



Detecção de Vazamentos em Tubulações de Água com Transformação de Séries Temporais por Convolução de Kernels Aleatórios

Matheus Medeiros Donatoni, Escola de Engenharia de São Carlos - USP, matheus.m.donatoni@usp.br
Leopoldo Pisanelli Rodrigues de Oliveira, Escola de Engenharia de São Carlos - USP, leopro@sc.usp.br

RESUMO EXPANDIDO

A prevenção de vazamentos em redes de distribuição de água é essencial para a boa gestão dos recursos hídricos, uma vez que as perdas geram prejuízos econômicos, energéticos e aumentam indesejavelmente a pegada de carbono (Okeya et al., 2014). Nesse contexto, o desenvolvimento de métodos eficazes para a detecção de vazamentos torna-se um requisito fundamental para a gestão de ativos hídricos e a redução de impactos ambientais (El-Zahab et al., 2018).

Essa necessidade motiva pesquisadores a desenvolverem diversas abordagens, que envolvem diferentes tipos de métodos e sensores. Por exemplo, Bohorquez et al. (2020) utilizaram dados de transientes de pressão para treinar uma Rede Neural Artificial (*Artificial Neural Network*), enquanto Shukla and Piratla (2020) empregaram dados de aceleração combinados com uma Rede Neural Convolucional (*Convolutional Neural Network*). De forma distinta, Ullah et al. (2023) aplicaram um sensor de emissão acústica e testaram vários algoritmos de aprendizado de máquina (*machine learning*), como *k*-vizinhos mais próximos (*k-nearest neighbors*), redes neurais (*neural networks*), árvores de decisão (*decision trees*) e florestas aleatórias (*random forests*).

Apesar de sua eficácia, os métodos existentes frequentemente possuem implementação complexa e carecem de abordagens multimodais (baseadas em múltiplos sensores), o que configura uma lacuna na literatura (Mahdi et al., 2024). Nesse contexto, a base de dados de Aghashahi et al. (2023) surge como uma oportunidade. Por oferecer sinais multimodais de vazamentos em uma rede de PVC sob diversas condições (arquiteturas, demandas e tipos de furos), ela viabiliza o desenvolvimento de novas metodologias que atendam a esse requisito de integração de dados.

Este trabalho avalia o uso da Transformação Convolucional de Kernels Aleatórios (ROCKET), proposta por Dempster et al. (2020), para a detecção de vazamentos de água. Trata-se de um método de alta acurácia e baixa complexidade computacional, facilmente aplicável a séries temporais multivariadas.

A base de dados utilizada contempla duas arquiteturas de rede: a ramificada (*Branched*), na qual os tubos não formam um circuito fechado, e a em laço (*Looped*), que constitui um circuito fechado. O sistema foi operado com diferentes demandas na linha de serviço (0 L/s, 0,18 L/s, 0,47 L/s e uma demanda transiente) e cinco condições distintas: sem vazamento, vazamento em flange, em orifício, em rachadura circunferencial e em rachadura longitudinal. Os dados foram coletados por pares de acelerômetros (25,6 kHz), sensores de pressão (25,6 kHz) e hidrofones (8 kHz), com cada registro resultando em séries temporais de aproximadamente 30 segundos.

Para este estudo, foram utilizados apenas os dados de aceleração e pressão. O ROCKET foi aplicado diretamente às séries temporais brutas. A técnica consiste em gerar aleatoriamente um grande número de *kernels* (filtros convolucionais) que são aplicados às séries de entrada. Da convolução de cada kernel com a série, extraem-se duas *features*: o valor máximo (MAX) e a proporção de valores positivos (PPV). As *features* resultantes de todos os *kernels* são então concatenadas e utilizadas como entrada para um classificador linear *Ridge*. A Figura 1 ilustra o fluxograma de implementação do ROCKET.

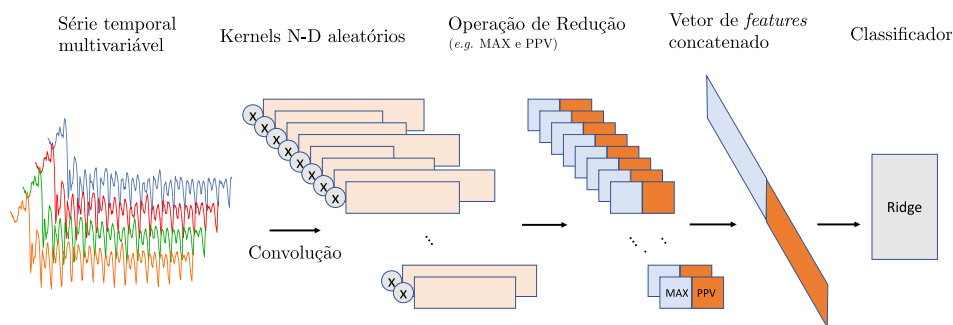


Figura 1. Fluxograma da aplicação do ROCKET a séries temporais multivariáveis. Adaptado de Middlehurst et al. (2024).

