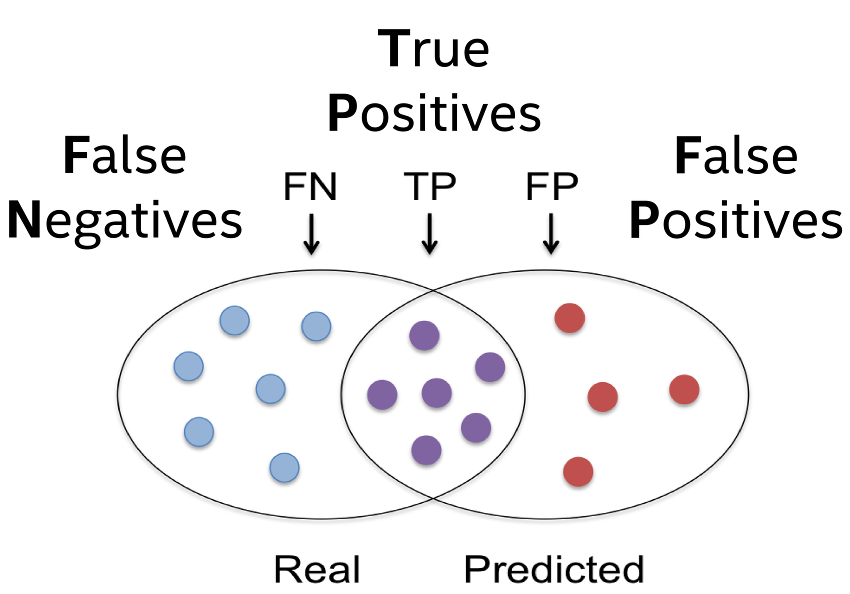
PRECISION AND RECALL FOR RANGE-BASED ANOMALY DETECTION

Tae Jun Lee (Microsoft), Justin Gottschlich (Intel Labs), Nesime Tatbul (Intel Labs and MIT), Eric Metcalf (Brown University), Stan Zdonik (Brown University)

The Problem

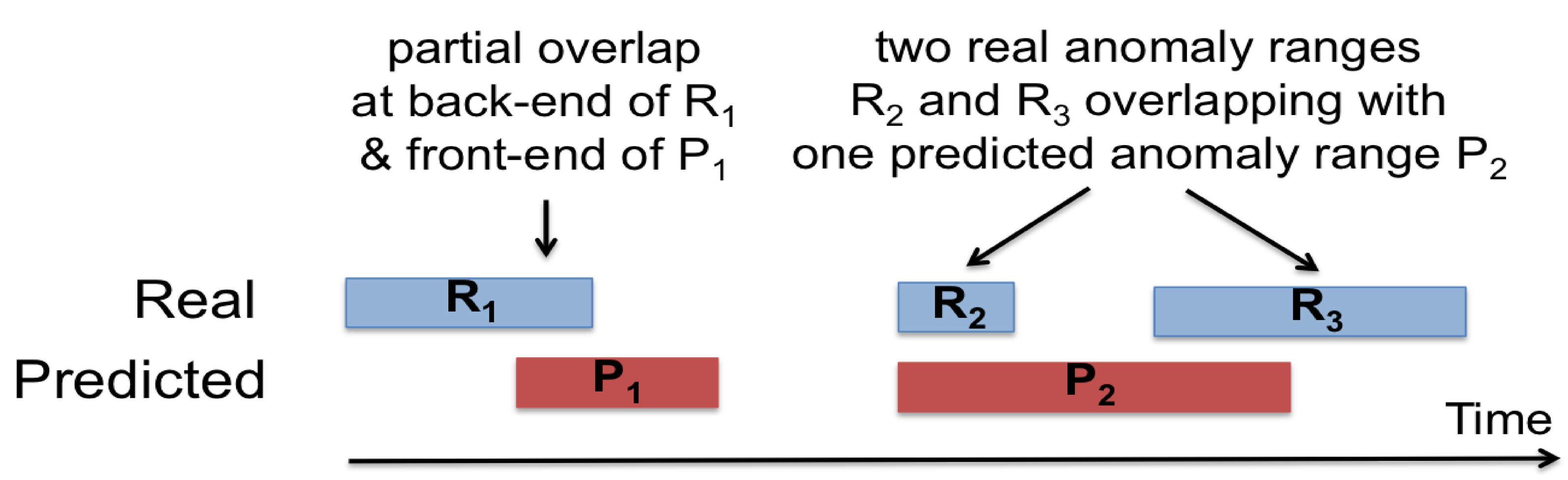
- Anomaly detection (AD) is the process of finding patterns that do not conform to expected behavior.
- Time series anomalies occur **over a range of time**. Time bias is **domain-specific**.
- **Goal: To design a model for evaluating, ranking, comparing the classification accuracy of time series AD algorithms.**

