

Localização de Falhas em Sistemas de Distribuição Radiais Baseado em Árvores de Decisão e em uma Alocação Otimizada de Medidores de Qualidade da Energia Elétrica

André Luís da Silva Pessoa

Mário Oleskovicz

Departamento de Engenharia Elétrica e de Computação

Universidade de São Paulo, USP

São Carlos, Brasil

alsp@usp.br, olesk@sc.usp.br

Resumo—Para não se alcançar uma qualidade da energia elétrica satisfatória está a descontinuidade do fornecimento de energia em Sistemas de Distribuição (SDs), normalmente causada por situações de curtos-circuitos. Assim, para caracterizar tais situações foi compilado um banco de dados através de simulações no sistema de distribuição do IEEE de 34 barras utilizando o software ATP (*Alternative Transients Program*). Este trabalho propõe o emprego da árvore de decisão J48 para identificar em qual zona do SD de 34 barras as faltas monofásicas ocorreram. Para o uso da árvore de decisão J48 foi empregado o software WEKA (*Waikato Environment for Knowledge Analysis*).

Palavras-chaves—Árvores de Decisão, Localização de Falhas, Medidores de Qualidade da Energia Elétrica, Sistemas de Distribuição.

I. INTRODUÇÃO

A Qualidade da Energia Elétrica (QEE) é uma área da engenharia elétrica na qual a sua abrangência é difícil de ser determinada, pois envolve desde a geração, transmissão e distribuição da energia elétrica, até todos os consumidores finais do produto “energia elétrica”. Assim, devido à QEE ser um tema que engloba a todos os usuários finais do sistema elétrico, as concessionárias de energia devem levar em consideração a relação entre os equipamentos dos usuários do sistema com os níveis de tensão que são disponibilizados. Os consumidores finais da energia, principalmente as indústrias, devem também respeitar os demais usuários conectados ao SD (Sistema de Distribuição), controlando as prováveis fontes dos distúrbios associados à QEE e mitigando os problemas que podem surgir ao ligar suas cargas lineares ou não lineares à rede [1]. Conforme apresentado em [1] a QEE é um aspecto importante para os sistemas elétricos de potência,

tendo impactos importantes na sua segurança, eficiência e reabilitação do sistema. Dentre as diversas definições, o termo QEE geralmente vem sendo utilizado para expressar a qualidade no fornecimento da tensão respeitando padrões pré-estabelecidos. Neste sentido, a QEE pode ser definida como a medida, análise e melhoria da tensão em um barramento para manter a sua forma próxima da onda senoidal, com amplitude e frequência fixas. Tanto as concessionárias quanto os usuários finais concordam que mais de 60% dos problemas de QEE são ocasionados por eventos imprevisíveis, sendo alguns desses eventos: os curtos-circuitos; as descargas atmosféricas; a ferroressonância; e as correntes geomagneticamente induzidas, dentre outros.

Dentre as várias causas das perturbações de QEE nos sistemas de energia elétrica, existem os curtos-circuitos. Neste cenário, para amenizar o dano que as cargas dos consumidores venham a sofrer, é essencial que a localização destes curtos-circuitos seja realizada de forma precisa e rápida [2]. Vale comentar que, a depender das filosofias desenvolvidas e aplicadas por muitas distribuidoras de energia, a localização precisa das situações de curtos-circuitos nos SDs, é realizada e baseada puramente em informações passadas pelos usuários do sistema e de testes para a delimitação da área de ocorrência do distúrbio [3].

Nas pesquisas iniciais direcionadas à localização de faltas em sistemas de transmissão, as técnicas desenvolvidas eram fundamentadas no cálculo da impedância aparente, definida a partir dos fasores de tensão de corrente da porção da linha em análise. Estas técnicas foram desenvolvidas e melhor aprimoradas a partir da introdução e difusão do uso de equipamentos micro processados [4]. Posteriormente, técnicas baseadas em inteligência artificial foram desenvolvidas visando localizar as faltas de forma tão eficiente, ou mais, do que os métodos tradicionais.

