Event Outlier Detection in Continuous Time

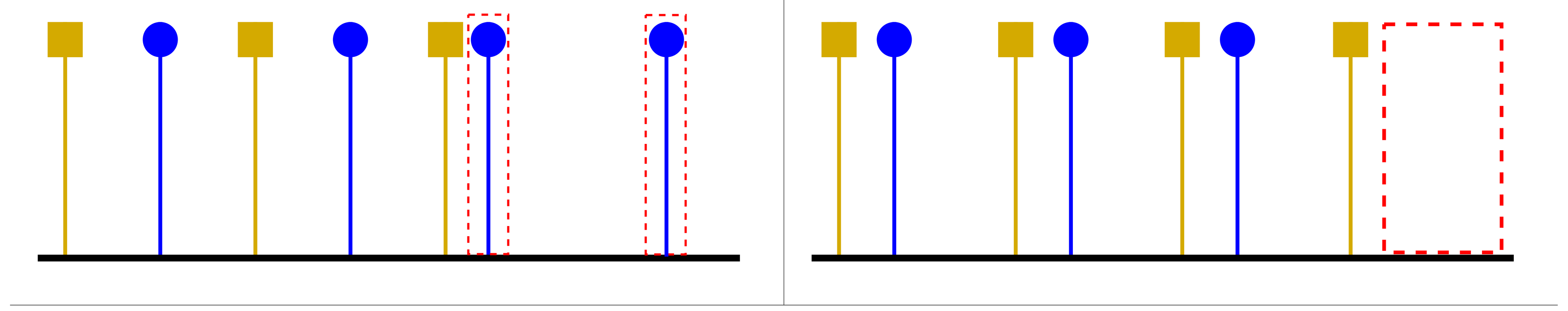
Siqi Liu[1,2] Milos Hauskrecht[1]

[1]University of Pittsburgh [2]Borealis AI

Introduction

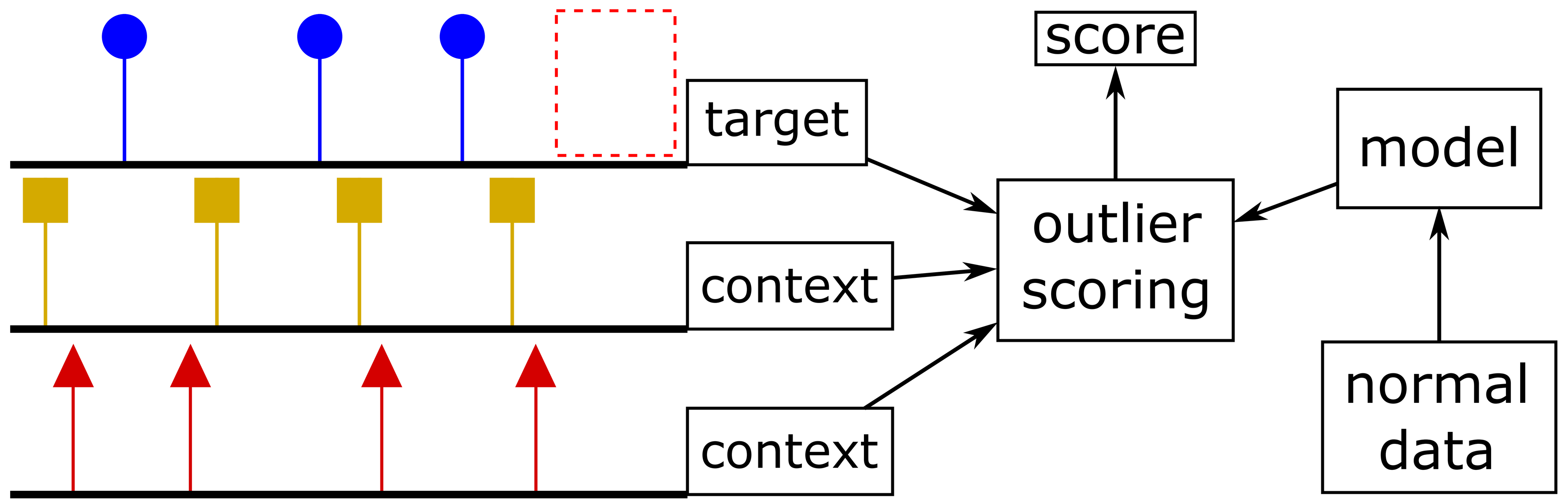
- Most previous works focus on solving the prediction problem
- **Prediction**: Given the history, what is the time and type of the next event?
- We formulate and solve new outlier detection problems
- **Outlier detection**: Given the history, is the recent **occurrence** or **absence** of events abnormal?

