

Classificação Inteligente de Falhas Multiestágio em Alimentadores Primários de Distribuição de Energia Elétrica

Marcelo Estevão da Silva* Joaquim Osterwald Frota Moura Filho**
Juan Carlos Pequeña Suni* Márcio André Baima Amora**

* Engenharia Elétrica, Campus Sobral, Universidade Federal do Ceará, CE, (e-mail: marcelo.estevao@alu.ufc.br, jcarlosps@ufc.br)

** Programa de Pós-Graduação em Engenharia Elétrica e Computação (PPGEEC), Campus Sobral, Universidade Federal do Ceará, CE, (e-mail: joaquim1905@alu.ufc.br, marcio@sobral.ufc.br)

Abstract:

Distribution systems, due to their complex topologies and configurations, present the challenge of maintaining the reliability and continuity of the energy supply. In this sense, one of the main faults in the electrical network is the emergence of multi-stage faults, which represent 20% of fault occurrences. Aiming at the context of smart grids, and considering electricity meters that will be optimally allocated, this work proposes a methodology for classifying multistage faults in primary radial and overhead distribution feeders, based on decision trees (DA), whose Input parameters are the currents of the primary distribution feeder under study, measured only at the substation. The current oscillographs were obtained from simulations with the *software* MATLAB/SIMULINK and the signal processing method adopted was the RMS (*Root Mean Square*). Therefore, the obtained results represent an accurate classification, superior to 97%, indicating efficiency of the proposed method for the classification of such defects.

Resumo:

Os sistemas de distribuição, devido as suas topologias e configurações complexas, apresentam o desafio de manter a confiabilidade e a continuidade do suprimento de energia. Neste sentido, um dos principais defeitos na rede elétrica é o surgimento de falhas multiestágio, que representam 20% das ocorrências de falhas. Visando o contexto das redes elétricas inteligentes, e considerando medidores de energia elétrica que serão alocados de forma otimizada, este trabalho propõe uma metodologia de classificação de falhas multiestágio em alimentadores primários de distribuição radiais e aéreos, baseada em árvores de decisão (AD), cujos parâmetros de entrada são as correntes do alimentador primário de distribuição em estudo, medidas apenas na subestação. As oscilografias de corrente foram obtidas a partir de simulações com o *software* MATLAB/SIMULINK e o método de processamento dos sinais adotado foi o RMS (*Root Mean Square*). Portanto, os resultados obtidos representam uma classificação acurada, superior a 97%, indicando eficiência do método proposto para a classificação de tais defeitos.

Keywords: Machine Learning; Fault Detection; Distribution Networks; Smart Grids; Power Systems.

Palavras-chaves: Aprendizado de Máquina; Detecção de Falhas; Redes de Distribuição; Redes Inteligentes; Sistemas de Potência.

1. INTRODUÇÃO

Os atuais sistemas elétricos de potência (SEPs), por serem interligados na maioria dos casos, constituem uma complexa rede de fornecimento de energia. Neste sentido, é vital que os SEPs disponham de elevada confiabilidade e continuidade no que concerne ao fornecimento de energia elétrica às suas cargas.

Um dos principais defeitos na rede elétrica é o surgimento de falhas, também conhecidos como curtos-circuitos (Kindermann, 1997). Esses defeitos ocorrem em pontos arbitrários no SEP e podem causar severos danos aos

consumidores quando tais eventos não são devidamente identificados e isolados através dos elementos de proteção, tais como relés.

Além disso, por muito tempo, os sistemas de geração empregaram máquinas síncronas para a geração de energia. Por outro lado, com o incremento das fontes de energias renováveis em sistemas de Geração Distribuída (GD) (Rai et al., 2021), elevou-se a ocorrência de distúrbios nos sistemas de distribuição, entre eles o curto-circuito. Essas fontes, tais como a Energia Eólica, utilizam uma diversidade de máquinas assíncronas conectadas à rede trifásica. Desta forma, o arranjo de sistemas de geração síncrona e

assíncrona aumentam a chance de ocorrência de distúrbios na rede, tais como faltas de curto-circuito (Tleis, 2008).

Os eventos de curto-circuito podem ocorrer em qualquer ponto do SEP. Estatisticamente, as ocorrências de curto-circuito são mais observadas nos sistemas de transmissão, pelo fato deste sistema ser extenso e percorrer diferentes regiões do país (Kindermann, 1997). Os sistemas de distribuição, assim como os de transmissão, estão expostos a tais eventos, e por serem responsáveis pela entrega de energia à maioria dos consumidores finais, merece especial atenção (Kagan et al., 2010). Isto posto, é evidente que a detecção e classificação das faltas no sistema de distribuição constitui-se-á uma valiosa ferramenta para a segura operação da rede, além de contribuir para a melhoria dos Indicadores Coletivos de Continuidade DEC (Duração Equivalente de Interrupção por Unidade Consumidora) e FEC (Frequência Equivalente de Interrupção por Unidade Consumidora).

Desta forma, com o advento das emergentes redes elétricas inteligentes, também conhecidas como *Smart Grids*, o monitoramento das incidências de curtos-circuitos tornar-se-á indispensável em virtude da maior necessidade de confiabilidade. Outrossim, tal monitoramento será mais preciso em virtude da grande variedade de sensores e medidores inteligentes instalados na rede, que proveem dados em tempo real para monitoramento, proteção e controle (Song et al., 2022).

