Tailored ensemble anomaly detection for Internet disruptions

Supplementary material

VANDERSANDEN, Mike; Beerts, Jelle; Kreitem, Hanna; Phokeer, Amreesh; LAMOTTE, Wim & QUAX, Peter (2026) Tailored ensemble anomaly detection for Internet disruptions. In: Shakshuki, Elhadi (Ed.). , p. 277 -285.

DOI: 10.1016/j.procs.2026.04.037
Handle: http://hdl.handle.net/1942/49400

Tailored ensemble anomaly detection for Internet disruptions

M. Vandersanden, J. Beerts, H. Kreitem,
A. Phokeer, W. Lamotte, P. Quax

DIGITAL FUTURE LAB
UHASSELT

Internet Society Pulse

The Internet: foundation for global communication

The Internet: object of disruptions

Examples

APPLE'S NEW IPHONE UPDATE IS RESTRICTING INTERNET FREEDOM IN THE UK

BIG BROTHER WATCH TEAM / APRIL 9, 2026

https://bigbrotherwatch.org.uk/blog/apples-new-iphone-update-is-restricting-internet-freedom-in-the-uk/

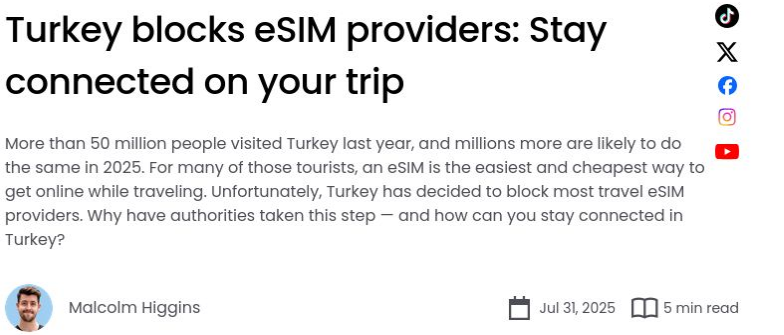
https://www.forbes.com/sites/kevinholdenplatt/2025/10/31/russia-wages-war-on-the-internet-in-ukraine-but-resistance-is-winning/

https://www.cnbc.com/2022/03/22/elon-musk-spacex-thousands-of-starlink-satellite-dishes-sent-to-ukraine.html

