

REDE BAYESIANA PARA ESTIMAÇÃO DE FALHAS INCIPIENTES EM TRANSFORMADORES DE POTÊNCIA UTILIZANDO DETECÇÃO DE DESCARGAS PARCIAIS POR EMISSÃO ACÚSTICA

Pedro Henrique da S. Palhares¹, Cacilda de J. Ribeiro¹, André P. Marques^{2,3}, Cláudio Henrique B. Azevedo²
e Leonardo da C. Brito¹

¹Escola de Engenharia Elétrica e de Computação, Universidade Federal de Goiás

²CELG Distribuição, ³Instituto Federal de Goiás

palhares@posgrad.ufg.br, cacilda@eeec.ufg.br, andre.pm@celg.com.br, claudio.hb@celg.com.br, brito@eeec.ufg.br

Resumo – É apresentada neste artigo uma metodologia para estimação de falhas incipientes no sistema isolante de transformadores de potência com base em resultados de ensaios de detecção de descargas parciais pelo método de emissão acústica, propiciando às equipes de engenharia de manutenção uma importante ferramenta de avaliação do estado destes equipamentos sob a ótica desta emergente técnica preditiva. Para esse propósito, é proposta uma abordagem utilizando uma Rede Bayesiana associada ao algoritmo *Hill Climbing* para a discretização dos parâmetros da rede. Os resultados mostram que o método é eficaz, apresentando empiricamente uma precisão de classificação de 89%, enquanto que uma abordagem alternativa, na qual uma Rede Neural *Perceptron* de Múltiplas Camadas foi aplicada ao mesmo problema, ofereceu uma precisão de 83%.

Palavras Chave – redes bayesianas, redes neurais, estimação, degradação, transformadores de potência.

Abstract – It is presented on this paper a methodology for estimating incipient faults in the insulating system of power transformers, based on tests results for detecting partial discharges by the acoustic emission method, providing to the maintenance engineering teams an important tool for evaluating the state of the equipment from the perspective of this emerging predictive technique. For this purpose, an approach using a Bayesian network associated with the Hill Climbing algorithm for discretization of network parameters is proposed. The results show that the method is effective, achieving a classification accuracy of 89%, while an alternative approach, in which a Neural Network Perceptron with Multiple Layers was applied to the same problem, offered a precision of 83%.

Keywords – bayesian networks, neural networks, estimation, degradation, power transformers.

1 Introdução

De fundamental importância para a transmissão e distribuição de energia elétrica, os transformadores de potência são equipamentos de custo elevado. Durante sua operação, podem surgir defeitos em sua isolação, provocando uma diminuição de sua capacidade dielétrica [1]. Segundo [2], a diminuição desta capacidade pode provocar o surgimento de descargas parciais (DPs) dentro do transformador de forma inesperada e de difícil detecção. Detectar defeitos incipientes de funcionamento nesses transformadores é de interesse econômico e operacional, uma vez que permite intervenções preventivas que evitem danos graves ao transformador o que, por consequência, garante maior continuidade e elevam a confiabilidade do fornecimento de energia elétrica aos – cada vez mais exigentes – consumidores.

Atualmente, têm-se utilizado vários métodos de detecção de descargas parciais: o elétrico, o químico (análise de gases dissolvidos ou AGD) e o acústico. Neste trabalho destacamos o método de detecção por emissão acústica que apresenta a vantagem de localização da descarga dentro do equipamento sem que haja a necessidade de se retirá-lo de operação [2], evitando-se descontinuidade no fornecimento da energia elétrica. Vários trabalhos na literatura abordam o problema da identificação de DPs. Dentre eles, destacam-se aqueles que tomam os resultados obtidos pelas técnicas AGD, conforme [3], [4] e [5]. Nesses trabalhos, apresentam-se métodos computacionais dirigidos à identificação dos possíveis defeitos no interior dos transformadores aplicando Redes Bayesianas (RB) e Redes Neurais (RN). Diferentemente dos trabalhos referenciados, este propõe prover ao engenheiro de manutenção uma metodologia capaz de auxiliá-lo na tomada de decisões quanto à manutenção preventiva dos transformadores, indicando provável severidade de degradação da isolação do transformador e, conseqüentemente, sugerindo uma ação preventiva. Neste caso, a tomada de decisão subjetiva, caracterizada como um processo de inferência probabilística, é realizada por meio de uma Rede de Crença Bayesiana, a qual apresenta as seguintes vantagens qualitativas frente à abordagem por RN: i) uma RN tem uma representação limitada, mapeando entrada-saída unidirecionalmente na forma de uma “caixa-preta”, impossibilitando a identificação de correlações entre as variáveis do problema de decisão; e ii) uma RB permite constatar essas correlações diretamente, além de explicitamente trabalhar com probabilidades (graus de crença) reais. Investiga-se também o desempenho quantitativo da RB frente a uma RN.