Na prática são gerados, aleatoriamente, 10.000 *kernels* que obedecem as seguintes regras: tamanho $l_k \in \{7, 9, 11\}$;



peso $w \in (0, 1]$; viés $b \in [-1, 1]$; dilatação $d = 2^x$ com $x \in [0, A]$ e $A = \log_2[(l_s - 1)/(l_k - 1)]$ em que l_s representa o tamanho da série temporal; e preenchimento com zeros, parâmetro booleano que se verdadeiro alinha o elemento “central” do *kernel* com todos os ponto da série. Cada convolução é realizada com passo 1. Como de cada um dos 10.000 *kernels* foram obtidas duas features (MAX e PPV), cada subsérie temporal foi transformada em um vetor de 20.000 características.

O procedimento de treinamento e teste foi realizado de forma independente para as arquiteturas *Branched* e *Looped*. Inicialmente, as séries temporais de 30 segundos foram segmentadas em subséries de 600 pontos cada, sem sobreposição, o que equivale a uma janela de tempo de aproximadamente 23,44 ms. Como foram considerados quatro sensores (dois acelerômetros e dois de pressão), cada subsérie foi representada por uma matriz multidimensional de formato 4×600 . Esse processo gerou 5.796 subséries para cada classe na arquitetura *Branched* e 6.004 na arquitetura *Looped*.

Para a tarefa de detecção, o problema foi modelado como uma classificação binária: as subséries da condição sem vazamento (NL) receberam o rótulo 0, enquanto todas as outras condições de vazamento (LK) receberam o rótulo 1.

O detector obteve uma acurácia de 0,949 na arquitetura *Branched* e de 0,999 na arquitetura *Looped*, resultando em uma acurácia média de 0,974. A Tabela 1 apresenta as matrizes de confusão normalizadas para cada arquitetura.

Tabela 1. Matrizes de confusão normalizadas de detecção para as arquiteturas *Branched* e *Looped*.

<i>Branched</i>				<i>Looped</i>			
		Previsto				Previsto	
		NL	LK			NL	LK
Real	NL	0.893	0.107	Real	NL	0.998	0.002
	LK	0.037	0.963		LK	0.000	1.000

Dado o excelente desempenho, combinado à sua simplicidade de implementação e alta acurácia, a abordagem ROC-KET com um classificador *Ridge* demonstra ser uma alternativa promissora e viável para a implementação de detectores multimodais de vazamentos em redes de água.

REFERÊNCIAS

- Aghashahi, M., Sela, L., and Banks, M. K. (2023). Benchmarking dataset for leak detection and localization in water distribution systems. *Data in Brief*, 48:109148.
- Bohorquez, J., Alexander, B., Simpson, A. R., and Lambert, M. F. (2020). Leak detection and topology identification in pipelines using fluid transients and artificial neural networks. *Journal of Water Resources Planning and Management*, 146.
- Dempster, A., Petitjean, F., and Webb, G. I. (2020). Rocket: exceptionally fast and accurate time series classification using random convolutional kernels. *Data Mining and Knowledge Discovery*, 34:1454–1495.
- El-Zahab, S., Abdelkader, E. M., and Zayed, T. (2018). An accelerometer-based leak detection system. *Mechanical Systems and Signal Processing*, 108:276–291.
- Mahdi, N. M., Jassim, A. H., Abulqasim, S. A., Basem, A., Ogaili, A. A. F., and Al-Haddad, L. A. (2024). Leak detection and localization in water distribution systems using advanced feature analysis and an artificial neural network. *Desalination and Water Treatment*, 320:100685.
- Middlehurst, M., Schäfer, P., and Bagnall, A. (2024). Bake off redux: a review and experimental evaluation of recent time series classification algorithms. *Data Mining and Knowledge Discovery*, 38:1958–2031.
- Okeya, I., Kapelan, Z., Hutton, C., and Naga, D. (2014). Online burst detection in a water distribution system using the kalman filter and hydraulic modelling. *Procedia Engineering*, 89:418–427.
- Shukla, H. and Piratla, K. (2020). Leakage detection in water pipelines using supervised classification of acceleration signals. *Automation in Construction*, 117:103256.
- Ullah, N., Ahmed, Z., and Kim, J.-M. (2023). Pipeline leakage detection using acoustic emission and machine learning algorithms. *Sensors*, 23:3226.

AGRADECIMENTOS

Agradecemos o financiamento fornecido pela Fundação para o Incremento da Pesquisa e do Aperfeiçoamento Industrial (FIPAI) sob o processo 010/24.

RESPONSABILIDADE PELAS INFORMAÇÕES

Os autores são os únicos responsáveis pelas informações incluídas neste trabalho.