

# Eventos como Pontos: Uma Simplificação

Este modelo pontual simplifica a teoria e a avaliação ao considerar eventos como instantes singulares no tempo.

Eduardo Ogasawara

[eduardo.ogasawara@cefet-rj.br](mailto:eduardo.ogasawara@cefet-rj.br)

<https://eic.cefet-rj.br/~eogasawara>

# Modelo Pontual: Fundamentos

$$\tau_i \in \mathbb{T}$$

Onde  $\tau_i$  representa o instante do i-ésimo evento e  $\mathbb{T}$  é o conjunto de instantes de tempo considerados.

Esta aproximação pontual pode falhar quando o evento ocupa um período significativo, como surtos, falhas prolongadas ou regimes anômalos. Comparar apenas instantes pode distorcer o que significa realmente acertar na detecção.

# Eventos como Intervalos Temporais

## Representação Intervalar

Em vez de pontos isolados, eventos reais possuem início e fim explícitos. Esta abordagem transforma eventos em objetos temporais completos, não apenas instantes.

$$e_i = [t_i^{\text{start}}, t_i^{\text{end}}]$$

Conjuntos de eventos reais e detectados:

$$\mathcal{E} = \{e_1, \dots, e_n\}, \quad \hat{\mathcal{E}} = \{\hat{e}_1, \dots, \hat{e}_m\}$$

### Evento Real

$e_i$ :  $i$ -ésimo evento com início  $t_i^{\text{start}}$  e fim  $t_i^{\text{end}}$

### Evento Detectado

$\hat{e}_j$ :  $j$ -ésimo evento detectado pelo sistema

### Conjuntos

$\mathcal{E}$  e  $\hat{\mathcal{E}}$ : eventos reais e detectados

# Tipologia Temporal dos Eventos

Eventos podem ser representados de três formas distintas, cada uma adequada para diferentes contextos de análise e detecção.



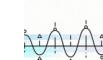
## Pontual

Instante único no tempo:  $e_i = \tau_i$



## Intervalar

Duração explícita:  $e_i = [t_i^{start}, t_i^{end}]$



## Estrutural (Padrão)

Subsequência com forma:  $e_i = X_{t:t+w}$ , onde  $w$  é o tamanho da janela

Esta organização em três níveis deixa claro que avaliação e detecção mudam fundamentalmente dependendo da representação adotada.

# Avaliação Além do Paradigma Pontual

A simples diferença entre instantes não resolve adequadamente o caso intervalar. Quando eventos têm duração, precisamos considerar sobreposição e deslocamento temporal.

$$\tau_i \neq \hat{\tau}_j$$

Onde  $\tau_i$  é o instante real e  $\hat{\tau}_j$  é o instante detectado no modelo pontual.

- **Limitação Crítica:** Um detector pode acertar quase toda a duração do evento e ainda assim "errar" o instante exato. Por isso, critérios de sobreposição e deslocamento são essenciais para avaliação intervalar.

# Interseção Temporal

## Conceito Fundamental

A interseção temporal mede quanto dois intervalos se sobrepõem. É zero quando não há sobreposição e serve como base para todas as métricas intervalares.

Considere um evento real e um detectado:

$$e_i = [t_i^{\text{start}}, t_i^{\text{end}}], \quad \hat{e}_j = [\hat{t}_j^{\text{start}}, \hat{t}_j^{\text{end}}]$$

## Fórmula da Sobreposição

$$|e_i \cap \hat{e}_j| = \max \left( 0, \min(t_i^{\text{end}}, \hat{t}_j^{\text{end}}) - \max(t_i^{\text{start}}, \hat{t}_j^{\text{start}}) \right)$$

A sobreposição é o "fim mais cedo" menos o "início mais tarde". Se negativo, não há interseção e o resultado é zero.

# União Temporal

A união temporal mede o intervalo total coberto pelos eventos real e detectado. É sempre pelo menos tão grande quanto cada intervalo individual e serve para tipificar a interseção.

$$|e_i \cup \hat{e}_j| = \max(t_i^{end}, \hat{t}_j^{end}) - \min(t_i^{start}, \hat{t}_j^{start})$$

## Intervalo Envelope

A união pega desde o começo mais cedo até o fim mais tarde, cobrindo ambos os eventos

## Referência de Tipicidade

Serve para determinar se a sobreposição foi grande ou pequena em proporção ao total coberto

# IoU Temporal: Métrica de Sobreposição

O Intersection over Union (IoU) transforma a sobreposição em uma medida padronizada entre 0 e 1, onde 1 significa coincidência perfeita de intervalos.

