Available as: Zoppi, T., Ceccarelli, A., Puccetti, T., & Bondavalli, A. (2023). Which Algorithm can Detect Unknown Attacks? Comparison of Supervised, Unsupervised and Meta-Learning Algorithms for Intrusion Detection. Computers & Security, 103107.

Which Algorithm can Detect Unknown Attacks?

Comparison of Supervised, Unsupervised and Meta-Learning Algorithms for Intrusion Detection

Tommaso Zoppi, Andrea Ceccarelli, Tommaso Puccetti, Andrea Bondavalli

RCL Group – University of Florence - Italy

e-mail: tommaso.zoppi@unifi.it

RCL
RESILIENT COMPUTING LAB

UNIVERSITÀ DEGLI STUDI FIRENZE
DIMAI
DIPARTIMENTO DI MATEMATICA E INFORMATICA "ULISSE DINI"

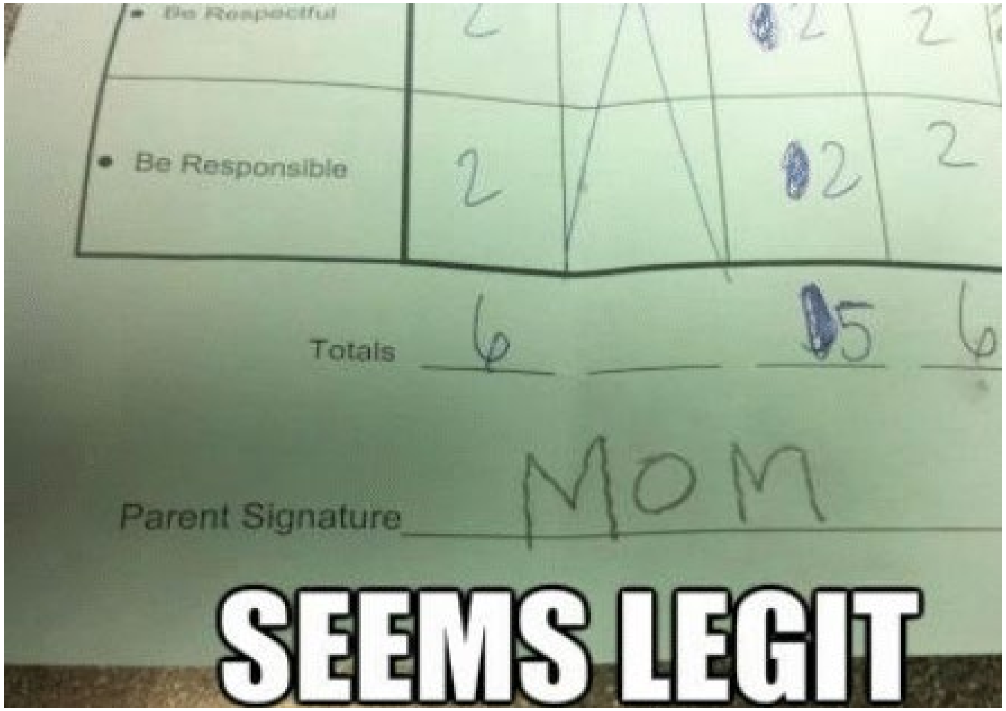
Traditional Intrusion Detectors

- Typical means to attain security mainly revolve around two main approaches:
  - Rule-based, Invariant-Based or
  - Signature-based

Images from https://blogs.vmware.com/security/2016/11/next-generation-antivirus-ngav.html

Signature-based Detection

- Network or host activity is analyzed to seek for matching attack patterns (signatures).
  - If the current behavior of the system matches one or more attack signatures (or rules), an alert is raised



