

Rede Bayesiana associada a Algoritmo Genético Compacto para classificação de Descargas Parciais em Transformadores de Potência

Pedro Henrique da Silva Palhares, Cacilda de Jesus
Ribeiro, Leonardo da Cunha Brito

Escola de Engenharia Elétrica, Mecânica e de Computação
Universidade Federal de Goiás
Goiânia, Goiás

André Pereira Marques, Cláudio Henrique Bezerra
Azevedo, José Augusto Lopes dos Santos

CELG Distribuição
Goiânia, Goiás

Resumo— Este trabalho apresenta um classificador capaz de estimar a gravidade de ocorrências de descargas parciais em transformadores de potência. Utilizando resultados de medições de emissões acústicas, é possível detectar a presença de Descargas Parciais (DPs) no interior de transformadores, o que permite avaliar o estado do sistema isolante dos mesmos sob esta ótica. Propõe-se, então, a utilização de uma Rede Bayesiana (RB) para classificação destes equipamentos. Associado à Rede Bayesiana, um Algoritmo Genético Compacto adaptado para solução de problemas de programação inteira mista é utilizado para discretização dos parâmetros contínuos que fazem parte da Rede Bayesiana. Compararam-se os resultados com Rede Neural Perceptron de Múltiplas Camadas e, após uma quantidade adequada de execuções do algoritmo, verificou-se que o método proposto apresentou resultado superior aos demais.

Palavras Chave— Descargas Parciais, Redes Bayesianas, Redes Neurais Artificiais, Sistemas de Apoio à Decisão, Transformadores de Potência.

I. INTRODUÇÃO

Prestar um serviço de qualidade, com mínimo custo e sem interrupções, são objetivos das concessionárias de energia elétrica. Para isso, almeja-se que equipamentos importantes e de custo elevado, como os transformadores de potência, operem satisfatoriamente e sem apresentar falhas ou defeitos que provoquem descontinuidade no fornecimento de energia elétrica, pois sua substituição envolve um procedimento caro e demorado [1].

Diversos trabalhos na literatura destacam a importância do desenvolvimento de técnicas preditivas para identificação de defeitos e falhas incipientes em transformadores de potência. Em [1], apresentam-se resultados de 33 anos de acompanhamentos de retirada de operação destes equipamentos e expõe-se a importância da implantação de métodos de detecção de falhas em evolução. Em [2], são apresentadas estatísticas de falhas em transformadores de potência, assim como descrevem-se métodos de diagnóstico e

avaliações sobre a qualidade do óleo isolante. Em [3], é apresentada uma revisão de defeitos e falhas em transformadores de potência e como os possíveis métodos de predição podem ajudar a identificá-los antes que venham ocorrer falhas, por vezes, catastróficas.

Visando evitar a indisponibilidade destes equipamentos e reduzir a quantidade de manutenções corretivas, técnicas preditivas vêm sendo desenvolvidas para detectar falhas incipientes decorrentes de descargas parciais (DPs), causadas, principalmente, pela diminuição da capacidade dielétrica do sistema isolante ou por flutuações de potenciais. Entre as técnicas utilizadas atualmente, às medições pelo método de emissão acústica tem se destacado, pois este não é invasivo e permite a localização das DPs [4], facilitando a identificação, portanto, do(s) ponto(s) de defeito com o equipamento em operação.

Para uma análise completa e eficiente do estado do equipamento, podem ser utilizados, em conjunto com as técnicas preditivas, métodos computacionais de classificação. Várias referências na literatura tratam este problema de classificação de transformadores utilizando dados de medições para identificação de DPs. Dentre elas, [5] e [6] trabalham com Análise de Gases Dissolvidos (AGD), utilizando Redes Bayesianas (RB) e Redes Neurais (RN). Em [7], utiliza-se uma Rede Neural de Múltiplas Camadas (MLP) associada a dados de medições de DPs por emissões acústicas para identificação do defeito.