Vlaamse Programmeer Wedstrijd 2018

https://www.vlaamseprogrammeerwedstrijd.be/2018/

https://restofworld.org/2026/iran-starlink-internet-shutdown/



https://saily.com/blog/turkey-esim-ban/

DIGITAL FUTURE LAB UHASSELT

Internet Society Pulse

Internet measurement platforms

Measuring for performance and accessibility

With
- Standardized tests
- Mimicking applications

Through
- Through data centers
- Through volunteers

May include disruption notifications

OONI

RIPE NCC
RIPE Atlas

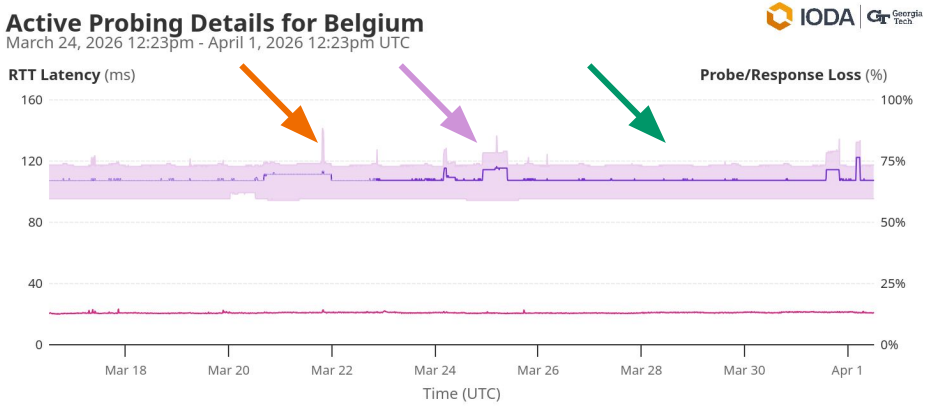
DIGITAL FUTURE LAB
UHASSELT

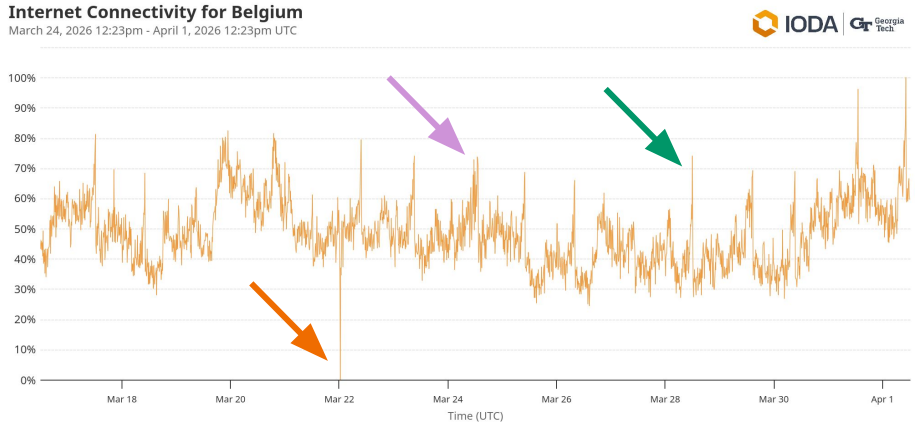
Internet Society
Pulse

What characterizes a disruption?

Measurement anomalies: disruption or irregularity?

Disruption scope: wide-spread or targeted?





DIGITAL FUTURE LAB UHASSELT

Internet Society Pulse

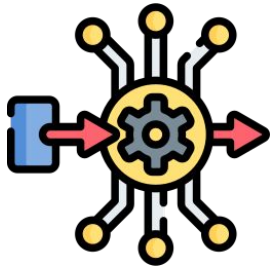
Existing detection methods

Heuristics:

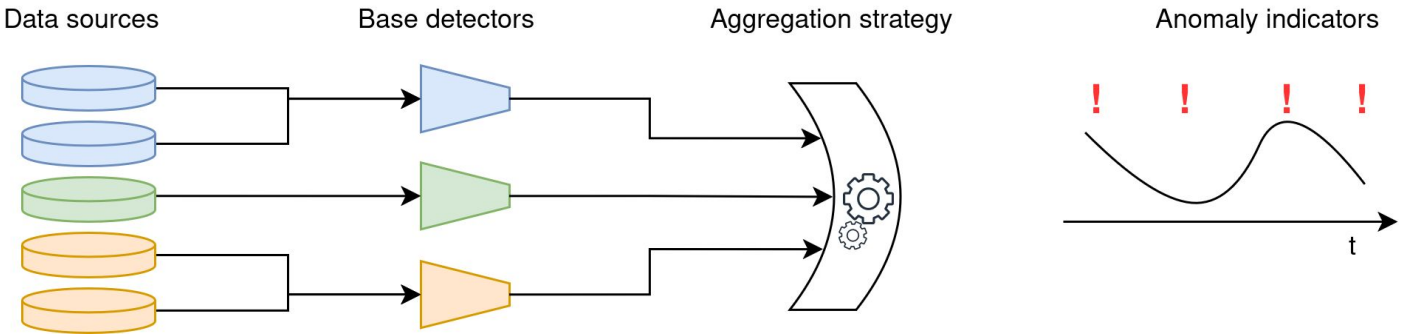
- Intuitive, straight-forward
- Not applicable to complex multivariate data

Machine Learning:

- Can be applicable to multivariate data sources
- Can be trained or plug-and-play
- Black box: explainable through post-hoc techniques



Proposed solution: Tailored Ensemble Approach



- Consider individual data sources
- Apply simple base detectors
- Use intuitive aggregation strategies

Detectors

Different univariate base detectors target different anomaly types

- Intuitive or default parameters
- Influenced by existing methods
- Binary results