Abnormal Occurrence	Abnormal Absence
Commission Outlier	**Omission Outlier**



Problem Formulation

- Contextual outlier detection
  - Whether there is an outlier in a specific (target) type of events can depend on other (context) types of events
- Outlier scoring
  - A score is assigned to an event or blank interval to indicate how likely it is to be an outlier
- Semi-supervised outlier detection [1]
  - A model trained on normal data is available



Assumptions

1. The outlier generating process is independent from the normal point process
2. The rate of the outlier generation is constant (can be relaxed to be stochastic)

Outlier Scoring Methods

- We develop the methods based on **Bayesian decision theory** and **hypothesis testing**
- Our outlier scoring methods use the **conditional intensity function** $\lambda_0(t)$ of the underlying point-process model

	Commission	Omission
Object	An event at t	A blank interval B
Score	$-\lambda_0(t)$	$\int_B \lambda_0(s)ds$

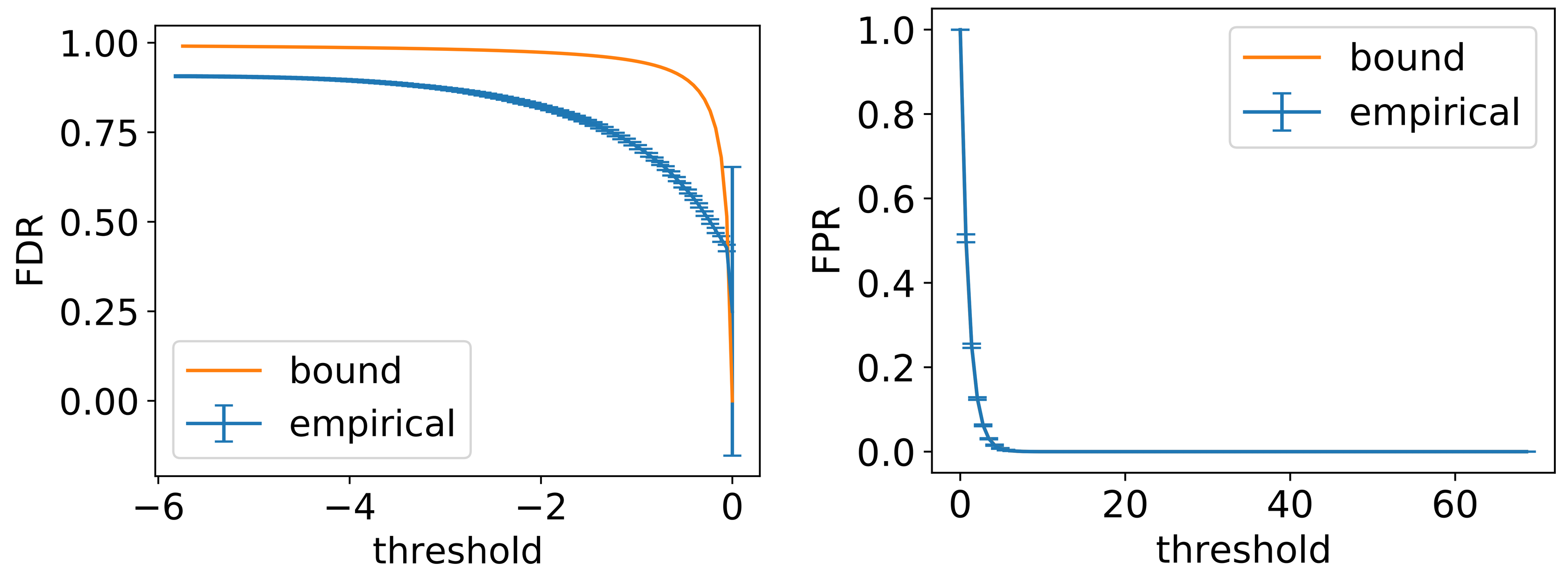
where

$$\lambda_0(t) = \lim_{dt \to 0^+} \frac{\mathbb{E}\left[N_0([t, t+dt))|\mathcal{H}_t\right]}{dt}$$

defines the rate of normal events given the history $\mathcal{H}_t$

- Our methods can be combined with any point-process model
  - In this work, we use a model adapted from the continuous-time LSTM [2]

Theoretical Guarantees



False Discovery Rate (for commission) and False Positive Rate (for omission) on Gamma process generated data