Assim como o trabalho [8], este artigo apresenta uma abordagem utilizando medições de DPs pelo método de emissões acústicas para classificação do estado do equipamento, onde uma Rede Bayesiana é treinada para classificação do estado do transformador. Diferentemente do trabalho anteriormente citado, que utiliza um algoritmo *Hill-Climbing* para discretização dos valores contínuos da RB, este trabalho apresenta uma metodologia utilizando um Algoritmo Genético Compacto [9] adaptado para solução de problemas de programação inteira mista (MIPCGA). A nova abordagem apresentada provê algumas vantagens sobre a anterior, a saber:

- a) o MIPCGA fornece faixas mais próximas do valor ótimo global;
b) os novos nós permitem a utilização de medições de duração variada, uma vez que, além da contagem de detecções pelos sensores envolvidos, valores médios e percentuais são utilizados para treinamento e classificação; e
c) não se utiliza um gerador de casos, mas somente casos reais para treinamento e validação da RB são aplicados.

II. FALHAS INCIPIENTES E DESCARGAS PARCIAIS

Segundo [8], as falhas incipientes são aquelas que geralmente desenvolvem através de uma gradual deterioração do sistema isolante. Por não serem observáveis a todo instante, são de difícil detecção. É, então, de grande importância a detecção prematura destas falhas ainda em estágio inicial, pois, com o tempo, quando a condição do equipamento se degrada em decorrência de *stress* elétrico, químico, mecânico e/ou térmico, estas evoluem e se intensificam até que ocorra de fato a falha [8]. Com este intuito, surgiram importantes técnicas para detecção de DPs, as quais, segundo [4] e [10], são como pequenas faíscas, decorrentes da diminuição da resistência elétrica do sistema isolante do transformador, sendo indicadores de possível falha em evolução.

Entre as diversas técnicas para detecção de DPs, optou-se, neste trabalho, pela utilização de detecção por emissão acústica, por ser um método não invasivo em que transdutores piezelétricos [11], posicionados de forma estratégica e segura na superfície externa do tanque do transformador, detectam ondas acústicas provenientes do interior do transformador produzidas por DPs. Os parâmetros mais importantes coletados pelo instrumento de detecção que registra os sinais acústicos gerados por DPs (captados pelos sensores) são apresentados na Tabela I.

TABELA I. DADOS COLETADOS POR MEIO DOS SENSORES

Grandeza	Descrição
Tempo (s)	Instante de tempo em que o evento foi detectado pelo sensor.
Amplitude (dB)	A amplitude máxima do sinal durante a detecção do evento
Energia	Energia acumulada durante a detecção.
Duração (μs)	Duração do evento.
Tempo de Subida (μs)	Tempo entre o início do evento e o pico da amplitude.

III. REDES BAYESIANAS

Redes Bayesianas são grafos acíclicos cujos nós possuem um conjunto finito de estados [8]. Para cada nó, existe uma tabela de probabilidades condicionais que relaciona cada valor do nó filho a cada valor possível de seus pais. Este trabalho utilizou os mesmos algoritmos de aprendizagem de estrutura e parâmetros e motor de inferência de [8], que estão apresentados nas subseções abaixo.

Estrutura

O método do K2 trata-se de um método de busca gulosa, que utiliza uma métrica própria (2) para avaliação da estrutura [12]. Seja N_{ijk} o número de casos na base de treinamento em que a variável u_i é instanciada com valor v_{ik} , $\pi(u_i)$ os pais de u_i , q_i a

quantidade de instanciações possíveis de $\pi(u_i)$ e r_{u_i} a quantidade valores possíveis discretos de u_i :

$$N_{ij} = \sum_{k=1}^{r_{u_i}} N_{ijk} \quad (1)$$

$$g(i, \pi(u_i)) = \prod_{j=1}^{q_i} \frac{(r_{u_i} - 1)!}{(N_{ij} + r_{u_i} - 1)!} \prod_{k=1}^{r_{u_i}} N_{ijk}! \quad (2)$$

Procedimento 1 - K2

Entrada: Conjunto U de nós da rede, lista de casos para treinamento, limite de pais para cada nó.