2 Detecção de Descargas Parciais por meio de Ensaio Acústico

Devido à grande dificuldade e custo da interrupção do funcionamento dos transformadores, tornou-se necessário o desenvolvimento de um método que fosse capaz de detectar descargas parciais (DPs) sem que houvesse o desligamento dos transformadores. Surgiu então o método da Emissão Acústica (EA), não invasivo, no qual são utilizados sensores acústicos

posicionados estrategicamente na superfície externa do transformador (levando em consideração o projeto do transformador), com o intuito de monitorar todo o interior do equipamento.

Cada DP age como uma fonte de ondas acústicas as quais propagam no interior do transformador, através do óleo, e podem ser detectadas nas paredes exteriores do tanque do equipamento [2]. Como são utilizados diversos sensores (neste trabalho foram utilizados 14 para cada transformador monitorado), é possível realizar a triangulação dos sinais de forma a localizar a fonte de emissões.

Os parâmetros mais importantes coletados pelo instrumento de ensaio que registra os sinais acústicos gerados por DPs (captados pelos sensores) são apresentados na Tabela 1, com ilustração na Fig. 1. Todos eles são medidos com referência ao limiar de recepção dos sensores.

Tabela 1 – 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 (J)	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.

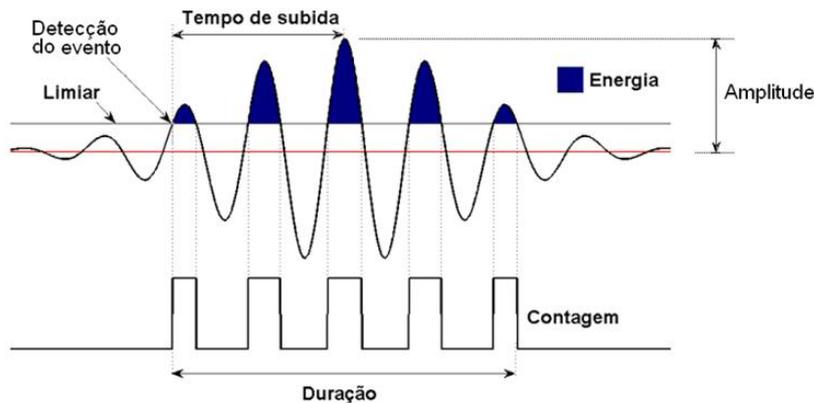


Figura 1 – Forma idealizada do sinal acústico.

De posse desses dados, é necessário estabelecer uma maneira de diferenciar os ruídos emitidos pela operação do transformador e as DPs. É sabido que as descargas mais intensas ocorrem nos picos e vales de tensão, ou seja, a 90° e a 270° da tensão senoidal (iniciada em zero) de fornecimento de energia elétrica, respectivamente. Sabendo-se que a frequência da tensão é de 60Hz e que t_1 e t_2 são os tempos de ocorrência do primeiro e do segundo evento, a diferença angular entre dois estes eventos consecutivos é pode ser calculada por:

$$\Delta t = (t_2 - t_1) \cdot 360^\circ \cdot 60 \quad (1)$$

$$\theta = \Delta t - \left\lfloor \frac{\Delta t}{360^\circ} \right\rfloor \cdot 360^\circ \quad (2)$$

Sabendo-se que as descargas parciais ocorrem nos picos e vales de 90° e 270° , respectivamente, pode-se ajustar o ângulo correto utilizando $180 - \theta/2$ para $\theta \leq 180^\circ$ e $\theta/2 + 180^\circ$ para $\theta > 180^\circ$. Tendo-se a quantidade de descargas ocorridas, o nível de ruído e o montante de energia, é possível então criar um método para identificação dos transformadores em melhor estado e os de estado mais precário, sob a ótica de descargas parciais.

