Detecção de Eventos em Séries Temporais

Eduardo Ogasawara[1]

[1]CEFET/RJ – Centro Federal de Educação Tecnológica Celso Suckow da Fonseca

eogasawara@ieee.org

Abstract. *This tutorial presents an integrated view of event detection in time series, covering anomalies, change points and motifs under a common conceptual framework. It connects event definitions, data structures, offline and online scenarios, detection strategies, and evaluation criteria, with special attention to temporal tolerance and benchmark design. It also includes a practical perspective based on the Harbinger framework for experimental organization, detector execution, and metric comparison.*
Keywords. *time series, event detection, anomalies, change points, motifs, evaluation*

Resumo. *Este tutorial apresenta uma visão integrada da detecção de eventos em séries temporais, cobrindo anomalias, pontos de mudança e* motifs *sob um quadro conceitual comum. A proposta articula definição de eventos, estruturas de dados, cenários offline e online, estratégias de detecção e critérios de avaliação, com atenção especial à tolerância temporal e ao desenho de benchmarks. Além da discussão teórica, o tutorial incorpora uma dimensão prática apoiada pelo framework Harbinger, que facilita a organização experimental, a execução de detectores e a comparação de métricas.*
Palavras-chave. *séries temporais, detecção de eventos, anomalias, pontos de mudança, motifs, avaliação*

1. Introdução

Séries temporais estão presentes em aplicações de finanças, saúde, ciência ambiental, observabilidade e monitoramento industrial. Nesses contextos, o interesse recai menos sobre o comportamento global da série e mais sobre instantes ou intervalos cuja ocorrência altera a interpretação do fenômeno observado. A esse conjunto de problemas damos o nome de detecção de eventos em séries temporais [Pang et al., 2021].

O tutorial adota a organização proposta por Ogasawara et al. [2025], que reúne três classes principais de eventos: anomalias, pontos de mudança e *motifs*. Essa divisão é útil porque cada classe responde a perguntas diferentes. Anomalias destacam observações improváveis ou comportamento fora do padrão. Pontos de mudança marcam alterações significativas na dinâmica do processo. *Motifs* correspondem a padrões recorrentes relevantes por si mesmos. Em aplicações reais, essas classes podem coexistir e podem ser confundidas quando a formulação do problema é imprecisa [Olteanu et al., 2023].

A taxonomia usada na apresentação foi construída a partir de um mapeamento sistemático da literatura, o que ajuda a organizar trabalhos que frequentemente aparecem dispersos em subáreas diferentes. O tutorial usa essa estrutura para conectar definições, cenários, estratégias e avaliação [Zhao, 2021].

Além da definição do evento, importa a forma como os dados são representados e processados. Uma mesma série pode ser analisada como sequência bruta, janela deslizante, subsequência, fluxo contínuo ou conjunto de eventos pontuais e intervalares. Cada escolha afeta o detector aplicável, o custo computacional e a interpretação do resultado. A representação faz parte da formulação do problema [Jensen et al., 2017].

Outro eixo importante é a avaliação. Em problemas temporais, a correspondência entre predição e rótulo depende de atraso de detecção, duração do evento e regime online. O material discute métricas com tolerância temporal, critérios para eventos intervalares e medidas tradicionais [Schmidl et al., 2022].