Saída: Grafo Acíclico Direcional.

Para todo $u_i \in U$ faça:

1. Encontre a variável u_i' ($i \neq i'$ e i' precede i na ordem dos nós) em que o valor de $g(i, \pi(u_i) \cup \{u_i'\})$ é máximo.
2. Adicione u_i' a $\pi(u_i)$.
3. Continue até que não seja possível encontrar a variável u_i' no passo 1, ou até que a quantidade de pais para este nó alcance o limite.

Para a execução do algoritmo, é necessário o conhecimento prévio da ordem dos nós, o que nem sempre é possível. Desta forma, sugere-se adicionalmente a utilização de um algoritmo *Hill-Climbing* [12]-[13] utilizando a métrica BIC (3) e assim permitir a escolha da melhor estrutura. Seja n a quantidade de nós e N a quantidade de exemplos da base de treinamento, D a base de dados conhecidos e G o GAD a ser pontuado, tem-se que:

$$BIC(G, D) = \sum_{i=1}^n \sum_{j=1}^{q_i} \sum_{k=1}^{r_{u_i}} N_{ijk} \log_2 \left(\frac{N_{ijk}}{N_{ij}} \right) - \frac{\log_2 N}{2} \sum_{i=1}^n q_i (r_{u_i} - 1) \quad (3)$$

A métrica BIC apresenta a vantagem de ser equivalente [12], isto é, tem a mesma pontuação para duas estruturas equivalentes. Desta forma, quando se gera uma nova estrutura, basta-se calcular a diferença desta com a anterior, sem necessidade de calcular a pontuação da estrutura toda. As operações válidas são “adicionar um arco”, “remover um arco” e “inverter a direção do arco”, procurando sempre por estruturas acíclicas.

Procedimento 2 - Hill-Climbing

Entrada: Estrutura Inicial S (caso não exista, consideram-se todos os nós desconectados).

Saída: Grafo Acíclico Direcional.

Repita enquanto S não for um ótimo local e o número máximo de iterações não tiver sido atingido:

1. Calcular a pontuação BIC relativa a cada operação válida na estrutura S .
2. Seja Δ^* a pontuação máxima obtida com as operações no passo 1. Seja A^* a operação responsável por Δ^* .
3. Caso $\Delta^* > 0$, persista A^* em S .

Neste trabalho, geraram-se duas estruturas, uma com cada método descrito (K2 e *Hill-Climbing*), escolhendo-se a melhor pontuada por (3).

Aprendizagem de Parâmetros

Uma vez que os sensores fornecem sempre todos os dados relativos à atividade presente no interior do equipamento (dados completos), sugere-se a utilização do MLE, que, segundo [14], trata-se de uma contagem, na forma $P(C|A) = N(C,A)/N(A)$, em que $N(C,A)$ é o número de casos em que C e A ocorrem, e $N(A)$ é o número de casos em que A ocorre.

Motor de Inferência

Vários motores de inferência em Redes Bayesianas estão disponíveis na literatura [12] e [14]. Porém, este trabalho utiliza o método da eliminação de variáveis (*Variable Elimination*), capaz de resolver consultas exatas (a probabilidade encontrada é exata e não aproximada) de probabilidades *a priori*, probabilidades *a posteriori*, explanação mais provável (MPE) e hipótese máxima *a posteriori* (MAP), através de sucessivas eliminações de variáveis no processo de construção da consulta.

Discretização de Valores Contínuos

A parametrização da RB exige o conhecimento das distribuições de probabilidade condicionais dos nós e seus nós pais, que, no caso de variáveis contínuas, não trataria apenas de uma tarefa de contagem. Alguns trabalhos na literatura abordam este problema, como [15], que utiliza somas de gaussianas ponderadas para aproximação das probabilidades *a priori* (nó raiz) e condicionais, e, em outros [14] e [16], as variáveis contínuas tem a necessidade de seguir a distribuição gaussiana.

