

Detecção de Novidades Baseada nas Métricas de Similaridades Dinâmicas: DTW, EDR e TWED, Aplicadas em Sinais de Qualidade de Energia

Paulo Victor Lopes Pires* Eder Barboza Kapisch*
Leandro Rodrigues Manso Silva* Carlos Augusto Duque*
Paulo Fernando Ribeiro**

* Departamento de Engenharia Elétrica, Universidade Federal de Juiz de Fora, MG, (e-mail: pires.paulo@engenharia.ufjf.br).

** Departamento de Engenharia Elétrica, Universidade Federal de Itajubá, MG

Abstract: This article presents the use of three dynamic similarity metrics: Dynamic Time Warping (DTW), Edit Distance on Real signal (EDR) and Time Warp Edit Distance (TWED) for detection of disturbances (novelties) in Electric Power Quality (EPQ) signals. DTW, EDR and TWED are mathematical tools for comparing two vectors, generating a measure of similarity between them. The detection of novelties happens by comparing a reference frame (vector) with the frame being analyzed. If the distance between the two frames exceeds a threshold value, a novelty is detected. To assess the quality of the proposed methods, the detection probability and the false alarm probability are determined, using the ROC (Receiver Operating Characteristic) curve. The tests are performed in a Matlab[®] environment. The three metrics showed a good performance in detecting novelty of signals, mainly TWED for signals with SNR = 60 dB and EDR for signals with SNR = 30 dB, attesting to the importance of applicability in detecting disturbances present in Electric Power Systems (EPS).

Resumo: O presente artigo apresenta a utilização de três métricas de similaridade dinâmicas: *Dynamic Time Warping* (DTW), *Edit Distance on Real signal* (EDR) e *Time Warp Edit Distance* (TWED) para detecção de distúrbios (novidades) em sinais de Qualidade de Energia Elétrica (QEE). DTW, EDR e TWED são ferramentas matemáticas de comparação entre dois vetores, gerando uma medida de similaridade entre eles. A detecção de novidades acontece pela comparação de um *frame* (vetor) de referência com o *frame* em análise. Caso a distância entre os dois *frames* ultrapasse um valor de limiar, detecta-se uma novidade. Para avaliar a qualidade dos métodos propostos, determina-se a probabilidade de detecção e a probabilidade de falso alarme, utilizando-se para isso a curva ROC (*Receiver Operating Characteristic*). Os testes são realizados em ambiente Matlab[®]. As três métricas apresentaram um bom desempenho em detecção de novidades de sinais, principalmente a TWED para sinais com SNR = 60 dB e, EDR para sinais com SNR = 30 dB, atestando a importância da aplicabilidade em detecção de distúrbios presentes em Sistemas Elétricos de Potência (SEP).

Keywords: Novelty Detection; Electric Power Quality; Electric Power Systems; Dynamic Time Warping; Edit Distance on Real signal; Time Warp Edit Distance.

Palavras-chaves: Detecção de Novidades; Qualidade de Energia Elétrica; Sistemas Elétricos de Potência; Dynamic Time Warping; Edit Distance on Real signal; Time Warp Edit Distance.

1. INTRODUÇÃO

Os Sistemas Elétricos de Potência (SEP) estão passando por uma evolução significativa, onde as estruturas convencionais estão sendo substituídas por sistemas inteligentes. Consequentemente, o conceito de Qualidade de Energia Elétrica (QEE) ganhou destaque em todo o mundo (Naderi et al., 2020). Com o desenvolvimento tecnológico, os esforços dos profissionais da área acadêmica e técnica procuram soluções e produtos que estejam de acordo com as normas QEE (IEEE, 2019). Os SEP's estão em constante evolução, tornando-se sistemas sofisticados, conhecidos como *Smart Grids* (SG) (Rezkallah et al., 2019).

Os SEP's são representados por três unidades básicas: geradores de energia, redes de transmissão e distribuição de energia e consumidores de energia, estes, representados pelas indústrias, comércios e residências. A energia é considerada um produto, e o mercado cada vez mais exigente, busca um produto de qualidade, ou seja, a QEE, tornou-se um parâmetro de extrema importância.

Antigamente, as cargas lineares, como os bancos de capacitores, eram responsáveis pelo controle de geração, transmissão e distribuição de energia. Atualmente, em sua maioria, as cargas não-lineares são as responsáveis pelo controle de energia, seja em ambiente industrial, comer-

cial ou residencial. Cargas não-lineares, inserem distorções harmônicas nos sinais de tensão e corrente, afetando a QEE (Jain, 2018). Dadas as condições dinâmicas do sistema (variações de carga e/ou geração) pode ocorrer um desequilíbrio entre geração e carga o que causa desvios da frequência do sistema (Phadke and Thorp, 2017).

Sabe-se que a presença de distúrbios elétricos afeta a QEE, resultando em vários efeitos negativos, tais como superaquecimento de cabos e transformadores, afetando sua vida útil, dispositivos de proteção instáveis e superaquecimento de máquinas, tais como motores. Se os distúrbios não forem detectados e tratados adequadamente, casos extremos podem ocorrer, como o colapso do SEP (Johnson and Hassan, 2016). Portanto, tornam-se necessárias técnicas de processamento de sinais, capazes de detectar os distúrbios em sinais de tensão (ou de corrente). As técnicas de detecção de novidades, têm ganhado destaque em pesquisas, tornando-se úteis na realização de tarefa de detecção de distúrbios (Kapisch et al., 2014).