## Definição do IoU

$$IoU(e_i, \hat{e}_j) = \frac{|e_i \cap \hat{e}_j|}{|e_i \cup \hat{e}_j|}$$

## Correspondência por Limiar

$$M(e_i, \hat{e}_j) = \begin{cases} 1, & \text{se } IoU(e_i, \hat{e}_j) \geq \gamma \\ 0, & \text{caso contrário} \end{cases}$$

O limiar  $\gamma$  define o que conta como acerto. IoU alto força o detector a acertar bem tanto início quanto fim.

Sem Sobreposição

Sobreposição Parcial

Coincidência Perfeita

# Métricas Intervalares

As métricas tradicionais de detecção são adaptadas para o contexto intervalar, onde verdadeiros positivos dependem do matching por IoU.



## Verdadeiros Positivos (TP)

$$TP = \sum_{i,j} M(e_i, \hat{e}_j)$$

Número de correspondências aceitas pelo limiar



## Falsos Positivos (FP)

$$FP = |\hat{\mathcal{E}}| - TP$$

Detecções sem correspondência aceita



## Falsos Negativos (FN)

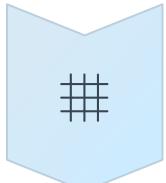
$$FN = |\mathcal{E}| - TP$$

Eventos reais sem correspondência aceita

As métricas resultantes — Precision( $\gamma$ ), Recall( $\gamma$ ) e  $F_1(\gamma)$  — tornam-se funções do limiar de IoU escolhido. Ao variar  $\gamma$ , você muda a exigência:  $\gamma$  baixo aceita sobreposições pequenas;  $\gamma$  alto exige alinhamento forte.

# Granularidade e Ambiguidade

A resolução temporal ( $\Delta t$ ) afeta diretamente a sobreposição medida e, consequentemente, todas as métricas intervalares.



## Granularidade Grossa

Passos de tempo grandes podem inflar artificialmente as sobreposições medidas



## Granularidade Fina

Passos muito pequenos podem fragmentar eventos e criar separações artificiais

Se o tempo é medido em passos grandes, dois eventos podem parecer mais sobrepostos do que realmente são. Se o tempo é muito fino, pequenas diferenças viram separações e um evento pode ser quebrado em vários.

# Paradigmas de Avaliação

1

## Avaliação Offline

Série completa  $X_{1:T}$  disponível para análise posterior. Avaliação com calma, olhando todos os eventos.

2

## Avaliação Online

Apenas observações  $X_t$  até o instante  $t$ . Decisões em tempo real sob informação parcial.

No online, as decisões e scores existem em tempo real. A avaliação precisa respeitar que o detector não via o futuro quando tomou cada decisão.

# Latência Temporal

Em sistemas online, o início detectado tende a atrasar em relação ao início real. A latência mede esse atraso e torna-se uma dimensão crítica de qualidade do detector.

A latência por evento é definida como:

$$\Delta t_i = \hat{t}_i^{start} - t_i^{start}$$

Onde  $\Delta t_i$  é a latência do evento  $i$ ,  $t_i^{start}$  é o início real e  $\hat{t}_i^{start}$  é o início detectado.

## Início Real

Evento efetivo começa no sistema.

## Início Detectado

Sistema registra o evento com atraso.

- ❑ Mesmo quando o detector acerta o evento, ele pode começar a sinalizar tarde. Em aplicações online, esse atraso é parte do erro: detectar "certo" mas tarde pode ser quase tão ruim quanto não detectar.

# Custo Temporal Acumulado

## Medindo o Impacto Total

A soma das latências mede o custo total de atrasos na detecção. A versão média facilita comparação entre detectores.

$$C = \sum_{i=1}^{|\mathcal{E}|} \Delta t_i$$

$$\bar{C} = \frac{1}{|\mathcal{E}|} \sum_i \Delta t_i$$

## Interpretação

- **Custo Total (C):** Quanto tempo de reação foi perdido no total
- **Custo Médio ( $\bar{C}$ ):** Permite comparar detectores em bases diferentes
- **Penalização Sistemática:** Detectores lentos são identificados objetivamente

Esta métrica transforma latência em algo agregável e comparável, essencial para avaliar sistemas em produção.