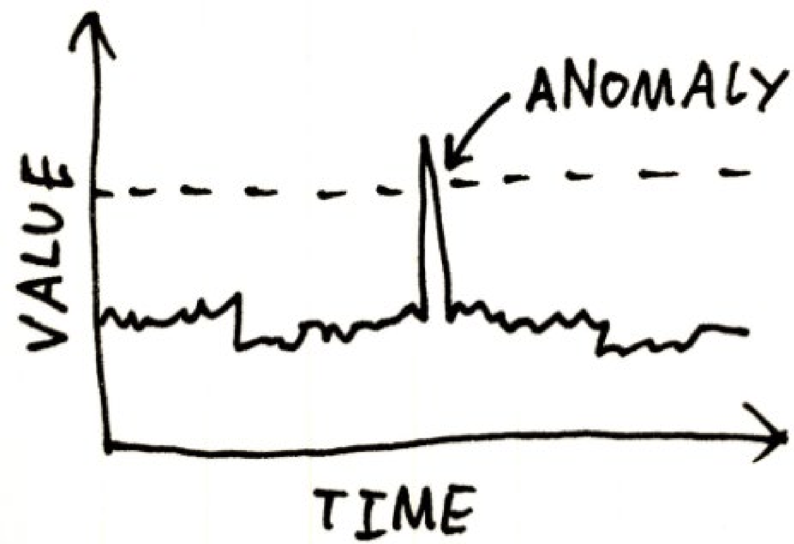
What about Unknown Threats?

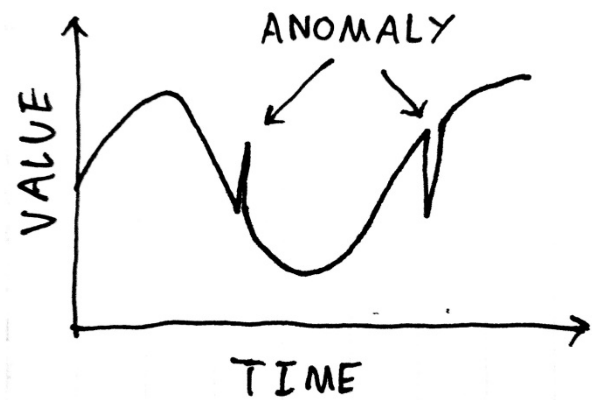
- Research and Practice found ways to defend against specific attacks
  - Mostly rule, signature-based or using supervised learning

- But what about unknown attacks or errors?
  - Unknown attacks: no rule / signature available yet

Anomaly-based Detection

- It allows to identify patterns in data streams and operations which are different from those expected, and label them as anomalies
  - They do not need signatures of anomalies
  - Instead, they characterize what is normal and act accordingly