Compare with multivariate approaches

Category	Anomaly conditions
Outlier	Deviating from regular measurements
Shift & Spike	Time-dependent irregularities
Seasonality	Pattern changes
Multivariate	Considers multiple data sources at once

Aggregation strategies

Voting strategies

- Intuitive
- Match binary detections

Strategy	Acceptance conditions
Majority voting	Votes cross threshold
Weighted voting	Votes are weighted and cross threshold
Time window voting	Votes are considered within a specified interval and cross threshold

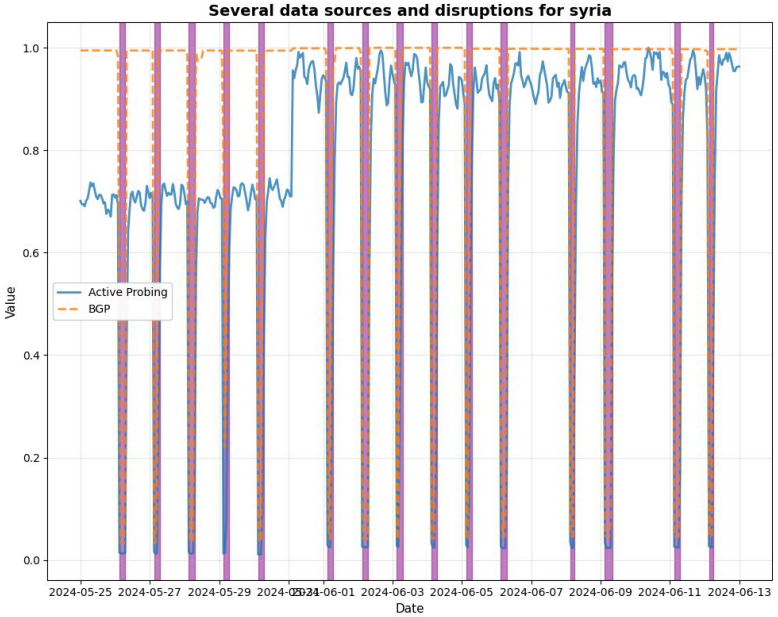
Real-world validation

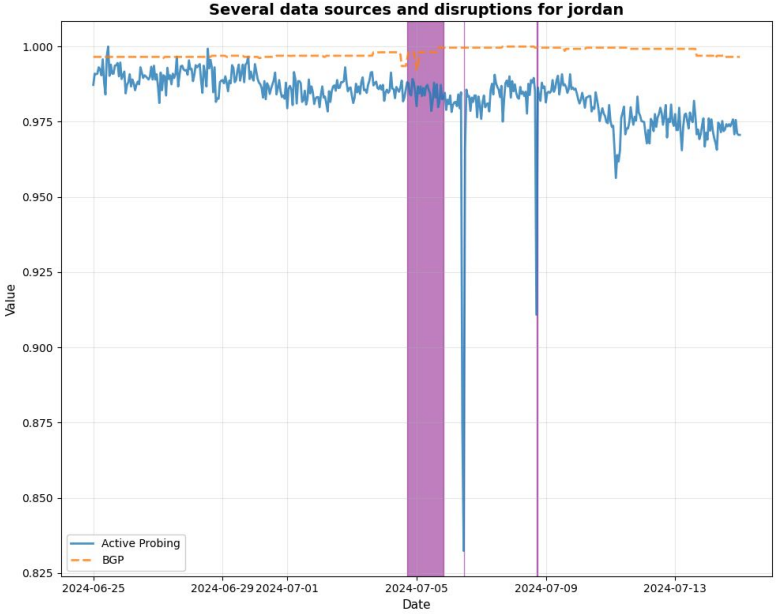
4 use cases: areas prone to intentional blocking during exam periods

varied characteristics

- Short and long periods
- Disruption-free periods
- Missing data sources

Use cases examples





Results: tailoring base detectors

Evaluating all base detectors against the cases

Set of tailored detectors, matching data source characteristics

- Overall above average performance for all cases
- Not always the best detector
- Not always the most intuitive option

Can be compared against set of best detectors

Results: tailored vs. best performance

- Tailored detectors perform similarly well to best detectors
- Weighted voting is more robust than majority voting