Em (Bollen and Gu, 2006) são definidos os conceitos de *triggering* que delimita os instantes de início e fim de um distúrbio em um sinal. Em (Musaruddin and Zivanovic, 2011), propõe-se a técnica de processamento de sinais denominadas de Decomposição em Modos Empíricos (DME) na tarefas de segmentação e identificação de distúrbios nos sinais. Em (Travassos et al., 2020) é utilizada a medida de dissimilaridade Ruzicka para detecção de novidades, através da comparação de dois *frames* de um sinal de potência (tensão ou corrente). Em (Silva, 2016) utilizou a técnica de Diferença de Energia entre Frames (DEF), ou seja, valor absoluto da diferença de energia entre os *frames* analisados, para detecção de novidades.

Nos últimos anos, os métodos de recuperação de formas de sinais e imagens, através das métricas de similaridades, onde cada ponto do contorno é representado por elementos de várias dimensões, tiveram uma presença significativa nas pesquisas técnicas e acadêmicas (Palazón-González and Marzal, 2012). As formas dos sinais e imagens, representadas por seus pontos do contorno, podem ser representadas como vetores.

Considerando-se duas sequências (vetores), compará-las é um processo que mede o quanto elas se diferem entre si. Às vezes existe um alinhamento natural entre os elementos de ambas as sequências e podemos compará-las usando métricas de similaridades clássicas, como por exemplo, a Distância Euclidiana. Entretanto, na maioria das situações, isso não acontece, pois as sequências podem sofrer algum tipo de distorção, ou terem comprimentos diferentes, tornando-se mais complexas, sendo necessária a utilização de métodos mais eficientes como os baseados em programação dinâmica (Cormen et al., 2009). Dentre eles, pode-se citar os métodos baseados em distâncias como a *Edit Distance on Real signal* (EDR) (Chen et al., 2005), a *Dinamic Time Warping* (DTW) (Sakoe and Chiba, 1978) e a *Time Warp Edit Distance* (TWED) (Marteau, 2008), entre outras.

A métricas de similaridades baseadas em programação dinâmica buscam encontrar o alinhamento ideal entre as duas sequências e podem ser utilizadas para determinação de similaridades entre as sequências (Salvador and Chan,

2007). O presente trabalho propõe uma nova técnica de detecção de novidades, através das métricas de similaridades dinâmicas, para comparação de *frames* (vetores) de sinais de tensão/corrente, através das formas de ondas dos respectivos sinais.

O artigo está dividido como a seguir. A seção 2 descreve os conceitos teóricos sobre detecção de novidades por comparação de *frames*, bem como as métricas de similaridades dinâmicas DTW, EDR e TWED. Os resultados de detecção de novidades através das respectivas métricas de similaridades dinâmicas estão expostos na seção 3. E por fim, a seção 4 traz as considerações finais do trabalho.

2. DETECÇÃO DE NOVIDADES BASEADA EM MÉTRICAS DE SIMILARIDADES

Nesta seção, é apresentada uma descrição teórica das métricas de similaridade dinâmicas DTW, EDR e TWED, utilizadas para detecção de novidades em sinais de tensão (ou de corrente). Cada vetor representa um determinado ciclo (*frame*) do sinal a ser analisado. Quanto mais semelhantes forem os vetores dos sinais em questão, menor a distância entre eles, ou seja, mais próxima de zero é a medida de similaridade. Por outro lado, quanto maior for essa medida, menor será a similaridade entre os respectivos vetores, indicando a presença de uma novidade.

2.1 Detecção de novidades por comparação de *frames*

O processo de detecção de novidade inicia-se pela segmentação do sinal analisado em (*frames*). Os *frames* podem ter comprimento fixo ou variável, ser sobrepostos, ou não. Os sinais de tensão (ou de corrente) em sistemas de potência são caracterizados pela presença predominante de uma componente senoidal de frequência fundamental, cuja duração do ciclo tem pouca ou nenhuma variação. Portanto, é natural escolher *frames* de comprimento fixo com uma duração de um número inteiro de ciclos N_c desta componente fundamental.

A comparação de *frames* para identificação de diferenças ou novidades entre eles, precisa de uma métrica que dimensione os *frames*. Desta forma, os *frames* são comparados através da utilização das métricas.

Um *frame* de referência é usado na comparação com outro *frame* presente no sinal. Se a comparação entre os *frames* exceder um certo limite (valor de limiar) pré-estabelecido, detecta-se uma novidade.

Nos métodos de detecção de novidade baseados em métricas de similaridade com um *frame* de referência fixo, o *frame* de referência é sempre o *frame* anterior ao *frame* analisado, independentemente da detecção de novidade. Nos métodos com um *frame* de referência dinâmico, este *frame* é atualizado somente a cada detecção de novidade. Nesse caso, o novo *frame* de referência se torna o último *frame* de novidade detectado. Em geral, os métodos que utilizam *frames* de referência dinâmicos apresentam melhor desempenho na detecção de novidades.

O limiar utilizado na comparação de *frames* pode ser fixo ou adaptável. Nos métodos de detecção de novidade com um limiar fixo, não há mudança no limiar aplicado à comparação de *frames*, independentemente da presença,

ou não, de novidades. Os métodos de detecção com limiar adaptativo, procuram adaptar ao limiar, caso ocorram variações no sinal, tais como mudanças no nível de ruído, presença de distúrbios elétricos, entre outros.

A segmentação adequada é essencial para a detecção de novidades de forma eficiente. *Frames* maiores têm um número maior de pontos, contribuindo para um aumento da quantidade de dados para o cálculo da métrica de similaridade. Além disso, com *frames* longos, a detecção de novidade é menos sensível, uma vez que a informação de distúrbios é diluída em *frames* longos. Por outro lado, *frames* mais curtos proporcionam maior precisão na localização e identificação de novidades, são mais sensíveis a distúrbios, e têm menos pontos para o cálculo da métrica de similaridade. Apesar dos *frames* curtos serem mais sensíveis a novidades, eles podem gerar um número excessivo de gatilhos, portanto, torna-se necessário uma segmentação adequada dos *frames* de um sinal.