Diversas técnicas de detecção e classificação de faltas em sistemas de distribuição foram propostas na literatura, empregando técnicas de processamento de sinais, algoritmos de aprendizado de máquina e inteligência computacional.

Em Bernardes et al. (2021) os autores apresentaram uma ferramenta para detecção e classificação de 7 tipos de curtos-circuitos em um sistema de distribuição baseado no uso de análise multi-resolução da Transformada de *Wavelet* (TW) associada com uma rede neural *Fuzzy-ARTMAP*, que permite ao classificador aprender continuamente ao mesmo tempo que mantém o conhecimento aprendido no passado. O módulo de aquisição de dados se baseia na leitura das correntes (I_a, I_b, I_c). Este método conferiu 100% de acurácia nas faltas simuladas.

Além disso, em Sharma et al. (2018), as 4 faltas simuladas foram submetidas à Transformada de *Stockwell* (TS). Neste método, as faltas são detectadas a partir de um índice de faltas proposto, com base na média obtida da matriz de *Stockwell* dos sinais de tensão e corrente da rede de distribuição.

Em Salim et al. (2008), os autores propuseram um sistema automático de detecção, classificação, localização das faltas e determinação da seção faltante. A detecção e classificação é fundamentada na TW, a localização é baseada na impedância e a determinação da seção faltante é embasada em uma Rede Neural (RN). O percentual de erro obtido foi aproximadamente 6%.

Ademais, em Rai et al. (2021) as faltas foram classificadas considerando uma rede de distribuição com geradores distribuídos. O algoritmo empregado foi uma Rede Neural Convolucional (CNN do inglês *Convolutional Neural Network*) cujas entradas foram as tensões e correntes da

rede. Devido o emprego da CNN, os sinais neste método dispensam o uso de uma técnica de pré-processamento, ou seja, a rede CNN recebe os sinais brutos para classificação. A técnica de validação cruzada também foi empregada e a acurácia obtida foi de aproximadamente 99%.

Portanto, este trabalho propõe uma nova abordagem de classificação de faltas, considerando não só as faltas simples, mas também as faltas multiestágio, baseadas nas oscilografias de corrente de um alimentador primário de um sistema de distribuição. Além disso, o emprego da Raiz do Valor Quadrático Médio (RMS inglês *Root Mean Square*) para o processamento das oscilografias constitui um ganho de desempenho computacional em comparação com os trabalhos supracitados, em virtude da simplicidade deste método.

A organização do trabalho segue a seguinte estrutura: no capítulo 2 é comentada a fundamentação teórica incluindo a natureza das faltas simples e multiestágio, a técnica RMS e a técnica de Árvore de Decisão (AD); as simulações desenvolvidas são apresentadas no capítulo 3, incluído os detalhes de cada elemento empregado e seus respectivos parâmetros; no capítulo 4 a metodologia com o procedimento de processamento das oscilografias e aplicações do algoritmo de classificação são detalhados; os resultados e discussão são discutidos no capítulo 5; e, finalmente, no capítulo 6 são apresentadas a conclusão e as projeções de trabalhos futuros.

2. FUNDAMENTAÇÃO TEÓRICA

2.1 Faltas em Sistemas de Potência

As faltas são mudanças inesperadas na configuração normal do sistema que resulta na dissipação de energia que não seja para o suprimento da carga desejada (Kindermann, 1997). As faltas podem apresentar-se quando da ocorrência de descargas atmosféricas, surtos de comutação, sobrecargas em cabos, transformadores, motores, geradores, entre outras formas. Assim, os curtos-circuitos podem ser classificados devido à sua origem (mecânica, elétrica ou erro de operação), sua localização (dentro ou fora do equipamento) e sua duração (auto-extinguível, transitória ou permanente) (Prévé, 2006). As principais causas de curto-circuitos são devido a (Kindermann, 1997):

- Falha de isolamento;
- Falha mecânica;
- Falha elétrica;
- Falha térmica;
- Falha de manutenção.

Além disso, verifica-se a existência de duas classes de faltas elétricas no tocante ao módulo e ângulo de fase da tensão e da corrente, que são as faltas simétricas e assimétricas. Enquanto a primeira apresenta sinal de corrente simétrico em relação ao eixo temporal e característica estacionária, a segunda apresenta comportamento exponencial de natureza transitória e com amortecimento até atingir o regime (Glover et al., 2012).

No contexto do SEP brasileiro, o percentual de ocorrência de curto-circuitos é distribuído segundo a Figura 1, onde se percebe a predominância do tipo $1\phi - terra$ (Kindermann, 1997).

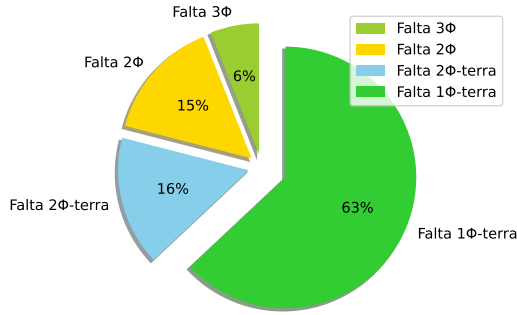


Figura 1. Ocorrência de Faltas em SEP.

Os tipos de faltas em relação às fases envolvidas de um sistema trifásico podem ser:

Falta Fase-Terra: (1ϕ -terra) É a mais frequente e mais grave no SEP. Ela ocorre quando há contato de qualquer uma das fases com o terra.