# Trade-off Latência vs Precisão

Reducir latência pode aumentar falsos positivos. Decisões mais cedo são decisões com menos evidência acumulada, criando um compromisso inevitável em sistemas online.

$$\Delta t \downarrow \Rightarrow \text{Precision} \downarrow$$

## Detecção Rápida

Menor latência ( $\Delta t$  baixo), mas decisão com pouca informação aumenta alarmes falsos

## Detecção Precisa

Espera por mais evidência melhora precisão, mas aumenta o atraso na resposta

Este é o dilema central de sistemas online: velocidade versus confiabilidade. A escolha do ponto de operação depende do custo relativo de cada tipo de erro no contexto da aplicação.

# Métricas Online Adicionais

Precisão e recall tradicionais não capturam adequadamente o desempenho temporal. Sistemas online requerem métricas que incorporem atraso e detecção precoce.



## Erro Médio de Latência

$MAE\Delta t$  mede o erro médio absoluto da latência, calculado sobre todos os  $\Delta t_i$

No contexto online, "acertar" precisa vir junto com "quando acertou". Estas métricas adicionais capturam a dimensão temporal da qualidade de detecção.



## Métricas de Detecção Precoce

Avaliam a capacidade do sistema de identificar eventos no início de sua ocorrência

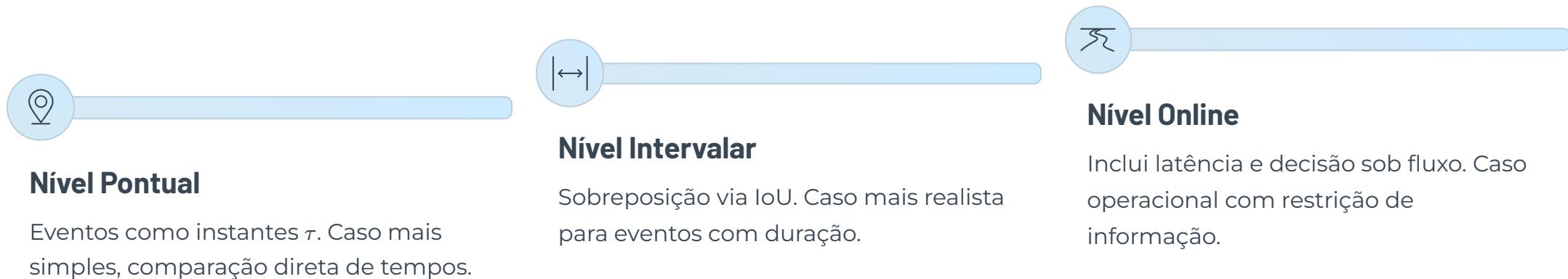


## Custo Ponderado

Penalizam atrasos com pesos maiores para eventos mais críticos ou urgentes

# Três Níveis de Avaliação

A avaliação de detecção de eventos pode ser organizada em camadas progressivas de complexidade e realismo.



Uma forma abstrata de representar a qualidade da avaliação:

$$Q = f(\text{tempo, estrutura, dinâmica})$$

Onde  $Q$  é a medida de qualidade e  $f(\cdot)$  é a regra que combina as dimensões de avaliação: temporal, estrutural e dinâmica.

# Referências Bibliográficas

Uma coleção cuidadosamente selecionada de obras fundamentais que abordam análise de séries temporais e mineração de dados.



## Event Detection in Time Series

**Ogasawara, E.; Salles, R.; Porto, F.; Pacitti,**

**E.** (2025). Publicação recente da Springer Nature Switzerland que explora técnicas avançadas de detecção de eventos em séries temporais.

## Time Series Analysis: With Applications in R

**Cryer, J. D.; Chan, K.-S.** (2008). Obra clássica da Springer que combina fundamentação teórica sólida com implementações práticas.

## Data Mining: Concepts and Techniques

**Han, J.; Pei, J.; Tong, H.** (2022). Quarta edição publicada pela Morgan Kaufmann que consolida conceitos fundamentais e técnicas avançadas de mineração de dados