Desta forma, neste artigo, utilizaram-se apenas valores discretos. De modo a estabelecer adequada e automaticamente os intervalos dos parâmetros, propõe-se um método de discretização de variáveis contínuas através da atribuição de valores discretos com limites inferior e superior, a intervalos contínuos, utilizando um algoritmo genético compacto adaptado para soluções de problemas de programação inteira mista, o MIPCGA.

IV. MIPCGA

Em um Algoritmo Genético tradicional (GA), os indivíduos competem entre si e o número de cada gene pode aumentar ou diminuir dependendo das decisões boas ou ruins tomadas pela GA [9]. No CGA, a seleção ocorre diferentemente da GA, onde os genes são avaliados no contexto do indivíduo como um todo. Suponha que duas soluções, *a* e *b*, estão competem entre si em um problema *onemax*:

TABELA II. APTIDÃO E REPRESENTAÇÃO BINÁRIA DE CADA SOLUÇÃO

Solução	Cromossomo	Aptidão
<i>a</i>	111011	5
<i>b</i>	110100	3

Durante a competição, a solução *a* irá vencer por ter uma maior aptidão. O problema nesta escolha ocorre na posição 4 do

cromossomo, onde ocorre um erro e um 0 é escolhido no lugar de um 1. Desta forma, [9] propõe um esquema semelhante à uma seleção por torneio: a cada iteração, a proporção de genes vencedores irá crescer $1/n$, sendo *n* o tamanho da população, beneficiando os genes vencedores do vencedor de cada torneio.

Utilizando-se de um procedimento adequado para codificação das variáveis inteiras e reais em bits e um procedimento baseado em [9] para manipulação das restrições, este artigo propõe a utilização do algoritmo CGA adaptado para solução de problemas inteiros mistos, o MIPCGA, na discretização das variáveis contínuas para serem utilizadas no processo de classificação da RB.

Procedimento 3 - MIPCGA

Entrada: *n*: tamanho da população,
nr: número de variáveis reais,
ni: número de variáveis inteiras,
nrbits: número de bits para variáveis reais,
maxi: número máximo de iterações;
Saída: *p*: solução final.

Repita enquanto *S* não for um ótimo local e o número máximo de iterações não tiver sido atingido:

```

1. tamvari = max (log2(maxx - minx) + 1);
2. nibits = log2(tamvari - 1) + 1;
3. pi = 0.5, 1 ≤ i ≤ ni * nibits + nr * nrbits;
4. para i = 1 até maxi
5.   (a, bitsa) = geraIndividuo(ni, nibits, nr, nrbits, minx, maxx);
6.   (b, bitsb) = geraIndividuo(ni, nibits, nr, nrbits, minx, maxx);
7.   (vencedor, bitswin, fwin, perdedor, bitloss, v) = compete(a, b);
8.   idx = 0;
9.   para j=1 até ni
10.    para k=1 até nibits
11.      idx = idx + 1;
12.      se bitswinidx ≠ 1
13.        se bitswinidx == 1
14.          pidx = pidx + 1/n;
15.        se não pidx = pidx - 1/n;
16.        fim se;
17.      fim se;
18.    fim para;
19.  fim para;
20.
21. para j=1 até nr
22.   para k=1 até nrbits
23.     idx = idx + 1;
24.     se bitswinidx ≠ 1
25.       se bitswinidx == 1
26.         pidx = pidx + 1/n;
27.       se não pidx = pidx - 1/n;
28.       fim se;
29.     fim se;
30.   fim para;
31. fim para;
32. continue = 0;
33. para i = 1 até ni * nibits + nr * nrbits
34.   se pi > 0 e pi < 1
35.     continue = 1;
36.     break;
37.   fim se;
38. fim para;
39. se continue = 0
40.   break;
```

```

41.     fim se;
42. fim para;