Falta Fase-Fase (2ϕ): Ocorre quando há contato entre quaisquer duas fases.

Falta Fase-Fase-Terra (2ϕ -terra): Se caracterizam pelo contato entre quaisquer duas fases e o terra.

Falta Fase-Fase-Fase-Terra (3ϕ - terra): É caracterizada pelo contato das três fases e o terra. Verifica-se que este tipo de falta é o mais raro.

2.2 Faltas Multiestágio

As faltas multiestágio são caracterizadas pelo surgimento de diferentes magnitudes de tensão antes que ela volte ao seu valor normal. Neste contexto, estima-se que pelo menos 20% das ocorrências de faltas sejam do tipo multiestágio (Styvaktakis and Bollen, 2003).

Desta forma, advém o surgimento de vários estágios dessas faltas até o completo cessamento da mesma. A Figura 2, a seguir, representa a oscilografia de uma corrente de falta multiestágio em uma das fases afetadas. A primeira seção corresponde à operação normal, seguida de uma falta assimétrica associada a um valor de impedância e, posteriormente, uma falta simétrica com um valor de impedância diferente da primeira.

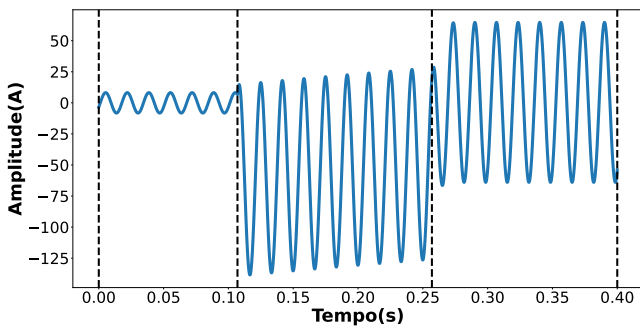


Figura 2. Oscilografia de falta multiestágio.

Por outro lado, as faltas multiestágio também podem ser definidas como múltiplas faltas que se sucedem em um curto intervalo de tempo e capturadas em um mesmo

registro. Também, pode-se verificar o surgimento de faltas multiestágio como um resultado de múltiplos mecanismos de eliminação de faltas, como relés de proteção, que não são síncronos entre si, de modo que a impedância do SEP e, conseqüentemente, a configuração da rede são modificadas (Styvaktakis and Bollen, 2003).

2.3 Classificador Árvore de Decisão

O algoritmo Árvore de Decisão (AD) é um método de aprendizado de máquina supervisionado não paramétrico empregado tanto para classificação quanto para regressão. A sua finalidade é estabelecer um modelo que aprenda regras de decisão simples inferidas a partir de um banco de dados (Breiman et al., 1984). Desta forma, considerando os vetores de treinamento $x_i \in R^n$, com $i = 1, \dots, l$ e um vetor de classes $y \in R^l$, uma AD divide recursivamente o espaço de atributos, para que as amostras de mesma classe sejam agrupadas. Então, seja um dado no nó m , representado por Q_m com N_m amostras. Para cada divisão possível $\theta = (j, t_m)$ consistindo de uma característica j e um limiar t_m , particiona-se os dados nos subconjuntos $Q_m^{esquerda}$ e $Q_m^{direita}$ (Pedregosa et al., 2011):

$$Q_m^{esquerda}(\theta) = (x, y) | x_j \leq t_m \quad (1)$$

$$Q_m^{direita}(\theta) = Q_m / Q_m^{esquerda}(\theta) \quad (2)$$

A qualidade de cada divisão no nó m é obtida pela função de impureza (Pedregosa et al., 2011):

$$G(Q_m, \theta) = \frac{N_m^{esquerda}}{N_m} H(Q_m^{esquerda}(\theta)) + \frac{N_m^{direita}}{N_m} H(Q_m^{direita}(\theta)) \quad (3)$$

Em seguida, otimizam-se os parâmetros que minimizam (3) (Pedregosa et al., 2011):

$$\theta^* = \operatorname{argmin}_{\theta} G(Q_m, \theta) \quad (4)$$

Desta forma, divide-se novamente os subconjuntos de (1) e (2) até atingir a máxima profundidade permitida (Pedregosa et al., 2011).

$$N_m < \min_{amostras} \quad (5)$$

As medidas de impureza mais comuns são o Índice Gini e Entropia. Para o Índice Gini (Pedregosa et al., 2011):

$$H(Q_m) = \sum_k p_{mk}(1 - p_{mk}) \quad (6)$$

Para a Entropia (Pedregosa et al., 2011):

$$H(Q_m) = - \sum_k p_{mk} \log(p_{mk}) \quad (7)$$

Onde p_{mk} é a proporção de observações da classe k no nó m (Pedregosa et al., 2011):

$$p_{mk} = \frac{1}{N_m} \sum_{y \in Q_m} I(y = k) \quad (8)$$

As principais vantagens das ADs são a interpretabilidade das regras de decisão e a possibilidade de visualização, ou seja, é baseada em um modelo “caixa branca”; requer mínimo pré-processamento dos dados; e é capaz de classificar múltiplas classes. Como desvantagens, destaca-se a tendência à sobreajuste (*overfitting*), que pode ser minimizada através do procedimento de poda; é propensa a ser enviesada caso o banco de dados não seja balanceado; e podem ser instáveis à variações nos dados resultando em árvores diferentes da obtida antes da variação (Pedregosa et al., 2011).