F1	Majority (tailored)	Majority (best)	Weighted (tailored)	Weighted (best)
Case 1	0.85	0.91	0.51	0.55
Case 2	0.87	1.00	0.71	0.86
Case 3	0.0	0.68	0.69	0.48
Case 4	0.0	0.50	0.28	0.25

Results: competitive performance

Ensemble approach performs competitive and robust compared to plug-and-play multivariate models

F1	Ensemble: weighted voting	Elliptic envelope	Isolation Forest
Case 1	0.51	0.52	0.56
Case 2	0.71	1.00	0.86
Case 3	0.69	0.64	0.55
Case 4	0.28	0.21	0.23

Results: competitive performance

F2 improves over F1: ensemble approach achieves better recall than precision

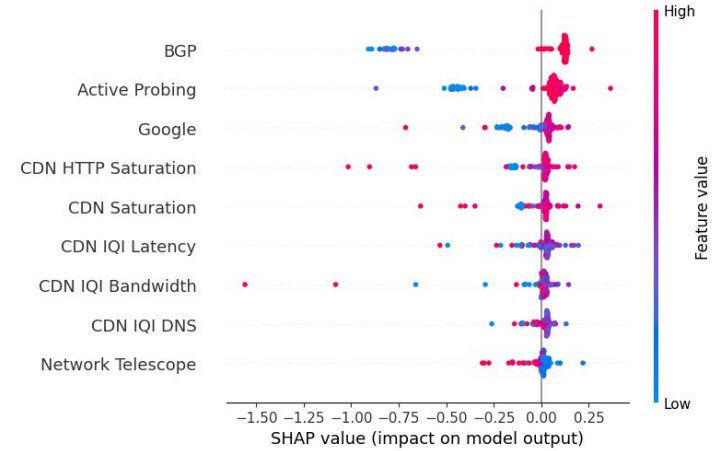
F2	Ensemble: weighted voting	Elliptic envelope	Isolation Forest
Case 1	0.73	0.69	0.76
Case 2	0.86	1.00	0.94
Case 3	0.79	0.55	0.51
Case 4	0.43	0.36	0.38

Results: ad hoc explainability

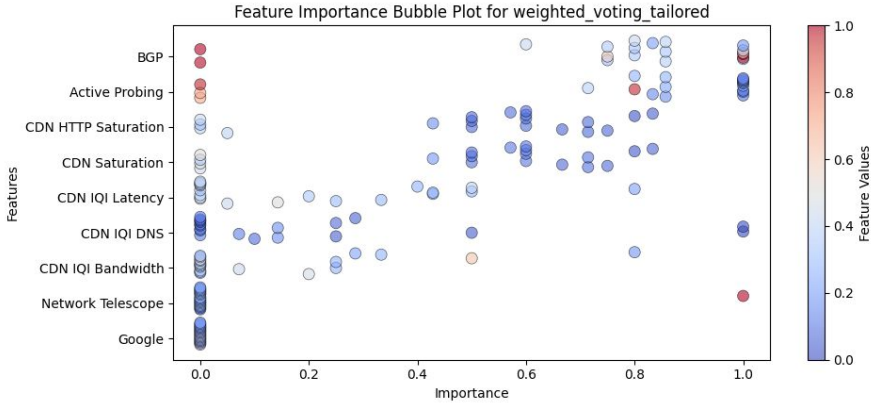
Exposed base detector results facilitate SHAP-inspired visualization