Com foco específico em SD, em [5] são utilizadas técnicas de inteligência artificial para a localização das faltas. Na pes-

Os autores gostariam de agradecer ao Laboratório de Sistemas de Energia Elétrica (LSEE), da Escola de Engenharia de São Carlos (EESC- USP) pela infraestrutura proporcionada, assim como à CAPES pelo suporte financeiro disponibilizado para a realização desta pesquisa.

quisa, os sinais decorrentes das situações de curtos-circuitos foram processados utilizando-se a Transformada Wavelet (TW) e então aplicados às Redes Neurais Artificiais (RNAs) e sistemas de Lógica Fuzzy (LF) para localizar as faltas. Em [6] é apresentada uma abordagem utilizando RNAs e a Transformada Wavelet Discreta (TWD) considerando o quarto nível de decomposição, tendo como wavelet mãe a Daubechies 4. [7] utilizaram de sistemas multiagentes para determinar em qual área do sistema as faltas ocorreram. [8] desenvolveram uma metodologia baseada em árvores de decisão C4.5, ID3, Gini e Chi, e também utilizaram de apenas um único ponto de medição na subestação do sistema para os testes. Quanto a uma abordagem baseada em uma alocação prévia de medidores de qualidade da energia elétrica, [9] desenvolveram um método analítico para determinar o local de ocorrência da falta.

Partindo da premissa de que as faltas precisam ser localizadas de forma rápida e precisa, este trabalho propõe o uso de árvores de decisão J48 [10] [11] para localizar a área e a zona de ocorrência das faltas, com base em dados extraídos dos sinais de tensão e corrente trifásicos registrados por medidores de QEE otimamente alocados no sistema. Outro ponto que esta pesquisa busca resolver é relacionado com o problema das múltiplas estimações da localização da falta a partir de um ponto específico, usualmente um ponto de medição no sistema elétrico de interesse [12].

Além desta introdução, este artigo está dividido em outras 4 seções. A seção 2 apresenta um breve descritivo sobre as árvores de decisão J48. A seção 3 apresenta a metodologia adotada nesta pesquisa. A seção 4 apresenta os resultados obtidos, e a seção 5 contém as conclusões pertinentes à pesquisa realizada.

II. ÁRVORE DE DECISÃO J48

A mineração de dados possibilita uma extração de informações implícitas, desconhecidas e úteis a partir de dados disponíveis. Dentre as diversas possibilidades de uso para os algoritmos de mineração de dados está a de classificar grupos de dados. Seguindo esta linha de classificar dados em classes, há o algoritmo conhecido como árvore de decisão C4.5 (conhecido como J48 no WEKA). Este algoritmo irá correlacionar variáveis independentes de entrada com um parâmetro de saída, através de regras SE-ENTÃO. A árvore de decisão vem demonstrando ser de grande utilidade para a descoberta de informações úteis a partir de um banco de dados. Outras vantagens deste algoritmo são: ser de fácil compreensão para o usuário final e estar disponível em diversas plataformas [10].

A árvore de decisão está dividida em 3 partes principais: nó principal (raiz) de onde partem todas as ramificações da árvore de decisão, os nós finais (folhas) que representam as classes, e os nós intermediários que representam as condições. Quanto ao algoritmo, vale salientar que é dividido em duas etapas principais: a construção da árvore e a poda. Tanto a construção da árvore de decisão, quanto a poda utilizam do cálculo da entropia para definir quais variáveis devem ser utilizadas, como estas variáveis estarão se dividindo (divisões

binárias) e quais nós poderão ser descartados de modo a reduzir a redundância da árvore de decisão [10] [11]. Desta forma, a partir de um banco de dados, o algoritmo da árvore de decisão correlaciona as variáveis de entrada com as classes, criando assim as regras SE-ENTÃO que serão utilizadas para por exemplo, no caso desta pesquisa, identificar em qual zona do sistema a falta ocorreu. Um ponto a destacar é a questão das árvores de decisão já realizarem uma prévia seleção de atributos, na qual ela identifica os parâmetros de entrada que mais estão correlacionados com a classificação que se deseja realizar, tornando-se em um verdadeiro especialista no problema em questão.

III. METODOLOGIA

Nesta pesquisa foi considerada a modelagem do sistema de distribuição do IEEE de 34 barras, onde os três pontos de medição são medidores de QEE otimamente alocados. Esta alocação otimizada de medidores de QEE resultou de um estudo prévio [13], e não será foco deste artigo. Após isto, o sistema foi dividido em três áreas. Na sequência, onze zonas específicas foram definidas, para a localização de possíveis curtos-circuitos.

Para o momento, apenas curtos-circuitos monofásicos foram simulados e considerados nesta pesquisa. Dependendo do tipo da falta, a área e a zona onde cada falta ocorreu foi determinada utilizando a árvore de decisão J48. Cada uma das etapas principais do algoritmo desenvolvido será apresentada e detalhada neste artigo.

Todo o processamento dos dados foi realizado em um PC core i7 @3.6GHz e 16GB RAM.

A. Sistema modelado e simulações

Para o emprego das árvores de decisão J48 para a localização das faltas, partiu-se de uma prévia modelagem do SD do IEEE de 34 barras [14]. Nas simulações considerou-se uma interface entre o ATP [15] e o MATLAB [16] para melhor tratar e processar as informações necessárias.