(Un)Supervised Algorithms

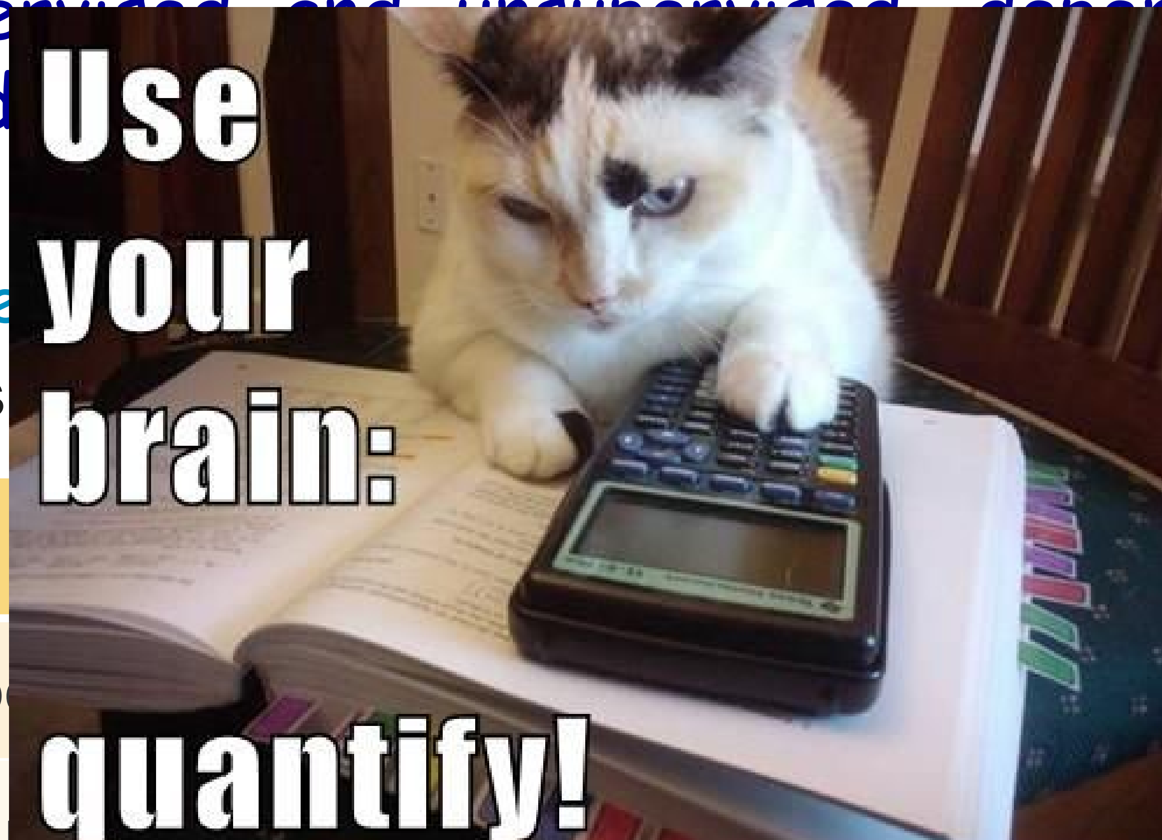
- ML Algorithms are usually partitioned as (semi)supervised and unsupervised, depending on their need of labels in the training data
  - Supervised Algorithms very well known
  - Unsupervised Algorithms
    - Do not assume any detailed knowledge of anomalous events

	Known Issue	Unknown Issue
Supervised	**Very Good!**	**Potentially Bad**
Unsupervised	**Average**	

(Un)Supervised Algorithms

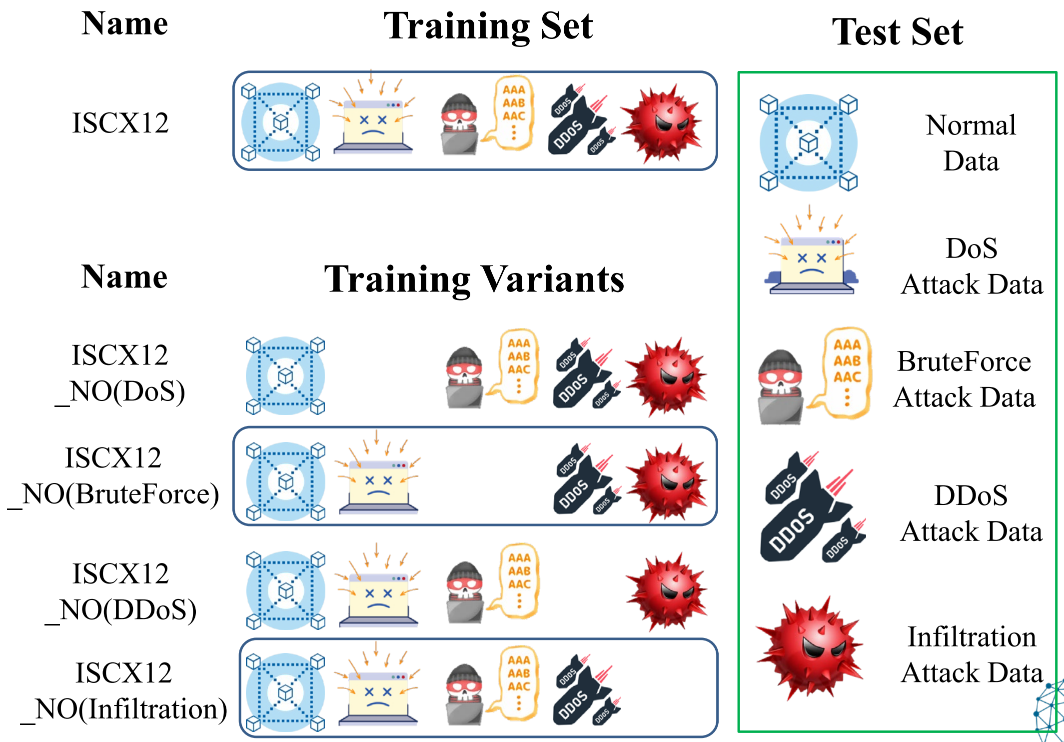
- ML Algorithms are usually partitioned as (semi)supe... ding on their need...
  - Supervised
  - Unsupervise...
    - Do not ass... us events