3 Rede Bayesiana

A utilização de Rede Bayesiana (RB) é particularmente adequada ao processo de identificação de transformadores defeituosos, pois se configura como uma ferramenta probabilística de representação do conhecimento. Em geral, uma RB é representada por sua estrutura, um grafo acíclico direcional (GAD), e por seus parâmetros que são constituídos de probabilidades condicionais. Cada nó do GAD representa uma variável estocástica e cada aresta representa dependências condicionais observáveis no mundo real. É possível aprender tanto a estrutura, quanto os parâmetros de uma RB, desde que se tenha uma base de treinamento representativa da mesma.

3.1 Teorema de Bayes

O Teorema de Bayes é a base da teoria de Redes Bayesianas, fornecendo em uma fórmula matemática uma maneira de se calcular probabilidades condicionais. Sua utilidade para Redes Bayesianas encontra-se na atualização de probabilidades de variáveis não observadas, em face de novas evidências. Sabendo que $P(x|e)$ é probabilidade de x dada a evidência e , $P(e|x)$ a verossimilhança de $P(x|e)$, $P(x)$ e $P(e)$ as probabilidades *a priori* de x e e , respectivamente, tem-se:

$$P(x|e) = \frac{P(e|x)P(x)}{P(e)} \quad (3)$$

3.2 Construção da Rede Bayesianas

O aprendizado da RB pode ser feito através do aprendizado de parâmetros em todas as possíveis estruturas e posterior escolha da estrutura que mais se encaixe com os dados usados para treinamento. Porém, o número de possíveis estruturas, $f(n)$ na equação (4), cresce super-exponencialmente com o aumento do número de nós n , como descrito por [6].

$$f(n) = \sum_{i=1}^n (-1)^{i+1} \frac{n!}{(n-i)!n!} 2^{i(n-i)} f(n-1) \quad (4)$$

Para contornar esse problema, [6] descreve o algoritmo K2 de busca gulosa, o qual utiliza a métrica da equação (6) para avaliar a estrutura. 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 :

$$N_{ij} = \sum_{k=1}^{r_{u_i}} N_{ijk} \quad (5)$$

$$g(i, \pi(u_i)) = \prod_{j=1}^{q_i} \frac{(r_{u_i}-1)!}{(N_{ij}+r_{u_i}-1)!} \prod_{k=1}^{r_{u_i}} N_{ijk} \quad (6)$$

A principal característica deste algoritmo é que ele supõe o conhecimento da ordem em que os nós são percorridos no GAD e impõe um limite de pais para cada nó, reduzindo-se a um problema de tempo polinomial. Os passos do algoritmo, conforme descrito em [7], são:

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.

O algoritmo K2, apesar de prover uma significativa melhora se comparado ao método de força bruta (busca exaustiva), ainda tem um problema: é necessário conhecer a ordem dos nós. Caso não esteja disponível um especialista para fornecer tal informação, a qualidade do GAD final pode ser comprometida. Por isso, este trabalho propõe a utilização de um método alternativo, o qual gera um GAD alternativo ao K2 e, por fim, escolhe-se o melhor dentre os dois. A alternativa escolhida foi o *Hill Climbing* ([6] e [7]) utilizando a métrica BIC (*Bayesian Information Criterion*) [6] como método de pontuação. 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:

$$BIC(G, D) = \sum_{i=1}^n \sum_{j=1}^{q_i} \sum_{k=1}^{r_{u_i}} N_{ijk} \log_2 \left(\frac{N_{ijk}}{N_{ij}} \right) - \frac{\log_2 N}{2} \sum_{i=1}^n q_i (r_{u_i} - 1) \quad (7)$$

Uma vantagem da métrica BIC sobre a métrica utilizada pelo K2 é a sua equivalência [6], isto é, estabelece a mesma pontuação para estruturas que, apesar de diferentes, são equivalentes. Assim, como o *Hill Climbing* é um algoritmo guloso no qual que é feita, a cada iteração, uma pequena modificação em uma proposta corrente de solução, não é necessário recalculer a pontuação da rede inteira, mas apenas calcular a diferença de pontuação causada por tal modificação. As operações válidas são “adicionar um arco”, “remover um arco” e “inverter a direção do arco”, desde que o grafo permaneça acíclico.