3. MODELAGEM DO SISTEMA

As duas principais categorias de sistemas de distribuição são: i) em anel; e ii) radial (Kagan et al., 2010). Enquanto o primeiro é o mais sofisticado e oferece múltiplas opções de suprimento ao consumidor, o segundo é o mais simples e barato, onde a energia flui em sentido único da fonte para o consumidor (Kindermann, 1997). Em virtude da maioria dos sistemas de distribuição serem do tipo radial, adotou-se este modelo no presente trabalho para as simulações no ambiente de simulação MATLAB/SIMULINK (MATLAB, 2021).

O sistema simulado é representado na Figura 3, cujos elementos constituintes são:

- Alimentador: 13,8 kV;
- Impedância de interligação: $0,2053+j0,3860 \Omega$;
- Linha π -equivalente: 5 km e 10 km;
- Transformador: 13,8/0,38 kV e 150 kVA;
- Carga em MT: 100 kVA e $fp = 0,92$ (indutivo);
- Carga em BT: 50 kVA e $fp = 0,92$ (indutivo).
- Blocos geradores de falta.
- Chaves seccionadoras.

As oscilografias de corrente são medidas no barramento 1 da Figura 3, ou seja, à montante das cargas.

3.1 Parâmetros de simulação

De modo a abranger diversos cenários de faltas, realizou-se 960 simulações, cada uma com parâmetros distintos, conforme a Tabela 1. Estes parâmetros são R_{on1} e R_{on2} , que representam a resistência de falta no estágio 1 e 2, respectivamente, além da resistência à terra R_g e o ângulo de inserção de falta.

Tabela 1. Parâmetros de simulação.

Parâmetro	Valor
R_{on1} (Ω)	0,01; 0,1; 0,2; 5; 10; 50; 100 e 150
R_{on2} (Ω)	0,01; 0,1; 0,2; 5; 10; 50; 100 e 150
R_g (Ω)	0,01
Ângulo de falta ($^\circ$)	0; 45; 90

3.2 Modelagem do alimentador

O alimentador considerado é do tipo trifásico equilibrado, com tensões entre linhas de 13,8 kV. As fases são conectadas em estrela (Y) com o neutro aterrado.

Ademais, potência de curto-circuito trifásico (P_{sc}) é definido como 100 MVA na tensão base (V_{base}) de 13,8 kV. Portanto, a indutância interna do alimentador (L), é dada por (Glover et al., 2012):

$$L = \frac{V_{base}^2}{P_{sc}} \times \frac{1}{2\pi f} \quad (9)$$

De modo análogo, a resistência interna (Ω) é dado por (Glover et al., 2012):

$$R = \frac{2\pi f L}{X/R} \quad (10)$$

3.3 Modelagem da linha

A seção de linha π -equivalente trifásica estabelece um modelo balanceado a parâmetros concentrados. Em oposição à abordagem de parâmetros distribuídos, este modelo de linha concentra seus parâmetros em uma única seção segundo a Figura 4.

A linha é composta por condutores de Alumínio Nu CAA (ACSR), 4/0 AWG (INDUSCABOS, 2022). Desta forma, os valores de resistências próprias (R_s) e mútuas (R_m), indutâncias próprias (L_s) e mútuas (L_m) dos indutores acoplados, além das capacitâncias de fase (C_p) e capacitâncias de terra (C_g), são definidas a partir dos parâmetros (RLC) de sequência positiva e zero que respeitam os acoplamentos indutivos e capacitivos entre os condutores, além dos parâmetros de terra (Glover et al., 2012). Portanto, para as seções de linha de 5 km e 10 km, adotaram-se os seguintes valores:

$$[R] = \begin{bmatrix} 0,4676 & 0,0982 & 0,0969 \\ 0,0982 & 0,0982 & 0,0954 \\ 0,0969 & 0,0954 & 0,4621 \end{bmatrix} (\Omega/km) \quad (11)$$

$$[L] = \begin{bmatrix} 0,0019 & 0,0007 & 0,0008 \\ 0,0007 & 0,0020 & 0,0006 \\ 0,0008 & 0,0006 & 0,0020 \end{bmatrix} (H/km) \quad (12)$$

$$[C] = \begin{bmatrix} 9,3933 & -1,7829 & -2,7864 \\ -1,7829 & 8,5371 & -1,0859 \\ -2,7864 & -1,0859 & 8,9411 \end{bmatrix} (\times 10^{-9})(F/km) \quad (13)$$

3.4 Modelagem do transformador

O transformador trifásico utilizado consiste de três enrolamentos ligados ao circuito primário e de três enrolamentos conectados ao circuito secundário. A configuração das conexões é Delta-Estrela ($\Delta-Y$) com neutro solidamente aterrado (ESPECIFICAÇÃO TÉCNICA N^o 942), sendo a ligação frequentemente empregada em transformadores de distribuição (Kagan et al., 2010).

Desta forma, os parâmetros empregados consideram especificações baseadas nos transformadores presentes nos alimentadores de teste IEEE (IEEE, 2022) (Tabela 2).