	...sue
Sup...	Bad
Unsupervised	Average



Evaluation Plan

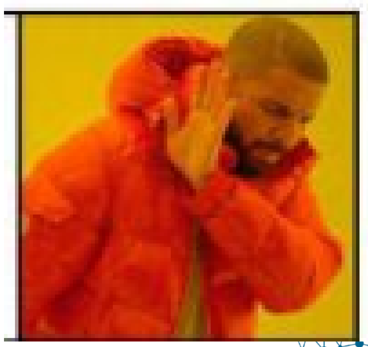
- How to evaluate detection of unknowns?
  - Unknown attack: not in the training set but in the test set
- Therefore we created variants of each dataset
  - where a given attack is unknown



Metrics For Benchmarking

- Accuracy / F-Measure (F1)

- Example: System where 95% of data is normal
  - "Optimistic silly detector": always outputs "normal"
  - TP = 0, TN = 95%, FP = 0, FN= 5%
  - Accuracy = 95%



RCL
RESILIENT COMPUTING LAB

Ital-IA
ITALIA INTELLIGENZA ARTIFICIALE
cini National Lab AIIS

Metrics For Benchmarking

- Accuracy / F-Measure (F1)
- Matthews Coefficient (MCC)
  - A bit complex, but fits also unbalanced datasets

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$$

- Example: System where 95% of data is normal
  - "Optimistic silly detector": always outputs "normal"
  - TP = 0, TN = 95%, FP = 0, FN= 5%
  - Accuracy = 95%
  - MCC = 0 (random guessing)

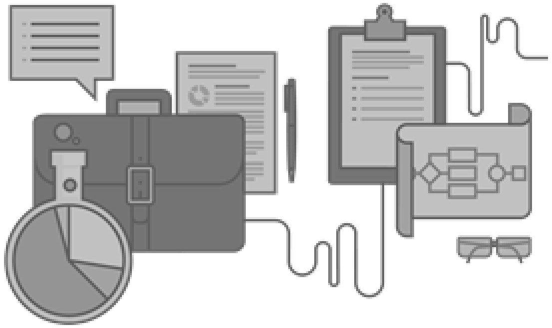
RCL
RESILIENT COMPUTING LAB

Ital-IA
ITALIA INTELLIGENZA ARTIFICIALE
cini National Lab AIIS

Metrics For Benchmarking

- Accuracy / F-Measure (F1)
- Matthews Coefficient (MCC)
  - A bit complex, but fits also unbalanced datasets

$$\mathrm{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$$



- Recall (or Coverage)
- Recall-Unknown
  - Recall, but considering only zero-day attacks