As informações utilizadas para a localização das faltas são registradas a partir dos três medidores de QEE alocados no SD de 34 barras. Para a alocação destes medidores de QEE partiu-se de uma alocação otimizada desenvolvida por [13]. Das propostas para a alocação dos medidores presentes em [13], foi adotado para esta pesquisa uma alocação contendo três medidores que possibilitam ter sensibilidade para observar no SD de 34 barras os afundamentos com tensão remanescente de no mínimo 0.9 p.u. Com base no posicionamento destes três medidores de QEE, o SD de 34 barras foi dividido em três regiões, conforme ilustrado pela Figura 1.

Em cada simulação dos curtos-circuitos aplicados, considerou-se uma taxa amostral de 256 amostras por ciclo. Com foco na estimação da distância da falta a partir dos três medidores alocados, foi considerado 1 ciclo de pós-falta dos sinais de tensão e corrente trifásicos. Quanto à posição das faltas, estas foram aplicadas nos ramos do SD com comprimento inferior a 200m a 1%, 25%, 50%, 75% e 99% do respectivo comprimento. Para os trechos maiores do que

800m, as situações de curtos-circuitos foram aplicadas a cada 200 metros. Vale comentar que este limiar de 200 metros foi adotado, pois, busca-se, para uma futura estimação da distância de falta, obter um erro médio de mais ou menos 200m.

Para esta pesquisa foi considerado que todo o processamento das informações advindas dos três medidores de QEE alocados será concentrado no sistema de localização de faltas. Além da variação da posição das faltas também foram considerados os valores das impedâncias de falta de 0,0001 Ω , 10 Ω , 20 Ω , 30 Ω , 40 Ω e 50 Ω . Para gerar o banco de dados de teste, foram aplicadas faltas a cada 50 m e consideradas impedâncias de falta de 5 Ω , 15 Ω , 25 Ω , 35 Ω e 45 Ω . Para esta pesquisa, é considerado que todo o processamento da informação dos medidores de QEE alocados será realizado no sistema de localização de faltas.

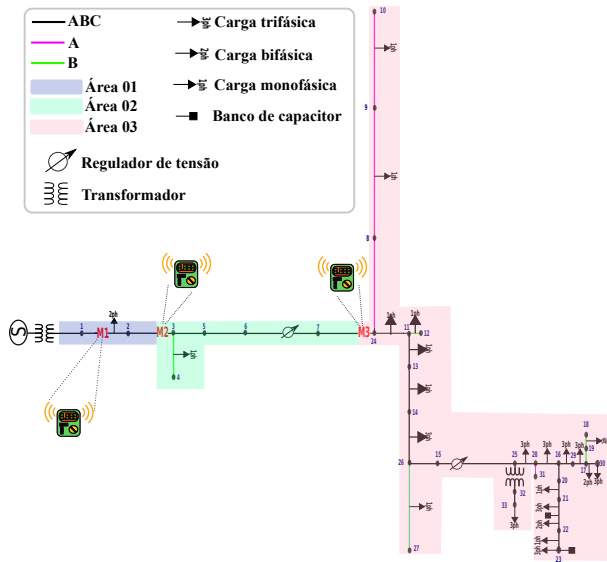


Fig. 1. SD do IEEE de 34 barras com três medidores de QEE posicionados.

Para cada um dos três medidores alocados foi delimitada uma área específica. Desta forma, dependendo da área em que a falta tiver ocorrido será considerado o medidor desta área em específico para a localização precisa do curto-circuito. Desta forma, após a detecção da ocorrência da falta e da sua localização em relação ao tipo do curto-circuito, será estimada a zona onde a falta ocorreu, com base nos dados registrados dos sinais de tensão e corrente trifásicas do medidor alocado na área, a qual a zona pertence.

Assim, as onze zonas foram caracterizadas seguindo o critério abaixo:

- em cada uma das três áreas delimitadas pelos três medidores de QEE, foi atribuída uma zona delimitando o alimentador principal; e
- cada ramal lateral foi considerado como sendo uma zona.

A Fig. 2 apresenta o SD de 34 barras dividido em onze zonas. A Tabela I apresenta o tipo de falta monofásica e as zonas aonde as mesmas podem ser simuladas. Por exemplo, a

zona 2 é caracterizada por um circuito monofásico envolvendo a fase B. Logo, somente um curto-circuito do tipo B-terra poderá ser simulado nesta zona.

