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?



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



Event Outlier Detection in Continuous Time Milos Hauskrecht¹ Sigi Liu^{1,2} ¹University of Pittsburgh ²Borealis Al

Assumptions



- 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

- and hypothesis testing
- Our outlier scoring methods use the **conditional intensity function** $\lambda_0(t)$ of the underlying point-process model

	Commissie
Object	An event a
Score	$-\lambda_0(t)$

where

$$\lambda_0(t) = \lim_{dt o 0^+} \frac{\mathbb{I}}{dt}$$

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

• We develop the methods based on **Bayesian decision theory**

Omission A blank interval B $\int_B \lambda_0(s) ds$

$\mathbb{E}\left[N_0([t, t+dt))|\mathcal{H}_t\right]$



• Siz	mulate (C tes $\alpha(t)$ re Constant rat Periodic rate Piecewise-co) ommission lative to the te (denoted as enstant rate (denoted as	h and (\mathbf{O}) r de normal p s $[0.1]$: $\alpha(t)$ denoted as $[\mathbf{p}]$	nission out oints) = 0.1 = $\alpha_0(1 + \sin c]$): $\alpha(t) = 1$	liers with $d^{2\pi t/p})/2$ $\alpha_0 g(t)$, when	re		
	$g(t): \mathcal{T} \to [0,1]$ is random piecewise-constant function							
• Evaluate performance with AUROC and compare								
• RN	• RND : random scoring • PPOD and CPPOD ; our metho							
• LE	LEN : scoring based on distribution W				without and with contextual			
of i	of inter-event time interval lengths information							
• Synthetic data: Generate event sequences using contextual								
SW	vitching Po	pisson pro	cess and G	amma pr	ocess			
Dataset	Poi (C) $[0.1]$	Poi (C) $[\sin]$	Poi (C) [pc]	Poi (O) [0.1]	Poi (O) [sin]	Poi (O) [pc]		
RND	$.500 (\pm .010)$	$.493 (\pm .007)$	$.512 (\pm .009)$	$.503 (\pm .008)$	$.498 (\pm .013)$	$.491 (\pm .007)$		
LEN	$.601 (\pm .008)$	$.575 (\pm .006)$	$.584 (\pm .011)$	$.650 (\pm .006)$	$.659 (\pm .007)$	$.652 (\pm .011)$		
PPOD	$.684 (\pm .010)$	$.661 (\pm .016)$	$.664 (\pm .009)$	$.737 (\pm .006)$	$.741 (\pm .012)$	$.734 (\pm .013)$		
CPPOD	$.711 (\pm .012)$	$.707 (\pm .017)$	$.697 (\pm .014)$	$.778 (\pm .005)$	$.791 (\pm .010)$	$.784 (\pm .010)$		
Dataset	Gam(C)[0.1]	$\frac{\text{Gam}(C)[\text{sin}]}{402(1-002)}$	$\frac{\text{Gam}(C)[pc]}{\text{Foc}(1,007)}$	Gam(O)[0.1]	$\frac{\text{Gam}(O)[\text{sin}]}{\text{For}(1-010)}$	$\frac{\text{Gam}(O)[pc]}{\text{F1F}(1,010)}$		
KND LEN	$.485 (\pm .007)$	$.493 (\pm .008)$ 769 (± .008)	$.500 (\pm .007)$ 757 (+ .005)	$.505 (\pm .012)$	$.503 (\pm .010)$	$.515 (\pm .010)$		
	$.734 (\pm .000)$ $.816 (\pm .008)$	$.702 (\pm .008)$ $.817 (\pm .006)$	$.131 (\pm .003)$ $.131 (\pm .005)$	$.799 (\pm .003)$ 001 (± .007)	$.809 (\pm .000)$ $.002 (\pm .006)$	$.815 (\pm .003)$ $.005 (\pm .006)$		
CPPOD	$.810 (\pm .008)$.871 (+ .006)	$.817 (\pm .000)$.886 (+ .004)	$.813 (\pm .003)$ $.870 (\pm .007)$	$.901 (\pm .007)$.956 (+ .003)	$.902 (\pm .000)$.956 (+ .004)	$.905 (\pm .000)$ $.955 (\pm .004)$		
						$\frac{1}{1}$		
• R	eal-woric		ktract sever	al larget (J		/ Iab test)		
ev	ents and as	SSOCIATED CO	ontext even	Its from M.	$\frac{100-111}{100}$			
Dataset	INR (C) [0.1]	$\frac{\text{INR}(C)[\sin]}{500}$	$\frac{\text{INR (C) [pc]}}{400}$	$\frac{\text{INR (O) [0.1]}}{400}$	$\frac{\text{INR (O) [sin]}}{510}$	$\frac{\text{INR (O) [pc]}}{\text{Equation (1 - 0.02)}}$		
RND	$.496 (\pm .010)$	$.508 (\pm .009)$	$.488 (\pm .010)$	$.498 (\pm .011)$	$.516 (\pm .012)$	$.508 (\pm .009)$		
LEN	$.596 (\pm .009)$	$.588 (\pm .010)$	$.607 (\pm .010)$	$.(26 (\pm .008))$	$.(1)((\pm .011))$	$.(20 (\pm .011)$		
CDDOD	$.082 (\pm .010)$ 687 (± .000)	$.073 (\pm .009)$ 680 ($\pm .009$)	$.073 (\pm .008)$ 681 (± .010)	$-748 (\pm .009)$ $746 (\pm .010)$	$.700 (\pm .010)$ 764 ($\pm .000$)	$.770 (\pm .009)$		
Dataset	$\frac{.001}{Cal} (C) [0 1]$	$\frac{.000}{\text{Cal}(\text{C})[\text{sin}]}$	$\frac{\mathbf{OOL} (\pm .010)}{\mathrm{Cal} (\mathrm{C}) [\mathrm{pc}]}$	$\frac{.740 (\pm .010)}{\text{Cal} (\Omega) [0.1]}$	$\frac{.104}{\text{Cal}(\Omega)} (\pm .009)$	$\frac{110 (\pm .009)}{\text{Cal} (\Omega) [\text{nc}]}$		
RND	504 (+ 013)	$\frac{600}{502} (+ 016)$	$\frac{\cos(0)}{508} (+ 011)$	$\frac{0.1}{493(+0.1)}$	$\frac{600}{518} (+ 017)$	$\frac{600}{496(+017)}$		
LEN	.739 (+ .012)	$.688(\pm .015)$.742 (+ .011)	.526 (+ .009)	.529 (+ .012)	$.541 (\pm .010)$		
PPOD	$(\pm .010)$.830 (± .010)	$.797 (\pm .010)$	$.837 (\pm .009)$	$.759 (\pm .008)$	$.758 (\pm .009)$	$.759 (\pm .011)$		
CPPOD	$.866(\pm .006)$	$.835(\pm .009)$	$.860(\pm .011)$	$.775(\pm .008)$	$.777(\pm .010)$	$.780(\pm .009)$		