ML Algorithms to be Benchmarked

- Supervised
  - Tree-based: Decision Tree, ADABoost, Gradient Boosting, XGBoost, Random Forests
  - Statistical: Naïve Bayes, LDA, Logistic Regression
  - Others: kNN, SVM, MLP
- Supervised - Deep learning
  - FastAI, AutoGluon, TabNet, Custom PyTorch
- Unsupervised
  - Clustering: K-Means, G-Means, LDCOF
  - Distance: ODIN, COF, LOF, SDO, FastABOD
  - Others: iForest, HBOS, One-Class SVM, SOM

Evaluation of Supervised Algs. (I)

- Best MCC scores of supervised algorithms
  - With respect to best unsupervised classifiers
  - Huge difference w.r.t. Unsupervised

Dataset	Supervised Deep Learning	Supervised (Non-Deep)	Unsupervised	Difference Sup - Unsup
ADFANet	0.9943	**0.9983**	0.9837	0.0146
AndMal	0.3895	**0.6458**	0.5503	0.0955
CICIDS17	0.9954	**0.9996**	0.6511	0.3485
CICIDS18	**0.9286**	0.9281	0.8277	0.1009
CIDDS	**0.9924**	0.9754	0.8026	0.1898
IoT_IDS	0.9965	**0.9998**	0.9739	0.0259
ISCX	0.8763	**0.8927**	0.7921	0.1006
NSLKDD	0.9830	**0.9888**	0.8384	0.1504
SDN20	0.9994	**0.9998**	0.8818	0.118
UGR	**0.9426**	0.9272	0.8161	0.1265
UNSW	0.8904	**0.9369**	0.8849	0.052

RESILIENT COMPUTING LAB

ITALIA INTELLIGENZA ARTIFICIALE
National Lab AIIS

Evaluation of Supervised Algs. (II)

- Now, let's look at Recall-Unk
  - Supervised against unsupervised classifiers
  - Negative value in the plot means that unsupervised classifiers outperform supervised in detecting unknowns

Need of Unsupervised Meta-Learning

- Overall, Supervised > Unsupervised
  - Except for Recall-Unk (detection of zero-days)

BUT BUT BUT

- Top-Performing Supervised Algorithms use complex learning strategies (meta-learning)
  - Random Forests -> Bagging
  - XGBoost -> (extreme) gradient boosting

why shouldnt we apply those to unsupervised algorithms?

