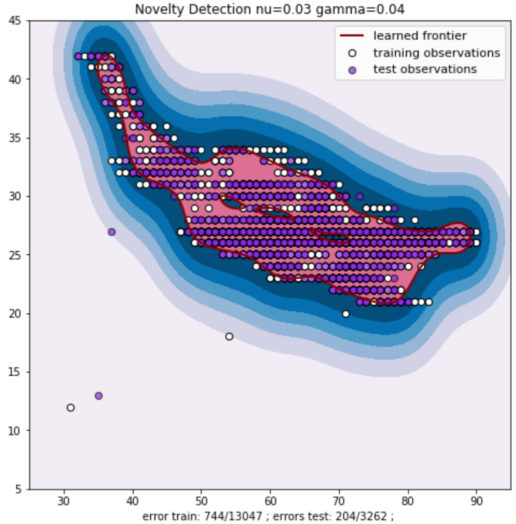

Fig.1 Learnt frontier D2

# Continuous Model Adaptation (CMA)

- Local training of the OCSVM model on the **window** of size **W**
- Each data point may have an **impact on** the underlying concepts
- **Retrain** the model *every* single data point
- Label the data vector as **novelty** (-1) or **inlier** (1)
- Monitor the number of **unnecessary re-trainings** by comparing the labels of data points *before* and *after* retraining

# Fixed-frequency Model Adaptation (FMA)

- Local Training of the OCSVM model on the **window** of size **W**
- **Retrain** the model with fixed-frequency H > 1
- **Case 1:** frequency H < W
  - During retraining some previous knowledge is still **available**
- **Case 2:** frequency H ≥ W
  - Previous model is completely **forgotten**
- FMA **reduces** model complexity and **increases** computational efficiency
- Initialization of parameter H requires knowledge of the data and the concepts they represent.

# Performance Evaluation

- **Unnecessary** re-trainings (*n*)
- **Total** re-trainings
- Novelty Counter (*c*)
- Average #SVs (E[m]≈v)
- Number of **marginal non-inliers** (*l*)

| Metric | CMA | FMA |
|---|---|---|
| Unnecessary retrainings $n$ | D1:12341; D2:11772 | D1:18; D2:7 |
| Total retrainings | D1:15309; D2:14698 | D1:30; D2:15 |
| Novelty counter $c$ | D1:2605; D2:3537 | D1.Case-1:5542<br>D1.Case-2:6013<br>D2.Case-1:6084<br>D2.Case-2:6480 |
| Average #SVs $\mathbb{E}[m] \approx \nu$ | D1:36; D2:32 | D1.Case-1:36<br>D1.Case-2:37<br>D2.Case-1:33<br>D2.Case-2: 33 |

# Performance Evaluation

School *of* Computing Science
Essence: Pervasive & Distributed Computing

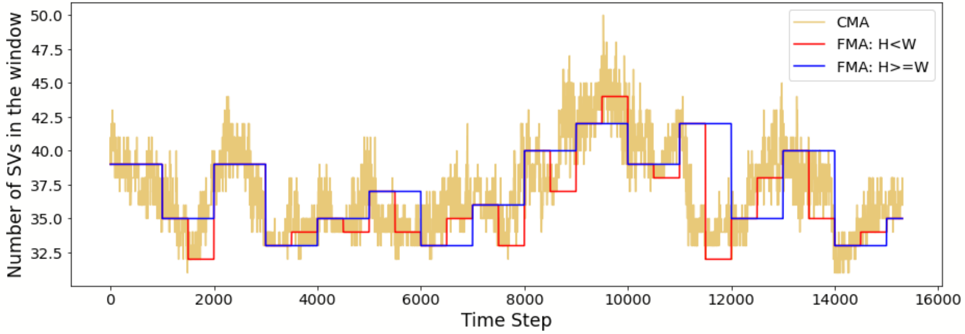


Fig.3 SVs in D2 dataset

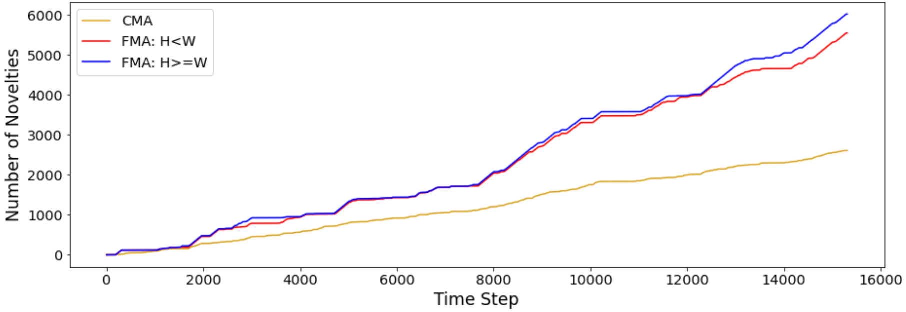


Fig.4 Novelties in D2 dataset

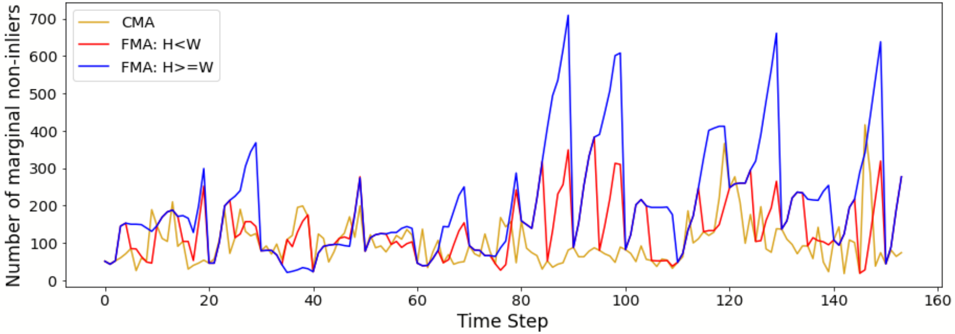


Fig.5 Marginal non-inliers in D2 dataset

# Conclusions

- **Higher number of novelties** identified by FMA occurs due to **sparse re-trainings**; CMA is better in incorporating novelty patterns into the model.
- Based on the number of SVs, the models have comparable complexity, but the amount of memory resources required by FMA are predictable.
- Low amount of marginal non-inliers in CMA indicates a**bility to represent the data accurately**; in FMA non-inliers indicate the **rate** at which the model **turns obsolete**

# Conclusion & Future Work

- Mechanism successfully identifies novelties, adapting to concept drifts in a resource-efficient way
- **Future Work**: investigation of other One-Class (SVM) variants and confidence-driven novelty detection and forecasting.

University of Glasgow

School *of* Computing Science

Essence: Pervasive & Distributed Computing

# Thank you

**Olga Jodelka**

http://www.dcs.gla.ac.uk/essence/