Point-based Anomalies



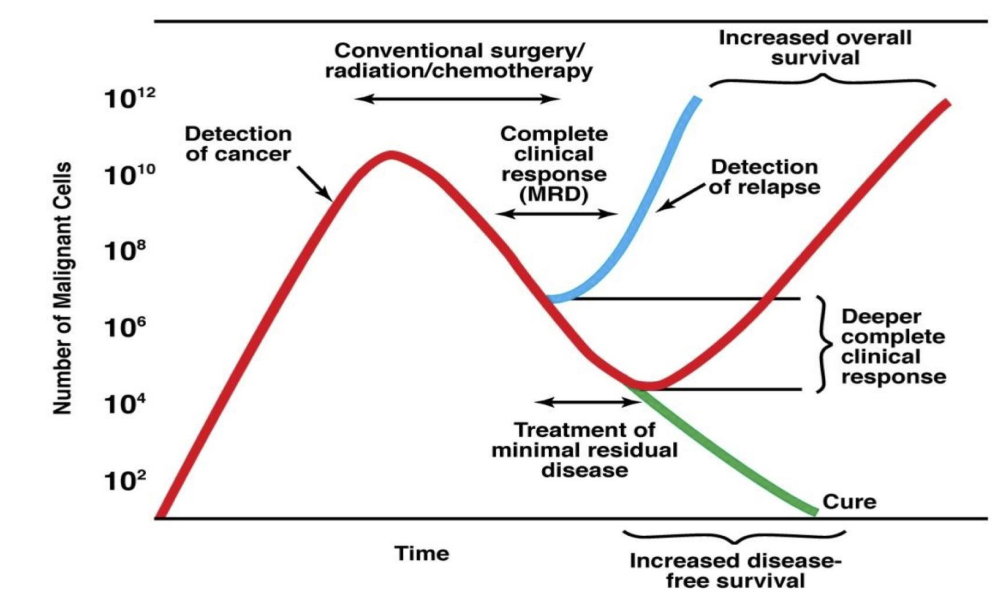
$Recall = TP \div (TP + FN)$
$Precision = TP \div (TP + FP)$

Range-based Anomalies



Recall = ?
Precision = ?

Domain-specific Time Bias



http://www.vaccinogeninc.com/oncovax/science-due-diligence/overview-part-1/

Cancer Detection: Respond early; no FN!

The Solution

Range-based Recall

$$Recall_T(R, P) = \frac{\sum_{i=1}^{N_r} Recall_T(R_i, P)}{N_r}$$

$$Recall_T(R_i, P) = \alpha * ExistenceReward(R_i, P) + \beta * OverlapReward(R_i, P)$$

$$ExistenceReward(R_i, P) = \begin{cases} 1 & \text{, if } \sum_{j=1}^{N_p} |R_i \cap P_j| \geq 1 \\ 0 & \text{, otherwise} \end{cases}$$

$$OverlapReward(R_i, P) = CardinalityFactor(R_i, P) * \sum_{i=1}^{N_p} \omega(R_i, R_i \cap P_j, \delta)$$

$$CardinalityFactor(R_i, P) = \begin{cases} 1 & \text{, if } R_i \text{ overlaps with at most one } P_j \in P \\ \gamma(R_i, P) & \text{, otherwise} \end{cases}$$

Range-based Precision

$$Precision_T(R, P) = \frac{\sum_{i=1}^{N_p} Precision_T(R, P_i)}{N_p}$$

$$Precision_T(R, P_i) = CardinalityFactor(P_i, R) * \sum_{j=1}^{N_r} \omega(P_i, P_i \cap R_j, \delta)$$

Notation	Description
R	set of real anomaly ranges
R_i	the i^{th} real anomaly range
P	set of predicted anomaly ranges
P_j	the j^{th} predicted anomaly range
N_r	number of real anomaly ranges
N_p	number of predicted anomaly ranges
α	relative weight of existence reward
β	relative weight of overlap reward
$\gamma()$	overlap cardinality function
$\omega()$	overlap size function
$\delta()$	positional bias function