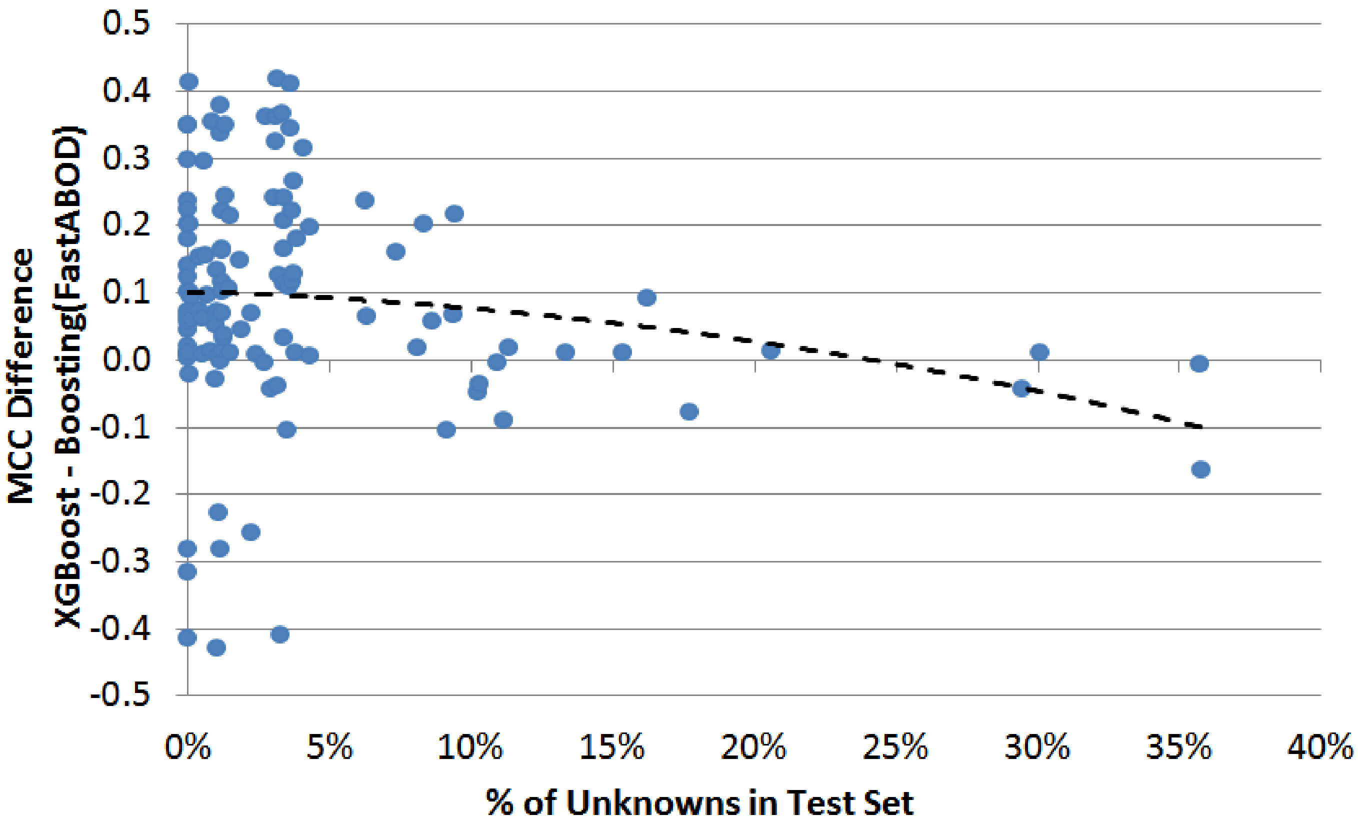
Unsupervised Meta-Learning

- As such, we built bagging and boosting ensembles of unsupervised algorithms

Dataset	Unsupervised Regular	Unsupervised Bagging	Unsupervised Boosting	Improvement w Boosting
ADFANet	0.9837	0.9867	**0.9916**	0.0079
AndMal	**0.5503**	0.4290	0.5277	-0.0226
CICIDS17	0.6511	0.6706	**0.8981**	0.247
CICIDS18	0.8277	0.8369	**0.8460**	0.0183
CIDDS	0.8026	0.8234	**0.9734**	0.1708
IoT_IDS	0.9739	**0.9902**	0.9896	0.0157
ISCX	0.7921	**0.8447**	0.8202	0.0281
NSLKDD	0.8384	0.8925	**0.9101**	0.0717
SDN20	0.8818	0.9441	**0.9481**	0.0663
UGR	0.8161	0.8445	**0.8745**	0.0584
UNSW	0.8849	0.8457	**0.8927**	0.0078