Onde a base da impedância, resistência e reatância são obtidas por (Glover et al., 2012):

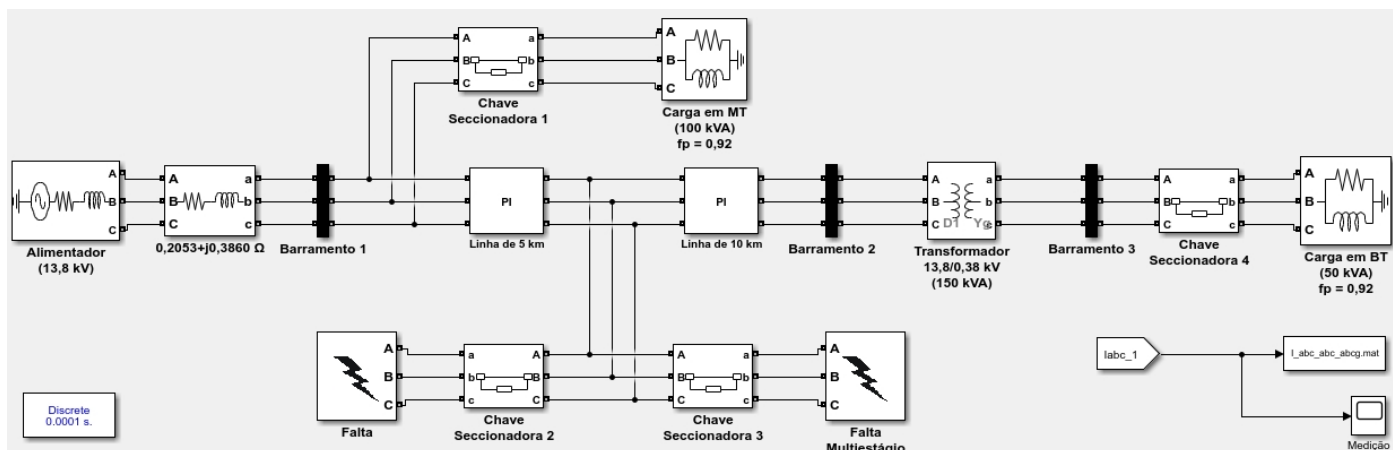


Figura 3. Alimentador Primário de Distribuição de Energia Simulado.

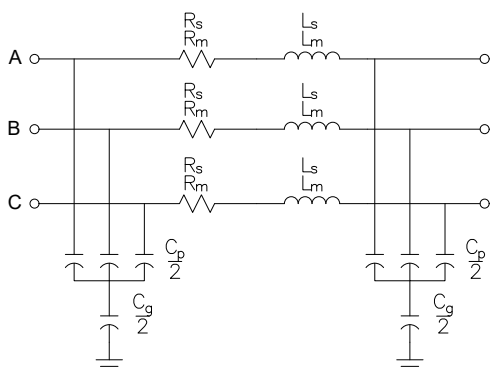


Figura 4. Modelagem da linha.

Tabela 2. Parâmetros do transformador.

Parâmetro	Descrição	Valor
P_n (kVA)	Potência nominal	150
f_n (Hz)	Frequência nominal	60
R_m (pu)	Resistência de magnetização	500
L_m (pu)	Indutância de magnetização	500
V_1 (kV)	Tensão nominal AT	13,8
V_2 (kV)	Tensão nominal BT	0,38
R (pu)	Resistência	0,011
L (pu)	Indutância	0,02

$$Z_{base} = R_{base} = X_{base} = \frac{V_n^2}{P_n} \quad (14)$$

E para a indutância (Glover et al., 2012):

$$L_{base} = \frac{X_{base}}{2\pi f_n} \quad (15)$$

3.5 Modelagem das cargas

O fornecimento de energia nos níveis de baixa (BT) e média tensão (MT) seguem normas próprias da concessionária da região em que estão instaladas as cargas simuladas. Tomando como base a norma ESPECIFICAÇÃO TÉCNICA N° 942, as cargas superiores a 75kW devem ser atendidas em tensão primária de distribuição de até 34,5 kV. Além disso, é definido no Brasil que o fator de

potência das cargas deve obedecer ao limite mínimo de 0,92 (ANEEL).

Portanto, adotaram-se os valores de 50 kW e 100 kW para as cargas de BT e MT, respectivamente, com o fator de potência de 0,92 para ambas as cargas.

3.6 Modelagem da falha

O elemento de falha implementa o chaveamento de curto-circuito, com duração programada, conforme o diagrama da Figura 5. Os valores de R_{on} e R_g para os estágios 1 e 2 são obtidos da Tabela 1. Para os casos em que não há simulação de falha envolvendo terra, define-se o valor de R_g como $1 \times 10^6 \Omega$.

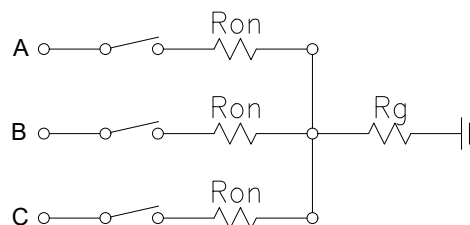


Figura 5. Modelagem da falha.

4. METODOLOGIA

Assim sendo, a metodologia deste trabalho (Figura 6) se resume em quatro etapas, sendo elas:

- i) Aquisição de dados;
- ii) Processamento;
- iii) Algoritmo de classificação;
- iv) Diagnóstico.

4.1 Aquisição de dados

As oscilografias de corrente obtidas de cada fase do sistema possuem duração de 0,4 s com frequência de amostragem

de 10 kHz. Desta forma, adotando-se a frequência de 60 Hz para o sistema, cada oscilografia obtida possui 4000 amostras, resultando em 166 amostras/ciclo. Além disso, cada oscilografia é constituída por um trecho de operação normal (duração de 0,1 s), seguido de um trecho relativo à falta no estágio 1 (duração de 0,15 s), seguido do trecho da falta no estágio 2 (duração de 0,15 s).