```

Procedimento 4 - GeraIndividuo

Entrada: ni: número de variáveis inteiras,
nibits: número de bits para variáveis inteiras,
nr: número de variáveis reais,
nrbits: número de bits para variáveis reais,
p: vetor de probabilidades,
minx: vetor com valores mínimos de cada variável,
maxx: vetor com valores máximos de cada variável,
bin2dec(b): converte um conjunto de bits para decimal.
Saída: vetor com os valores gerados,
b: vetor com a representação binária dos elementos de x .

```

1.     idx = 0;
2.     i = 1;
3.     enquanto i ≤ ni
4.         para j = 1 até nibits
5.             idx = idx + 1;
6.             se random() < pidx
7.                 tmpi,j = 1;
8.             se não tmpi,j = 0;
9.             fim se;
10.        fim para;
11.        dx = (maxx(i) - minx(i))/(2k-1);
12.        dec = bin2dec(tmpi);
13.        f = dx * dec + minx(i);
14.        xi = ⌊f⌋;
15.        bi = tmpi;
16.    fim enquanto;
17.    para i=1 até nr
18.        para j = 1 até nrbits
19.            idx = idx + 1;
20.            se random() < pidx
21.                tmpr,j = 1;
22.            se não tmpr,j = 0;
23.            fim se;
24.            fim para;
25.            dx = (maxx(ni + i) - minx(ni + i))/(2k-1);
26.            dec = bin2dec(tmpr);
27.            f = dx * dec + minx(ni + i);
28.            xni+i = f;
29.            bni+i = tmpr;
30.        fim para;

```

Procedimento 5 - Compete

Entrada: x_i: soluções que vão competir,
b_i: apresentação binária de cada solução,
f_i: avaliação da função objetivo de cada solução,
restricao(x_i): função que retorna a soma das violações das restrições. Um valor negativo ou zero, caso não haja violação.
Saída: vencedor: melhor solução,
bitswin: apresentação binária da melhor solução,
fwin: avaliação da função objetivo da melhor solução,
perdedor: pior solução,
bitsloss: apresentação binária da pior solução,
v: violação de restrição pela melhor solução;

```

1.     v1 = restricao(x1);

```

```

2.     v2 = restricao(x2);
3.     se v1 > v2
4.         vencedor = v1;
5.         bitswin = b1;
6.         fwin = f1;
7.         perdedor = v2;
8.         bitsloss = b2;
9.         v = v1;
10.    senão se v1 < v2
11.        vencedor = v2;
12.        bitswin = b2;
13.        fwin = f2;
14.        perdedor = v1;
15.        bitsloss = b1;
16.        v = v2;
17.    senão
18.        se f1 ≤ f2
19.            vencedor = v1;
20.            bitswin = b1;
21.            fwin = f1;
22.            perdedor = v2;
23.            bitsloss = b2;
24.            v = v1;
25.        senão
26.            vencedor = v2;
27.            bitswin = b2;
28.            fwin = f2;
29.            perdedor = v1;
30.            bitsloss = b1;
31.            v = v2;
32.        fim se;
33.    fim se;

```

É importante destacar que o Procedimento 4 seleciona o candidato que menos viola as restrições ou, caso esse valor seja igual para as duas, seleciona aquela com o melhor resultado da função objetivo, neste caso o algoritmo está minimizando. O procedimento tem resultado idêntico ao apresentado em [17].

A aplicação do MIPCGA para discretização dos valores da Rede Bayesiana envolve uma solução de um problema de $2n$ variáveis, onde n é a quantidade máxima de limites. Um exemplo de solução válida está apresentado na Tabela III e sua representação discreta na Tabela IV, onde 0 é o menor valor e 1756 é o maior valor possível para a variável.

TABELA III. EXEMPLO DE VALORES VÁLIDOS PARA LIMITES.

Variável Inteira	Variável Real
1	301
0	9874
0	1000
1	455

TABELA IV. VALORES DISCRETOS EQUIVALENTES.