Customizable weights & functions

ω() Example:

```
function ω(AnomalyRange, OverlapSet, δ)
   MyValue ← 0
   MaxValue ← 0
   AnomalyLength ← length(AnomalyRange)
   for i ← 1, AnomalyLength do
      Bias ← δ(i, AnomalyLength)
      MaxValue ← MaxValue + Bias
      if AnomalyRange[i] in OverlapSet then
         MyValue ← MyValue + Bias
   return MyValue/MaxValue
```

δ() Examples:

```
function δ(i, AnomalyLength)
   return 1
```
(flat)

```
function δ(i, AnomalyLength)
   return AnomalyLength - i + 1
```
(front-end)

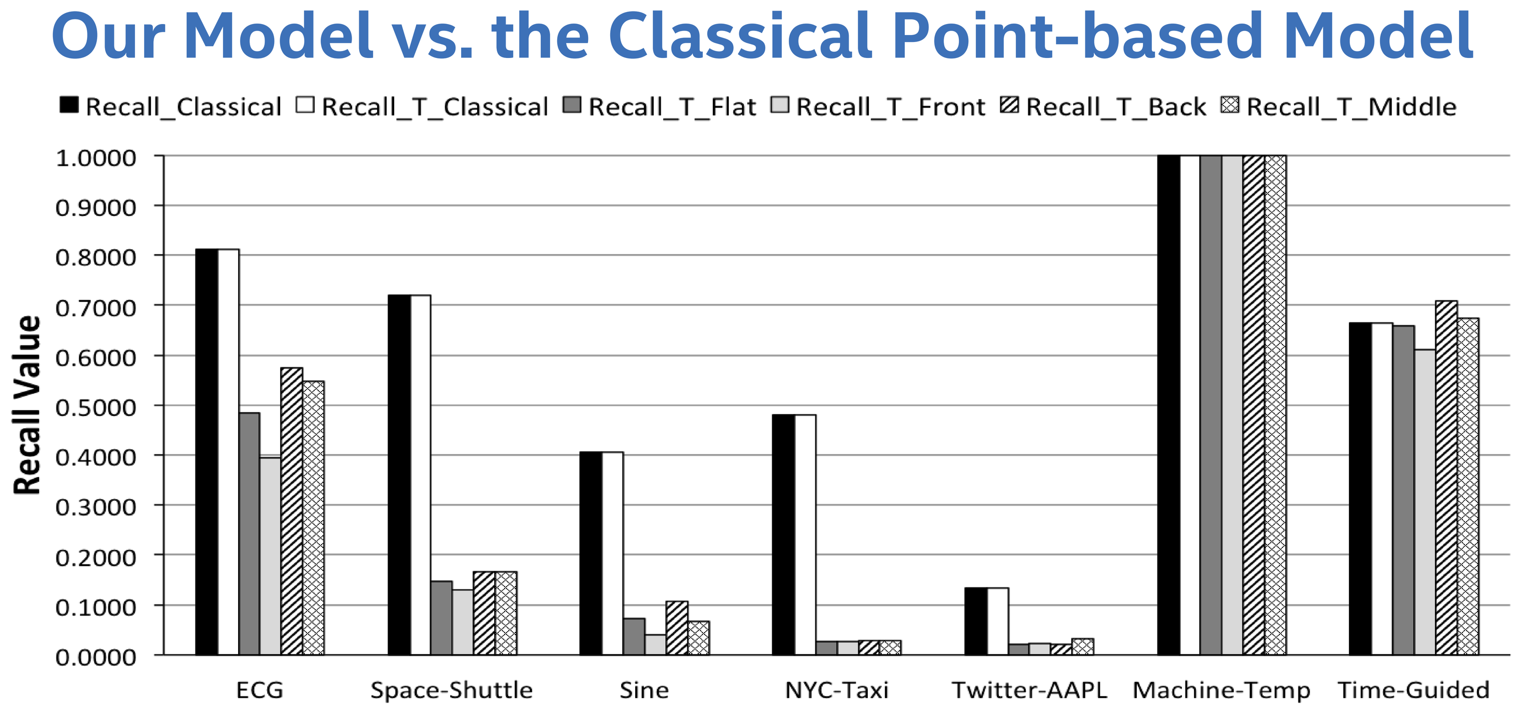
- Our model **subsumes** the classical point-based model, when:
  - all R_i and P_j are represented as unit-size ranges, and
  - $\alpha=0, \beta=1, \gamma()=1, \omega()$ is as defined above, and $\delta()$ is flat.