O presente trabalho adotou o tamanho da segmentação do sinal de tensão (ou de corrente) de um ciclo do componente fundamental. Este é independente do tipo de distúrbio, que porventura, pode estar presente no sinal. Portanto, a detecção de novidades será através da comparação entre os dois *frames* do sinal.

Para ilustrar os conceitos de *frame* de referência dinâmico apresentados anteriormente, considere o sinal mostrado na Figura 1, que foi segmentado em *frames* não sobrepostos com $N_c = 1$. Este sinal tem um distúrbio Transitório Oscilatório, que ocorre durante o quarto ciclo (ou quarto *frame*) do sinal cosseno.

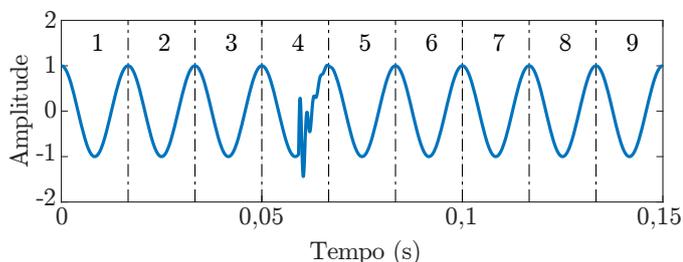


Figura 1. *Frame* de novidade detectada para Transitório Oscilatório.

Inicialmente, o primeiro *frame* é considerado como *frame* de referência. Quando o segundo *frame* é comparado com o *frame* de referência, não há diferença entre eles, conseqüentemente, não há detecção de novidade, portanto, o *frame* de referência permanece o mesmo. De modo análogo, o mesmo ocorre com o terceiro *frame*, comparado com o *frame* de referência, não há detecção de novidade.

Ao comparar o quarto *frame* com o *frame* de referência, uma diferença é notada, devido ao Transitório Oscilatório presente no quarto ciclo. Desta vez, uma novidade é detectada, portanto, o quarto *frame* será o novo *frame* de referência.

Quando o quinto *frame* é comparado com o novo *frame* de referência, uma diferença é notada, pois o quinto *frame* retorna ao seu estado normal sem perturbações. Portanto, o quinto *frame* é o *frame* de novidade e se torna o novo *frame* de referência.

Ao comparar o sexto *frame* com o novo *frame* de referência, agora o quinto *frame*, não há diferença entre eles, portanto, não há detecção de novidade. O quinto *frame* permanece como o *frame* de referência e a comparação passa para o *frame* seguinte. Como o sétimo *frame* também não mostra nenhuma diferença em relação ao quinto *frame*, o *frame* de referência permanece o mesmo, pois nenhuma novidade é detectada. O mesmo resultado se repete até o último *frame*.

2.2 Dynamic Time Warping – DTW

Para fins de análise de detecção de novidades, utilizou-se um sinal de tensão (ou de corrente) de amplitude normalizada. Cada ciclo do respectivo sinal é um *frame* de pontos representado matematicamente por um vetor unidimensional. A identificação de novidade será através da comparação de um *frame* de referência \mathbf{x} e, um *frame* em análise \mathbf{y} (1):

$$\begin{aligned} \mathbf{x} &= [x_1 \ x_2 \ \cdots \ x_m \ \cdots \ x_M] \\ \mathbf{y} &= [y_1 \ y_2 \ \cdots \ y_n \ \cdots \ y_N] \end{aligned} \quad (1)$$

onde $x_m, m = 1, 2, \dots, M$ e $y_n, n = 1, 2, \dots, N$ são as respectivas amostras dos vetores \mathbf{x} e \mathbf{y} .

Em relação às métricas clássicas de similaridade de distância, como a Distância Euclidiana, a análise de similaridade entre os dois vetores, ocorre através da correspondência “ponto a ponto”. Já na métrica de similaridade dinâmica DTW, a análise utiliza a correspondência “um ponto para muitos pontos” ou “muitos pontos para um ponto”, ou seja, correspondência multiponto, como pode-se observar na Figura 2 (Tavenard, 2021). Caso os sinais sejam idênticos, o valor da métrica DTW será nulo (Chen et al., 2012; Salvador and Chan, 2007).

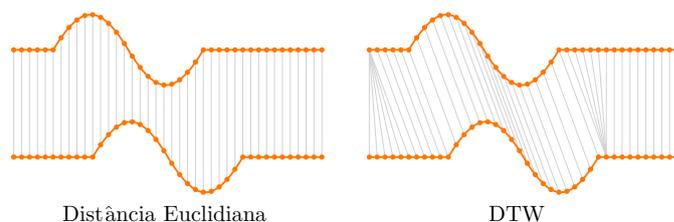


Figura 2. Distância Euclidiana e DTW.