Procedimento 3 – Hill Climbing para geração do GAD

Entrada: Estrutura Inicial (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 do passo 1. Seja A^* a operação responsável por Δ^* .
3. Caso $\Delta^* > 0$, persista A^* em S .

Dessa maneira é possível gerar duas estruturas, uma com K2 e outra com *Hill Climbing*, e utilizar a melhor pontuada pela métrica BIC.

3.3 Aprendizado de Parâmetros

O aprendizado de parâmetros com dados completos utilizando o método MLE (*Maximum Likelihood Estimation*) se trata basicamente de uma tarefa de contagem, onde $P(C|A) = N(C,A)/N(A)$, na qual $N(C,A)$ é a quantidade de casos no banco de treinamento em que ocorrem C e A [6].

3.4 Motor de Inferência

O motor de inferência utilizado pela RB deste trabalho foi implementado utilizando o processo de eliminação de variáveis (*Variable Elimination*). É um dos métodos mais simples, porém poderoso, em que as consultas são feitas através de sucessivas eliminações de variáveis da RB [8]. Esse é capaz de responder consultas de probabilidade *a priori*, probabilidade *a posteriori*, explanação mais provável (MPE) e hipótese máxima *a posteriori* (MAP).

3.5 Geração de Casos de Treinamento e Validação

Para construir um Classificador Bayesiano para avaliação das condições dos transformadores, é necessária a definição dos nós a serem utilizados e uma base de treinamento, para seja possível aprender a estrutura e os parâmetros. Como é desejado detectar a possibilidade de falha nos transformadores, é importante levar em consideração a quantidade de descargas parciais, a quantidade de energia e a quantidade de ruído. A quantidade de descargas parciais em conjunto com a energia funciona como elemento indicador de defeito, enquanto que o ruído é um elemento penalizador. Assim, são definidos quatro nós: Classificação, Quantidade de Descargas Parciais, Ruído e Energia, sendo alimentados com as informações do canal com maior atividade (descargas parciais e ruído), pois é imprescindível prever o pior caso.

A maior dificuldade é dispor de uma base de treinamento suficientemente grande e diversificada para a realização de classificações confiáveis, o que pode ser amenizado com um gerador de casos fictícios subsidiado por casos reais. Assim sendo, a utilização de um gerador de casos é útil para a realização de testes e avaliação da RB.

Procedimento 3 – Gerador de Casos

Entrada: Casos reais utilizados para treinamento, quantidade de casos a serem gerados.

Saída: Casos para a base de treinamento.

1. Treine uma Rede Bayesiana D com os casos reais.
Repita enquanto a quantidade desejada de casos não seja atingida:
 2. Crie instâncias v_{ik} das variáveis randômicas u_i que possuem $\pi(u_i) = \{ \}$ utilizando as probabilidades *a priori* dessas variáveis da Rede D .
 3. Instancie as demais variáveis u_i com probabilidades *a posteriori* de D .
-

3.6 Discretização dos Parâmetros da Rede Bayesiana

De modo a estabelecer adequada e automaticamente os intervalos dos parâmetros, este trabalho propõe um método de discretização de variáveis contínuas através da atribuição de valores discretos (limites inferior e superior) a intervalos contínuos.

A discretização de valores contínuos levanta a dificuldade de se estabelecer intervalos de valores que representem adequadamente o problema. Para essa tarefa, a utilização de um otimizador *Hill Climbing* mostra-se apropriada, uma vez que, partindo de um ponto inicial, ajustar-se-ão o número de faixas e seus limites de tal forma que a taxa de classificação correta aumente. A referência [9] descreve uma regra conhecida como “regra do 1/5 de sucesso”, em que é utilizada uma distribuição gaussiana com desvio σ , que ajusta os valores aleatórios gerados. O desvio σ é ajustado a cada k iterações. O Procedimento *Hill Climbing* de otimização dos limites é mostrado a seguir.