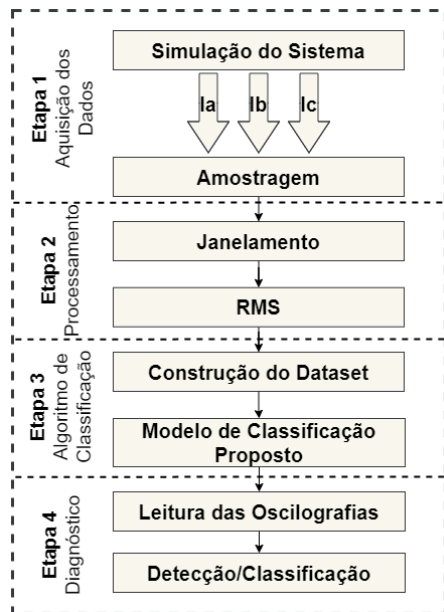


Figura 6. Metodologia.

Da Tabela 1, têm-se oito valores possíveis para R_{on1} e R_{on2} e três para o ângulo de inserção de falta, totalizando 192 configurações possíveis para cada categoria de falta multiestágio.

4.2 Processamento

O processamento das oscilografias é realizado via aplicação de RMS, que se refere a uma medida de carácter estatístico da magnitude de uma variável temporal (Glover et al., 2012).

Desta forma, o valor RMS pode ser obtido tanto para variáveis contínuas quanto discretas. Para o caso discreto (Glover et al., 2012):

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} = \sqrt{\frac{x_1^2 + x_2^2 + \dots + x_N^2}{N}} \quad (16)$$

onde N é um conjunto de elementos (x_1, x_2, \dots, x_N) .

Os valores RMS são obtidos a partir de janelas com comprimento de 2 ciclos cada, visto que as janelas de 1 ou 3 ciclos não forneceram resultados superiores. Desta forma, extraiu-se uma janela RMS do trecho de operação normal, outra janela do estágio 1 (para falta simples) e outra no limiar dos estágios 1 e 2 (para falta multiestágio).

4.3 Algoritmo de classificação e diagnóstico

Banco de dados: Cada janela RMS é rotulada com sua respectiva classe e armazenada no banco de dados. Em

resumo, cada uma das 9 classes (Tabela 3) possui 192 amostras.

Tabela 3. Designação das faltas.

Tipo	Falta	Designação
Simple	AB	a
Simple	ABC	b
Multiestágio	ABC-ABCG	c
Multiestágio	AB-ABC	d
Multiestágio	AB-ABG	e
Simple	AG	f
Multiestágio	AG-ABCG	g
Multiestágio	AG-ABG	h
Simple	NORMAL	i

Modelo de classificação proposto: Para a criação do algoritmo, dividiu-se o banco de dados em 70% para treinamento, 10% para validação e 20% para teste. Na etapa de validação realizou-se uma busca exaustiva para encontrar os melhores hiperparâmetros (Tabela 4).

Tabela 4. Designação dos hiperparâmetros.

Hiperparâmetro	Descrição
Hiperparâmetro 1	Profundidade
Hiperparâmetro 2	Número máximo de atributos para a divisão de um nó
Hiperparâmetro 3	Número mínimo de amostras para a divisão de um nó

Na Figura 7 são apresentadas as etapas adotadas para a criação do modelo de classificação proposto.

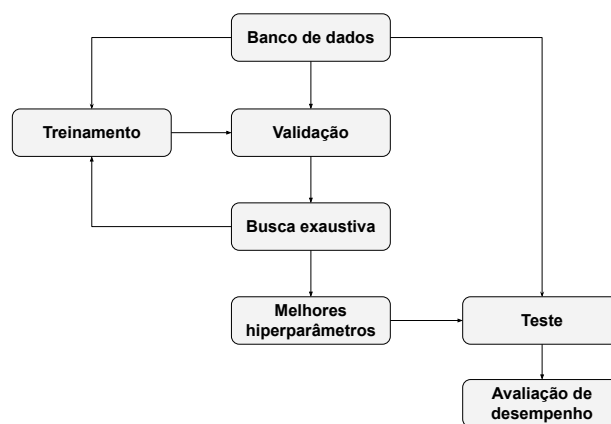


Figura 7. Modelo de Classificação Proposto.

As métricas de avaliação de desempenho utilizadas foram a acurácia (ACC) e a matriz de confusão. A acurácia é calculada como segue (Witten and Frank, 2005):

$$ACC = \frac{VP + VN}{VP + VN + FP + FN} \quad (17)$$

onde VP, VN, FP, FN, são, respectivamente, as taxas de verdadeiro positivo, verdadeiro negativo, falso positivo e falso negativo.

Por outro lado, na matriz de confusão, cada linha da matriz representa instâncias de uma classe prevista enquanto cada coluna representa instâncias da classe (Witten and Frank, 2005).

5. RESULTADOS

Neste capítulo serão apresentados os resultados encontrados para a simulação e para o classificador. Na primeira parte é mostrado os experimentos realizados nas simulações. Posteriormente é analisada a criação do algoritmo e os seus hiperparâmetros.