Percebe-se que a métrica de similaridade dinâmica DTW é uma métrica flexível, pois, além de trabalhar com a correspondência multipontos, suporta vetores de comprimentos diferentes também. Seguem-se as equações descritas pelo algoritmo DTW (2), salientando que o algoritmo é recursivo, ou seja, a métrica DTW é uma programação dinâmica, cujo objetivo, encontrar o menor valor DTW entre os dois vetores.

$$D_{m,n} = d_{m,n}(\mathbf{x}, \mathbf{y}) + \min \begin{cases} D_{m-1,n}, & \text{ou} \\ D_{m,n-1}, & \text{ou} \\ D_{m-1,n-1} \end{cases} \quad (2)$$

O primeiro termo, após o sinal de igualdade (2), é conhecido como função custo, representada pelas métricas clássicas de similaridade de distância. O presente trabalho utilizou a Distância Euclidiana (3) :

$$d_{m,n}(\mathbf{x}, \mathbf{y}) = \sqrt{(x_m - y_n)^2} \quad (3)$$

O processamento do algoritmo DTW, funciona da seguinte forma: os vetores (1) serão responsáveis pela construção da matriz D de ordem $m \times n$. Cada amostra da m -ésima linha e n -ésima coluna, será representada matematicamente como $D_{m,n}$, que será a soma entre a função custo (3) e ao menor valor encontrado das amostras adjacentes anteriores, ou seja, amostra linha anterior $D_{m-1,n}$, amostra coluna anterior $D_{m,n-1}$, ou amostra diagonal anterior $D_{m-1,n-1}$, pode ser visto na Figura 3.

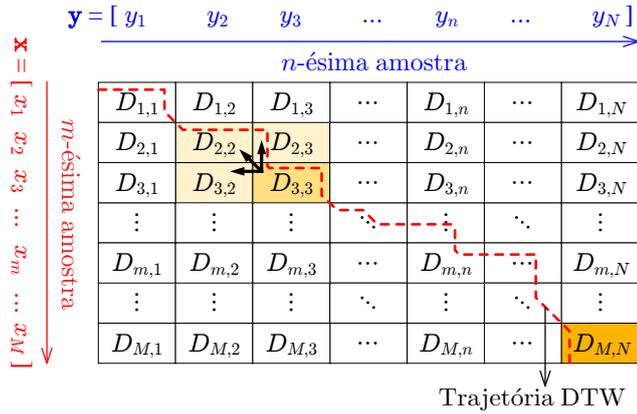


Figura 3. Matriz DTW.

Como algoritmo DTW inicia-se em $D_{M,N}$ e termina em $D_{1,1}$ ou vice-versa. A trajetória, representada pelas amostras $D_{m,n}$, serão os possíveis caminhos, durante o processamento do algoritmo DTW. Quanto mais próximo a trajetória for da diagonal principal da matriz, mais similares serão os vetores \mathbf{x} e \mathbf{y} . Caso os vetores sejam idênticos, a trajetória é a própria diagonal principal da matriz. O valor da medida de similaridade é a soma de todos os elementos da trajetória. Mais detalhes do processamento do algoritmo DTW podem ser encontrados em (Sakoe and Chiba, 1978).

Para a determinação do valor da métrica de similaridade DTW nos testes, utilizou-se o software Matlab[®], através da função `dtw(x,y,metric)`, onde os parâmetros \mathbf{x} e \mathbf{y} são os vetores *frames* do sinal em análise, e o parâmetro *metric*, é responsável pela especificação da função custo.

2.3 Edit Distance on Real signal – EDR

Utilizando-se do mesmo raciocínio do primeiro parágrafo da subseção 2.2, ou seja, cada *frame* do sinal é representado por um vetor, a métrica de similaridade EDR, além de ter uma correspondência multipontos, tem como objetivo, converter um determinado vetor em outro vetor, através de três operações: adição, substituição e remoção de pontos do vetor em análise. Em relação à métrica DTW, a métrica EDR é mais robusta a ruídos. (Chen et al., 2005).

A métrica EDR é uma extensão da métrica de distância Levenshtein (Levenshtein et al., 1966). Seguem-se as equações descritas pelo algoritmo EDR (4), para a determinação do grau de similaridade entre os vetores:

$$D_{m,n} = \min \begin{cases} D_{m-1,n} + 1, & \text{ou} \\ D_{m,n-1} + 1, & \text{ou} \\ D_{m-1,n-1} + \begin{cases} 0 \Leftarrow d_{m,n}(x,y) \leq \varepsilon \\ 1 \Leftarrow d_{m,n}(x,y) > \varepsilon \end{cases} \end{cases} \quad (4)$$

onde ε é a tolerância, também conhecida como parâmetro de elasticidade, e $d_{m,n}(x,y)$ é a Distância Euclidiana (3).

O processamento do algoritmo EDR funciona da seguinte forma: os vetores (1) serão responsáveis pela construção da matriz D de ordem $m \times n$. Cada amostra da m -ésima linha e n -ésima coluna, será representada matematicamente como $D_{m,n}$, que será o menor valor encontrado das amostras adjacentes anteriores, ou seja, amostra linha anterior $D_{m-1,n} + 1$, amostra coluna anterior $D_{m,n-1} + 1$, ou amostra diagonal anterior, esta, dependerá do parâmetro de elasticidade ε , ou seja, $D_{m-1,n-1} + 0$, ou $D_{m-1,n-1} + 1$, como pode ser visto na Figura 4. O valor da métrica de similaridade EDR será $D_{M,N}$. Mais detalhes sobre o processamento do algoritmo EDR podem ser encontrados em (Chen et al., 2005).

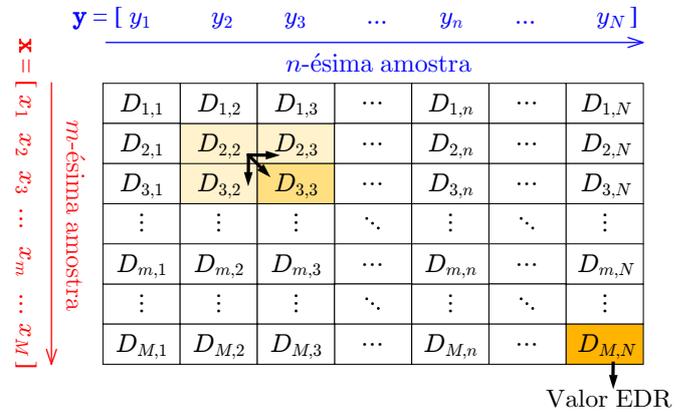


Figura 4. Matriz EDR.