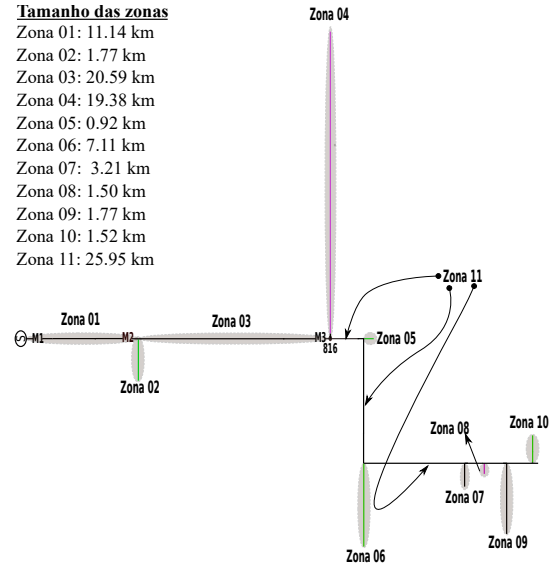


Fig. 2. SD do IEEE de 34 barras dividido em 11 zonas.

Tabela I. ZONAS RELACIONADAS ÀS FALTAS FASE-TERRA QUE OCORRERAM NAS ÁREAS 01, 02 E 03.

Tipo da falta	Zonas Relacionadas
A-terra	01 , 03, 04, 07, 08, 09 e 11
B-terra	01, 02, 03, 05, 06, 07, 09, 10 e 11
C-terra	01, 03, 07, 09 e 11

B. Parâmetros selecionados

Para o contexto desta pesquisa, foram calculados diversos parâmetros a partir do ciclo de amostras dos sinais de tensão e de corrente trifásicos extraídos após a detecção da falta. Estes parâmetros foram escolhidos a partir da revisão bibliográfica realizada.

Ao todo foram calculados para cada situação de falta 114 parâmetros, são eles:

- RMS_i para $i \in \{IA, IB, IC, VA, VB, VC\}$ indica o valor RMS dos sinais de tensão e corrente trifásicos;
- Amp_i e Fas_i para $i \in \{IA, IB, IC, VA, VB, VC\}$ indicam, respectivamente, os valores das amplitudes e fases dos sinais de tensão e corrente trifásicos; e
- $\{En01_i, En02_i, En03_i, \dots, En16_i\}$, no qual $i \in \{IA, IB, IC, VA, VB, VC\}$, indicam a energia dos coeficientes do 4 nível de decomposição da *Wavelet Packet Transform*, com wavelet mãe a Daubechies com suporte 4, dos sinais de tensão e corrente trifásicos.

Após apresentar o banco de dados de treinamento para as árvores de decisão correlacionarem estes parâmetros calculados com as zonas de ocorrência das faltas na área 03, a árvore de decisão referente a cada tipo de falta utilizou (devido à seleção de atributos inerente ao próprio algoritmo) os parâmetros conforme a Tabela II.

Tabela II. PARÂMETROS UTILIZADOS PELAS ÁRVORES DE DECISÃO PARA LOCALIZAR AS FALTAS MONOFÁSICAS QUE OCORREREM NA ÁREA 03.

Tipo da falta	Parâmetros
A-terra	Amp_{IB} , Fas_{VA} , RMS_{IC} , $En11_{IA}$, $En08_{IC}$, $En09_{IB}$, $En14_{IC}$, $En02_{VC}$, $En02_{IA}$, $En02_{VB}$, $En04_{IC}$, $En16_{VB}$, $En06_{VC}$ e $En11_{IB}$
B-terra	RMS_{IA} , Fas_{IB} , $En11_{IA}$, $En04_{VA}$, $En07_{IB}$, $En13_{IA}$, $En05_{IC}$, $En04_{IA}$, $En14_{VA}$, $En02_{VA}$, $En13_{VC}$, $En13_{IC}$, $En16_{IC}$, $En13_{IB}$, $En14_{IC}$, $En16_{IA}$, $En16_{IB}$, $En06_{VC}$, $En14_{IA}$, $En09_{IC}$, $En09_{IA}$, $En11_{IA}$, $En15_{IA}$, $En08_{VA}$ e $En10_{IB}$
C-terra	RMS_{IA} , Amp_{VA} , Fas_{IB} , $En08_{IA}$, $En14_{IA}$, $En04_{VA}$, $En07_{IA}$, $En13_{VC}$, $En05_{VA}$, $En05_{VC}$, $En03_{IB}$, $En13_{IB}$, $En03_{VA}$, $En01_{IA}$, $En05_{VC}$, $En05_{IB}$, $En03_{VB}$, $En04_{VA}$

C. Identificação da zona de ocorrência da falta

Para o uso da árvore de decisão J48 foi empregado o software WEKA [18]. Para o desenvolvimento desta pesquisa, foram construídas árvores de decisão para as áreas 01, 02 e 03 para determinar se as faltas monofásicas ocorreram em alguma das 11 zonas previamente descritas. Para ilustrar, a Tabela III apresenta as regras para determinar se as faltas A-terra da área 03 ocorreram na zona 09. Então, pelas regras ilustradas para a zona 9, verifica-se que o algoritmo, para indicar a zona de ocorrência da falta, pode ser compilado a partir de um conjunto de condições SE-ENTÃO.