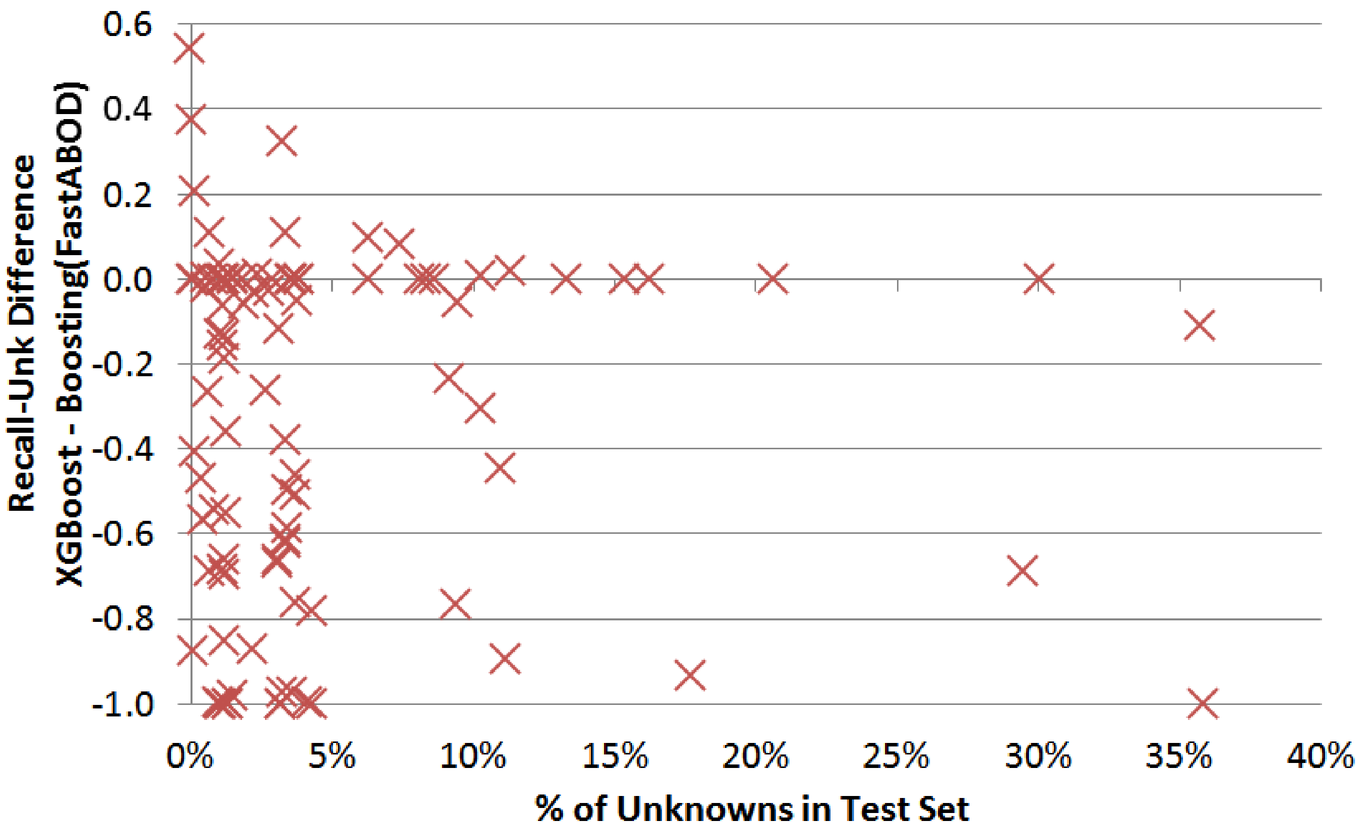
Compare w XGB - MCC

- Scores of XGB (sup) vs FastABOD (w boosting)
- MCC of XGB decays the more unknowns happen
  - Up to a point in which Unsup > Sup



VS – Recall-Unk Comparison

- Also, Recall-Unk of FastABOD is far better than those of XGBoost
  - And this gets more evident the more zero-days appear



Takeovers

- There is no "silver bullet" algorithm to plug into a system for excellent intrusion detection capabilities
- Deep Learning algorithms do not really fit the analysis of tabular data coming from network monitoring
  - XGBoost > Deep Learners (FastAI, TabNet, Autogluon …)
- XGBoost (sup) shows good overall detection capabilities
- Applying meta-learning dramatically reduces misclassifications of unsupervised algorithms
  - Up to a point in which FastABOD > XGBoost
  - But only if we expect zero-days to happen very frequently!

Q&A Time

Paper available as: Zoppi, T., Ceccarelli, A., Puccetti, T., & Bondavalli, A. (2023). **Which Algorithm can Detect Unknown Attacks? Comparison of Supervised, Unsupervised and Meta-Learning Algorithms for Intrusion Detection**. Computers & Security, 103107.