Para a determinação do valor da métrica EDR nos testes, utilizou-se o software Matlab[®], através da função `edr(x,y,tol)`, onde os parâmetros \mathbf{x} e \mathbf{y} são os vetores *frames* do sinal em análise, e o parâmetro *tol*, é tolerância, responsável pela especificação do parâmetro de elasticidade.

2.4 Time Warp Edit Distance – TWED

A métrica *Time Warp Edit Distance* TWED, foi desenvolvida por Marteau (Marteau, 2008). Em certo sentido, é uma combinação das métricas DTW e EDR. Além das vantagens discutidas anteriormente para DTW e EDR, presentes também para TWED, esta também é capaz de lidar com séries temporais de diferentes taxas de amostragem, incluindo as séries temporais de amostragens reduzidas (Serra and Arcos, 2014). Seguem-se as equações descritas pelo algoritmo TWED (5):

$$D_{m,n} = \min \begin{cases} D_{m-1,n} + \Gamma_x, & \text{ou} \\ D_{m,n-1} + \Gamma_y, & \text{ou} \\ D_{m-1,n-1} + \Gamma_{xy} \end{cases} \quad (5)$$

onde,

$$\begin{aligned} \Gamma_x &= D_{m-1,n} + d_{m,m-1}(x,x) + \nu + \lambda \\ \Gamma_y &= D_{m,n-1} + d_{n,n-1}(y,y) + \nu + \lambda \\ \Gamma_{x,y} &= D_{m-1,n-1} + d_{m,n}(x,y) + d_{m-1,n-1}(x,y) + 2\nu|m-n| \end{aligned} \quad (6)$$

onde ν e λ são parâmetros de elasticidade e penalidade, respectivamente e, $d_{m,n}(x,y)$, $d_{m-1,n-1}(x,y)$, $d_{m,m-1}(x,x)$ e $d_{n,n-1}(y,y)$, são as funções custos. O presente trabalho as adotou as mesmas, ou seja, Distância Euclidiana (3).

O processamento do algoritmo TWED é semelhante ao algoritmo EDR, o que muda, são os cálculos em relação as m -ésima linha com a n -ésima coluna, presentes no algoritmo TWED. Como o parâmetro de penalidade λ não interfere na dinâmica do processamento do algoritmo, o parâmetro foi desconsiderado, ou seja, adotado igual a zero. Em relação ao parâmetro de elasticidade ν , o valor adotado será igual ao parâmetro de elasticidade do algoritmo EDR.

Para a determinação do valor da métrica TWED, através do software Matlab[®], utilizou-se a função `twed(A, timeSA, B, timeSB, lambda, nu)`, onde os parâmetros A e B são os vetores *frames* do sinal em análise, `timeSA` e `timeSB`, são os números de pontos que os respectivos vetores têm, os parâmetros `lambda` e o `nu`, são os parâmetros de penalidade e elasticidade, respectivamente. O código completo encontra-se em (Wikipedia contributors, 2022).

Em relação aos custos computacionais, a Distância Euclidiana tem $O(n)$, ou seja, o custo computacional aumenta linearmente à medida que a quantidade de dados cresce. As métricas DTW, EDR e TWED têm custo computacional $O(n^2)$, ou seja, aumenta exponencialmente com a quantidade de dados (Rojas et al., 2016).

3. RESULTADOS

Esta seção mostra os resultados das métricas DTW, EDR e TWED na detecção de novidades. Também são mostrados os testes estatísticos para avaliar o desempenho das métricas. O desempenho é avaliado em termos das curvas ROC e, da área abaixo da curva (AUC – *Area Under Curve*), para ambas métricas, em dois cenários de ruído, SNR 60 dB e SNR 30 dB, no sinal de tensão (ou de corrente).

3.1 Detecção de novidade com as métricas de similaridades

Os resultados das métricas e da detecção de novidades são mostrados na Figura 5, Figura 6 e Figura 7, para as métricas DTW, EDR e TWED, respectivamente. À esquerda de cada figura, está o sinal composto de dez *frames* com um determinado tipo de distúrbio. Nesse gráfico, os *frames* de novidades estão destacados pelo símbolo \boxed{N} . À direita de cada figura, estão as métricas de similaridade dinâmicas, calculadas para cada ciclo, representada pela linha em preto. A linha em vermelho é o gabarito para detecção de novidades, ou seja, os *frames* em que as novidades deveria ser detectadas. Esse gabarito apresenta valor 1 quando uma novidade deve ser detectada, e 0 caso contrário.

Pode-se observar que a detecção de novidade pelas métricas está de acordo com o gabarito, consequentemente, maior a probabilidade de detecção. Caso contrário, maior seria a probabilidade de falso alarme.

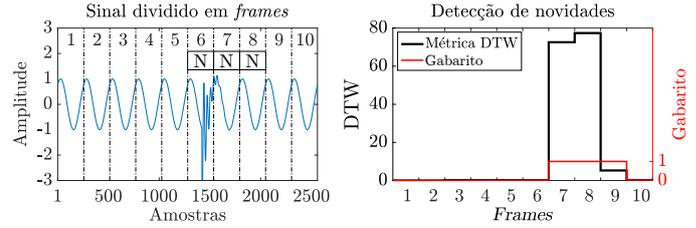


Figura 5. Detecção de novidade – DTW, para Transitório Oscilatório.