without computationally intensive post-hoc calculations

limited to positive influences



SHAP feature importance

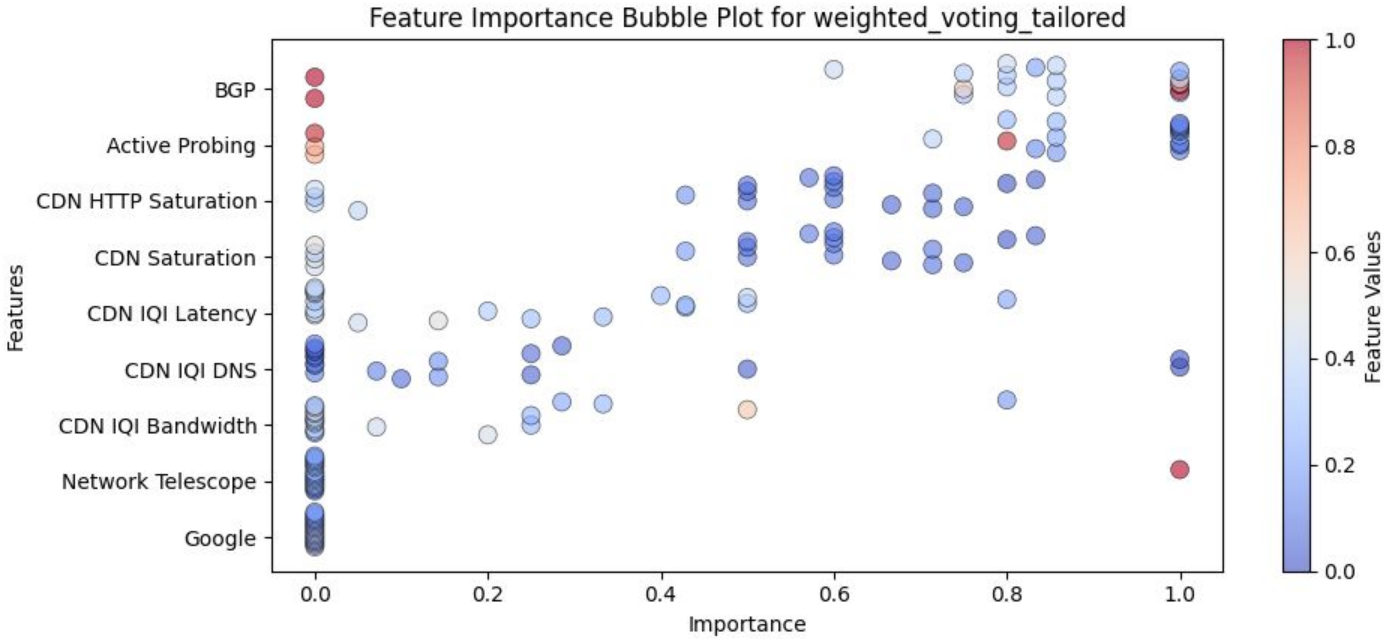


Feature importance of proposed approach

DIGITAL FUTURE LAB UHASSELT

Internet Society Pulse

Results: ad hoc explainability



Conclusion

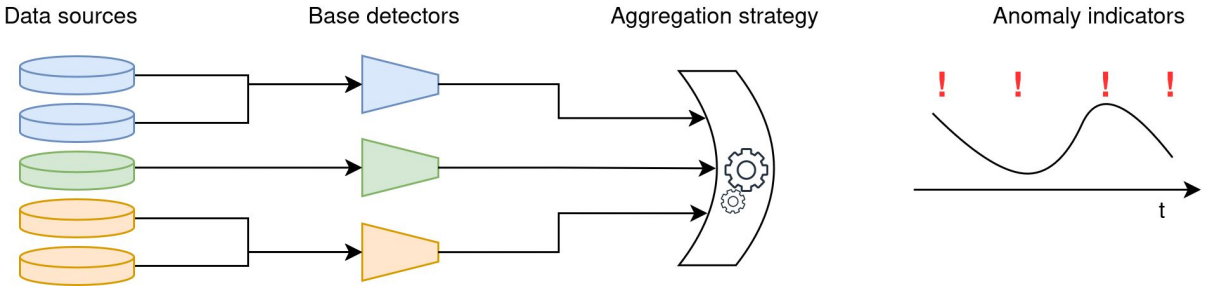
Tailored ensemble anomaly detection approach

- Plug-and-play
- Similar performance
- Robustness
- Inherent explainability

Next steps

- Fingerprinting disruptions
- Advanced or adaptive aggregation strategies

Tailored ensemble anomaly detection for Internet disruptions



Simple base detectors that consider individual data sources

Intuitive aggregation strategies

Competitive and robust performance

Ad hoc explainability

Code and data @ https://github.com/EDM-Research/ExplainableInternetMeasurements

Internet Society Pulse

Thank you!