Valor Discreto	Aptidão
1	0-301
2	301-1301
3	1301-1756

Para avaliação das soluções, neste trabalho, considerou-se o seguinte procedimento para cálculo da função objetivo:

Procedimento 6 – Avaliação da Função Objetivo

1. Separam-se as medições em dois grupos: 70% para treinamento da RB e 30% para validação escolhidos ao acaso.
2. Utilizando os casos de treinamento e considerando-se os limites da solução a ser avaliada:
 1. Aprende-se a estrutura da RB utilizando os algoritmos descritos na seção 3.B.
 2. Utiliza-se o método MLE para aprender os parâmetros da RB.
3. Verifica-se o desempenho da RB através da medição dos acertos da classificação do grupo de validação:
 1. Para cada caso i do grupo de validação:

$$\text{score}_i = \text{score}_i + \text{prob}(i_c) + \text{acerto}(i),$$
 onde $\text{prob}(i_c)$ retorna a probabilidade de ocorrer a classificação correta de i e $\text{acerto}(i)$ retorna 0 em caso de classificação errada e 1 em caso de classificação correta.
4. O valor da função objetivo a ser maximizado é:

$$f(x) = \sum \text{score}_i$$

V. METODOLOGIA PARA CLASSIFICAÇÃO DE DESCARGAS PARCIAIS EM TRANSFORMADORES DE POTÊNCIA

Propõe-se então, neste trabalho, uma metodologia para classificação de descargas parciais em transformadores de potência, utilizando os procedimentos descritos neste e medições em instrumentos utilizando sensores de detecção de emissões acústicas.

A Tabela V descreve as informações utilizadas para classificação e a Tabela VI descreve as possíveis classificações.

TABELA V. VALORES DISCRETOS EQUIVALENTES.

Nó	Descrição
Contagem	Contagem de atividades registradas pelos sensores.
Média	Valor médio da energia registrada na medição.
Porcentagem	Contabiliza-se a porcentagem de minutos durante a medição em que houve atividades.

TABELA VI. VALORES PARA O NÓ DE CLASSIFICAÇÃO.

Classificação	Ação
A	Continuar a operar o equipamento normalmente.
B	Continuar a operar normalmente estando atento à evolução de DPs nos próximos registros.
C	Investigar e realizar outros ensaios o mais breve possível para confirmar resultados e tendências de DPs.
D	Programar a retirada do equipamento de operação em caráter de urgência para uma inspeção interna, localização e correção de defeito.

Procedimento 7 – Metodologia

1. Utiliza-se o algoritmo MIPCGA da seção 4 para discretização dos nós da RB.
2. Escolhidos ao acaso, separa-se as medições em dois grupos: 70% para treinamento e 30% para validação.
3. Utilizando os casos de treinamento:
 1. Aprende-se a estrutura da RB utilizando os algoritmos descritos na seção 3. A ordem dos nós definidas para o K2 é: Classificação, Contagem, Média e Percentual. A quantidade máxima de pais para K2 é 4. Para o *Hill Climbing*, são considerados todos os nós desconectados.
 2. Utiliza-se o método MLE para aprender os parâmetros da RB.
4. Verifica-se o desempenho da RB através da medição dos acertos da classificação do grupo de validação.

VI. RESULTADO E DISCUSSÃO

Foram utilizadas quarenta medições realizadas em transformadores de potência. Logo, têm-se 28 amostras para treinamento da RB e 12 para validação. Aplicou-se a metodologia (Procedimento 7) 30 vezes e verificou-se o desempenho de cada RB 100 vezes, selecionando os grupos de treinamento e validação ao acaso. Utilizou-se $\log_{10}$ nos valores contínuos.

Utilizando aprendizagem de estrutura, entre as 30 RBs criadas, obteve-se um valor médio de acerto de 79.94%, com a média máxima de 85.67% e a média mínima de 72.25%. Na Fig. 1 estão apresentadas todas as execuções do algoritmo.