Tabela III. REGRAS OBTIDAS APÓS A APLICAÇÃO DA ÁRVORE DE DECISÃO J48 PARA IDENTIFICAR SE AS FALTAS FASE A-TERRA APLICADAS À ÁREA 03 OCORRERAM NA ZONA 9.

Condição	Regra	Zona
1	$Amp_{IB} > 61,49$ e $En02_{VC} \leq 0,050038$ e $En02_{IA} \leq 0,048028$ e $En02_{VB} \leq 0,009954$ e $En04_{IC} \leq 0,023392$ e $RMS_{IC} \leq 43,562$	09
2	$Amp_{IB} > 61,49$ e $En02_{VC} \leq 0,050038$ e $En02_{IA} \leq 0,048028$ e $En02_{VB} \leq 0,009954$ e $En04_{IC} > 0,023392$	09
3	$Amp_{IB} > 61,49$ e $En02_{VC} \leq 0,050038$ e $En02_{IA} \leq 0,048028$ e $En02_{VB} > 0,009954$ e $En16_{VB} > 0,004245$ e $En14_{IC} \leq 0,002428$ e $Fas_{VA} \leq 0,73513$	09
4	$Amp_{IB} > 61,49$ e $En02_{VC} \leq 0,050038$ e $En02_{IA} \leq 0,048028$ e $En02_{VB} > 0,009954$ e $En16_{VB} > 0,004245$ e $En14_{IC} > 0,002428$	09

IV. RESULTADOS

Nesta seção serão apresentados os resultados considerando a localização de faltas monofásicas no sistema de distribuição do IEEE de 34 barras.

Para as áreas 02 e 03, devido a presença de ramais laterais, é observada a possibilidade de múltiplas estimações de faltas

para os medidores alocados nestas respectivas áreas. Como na área 01 não há ramais laterais, consequentemente, não existe a possibilidade de haver o problema de múltiplas estimações das faltas a partir das medições advindas do medidor 01. Visto que o objetivo desta pesquisa é identificar a zona faltosa e buscar responder ao problema das múltiplas estimações das faltas, o foco deste artigo será na localização de faltas B-terra que ocorreram na área 02, e para as faltas A, B e C-terra que ocorreram na área 03.

Inicialmente, será dada ênfase à localização das faltas A-terra que ocorreram na área 03. Considerando a árvore de decisão J48, e usando dados de treinamento e teste de faltas fase A-terra aplicadas na área 03, o tempo para construir o modelo da árvore de decisão J48 foi de 0,15 s, e o tempo para avaliar o conjunto de teste foi de 0,18 s. Outro ponto que deve ser enfatizado é que dos 114 parâmetros apresentados ao algoritmo da árvore de decisão J48 para fornecer todas as regras para a localização de faltas fase A-terra, para a área 03, apenas 14 desses parâmetros foram utilizados pelo algoritmo.

A escolha inicial de apresentar resultados da área 3 é justificada pela extensão e complexidade (vários ramos laterais) que esta área apresenta. Em seguida, as outras situações de faltas monofásicas (B e C-terra) também serão analisadas para área 3. Apenas para ilustrar, o conjunto de dados de treinamento para a área 03, considerando faltas monofásicas (A, B e C-terra), é composto de 1.956 casos, 1.638 casos e 1.314 casos, respectivamente. O conjunto de dados de teste para a mesma condição e fases tem 5.155, 4.135 e 3.195 casos, respectivamente.

Na área 2 somente as faltas B-terra serão avaliadas. Uma vez que esta é a única situação de falta que pode resultar em múltiplas localizações na área 2, devido à presença da zona 2 (Fig. 3). O conjunto de dados de treinamento para a área 02, considerando apenas a fase B-terra, tem 690 casos. O conjunto de dados de teste para a mesma condição tem 2.688 casos.

Para demonstrar inicialmente o desempenho das árvores de decisão J48 para identificar faltas no SD de 34 barras, foi aplicada uma falta A-terra na zona 9 da Fig. 2. A oscilografia resultante é mostrada na Fig. 3. A partir desta situação de curto-circuito, foram obtidos os seguintes valores: 63,0041 para Amp_{IB} ; 0,0446 para E_{02VC} ; 0,0426 para E_{02IA} ; 0,0070 para E_{02VB} e 0,0268 para E_{04IC} . A partir dos valores desses parâmetros e das regras mostradas na Tabela III, pode-se concluir pela condição 2, que essa falta ocorreu na zona 09 da área 03. Para avaliação do desempenho do algoritmo foram utilizadas matrizes de confusão e curvas ROC (*Receiver Operating Characteristic*) [19].