Procedimento 4 – *Hill Climbing* para Discretização dos Parâmetros

Entrada: Limites atuais, Número de vizinhos V , Número máximo de iterações, número k de iterações com que σ é ajustado.

Saída: Limites quase-ótimos.

1. Calcule a pontuação dos limites atuais. Coloque esta pontuação em P_{atual} .
Repita enquanto o número de iterações não for máximo:
2. Faça uma cópia dos limites atuais em L_{atual} .
3. Para cada intervalo existente, crie uma cópia, divida em 2 intervalos em proporção aleatória (ajustada por σ) e verifique se a pontuação é melhor que P_{atual} . Se sim, substitua P_{atual} e L_{atual} .
4. Para cada intervalo existente, crie uma cópia, remova o intervalo e verifique sua pontuação. Se melhor ou igual a P_{atual} , substitua P_{atual} e L_{atual} .
5. Faça uma cópia dos limites e V tentativas aleatórias (ajustada por σ) de alteração. Se a pontuação de algum deles for melhor que P_{atual} , substitua P_{atual} e L_{atual} .

6. Substitua os limites atuais por L_{atual} .
7. A cada k iterações, ajuste σ de acordo com a regra do 1/5 de sucesso.

O Procedimento 4 utiliza um método de pontuação para verificar o melhoramento do algoritmo. É possível a utilização de um método de pontuação suave (*soft scoring*) ou de pontuação rígida (*hard scoring*). Na pontuação rígida, conta-se a quantidade de acertos de classificação que a RB provê. Cada acerto corresponde a um ponto. Na pontuação suave, conta-se com um mecanismo de penalização, que subtrai, a cada erro, a porcentagem que falta para o acerto ideal de 100%. Logo, a pontuação suave é menos gulosa. Para os parâmetros referentes às quantidades de DPs e de ruído, assumiram-se os mesmos intervalos, uma vez que possuem a mesma ordem de grandeza e unidade (são contagens feitas a partir da ocorrência de eventos).

4 Metodologia para Estimação de falhas incipientes em Transformadores de Potência

Propõem-se então neste trabalho, aplicando-se em conjunto os procedimentos descritos, uma metodologia para estimativa de falhas incipientes em transformadores de potência.

É descrito na Tabela 2 possíveis classificações que o transformador poderia receber.

Tabela 2 – Possíveis Classificações do estado do Transformador.

Classificação	Avaliação	Ação Recomendada
Ótima	<i>Resultados Satisfatórios</i>	<i>Continuar a operar normalmente.</i>
Boa	<i>Resultados dentro da normalidade para o equipamento</i>	<i>Continuar a operar normalmente estando atento à evolução de DPs nos próximos registros.</i>
Ruim	<i>Resultados Preocupantes</i>	<i>Dar continuidade na investigação e realizar outros ensaios o mais brevemente possível para confirmar medições e tendências de DPs.</i>
Péssima	<i>Urgência</i>	<i>Planejar uma retirada do equipamento de operação em caráter de urgência para uma inspeção interna, localização e correção de defeito.</i>

A metodologia aplicada para estimação é descrita no Procedimento 5.

Procedimento 5 – Metodologia para Estimação

Entrada: Limites atuais, número de vizinhos V , número máximo I_{max} de iterações, número k de iterações com que σ é ajustado, quantidade N_1 de casos para treinamento e quantidade N_2 de casos para validação.

Saída: Limites quase-ótimos.

1. Definem-se limites L para treinamento da RB. Caso existam casos suficientes para treinamento e validação da RB, salte para o passo 4.
2. Cria-se uma RB com amostras reais. (Utiliza-se a de um classificador ingênuo.) Para parâmetros, utiliza-se o método do MLE (subseção 3.2) para aprendê-los.
3. Utiliza-se o Gerador de Casos (Procedimento 3) para criação de N_1 casos para treinamento da RB e criação de casos, que em conjunto com os casos reais, formarão N_2 casos que serão utilizados para validação.
4. Utiliza-se o método da subseção 3.1 para aprender a estrutura da RB. A ordem dos nós definidas para o K2 é: Classificação, Descarga Parcial, Ruído e Energia. A quantidade máxima de pais para K2 é 4. Para o *Hill Climbing*, são considerados todos os nós desconectados.
5. Utiliza-se o método MLE (subseção 3.2) para aprender os parâmetros da RB.
6. Mede-se a quantidade de acertos utilizando os casos para validação.
7. Definem-se os limites de acordo com as Tabelas 3 e 4.
8. Utiliza-se o algoritmo para otimização dos limites (subseção 3.5), utilizando a pontuação rígida e a pontuação suave, utilizando V vizinhos, I_{max} iterações e k iterações para redefinição do valor do desvio σ (regra do 1/5 de sucesso). Repete-se este passo N_1 vezes, a fim de encontrar limites quase-ótimos.
9. Verifica-se o desempenho da rede com os limites criados.