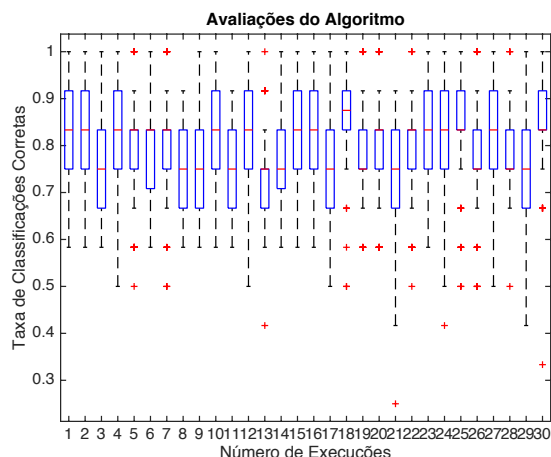


Figura 1. Resultado da avaliação do algoritmo.

Para efeito de comparação, utilizou-se uma Rede Neural *Multilayer Perceptron* (MLP). Executou-se o treinamento e a validação 100 vezes, dividindo-se os dados das medições em grupos de treinamento e validação de 70% e 30% dos dados, respectivamente. Variou-se o número de neurônios e camadas da MLP, verificando que, para este problema, uma camada interna com cinco neurônios, foi a melhor configuração. A média de acertos da MLP foi de 73.67%, tendo seu valor mínimo de 8.33% e máximo de 100%. A aplicação de $\log_{10}$ nos valores contínuos não apresentou melhoria nos resultados da MLP.

Na Tabela VII, é mostrado um comparativo de taxa de acertos dos algoritmos. Para se escolher a melhor RB, utilizaram-se os limites que apresentaram a melhor taxa média de acertos.

TABELA VII. COMPARATIVOS DOS ALGORITMOS SELECIONANDO OS MELHORES LIMITES ENCONTRADOS PELO MIPCGA.

Quantidade de Acertos	RB	MLP
Menor	50%	8.33%
Média	85.67%	73.67%
Maior	100%	100%

TABELA VIII. VALORES DISCRETOS DE UMA RBs ESCOLHIDA AO ACASO.

Valor Discreto	Contagem	Percentual	Média
1	0- 4.5037	0-1.72%	0-0.10
2	4.5037-7.4667	1.72%-1.91%	0.10-3.97
3	-	1.91%-100%	3.97-12.2

VII. CONCLUSÃO

Este artigo apresentou uma metodologia capaz de criar uma Rede Bayesiana adequada à classificação do estado de transformadores de potência, levando-se em consideração a detecção de descargas parciais pelo método de emissão acústica. Tal rede é capaz de indicar não só a classificação do equipamento, mas também a porcentagem de crença no resultado, dadas as evidências apresentadas. Propôs-se, para discretização dos valores contínuos que alimentam a RB, um otimizador baseado no Algoritmo Genético Compacto para solução de problemas de otimização inteira mista com restrições.

A metodologia proposta, que envolve treinamento da RB e discretização utilizando o MIPCGA, possui potencial para aplicação em outros problemas, visto que sendo alimentada com um conjunto de dados de treinamento, contínuos ou discretos, constrói-se uma Rede Bayesiana que pode ser utilizada para classificação. A execução de uma metaheurística é um processo que pode ser demorado, porém o algoritmo só necessita ser utilizado à medida que a RB necessite ser treinada novamente. Desta forma, após treinada a RB, pode-se executar classificações rápidas sem necessidade de execução do otimizador.

Como critério de comparação, criaram-se *Perceptrons* de Múltiplas Camadas, método adequado para o problema de classificação, e verificou-se que a RB alcançou desempenho superior. Quando observados os piores casos de classificação, obteve-se uma taxa de 50% da RB, em face de 8.33% da MLP. Em termos de valores médios, a RB acertou 85.67% dos casos, enquanto que a MLP acertou 73.67%. Desta forma, a metodologia proposta apresentou resultados melhores, tanto no pior caso, quanto em média, em relação à MLP.