Want to reach out?

Supported by the Hasselt University Special Research Fund (BOF22OWB17).

This project was supported by the Internet Society Pulse Research Fellowship.

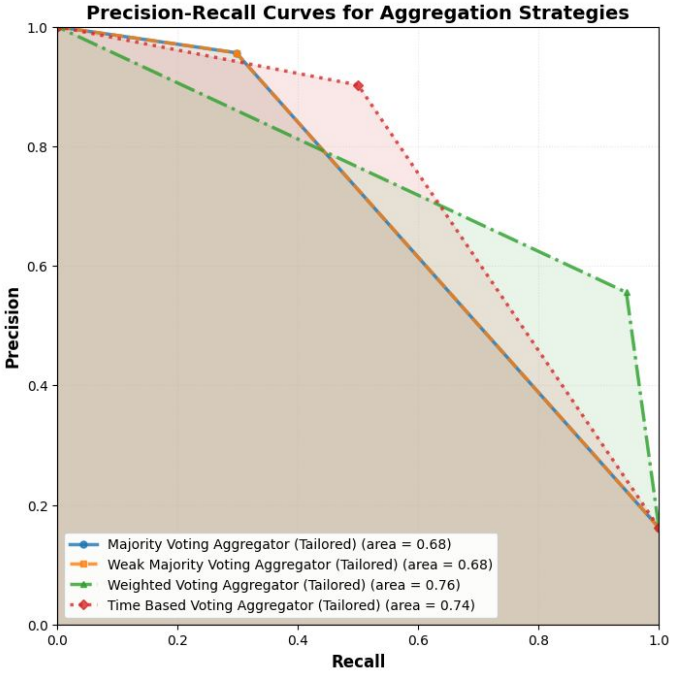
Icons sourced from Flaticon.com

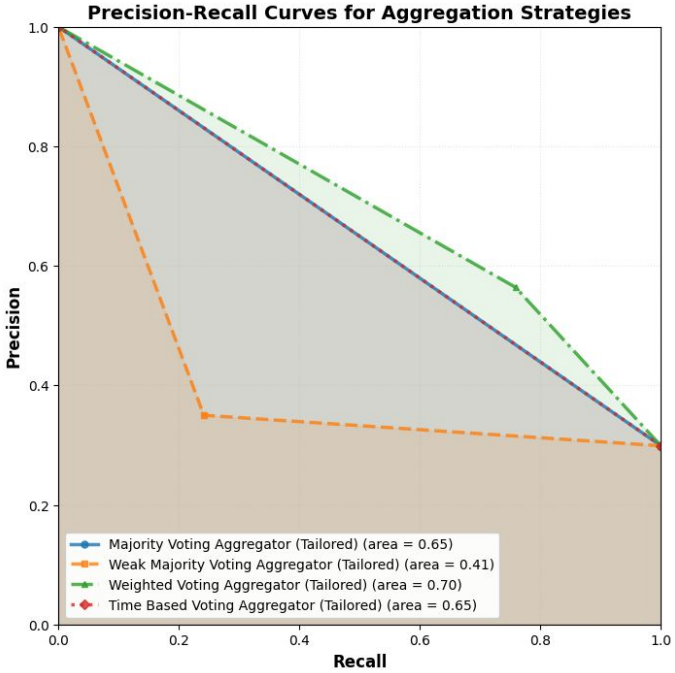
DIGITAL FUTURE LAB

UHASSELT

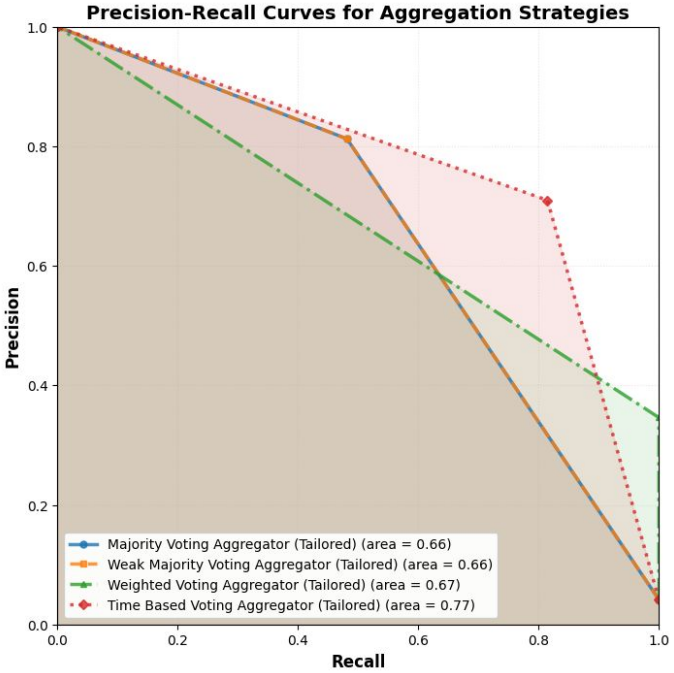
Internet Society Pulse

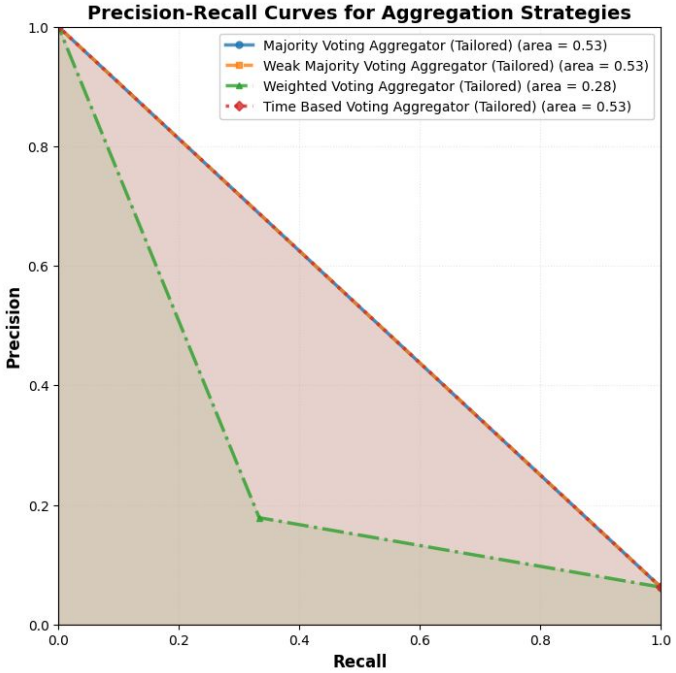
AUC-PR


Precision-Recall Curves for Aggregation Strategies
Precision
Recall
Majority Voting Aggregator (Tailored) (area = 0.68)
Weak Majority Voting Aggregator (Tailored) (area = 0.68)
Weighted Voting Aggregator (Tailored) (area = 0.76)
Time Based Voting Aggregator (Tailored) (area = 0.74)


Precision-Recall Curves for Aggregation Strategies
Precision
Recall
Majority Voting Aggregator (Tailored) (area = 0.65)
Weak Majority Voting Aggregator (Tailored) (area = 0.41)
Weighted Voting Aggregator (Tailored) (area = 0.70)
Time Based Voting Aggregator (Tailored) (area = 0.65)

AUC-PR


Precision-Recall Curves for Aggregation Strategies
Precision
Recall
Majority Voting Aggregator (Tailored) (area = 0.66)
Weak Majority Voting Aggregator (Tailored) (area = 0.66)
Weighted Voting Aggregator (Tailored) (area = 0.67)
Time Based Voting Aggregator (Tailored) (area = 0.77)


Precision-Recall Curves for Aggregation Strategies
Precision
Recall
Majority Voting Aggregator (Tailored) (area = 0.53)
Weak Majority Voting Aggregator (Tailored) (area = 0.53)
Weighted Voting Aggregator (Tailored) (area = 0.28)
Time Based Voting Aggregator (Tailored) (area = 0.53)