Vale ressaltar que uma taxa de acerto de 99,42 % foi obtida para os 5.155 casos simulados de faltas A-terra, aplicando a árvore de decisão J48 na área 3. Assim, em apenas 30 curtos-circuitos aplicados, não foi possível localizar corretamente, a partir das regras geradas, a zona de ocorrência. Para todos os casos analisados o erro quadrático médio obtido foi de 0,0477.

As Fig. 4 a 6 mostram as matrizes de confusão sobre as faltas fase-terra que ocorreram na área 3. A matriz de confusão permite mostrar onde houve uma localização correta e permite

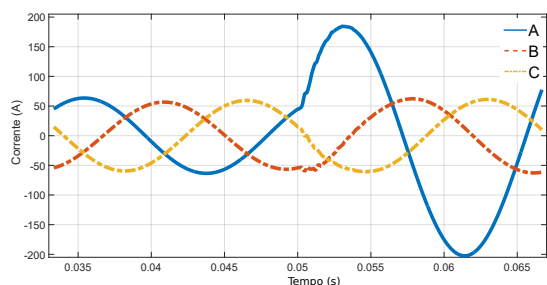


Fig. 3. Oscilografia obtida a partir de uma falta A-terra aplicada na zona 09.

Zona estimada \ Zona alvo	4	7	8	9	11	
4	1912 37,1%	0 0,0%	0 0,0%	0 0,0%	8 0,2%	99,6% 0,4%
7	0 0,0%	320 6,2%	0 0,0%	0 0,0%	0 0,0%	100% 0,0%
8	0 0,0%	0 0,0%	40 0,8%	0 0,0%	4 0,1%	90,9% 9,1%
9	0 0,0%	0 0,0%	0 0,0%	221 4,3%	6 0,1%	97,4% 2,6%
11	8 0,2%	0 0,0%	0 0,0%	4 0,1%	2632 51,1%	99,5% 0,5%
	99,6% 0,4%	100% 0,0%	100% 0,0%	98,2% 1,8%	99,3% 0,7%	99,4% 0,6%

Fig. 4. Matriz de confusão para localização de faltas fase A-terra que ocorreram na área 3

identificar os erros cometidos pelo algoritmo. A linha da matriz de confusão mostra a zona estimada (zona prevista) e a coluna é a zona alvo (zona verdadeira). Por exemplo, na Fig. 4, 2.632 faltas foram corretamente localizadas como ocorrendo na zona 11, 8 faltas ocorridas na zona 4 foram localizadas incorretamente na zona 11, e 4 faltas ocorridas na zona 9 foram localizadas incorretamente na zona 11. Por essas matrizes de confusão a taxa de acerto para as fases A, B e C-terra, respectivamente, foi de 99,4 %, 98,3 % e 99 %.

As Fig. 7, 8 e 9 são curvas ROC usadas para avaliar o desempenho do algoritmo J48 na localização das faltas aplicadas na área 3. Estes gráficos permitem observar o desempenho do algoritmo para cada zona e identificar em que zona há uma maior dificuldade em localizar as faltas.

Com base na *Area Under the Curve* (AUC), pela análise das curvas ROC é possível determinar a eficiência do localizador de faltas. Desta forma é considerada uma localização excepcional quando $AUC > 0,9$, uma localização excelente quando $0,8 \leq AUC \leq 0,9$, uma localização aceitável quando $0,7 \leq AUC \leq 0,8$, uma localização razoável quando $0,6 \leq AUC \leq 0,7$, uma localização pobre quando $0,5 \leq AUC \leq 0,6$ e uma localização errada quando $AUC = 0,5$ [19]. As Fig. 7, 8 e 9 apresentam o desempenho do algoritmo J48 para a localização de faltas monofásicas aplicadas na área 3. Foram

Zona estimada \ Zona alvo	5	6	7	9	10	11	
5	84 2,0%	0 0,0%	0 0,0%	0 0,0%	0 0,0%	0 0,0%	100% 0,0%
6	0 0,0%	688 16,6%	0 0,0%	0 0,0%	0 0,0%	0 0,0%	100% 0,0%
7	0 0,0%	0 0,0%	319 7,7%	0 0,0%	0 0,0%	0 0,0%	100% 0,0%
9	0 0,0%	0 0,0%	0 0,0%	218 5,3%	0 0,0%	6 0,1%	97,3% 2,7%
10	0 0,0%	0 0,0%	1 0,0%	0 0,0%	112 2,7%	1 0,0%	98,2% 1,8%
11	6 0,1%	22 0,5%	0 0,0%	7 0,2%	28 0,7%	2643 63,9%	97,7% 2,3%
	93,3% 6,7%	96,9% 3,1%	99,7% 0,3%	96,9% 3,1%	80,0% 20,0%	99,7% 0,3%	98,3% 1,7%