5.1 Resultados da Simulação

As oscilografias de correntes obtidas, para o caso da falta multiestágio AB-ABCG está representada na Figura 8, onde também estão indicadas as janelas de extração do valor RMS. Na primeira janela é obtida o valor associado à operação normal, na janela seguinte o valor associado à falta simples AB e, finalmente, na última janela o valor associado à falta multiestágio AB-ABCG. Para as demais faltas da Tabela 3, resultados análogos são obtidos.

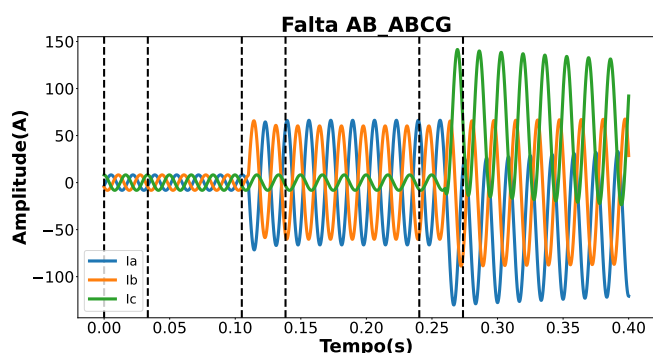


Figura 8. Falta Multiestáfio AB-ABCG.

As amostras adquiridas de cada janela são processadas (extração do valor do RMS) e posteriormente rotuladas com as respectivas faltas. Todos os elementos rotulados são unidos para formar o banco de dados.

5.2 Resultados do modelo de classificação proposto

Na etapa de validação, definiu-se os hiperparâmetros que possuem maior influência na acurácia final do modelo.

As Figuras 9-11 representam a relação entre a acurácia e o hiperparâmetro estudado, de acordo com a Tabela 4.

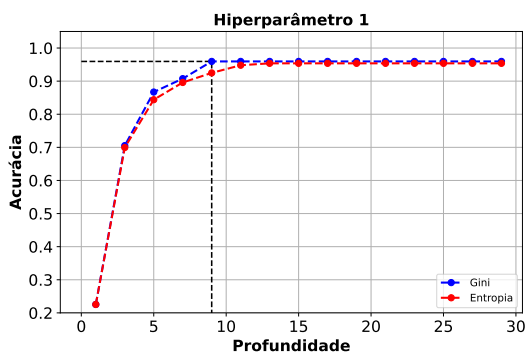


Figura 9. Influência da profundidade na acurácia.

Analisando a Figura 9 nota-se a relação entre a acurácia e a profundidade da AD. Essas duas variáveis são diretamente proporcionais até chegar a um ponto de saturação. A

medida de impureza com melhor performance foi o índice Gini, com uma acurácia de 0,9595 na validação, chegando a esse resultado com 9 níveis de profundidade.

Por outro lado, na Figura 10, é possível analisar a influência do número de atributos para a divisão de um nó. Portanto, o ponto ótimo de acurácia foi obtido com 3 atributos.

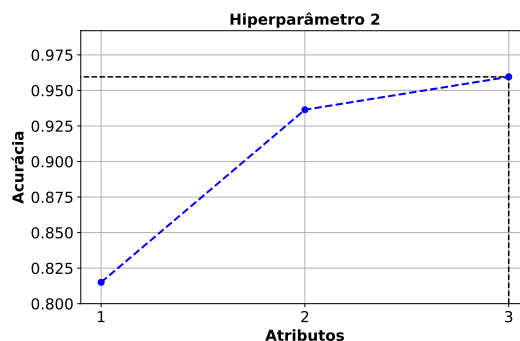


Figura 10. Influência do número de atributos na divisão de um nó.

Seguidamente, para a análise do número mínimo de amostras na divisão de cada nó (Figura 11) o melhor resultado encontrado foi obtido com o valor 3.

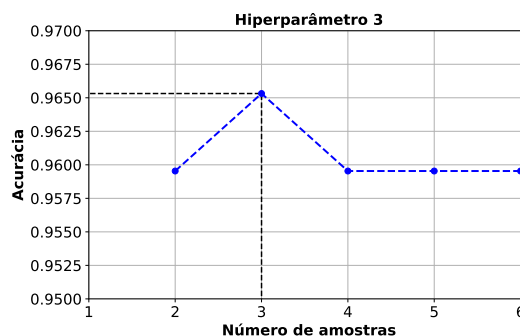


Figura 11. Influência do número de divisões em um nó.

Em suma, a AD foi criada utilizando como medida de impureza o índice Gini, profundidade de 9 níveis, cada nó utilizando no mínimo 3 amostras para a sua divisão e são considerados 3 atributos para tomar a decisão de divisão em cada nó. Portanto, o algoritmo logrou uma acurácia de 97,10% na classificação do banco de dados de teste. Na Tabela 5 é apresentada a matriz de confusão obtida do modelo, cujas designações de falhas obedecem à Tabela 3.

Verificando a matriz de confusão (Tabela 5) percebe-se que a AD conseguiu classificar corretamente todos os elementos nas classes AB, ABC, AG, que são faltas do tipo simples; a classe Normal, que se refere à operação normal do sistema, que por possuir um comportamento em regime, sempre apresentará valores muito próximos entre si; e por fim a classe multiestáfio AG-ABCG.

Por outro lado, os maiores erros ocorreram nas classes ABC-ABCG e AB-ABC com 3 e 4 erros respectivamente. Uma explicação para este comportamento é a semelhança entre faltas ABCG e ABC, que diferem apenas do elemento terra. Nestes casos, a presença ou não do terra reflete em pouca influência na característica das oscilografias de

correntes, motivo pelo qual alguns autores preferem juntar as duas classes, como é o caso de Bernardes et al. (2021).