Experiments

- Simulate **(C)**ommission and **(O)**mission outliers with different rates $\alpha(t)$ relative to the normal points
  - Constant rate (denoted as **[0.1]**): $\alpha(t) = 0.1$
  - Periodic rate (denoted as **[sin]**): $\alpha(t) = \alpha_0(1 + \sin(2\pi t/p))/2$
  - Piecewise-constant rate (denoted as **[pc]**): $\alpha(t) = \alpha_0 g(t)$, where $g(t) : \mathcal{T} \to [0, 1]$ is random piecewise-constant function
- Evaluate performance with AUROC and compare
- **RND**: random scoring
- **LEN**: scoring based on distribution of inter-event time interval lengths
- **PPOD** and **CPPOD**: our methods without and with contextual information
- **Synthetic data**: Generate event sequences using contextual switching **Poisson** process and **Gamma** process

Dataset	Poi (C) [0.1]	Poi (C) [sin]	Poi (C) [pc]	Poi (O) [0.1]	Poi (O) [sin]	Poi (O) [pc]
RND	.500 (± .010)	.493 (± .007)	.512 (± .009)	.503 (± .008)	.498 (± .013)	.491 (± .007)
LEN	.601 (± .008)	.575 (± .006)	.584 (± .011)	.650 (± .006)	.659 (± .007)	.652 (± .011)
PPOD	.684 (± .010)	.661 (± .016)	.664 (± .009)	.737 (± .006)	.741 (± .012)	.734 (± .013)
CPPOD	**.711** (± .012)	**.707** (± .017)	**.697** (± .014)	**.778** (± .005)	**.791** (± .010)	**.784** (± .010)
Dataset	Gam (C) [0.1]	Gam (C) [sin]	Gam (C) [pc]	Gam (O) [0.1]	Gam (O) [sin]	Gam (O) [pc]
RND	.485 (± .007)	.493 (± .008)	.506 (± .007)	.505 (± .012)	.503 (± .010)	.515 (± .010)
LEN	.754 (± .006)	.762 (± .008)	.757 (± .005)	.799 (± .005)	.809 (± .006)	.813 (± .005)
PPOD	.816 (± .008)	.817 (± .006)	.813 (± .005)	.901 (± .007)	.902 (± .006)	.905 (± .006)
CPPOD	**.871** (± .006)	**.886** (± .004)	**.870** (± .007)	**.956** (± .003)	**.956** (± .004)	**.955** (± .004)

- **Real-world data**: Extract several target (medication / lab test) events and associated context events from MIMIC-III [3]

Dataset	INR (C) [0.1]	INR (C) [sin]	INR (C) [pc]	INR (O) [0.1]	INR (O) [sin]	INR (O) [pc]
RND	.496 (± .010)	.508 (± .009)	.488 (± .010)	.498 (± .011)	.516 (± .012)	.508 (± .009)
LEN	.596 (± .009)	.588 (± .010)	.607 (± .010)	.726 (± .008)	.717 (± .011)	.720 (± .011)
PPOD	.682 (± .010)	.675 (± .009)	.673 (± .008)	**.748** (± .009)	.760 (± .010)	**.773** (± .009)
CPPOD	**.687** (± .009)	**.680** (± .009)	**.681** (± .010)	.746 (± .010)	**.764** (± .009)	.770 (± .009)
Dataset	Cal (C) [0.1]	Cal (C) [sin]	Cal (C) [pc]	Cal (O) [0.1]	Cal (O) [sin]	Cal (O) [pc]
RND	.504 (± .013)	.502 (± .016)	.508 (± .011)	.493 (± .016)	.518 (± .017)	.496 (± .017)
LEN	.739 (± .012)	.688 (± .015)	.742 (± .011)	.526 (± .009)	.529 (± .012)	.541 (± .010)
PPOD	.830 (± .010)	.797 (± .010)	.837 (± .009)	.759 (± .008)	.758 (± .009)	.759 (± .011)
CPPOD	**.866** (± .006)	**.835** (± .009)	**.860** (± .011)	**.775** (± .008)	**.777** (± .010)	**.780** (± .009)

[1] Chandola et al. Anomaly detection: A survey. *ACM Comput. Surv.*, 41(3), 2009.

[2] Mei and Eisner. The neural Hawkes process: A neurally self-modulating multivariate point process. In *Advances in Neural Information Processing Systems*, 2017.

[3] Johnson et al. MIMIC-III, a freely accessible critical care database. *Scientific Data*, 3, 2016.