Fig. 5. Matriz de confusão para localização de faltas fase B-terra que ocorreram na área 3

Zona estimada \ Zona alvo	7	9	11	
7	320 10,0%	0 0,0%	0 0,0%	100% 0,0%
9	0 0,0%	204 6,4%	10 0,3%	95,3% 4,7%
11	0 0,0%	21 0,7%	2640 82,6%	99,2% 0,8%
	100% 0,0%	90,7% 9,3%	99,6% 0,4%	99,0% 1,0%

Fig. 6. Matriz de confusão para localização de faltas fase C-terra que ocorreram na área 3.

feitos testes localizando faltas B-terra na área 2 e foi obtido um valor de AUC de 1 na localização de faltas nas zonas 2 e 3, o desempenho do algoritmo para identificar faltas que ocorreram no ramo principal (zona 3) é semelhante à localização de faltas que ocorreram no ramo lateral (zona 2).

V. CONCLUSÕES

Esta pesquisa apresentou resultados promissores com o uso de árvores de decisão para estimar a localização de faltas monofásicas no sistema de distribuição IEEE de 34 barras. Na área 3, definida no sistema de distribuição e, no momento, considerando somente faltas monofásicas, o algoritmo de árvore de decisão J48 apresentou excelentes resultados na identificação da zona de ocorrência das faltas no SD do IEEE de 34 barras. Os resultados foram explorados principalmente utilizando matrizes de confusão, assim como curvas ROC. A matriz de confusão possibilita mostrar onde existe uma localização correta e identificar os erros cometidos pelo algoritmo. Com base na *Area Under the Curve* (AUC), das curvas

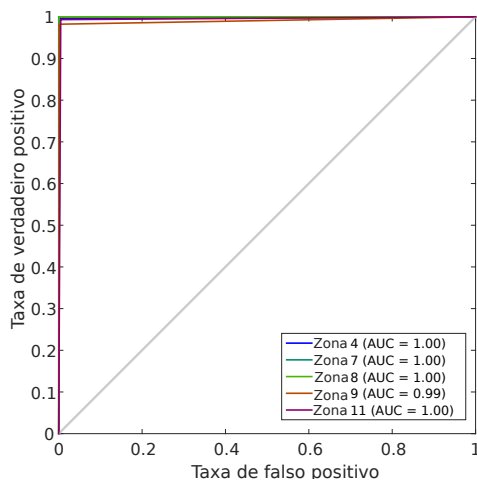


Fig. 7. A curva ROC para localização de faltas fase A-terra que ocorreram na área 3.

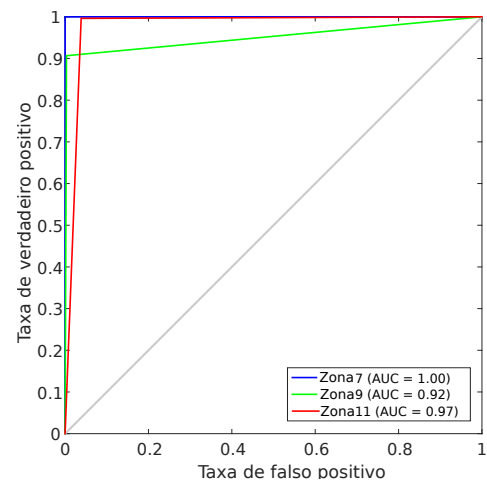


Fig. 9. A curva ROC para localização de faltas fase C-terra que ocorreram na área 3.

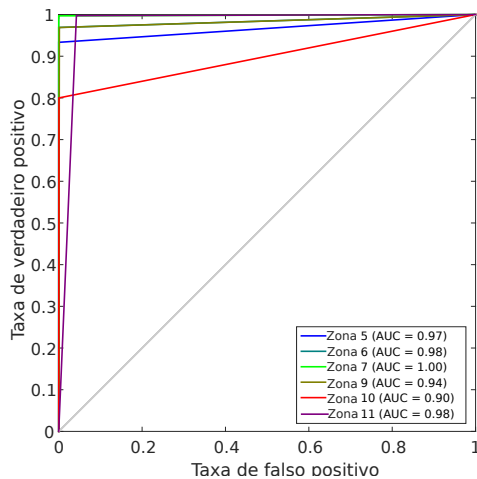


Fig. 8. A curva ROC para localização de faltas fase B-terra que ocorreram na área 3.