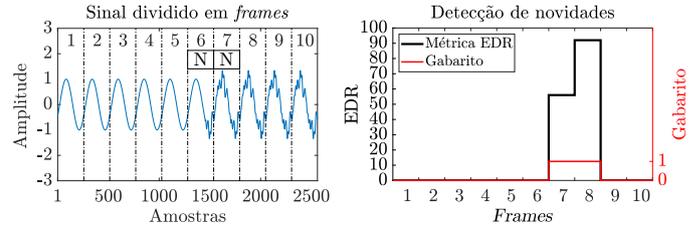


Figura 6. Detecção de novidade – EDR, para distorção harmônica.

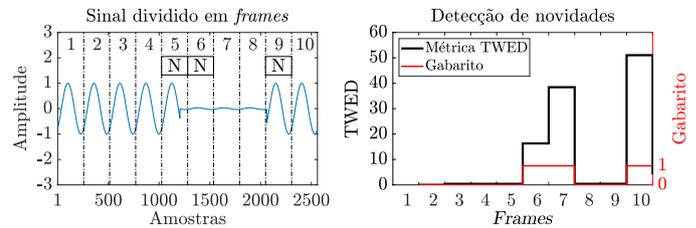


Figura 7. Detecção de novidade – TWED, para Interrupção.

3.2 Curvas ROC

Para analisar a eficiência do detector de novidade baseado nas métricas, foi construída a curva *Receiver Operating Characteristic* (ROC) (Van Trees, 2004), que depende de dois parâmetros, a probabilidade de detecção P_D e a probabilidade de falso alarme P_{FA} , representadas em (7):

$$P_D = \frac{PV}{TP}, \quad P_{FA} = \frac{NV}{TN}, \quad (7)$$

onde PV é o número de *frames* verdadeiros positivos (ciclos) detectados pelo detector de novidade, TP é o número total de *frames* positivos no sinal, NV é o número de *frames* verdadeiros negativos não detectados pelo detector de novidade, TN é o número total de *frames* negativos no sinal. Quanto mais próximo o P_D estiver de 1 e P_{FA} de 0, melhor será a eficiência do detector.

A curva ROC é gerada pela variação da sensibilidade constante (o limiar) do detector. Através da curva ROC, é possível obter informações sobre a eficácia do detector. Um dos parâmetros indispensáveis, é a área abaixo da curva, (*Area Under Curve* – AUC), que em um detector ideal tem o valor igual a 1.

Para determinar o limiar ótimo, foi utilizado o índice de Soma Produto (SP) (Simas Filho et al., 2009), e o limiar que originou o ponto com o SP mais alto é o limiar ótimo. O índice de SP é calculado de acordo com (8):

$$SP = \sqrt{\sqrt{P_D \times (1 - P_{FA})} \times \frac{P_D + (1 - P_{FA})}{2}} \quad (8)$$

Para os testes, utilizou-se 200 sinais com determinados distúrbios, cada sinal contém 40 ciclos. Os sinais foram gerados com os seguintes parâmetros: amplitudes normalizadas, frequência fundamental $f = 60$ Hz, e frequência de amostragem $F_s = 15360$ Hz. O presente trabalho utilizou os seguintes distúrbios: Transitório Oscilatório, elevação de tensão (também conhecida como *Swell*), afundamento de tensão (também conhecida como *Sag*), Interrupção, Harmônicos variantes no tempo, *Spikes* e *Notches*. Foram adicionados ruídos no sinal com SNR = 60 dB e SNR = 30 dB, com o objetivo de analisar a eficiência das métricas DTW, EDR e TWED. Para as métricas EDR e TWED, o valor do parâmetro de elasticidade adotado foi de 0,12. Utilizou-se também o detector de novidades, baseado em Diferença de Energia entre Frames (DEF) (Silva, 2016), para fins de comparação com as técnicas de similaridades dinâmicas citadas. As Figuras 8 a 14 mostram os resultados para os distúrbios testados.

Os resultados para o distúrbio Transitório Oscilatório estão mostrados na Figura 8. Percebe-se que para este distúrbio, as métrica DTW e TWED apresentaram desempenhos melhores para SNR = 60 dB. Para SNR = 30 dB, a métrica TWED foi melhor.

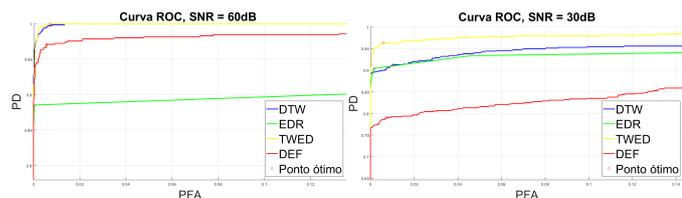


Figura 8. Curva ROC para Transitório Oscilatório.

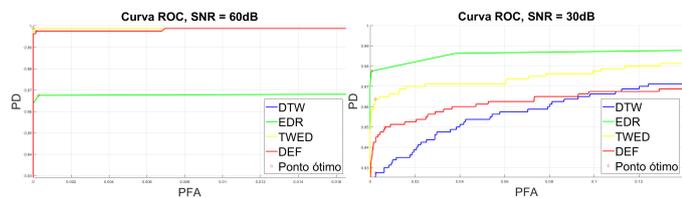


Figura 9. Curva ROC para elevação de tensão.

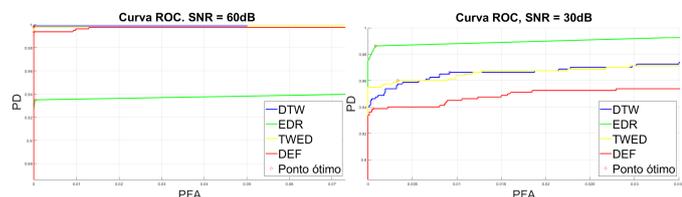


Figura 10. Curva ROC para afundamento de tensão.