Experimental Results

- Datasets with labels:
  - Real: NAB Data Corpus[1]
  - Synthetic: Paranom Tool[2]
- Time series AD system:
  - LSTM on TensorFlow
- Our range-based metrics can be computed efficiently (see paper for cost analysis).

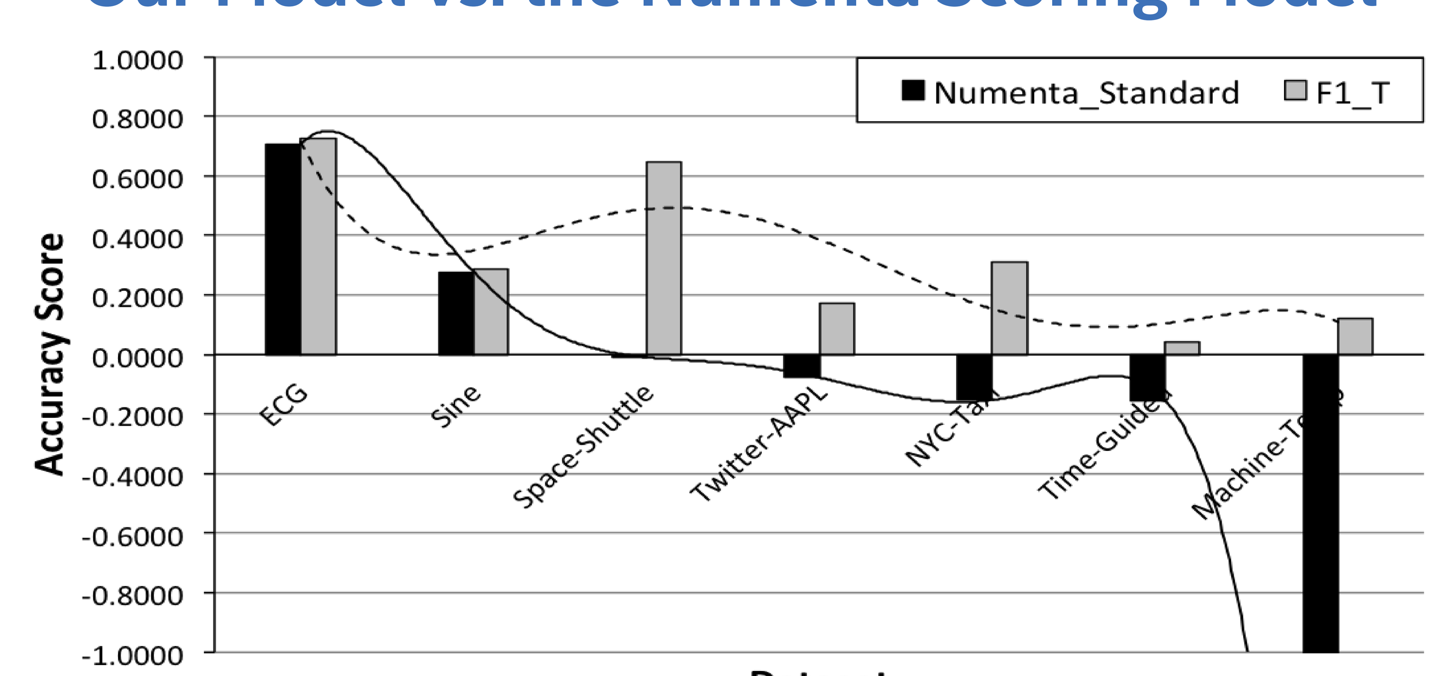
[1] Numenta Anomaly Benchmark (NAB): https://github.com/numenta/NAB/
[2] Paranom: https://arxiv.org/abs/1801.03164/

Our Model vs. the Classical Point-based Model



Range-based recall subsumes classical point-based recall and is sensitive to positional bias.

Our Model vs. the Numenta Scoring Model



Our model can be tuned to mimic Numenta as well as catching additional intricacies.

Conclusion

- Our new accuracy model for time series AD is **expressive**, **flexible**, and **extensible**.
- Ongoing work includes:
  - developing new ML training strategies optimized for our model (see **Greenhouse [SysML'18]**)
  - applying our model on real-world use cases (e.g., **autonomous driving**)
  - creating an open-source **benchmarking suite** for time series AD

BROWN