Tabela 5. Matriz de confusão da AD.

Classe	a	b	c	d	e	f	g	h	i
a	34	0	0	0	0	0	0	0	0
b	0	44	0	0	0	0	0	0	0
c	0	3	33	0	0	0	0	0	0
d	0	2	0	36	0	0	2	0	0
e	0	0	0	0	31	0	0	1	0
f	0	0	0	0	0	42	0	0	0
g	0	0	0	0	0	0	42	0	0
h	0	0	0	0	2	0	0	36	0
i	0	0	0	0	0	0	0	0	38

6. CONCLUSÃO

Neste trabalho apresentou-se uma metodologia para a classificação de faltas multiestágio em um alimentador primário de distribuição de energia. As oscilografias de correntes retiradas do sistema, foram utilizadas para classificação de 9 categorias de faltas.

Também foi apresentado uma abordagem para o processamento e classificação destas faltas. Para tal, empregou-se o cálculo do RMS e o algoritmo Árvore de decisão. Os resultados obtidos são satisfatórios, alcançando valor de acurácia de 97,10%, que evidencia a aplicabilidade desta metodologia na detecção de faltas desta natureza. Outra vantagem da abordagem proposta é a facilidade de implementação em sistemas embarcados com dispositivos lógicos programáveis FPGAs (Matriz de Portas Programáveis) e VHDLs (Linguagem de Descrição de *Hardware*), devido ao processamento RMS e ao algoritmo AD que oferecem baixo custo computacional, característica necessária para a implantação em medidores instalados estrategicamente ao longo da rede de distribuição.

Como trabalhos futuros espera-se empregar novos tipos de modelos de aprendizado de máquina, associados a outras técnicas de processamento dos sinais, com o propósito de alcançar maiores acurácias e análises mais detalhadas sobre a incidência das faltas, além de testá-las em sistemas de distribuição mais complexos, investigando a influência da inserção de fontes renováveis no contexto da Geração Distribuída.

AGRADECIMENTOS

Este trabalho foi realizado com o apoio da Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP) através da concessão de bolsa de estudo de Mestrado.

REFERÊNCIAS

ANEEL (2021). Resolução Normativa ANEEL Nº 1.000, de 7 de dezembro de 2021. Resolução normativa, Ministério de Minas e Energia/Agência Nacional de Energia Elétrica, Brasília, DF.

Bernardes, H.R.S., Tonelli-Neto, M.S., and Minussi, C.R. (2021). Fault classification in power distribution systems using multiresolution analysis and a fuzzy-artmap neural network. *IEEE Latin America Transactions*, 19, 1824–1831.

Breiman, L., Friedman, J., Stone, C.J., and Olshen, R. (1984). *Classification and Regression Trees*. Taylor Francis, Belmont.

ESPECIFICAÇÃO TÉCNICA Nº 942 (2021). Fornecimento de Energia Elétrica em Tensão Primária de Distribuição até 34,5 kV. Norma técnica, Enel Distribuição Ceará, Fortaleza, CE.

Glover, J.D., Sarma, M.S., , and Overbye, T.J. (2012). *Power System Analysis and Design*. Cengage Learning, Stamford, 5 edition.

IEEE (2022). Ieee pes test feeder. URL <https://cmte.ieee.org/pes-testfeeders/resources/>. Acesso em 20/08/2022.

INDUSCABOS (2022). Cabos de alumínio nu caa (acsr). URL <https://www.induscabos.com.br/portfolio-item/fios-e-cabos-de-aluminio-nu-caa-acsr/>. Acesso em 20/08/2022.

Kagan, N., de Oliveira, C.C.B., and Robba, E.J. (2010). 2 edition.

Kindermann, G. (1997). *Curto-Circuito*. Sagra Luzzatto, Porto Alegre, 2 edition.

MATLAB (2021). *version 9.11.0.1837725 (R2021b)*. The MathWorks Inc., Natick, Massachusetts.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

Prévé, C. (2006). *Protection of electrical networks*. ISTE Ltd, London.

Rai, P., Londhe, N.D., and Raj, R. (2021). Fault classification in power system distribution network integrated with distributed generators using cnn. *Electric Power Systems Research*.

Salim, R.H., de Oliveira, K.R.C., Filomena, A.D., Resener, M., and Bretas, A.S. (2008). Hybrid fault diagnosis scheme implementation for power distribution systems automation. *IEEE Transactions On Power Delivery*, 23, 1846–1856.

Sharma, R., Mahela, O.P., and Agarwal, S. (2018). Detection of power system faults in distribution system using stockwell transform. *2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*.

Song, E.Y., FitzPatrick, G.J., Lee, K.B., and Griffor, E. (2022). A methodology for modeling interoperability of smart sensors in smart grids. *IEEE Transactions on Smart Grid*, 13(1), 555–563.

Styvaktakis, E. and Bollen, M.H.J. (2003). Signatures of voltage dips: Transformer saturation and multistage dips. *IEEE Transactions On Power Delivery*, 18, 265–270.

Tleis, A.N.D. (2008). *Power Systems Modelling and Fault Analysis: Theory and Practice*. Elsevier, Oxford.

Witten, I.H. and Frank, E. (2005). *Data mining : practical machine learning tools and techniques*, volume 2. Elsevier.