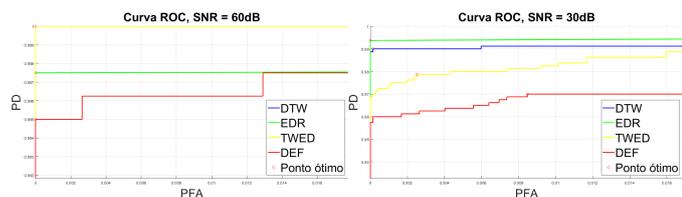


Figura 11. Curva ROC para Interrupção.

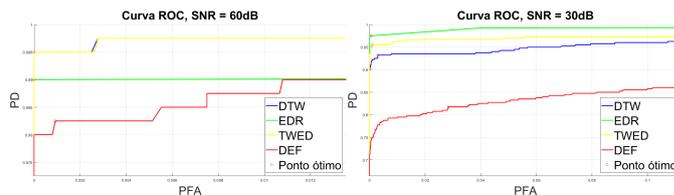


Figura 12. Curva ROC para distorção harmônica.

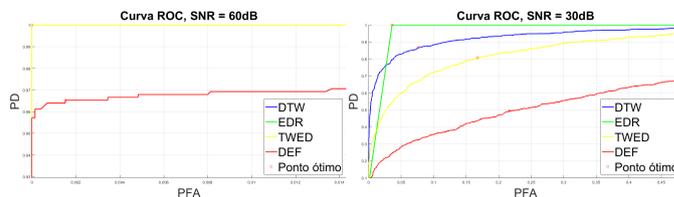


Figura 13. Curva ROC para Spikes.

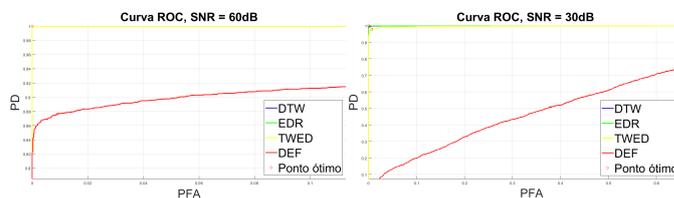


Figura 14. Curva ROC para Notches.

Os resultados para elevação de tensão, afundamento de tensão, e Interrupção são mostrados na Figura 9, Figura 10 e Figura 11, respectivamente. Para elevação de tensão, DTW, TWED e DEF tiveram ótimos desempenhos para SNR = 60 dB. Para afundamento de tensão, a TWED teve um desempenho um pouco melhor do que a DTW e DEF para SNR = 60 dB. No caso de Interrupções, as métricas DTW e TWED tiveram o mesmo desempenho, ambos muito bons, para SNR = 60dB. Observa-se que para ambos os distúrbios, com SNR = 30 dB, a métrica EDR foi a que apresentou um desempenho melhor.

Na Figura 12, para distorção harmônica, mostra-se que a métrica DTW e a métrica TWED tiveram praticamente o mesmo desempenho, ambos melhores, para SNR = 60 dB. Ambas as métricas apresentaram um desempenho inferior para SNR = 30 dB, onde a métrica EDR apresentou-se melhor.

O resultado para *Spikes* e *Notches* são apresentados nas Figuras 13 e 14, respectivamente. Todas as métricas, exceto a DEF, apresentaram bom desempenho para SNR = 60 dB. Devido às características destes distúrbios, principalmente a *Spikes*, que são parecidas aos ruídos com SNR = 30 dB, pode ser difícil identificar as novidades. A métrica EDR conseguiu apresentar um desempenho melhor.

Através das curvas ROC, pode-se observar que as métricas DTW e TWED apresentaram bons desempenhos, para SNR = 60 dB e um desempenho ligeiramente bom para SNR = 30 dB. Neste, a métrica EDR foi melhor, exceto para o distúrbio Transitório Oscilatório, no qual a métrica TWED apresentou-se melhor. Os resultados da AUC para as curvas ROC foram extraídos e organizadas na Tabela 1.

Tabela 1. Área abaixo da curva – AUC

Métricas →	DTW		EDR		TWED		DEF	
	60 dB	30 dB						
SNR →	AUC							
Distúrbios ↓	AUC							
Transitório Oscilatório	0.9999	0.9782	0.9427	0.9650	0.9999	0.9926	0.9933	0.9189
Elevação de tensão	0.9992	0.9868	0.9844	0.9927	0.9992	0.9901	0.9989	0.9800
Afundamento de tensão	0.9991	0.9904	0.9675	0.9960	0.9993	0.9933	0.9980	0.9844
Spikes	0.9999	0.9572	0.9999	0.9819	0.9999	0.8987	0.9923	0.6717
Notches	0.9999	0.9999	0.9999	0.9999	0.9999	0.9984	0.9580	0.5954
Interrupção	0.9999	0.9954	0.9988	0.9974	0.9999	0.9970	0.9991	0.9878
Distorção harmônica	0.9977	0.9850	0.9950	0.9958	0.9976	0.9884	0.9929	0.9362

4. CONCLUSÃO

O presente trabalho apresentou um detector de novidades aplicado a sinais de Qualidade de Energia Elétrica (QEE), baseados nas métricas de similaridade dinâmicas, DTW, EDR e TWED. Comparou-se os desempenhos das métricas citadas com o detector de novidades Diferença de Energia entre Frames (DEF) para vários distúrbios de QEE: Transitório Oscilatório, elevação de tensão, afundamento de tensão, Interrupção, distorção harmônica, *Notches* e *Spikes*. Dois valores de SNR foram utilizados: 60 dB e 30 dB, no sinal de tensão (ou corrente).