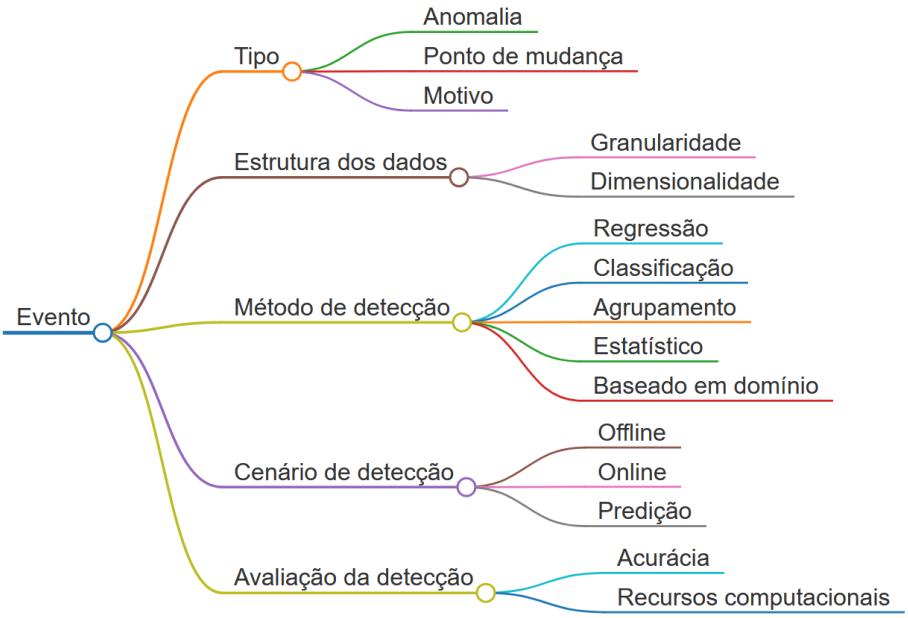

Figura 1. Taxonomia usada para estruturar a discussão do tutorial.

2. Escopo do tutorial

O tutorial é organizado em torno de cinco blocos conceituais: tipo de evento, estrutura dos dados, cenário de processamento, estratégia de detecção e avaliação. Esses blocos aparecem de forma integrada ao longo da apresentação, porque a escolha em um deles afeta diretamente os demais. Uma definição de anomalia para dados offline exige adaptação para monitoramento em fluxo, e uma métrica clássica pode penalizar um detector que identifica corretamente o evento, mas com atraso tolerável [Zhao, 2021].

O primeiro bloco trata da semântica do evento. A noção de anomalia é contextual e depende da unidade de análise, do período observado e do que se considera comportamento normal. Pontos de mudança abrangem rupturas abruptas, mudanças graduais, instabilidades persistentes e transições entre regimes. Já os *motifs* são apresentados como padrões recorrentes úteis por repetição ou por associação com atividade normal [Truong et al., 2020].

O segundo bloco aborda a representação dos dados e suas consequências para o problema. Serão discutidas granularidade temporal, janelas deslizantes, subsequências,

eventos pontuais e intervalares, além de cenários univariados e multivariados. A intenção é mostrar que o mesmo evento pode assumir formas diferentes conforme o nível de agregação e a forma como a série é particionada. Essa parte do tutorial também evidencia como a representação influencia indexação, custo de processamento e ruído [Jensen et al., 2017].

A representação também orienta a escolha das hipóteses que o detector pode sustentar. Em uma série com muitos picos curtos, por exemplo, uma formulação orientada a ponto tende a simplificar o comportamento do processo, enquanto uma formulação intervalar preserva a duração do episódio e a sua persistência. Essa distinção é importante porque altera o tipo de rotulagem possível, a interpretação do erro e o modo como o resultado será consumido por especialistas ou sistemas automáticos [Tatbul et al., 2018].

O terceiro bloco diferencia os cenários offline, online e preditivo. Em detecção offline, o objetivo é reconstruir e analisar eventos em dados já observados. Em detecção online, o foco passa a ser resposta em fluxo, com restrições de latência, memória e atualização incremental. Na formulação preditiva, a questão inclui antecipação, horizonte de previsão e custo de falso alarme. Esses três cenários representam formulações distintas de um problema comum, com exigências operacionais diferentes [Ahmad et al., 2017].

O quarto bloco compara famílias de métodos. A exposição percorre abordagens estatísticas, supervisionadas, não supervisionadas e guiadas por conhecimento de domínio, sempre indicando quais hipóteses cada família assume sobre a série. Serão discutidos regressão, classificação, agrupamento, técnicas estatísticas e estratégias baseadas em domínio, com destaque para requisitos de dados, custo computacional e interpretabilidade [Zhao, 2021].

O quinto bloco trata da avaliação e do desenho experimental. Aqui entram tolerância temporal, eventos intervalares e métricas tradicionais. O objetivo é capturar um episódio relevante no tempo certo. Serão discutidos janelas de tolerância, medidas orientadas a evento e a diferença entre performance em benchmark e utilidade operacional [Wenig et al., 2022]. Essa discussão é central para interpretar resultados de forma consistente.