• Simulate (C) ommission and (O) mission outliers with different						
rates $\alpha(t)$ relative to the normal points						
• Constant rate (denoted as $[0,1]$): $\alpha(t) = 0.1$						
Poriodic rate (denoted as [sin]): $\alpha(t) = \alpha(1 \pm \sin(2\pi t/n))/2$						
• Terroute rate (denoted as $[\operatorname{Sin}]$). $\alpha(\iota) = \alpha_0(\iota + \operatorname{Sin}(2\pi\iota/p))/2$						
• Flecewise-constant rate (denoted as $[\mathbf{pc}]$). $\alpha(\iota) = \alpha_0 g(\iota)$, where						
$g(t): \mathcal{T} \to [0, 1]$ is random piecewise-constant function						
• Evaluate performance with AUROC and compare						
• RND : random scoring • PPOD and CPPOD : our methods						
• LEN: scoring based on distribution without and with contextual						
of inter_event time interval lengths information						
Supplied the fille						
• Sympletic 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 $.500 (\pm .010) .493 (\pm .007) .512 (\pm .009) .503 (\pm .008) .498 (\pm .013) .491 (\pm .007)$						
LEN $.601 (\pm .008) .575 (\pm .006) .584 (\pm .011) .650 (\pm .006) .659 (\pm .007) .652 (\pm .011)$						
PPOD $.684 (\pm .010) .661 (\pm .016) .664 (\pm .009) .737 (\pm .006) .741 (\pm .012) .734 (\pm .013)$						
$CPPOD \left .711 (\pm .012) .707 (\pm .017) .697 (\pm .014) \right .778 (\pm .005) .791 (\pm .010) .784 (\pm .010) \\$						
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 $1.754 (\pm .006) .762 (\pm .008) .757 (\pm .005) .799 (\pm .005) .809 (\pm .006) .813 (\pm .005)$						
PPOD $.816 (\pm .008) .817 (\pm .006) .813 (\pm .005) .901 (\pm .007) .902 (\pm .006) .905 (\pm .006)$						
$CPPOD .871 (\pm .006) .886 (\pm .004) .870 (\pm .007) .956 (\pm .003) .956 (\pm .004) .955 (\pm .004) .$						
• 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 $.596 (\pm .009) .588 (\pm .010) .607 (\pm .010) .726 (\pm .008) .717 (\pm .011) .720 (\pm .011)$						
PPOD $.682 (\pm .010) .675 (\pm .009) .673 (\pm .008) $.748 $(\pm .009) .760 (\pm .010) .773 (\pm .009)$						
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 $.739 (\pm .012) .688 (\pm .015) .742 (\pm .011) .526 (\pm .009) .529 (\pm .012) .541 (\pm .010)$						
PPOD $.830 (\pm .010) .797 (\pm .010) .837 (\pm .009) .759 (\pm .008) .758 (\pm .009) .759 (\pm .011)$						
$CPPOD .866 (\pm .006) .835 (\pm .009) .860 (\pm .011) .775 (\pm .008) .777 (\pm .010) .780 (\pm .009) .009 (\pm .009$						

Experiments

[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.