Analisando os resultados, é possível notar que a métrica TWED, em relação a métrica DTW, obteve um bom desempenho para todos os distúrbios com SNR = 60 dB. Para SNR = 30 dB, a métrica EDR apresentou-se melhor desempenho em praticamente todos os distúrbios. Portanto, atesta-se a aplicabilidade das métricas de similaridades dinâmicas, nas aplicações de detecção de novidades, presentes em Sistemas Elétricos de Potência (SEP).

AGRADECIMENTOS

Os autores agradecem ao CNPq, CAPES e INERGE que financiaram este trabalho.

REFERÊNCIAS

Bollen, M.H. and Gu, I.Y. (2006). *Signal processing of power quality disturbances*, volume 30. John Wiley & Sons.

Chen, L., Özsü, M.T., and Oria, V. (2005). Robust and fast similarity search for moving object trajectories. In *Proceedings of the 2005 ACM SIGMOD international conference on Management of data*, 491–502.

Chen, Q., Hu, G., Gu, F., and Xiang, P. (2012). Learning optimal warping window size of dtw for time series classification. In *2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA)*, 1272–1277. IEEE.

Cormen, T.H., Leiserson, C.E., Rivest, R.L., and Stein, C. (2009). *Introduction to algorithms*. MIT press.

IEEE (2019). IEEE recommended practice for monitoring electric power quality - redline. *IEEE Std 1159-2019 (Revision of IEEE Std 1159-2009) - Redline*, 1–180.

Jain, S. (2018). Power quality: An introduction. In *Modeling and Control of Power Electronics Converter System for Power Quality Improvements*, 1–29. Elsevier.

Johnson, D.O. and Hassan, K.A. (2016). Issues of power quality in electrical systems. *International Journal of Energy and Power Engineering*, 5(4), 148–154.

Kapisch, E.B., Silva, L.R.M., Martins, C.H.N., Barbosa, A.S., Duque, C.A., de Andrade Filho, L.M.M., and Cerqueira, A.S. (2014). An electrical signal disturbance detector and compressor based on fpga platform. In

2014 16th International Conference on Harmonics and Quality of Power (ICHQP), 278–282. IEEE.

Levenshtein, V.I. et al. (1966). Binary codes capable of correcting deletions, insertions, and reversals. In *Soviet physics doklady*, volume 10, 707–710. Soviet Union.

Marteau, P.F. (2008). Time warp edit distance with stiffness adjustment for time series matching. *IEEE transactions on pattern analysis and machine intelligence*, 31(2), 306–318.

Musaruddin, M. and Zivanovic, R. (2011). Signal segmentation of fault records based on empirical mode decomposition. In *TENCON 2011-2011 IEEE Region 10 Conference*, 138–143. IEEE.

Naderi, Y., Hosseini, S.H., Ghassemzadeh, S., Mohammadi-Ivatloo, B., Savaghebi, M., Vasquez, J.C., and Guerrero, J.M. (2020). Power quality issues of smart microgrids: applied techniques and decision making analysis. In *Decision making applications in modern power systems*, 89–119. Elsevier.

Palazón-González, V. and Marzal, A. (2012). On the dynamic time warping of cyclic sequences for shape retrieval. *Image and Vision Computing*, 30(12), 978–990.

Phadke, A.G. and Thorp, J.S. (2017). Phasor estimation at off-nominal frequency inputs. In *Synchronized phasor measurements and their applications*, 47–72. Springer.

Rezkallah, M., Chandra, A., Hamadi, A., Ibrahim, H., and Ghandour, M. (2019). Power quality in smart grids. In *Pathways to a smarter power system*, 225–245. Elsevier.

Rojas, I., Pomares, H., and Valenzuela, O. (2016). Advances in time series analysis and forecasting selected contributions from itise 2016. In *Conference proceedings ITISE*, 15. Springer.

Sakoe, H. and Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE transactions on acoustics, speech, and signal processing*, 26(1), 43–49.

Salvador, S. and Chan, P. (2007). Toward accurate dynamic time warping in linear time and space. *Intelligent Data Analysis*, 11(5), 561–580.

Serra, J. and Arcos, J.L. (2014). An empirical evaluation of similarity measures for time series classification. *Knowledge-Based Systems*, 67, 305–314.

Silva, L.R.M. (2016). *Estudo e implementação de técnicas de detecção e compressão de distúrbios elétricos*. Ph.D. thesis, Federal University of Juiz de Fora, Juiz de Fora, Brazil. Edited by Editora UFJF, Juiz de Fora, 2016.

Simas Filho, E.F., Seixas, J.M., and Calôba, L.P. (2009). Optimized calorimeter signal compaction for an independent component based atlas electron/jet second-level trigger. *Proceedings of Science*, 8, 1–10.

Tavenard, R. (2021). An introduction to dynamic time warping. URL <https://rtavenar.github.io/blog/dtw.html>. [Online; accessed 4-March-2022].

Travassos, F.C., Kapisch, E.B., Silva, L.R.M., Lima, M.A.A., and Duque, C.A. (2020). Detecção adaptativa de novidades baseada em métrica de similaridade aplicada a sinais de qualidade de energia. In *Congresso Brasileiro de Automática-CBA*, volume 2.

Van Trees, H.L. (2004). *Detection, estimation, and modulation theory, part I: detection, estimation, and linear modulation theory*. John Wiley & Sons.

Wikipedia contributors (2022). Time warp edit distance — Wikipedia, the free encyclopedia. URL https://en.wikipedia.org/w/index.php?title=Time_Warp_Edit_Distance&oldid=1064784948. [Online; accessed 4-March-2022].