O componente prático com Harbinger aparece como apoio à experimentação ao longo de todos os blocos. A proposta é usar um mesmo conjunto de dados para organizar pipelines, executar detectores distintos e comparar saídas com critérios consistentes. Nesse contexto, o Harbinger entra como um pacote em R voltado à detecção de eventos em séries temporais, usado para exemplificar o fluxo experimental desde a preparação dos dados até a leitura comparativa das saídas [Ogasawara et al., 2025].

Na prática, essa parte fecha o ciclo conceitual do tutorial. A mesma definição de evento que orienta a taxonomia também determina qual medida é aceitável para julgar um detector. Isso fica mais claro quando se compara uma métrica rígida, adequada a ocorrências pontuais, com uma métrica que aceita atraso, tolerância e superposição parcial. Em vez de medir apenas acerto e erro no instante exato, o tutorial apresenta o raciocínio necessário para avaliar utilidade temporal, estabilidade dos resultados e aderência ao contexto de uso [Paparrizos et al., 2022].

3. Biografia

Eduardo Ogasawara é professor do Departamento de Ciência da Computação do CEFET/RJ desde 2010. Possui mestrado e doutorado em Engenharia de Sistemas e Computação pela COPPE/UFRJ. Atua em banco de dados, ciência de dados, mineração de dados e análise de séries temporais, com experiência prévia na indústria de TI e orientação de trabalhos na interface entre dados, aprendizado de máquina e sistemas. É membro sênior da IEEE, membro da ACM e da SBC, e autor de livro sobre detecção de eventos em séries temporais.

Referências

Ahmad, S., Lavin, A., Purdy, S., e Agha, Z. (2017). Unsupervised real-time anomaly detection for streaming data. *Neurocomputing*, 262:134 – 147.

Jensen, S. K., Pedersen, T. B., e Thomsen, C. (2017). Time Series Management Systems: A Survey. *IEEE Transactions on Knowledge and Data Engineering*, 29(11):2581 – 2600.

Ogasawara, E., Salles, R., Porto, F., e Pacitti, E. (2025). *Event Detection in Time Series*. Synthesis Lectures on Data Management. Springer Nature Switzerland, Cham, 1 edition.

Olteanu, M., Rossi, F., e Yger, F. (2023). Meta-survey on outlier and anomaly detection. *Neurocomputing*, 555.

Pang, G., Shen, C., Cao, L., e Van Den Hengel, A. (2021). Deep Learning for Anomaly Detection: A Review. *ACM Computing Surveys*, 54(2).

Paparrizos, J., Boniol, P., Palpanas, T., Tsay, R. S., Elmore, A., e Franklin, M. J. (2022). Volume Under the Surface: A New Accuracy Evaluation Measure for Time-Series Anomaly Detection. *Proceedings of the VLDB Endowment*, 15(11):2774 – 2787.

Salles, R., Lima, J., Coutinho, R., Pacitti, E., Masseglia, F., Akbarinia, R., Chen, C., Garibaldi, J., Porto, F., e Ogasawara, E. (2023). SoftED: Metrics for Soft Evaluation of Time Series Event Detection.

Schmidl, S., Wenig, P., e Papenbrock, T. (2022). Anomaly Detection in Time Series: A Comprehensive Evaluation. *Proceedings of the VLDB Endowment*, 15(9):1779 – 1797.

Tatbul, N., Lee, T. J., Zdonik, S., Alam, M., e Gottschlich, J. (2018). Precision and recall for time series. In *Advances in Neural Information Processing Systems*, volume 2018-December, pages 1920 – 1930.

Truong, C., Oudre, L., e Vayatis, N. (2020). Selective review of offline change point detection methods. *Signal Processing*, 167.

Wenig, P., Schmidl, S., e Papenbrock, T. (2022). TimeEval: A Benchmarking Toolkit for Time Series Anomaly Detection Algorithms. *Proceedings of the VLDB Endowment*, 15(12):3678 – 3681.

Zhao, L. (2021). Event Prediction in the Big Data Era: A Systematic Survey. *ACM Computing Surveys*, 54(5).