Os resultados indicaram a RB como mais adequada para o processo de classificação. É interessante observar que a metodologia proposta é adequada a outros problemas de

classificação, uma vez que se podem utilizar tanto parâmetros discretos, quanto parâmetros contínuos, desde que exista um conjunto de dados adequado para treinamento.

AGRADECIMENTOS

Os autores agradecem à FAPEG (Fundação de Amparo à Pesquisa do Estado de Goiás) e à ANEEL (Agência Nacional de Energia Elétrica) pelos apoios financeiros no desenvolvimento deste trabalho.

REFERÊNCIAS

- [1] A. P. Marques et al, "Power Transformer Disruptions – A Case Study", *IEEE Electrical Insulation Magazine*, vol. 30, no. 2, pp. 17-21, March, 2014.
- [2] W. Wang, A. J. Vandermaar and K. D. Srivasta, "Review of Condition Assessment of Power Transformers in Service", *IEEE Electrical Insulation Magazine*, vol. 18, no. 6, pp. 12-25, Nov., 2002.
- [3] H. F. Júnior, J. G. S. Costa and J. L. M. Olivas, "A review of monitoring methods for predictive maintenance of electric power transformers based on dissolved gas analysis", *Renewable and Sustainable Energy Reviews*, vol. 46, pp. 201-209, 2015.
- [4] E. Mohammadi et al, "Partial Discharge Localization and Classification Using Acoustic Emission Analysis in Power Transformer", in *INTELEC*, Incheon, South Korea., 2009, pp. 1-6.
- [5] W. H. Tang, Z. Lu, Q. H. Wu, "A Bayesian Network Approach to Power System Asset Management for Transformer Dissolved Gas Analysis", in *Proc. 2008 Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, 2008.
- [6] X. Hao, S. Cai-xin, "Immune Network Classification Algorithm for Fault Diagnosis of Power Transformer", *IEEE Trans. Power Delivery*, vol. 22, No.2, 2007.
- [7] C. Kuo, "Artificial recognition system for defective types of transformer by acoustic emission", *Expert Systems with Applications*, vol. 36, 2009.
- [8] P.H.S. Palhares et al, "Classificador Bayesiano Associado à Técnica de Ensaios de Emissão Acústica para a Análise de Descargas Parciais em Transformadores de Potência", no 4º *Simpósio Brasileiro de Sistemas Elétricos*, 2012.
- [9] G. R. Harik, F. G. Lobo, D. E. Goldberg, "The Compact Genetic Algorithm", *IEEE Transactions on Evolutionary Computation*, vol 3., pp. 287-297, Nov. 1999.
- [10] C. H. Azevedo et al, "Methodology for the Detection of Partial Discharges in Power Transformers", in *EUROCON, Saint-Petersburg, Russia*, 2009.
- [11] IEEE Guide for the Detection and Location of Acoustic Emissions from Partial Discharges in Oil-Immersed Power Transformers and Reactors, IEEE Standard C57.127, 2007
- [12] F. V. Jensen and T. D. Nielsen, "Bayesian Networks and Decision Graphs", 2nd ed., Springer, 2007,
- [13] T. Pansombut, "Learning Bayesian Belief Networks". North Carolina State University, North Carolina, EUA. [Online]. Available: http://www4.ncsu.edu/~tpansom/pansombut_tatdow_890.pdf
- [14] A. Darwiche, "Modeling and Reasoning with Bayesian Networks", 1st ed., Cambridge University Press, 2009.
- [15] E. Driver and D. Morrell, "Implementation of Continuous Bayesian Networks Using Sums of Weighted Gaussians", in *Proc. 11th Conference on Uncertainty in Artificial Intelligence*, 1995.
- [16] P. J. F. Lucas and A. Hommersom, "Modeling the Interactions between Discrete and Continuous Causal Factors in Bayesian Networks", *International Journal of Intelligent Systems*, vol. 30, pp. 209-235, 2015.
- [17] K. Deb, "An efficient constraint handling method for genetic algorithm", *Computer methods in applied mechanics and engineering*, vol 186, 2000.