ROC, foi possível determinar a eficiência do localizador. Outro ponto de destaque é que esta metodologia pode ser replicada para quaisquer sistemas em que seja possível extrair características das faltas no sistema através de simulações. Outro ponto de destaque é a seleção de atributos realizada pelas árvores de decisão, na qual foi evidenciada a importância da *Wavelet Packet Transform* para a localização de faltas em SD.

V. REFERÊNCIAS

- [1] E. Fuchs and M. A. Masoum, "Power quality in power systems and electrical machines". Academic press, 2011.
- [2] P. Farzan, M. Izadi, C. Gomes, M. Kadir, M. Hesamian, M. Radzi et al., "Short circuit power based fault location algorithm in distribution networks," in *Power Engineering and Optimization Conference (PE-OCO), 2014 IEEE 8th International*. IEEE, 2014, pp. 105–109.
- [3] J. Zhu, D. L. Lubkeman, A. Girgis et al., "Automated fault location and diagnosis on electric power distribution feeders," *Power Delivery, IEEE Transactions on*, vol. 12, no. 2, pp. 801–809, 1997.
- [4] L. Sousa Martins, J. Martins, V. F. Pires, and C. Alegria, "The application of neural networks and clarke-concordia transformation in fault location on distribution power systems," in *Transmission and Distribution Conference and Exhibition 2002: Asia Pacific. IEEE/PES*, vol. 3. IEEE, 2002, pp. 2091–2095.
- [5] A. Rafinia and J. Moshagh, "A new approach to fault location in three-phase underground distribution system using combination of wavelet analysis with ann and fls," *International Journal of Electrical Power & Energy Systems*, vol. 55, pp. 261–274, 2014.
- [6] K. Lout and R. K. Aggarwal, "Current transients based phase selection and fault location in active distribution networks with spurs using artificial intelligence," in *2013 IEEE Power & Energy Society General Meeting*. IEEE, 2013, pp. 1–5.
- [7] J. Ghorbani, M. A. Choudhry, and A. Feliachi, "Fault location and isolation using multi agent systems in power distribution systems with distributed generation sources," in *Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES*. IEEE, 2013, pp. 1–6.
- [8] A. Zapata-Tapasco, J. Mora-Flórez, and M. C. de Almeida, "Fault location in power distribution systems using a learning approach based on decision trees," in *Transmission & Distribution Conference and Exposition-Latin America (PES T&D-LA), 2014 IEEE PES*. IEEE, 2014, pp. 1–6.
- [9] F. C. Trindade, W. Freitas, and J. C. Vieira, "Fault location in distribution systems based on smart feeder meters," *IEEE Transactions on Power Delivery*, vol. 29, no. 1, pp. 251–260, 2014.
- [10] N. Bhargava, G. Sharma, R. Bhargava, and M. Mathuria, "Decision tree analysis on j48 algorithm for data mining," *Proceedings of International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 6, 2013.
- [11] T. R. Patil and S. Sherekar, "Performance analysis of naive bayes and j48 classification algorithm for data classification," *International Journal of Computer Science and Applications*, vol. 6, no. 2, pp. 256–261, 2013.
- [12] M. M. Saha, J. J. Izykowski, and E. Rosolowski, *Fault location on power networks*. Springer Science & Business Media, 2009.
- [13] D. P. S. Gomes, "Metodologia e simplificação do espaço de busca para a alocação do monitor de qualidade de energia voltada para variações de tensão variações," Master's thesis, Universidade de São Paulo, 2016.
- [14] D. T. Feeders, "Ieee pes distribution system analysis subcommittee's, distribution test feeder working group," 2013.
- [15] L. Prikler and H. Hoidalén, "Atpdraw manual," 2002.
- [16] M. MatLab, "The language of technical computing," *The MathWorks, Inc.* <http://www.mathworks.com>, 2012.
- [17] P. Ray, D. P. Mishra, and D. D. Panda, "Hybrid technique for fault location of a distribution line," in *2015 Annual IEEE India Conference (INDICON)*. IEEE, 2015, pp. 1–6.
- [18] G. Holmes, A. Donkin, and I. H. Witten, "Weka: A machine learning workbench," in *Intelligent Information Systems, 1994. Proceedings of the 1994 Second Australian and New Zealand Conference on*. IEEE, 1994, pp. 357–361.
- [19] R. Shatnawi, W. Li, J. Swain, and T. Newman, "Finding software metrics threshold values using roc curves," *Journal of software maintenance and evolution: Research and practice*, vol. 22, no. 1, pp. 1–16, 2010.