Tabela 3 – Intervalos de Valores de Energia.

Valor	Intervalo
1	$0 - \infty$

Tabela 4 – Intervalos de Valores de Quantidade de DPs e Ruído.

Valor	Intervalo
1	$0 - \infty$

5 Resultados e Discussão

Para se construir o classificador, foram utilizados resultados de seis ensaios realizados transformadores, sendo que, em cada ensaio, foram monitorados dois transformadores simultaneamente. Logo, têm-se 12 amostras para se construir a RB. Foi aplicada a metodologia descrita no Procedimento 5 para a verificação do seu desempenho utilizando os limites L das Tabelas 5 e 6, $N_1 = 100$, $N_2 = 100$, $V = 5$, $I_{max} = 100$, $N_i = 2$ e $k = 5$.

Os limites encontrados através do Procedimento 4, utilizando a pontuação suave e rígida são apresentados nas Tabelas 7 a 10. As linhas correspondem aos índices das faixas e as colunas às iterações do Procedimento 4.

Tabela 5 – Valores discretos para valores de Energia.

Valor	Intervalo
1	$0 - 50.000$
2	$50.001 - 7.000.000$
3	$7.000.001 - \infty$

Tabela 6 – Valores discretos para quantidade de DPs e Ruído.

Valor	Intervalo
1	$0 - 1.000$
2	$1.001 - 5.000$
3	$5.001 - 100.000$
4	$100.001 - \infty$

Tabela 7 – Faixas de Energia ao longo das iterações (suave).

	0	100	200
1	$0 - \infty$	$0 - 50.269$	$0 - 50.269$
2		$50.270 - 1.826.677$	$50.270 - 238.793$
3		$1.826.678 - \infty$	$238.794 - 1.826.677$
4			$1.826.678 - 6.737.032$
5			$6.737.033 - \infty$

Tabela 8 – Faixas das Quantidades de DPs e Ruído ao longo das iterações (suave).

	0	100	200
1	$0 - \infty$	$0 - 10$	$0 - 10$
2		$11 - 5.115$	$11 - 5.115$
3		$5.116 - 729.902$	$5.116 - 729.902$
4		$729.903 - 131.863.007$	$729.903 - 131.863.007$
5		$131.863.008 - \infty$	$131.863.008 - \infty$

Tabela 9 – Faixas de Energia ao longo das iterações (rígida).

	0	100	200
1	$0 - \infty$	$0 - 851.400$	$0 - 321.179$
2		$851.401 - \infty$	$321.180 - \infty$

Tabela 10 – Faixas das Quantidades de DPs e Ruído ao longo das iterações (rígida).

	0	100	200
1	$0 - \infty$	$0 - 17$	$0 - 17$
2		$18 - 4.772$	$18 - 4.772$
3		$4.773 - 257.013$	$4773 - 49.333$
4		$257.014 - \infty$	$49.334 - 257.013$
5			$257.014 - \infty$

A Tabela 11 compara as taxas de acerto da base de validação da Rede Bayesiana utilizando cada limite.

Tabela 11 – Taxa de acerto da Rede Bayesiana.

Limite	Taxa
<i>L</i>	83%
Pontuação Rígida	89%
Pontuação Suave	89%

O procedimento 5 proposto neste trabalho foi executado 60 vezes, metade utilizando pontuação suave e metade utilizando pontuação rígida, sendo o melhor caso detalhado acima. Com a pontuação rígida se obteve uma média de acertos de 86,5% e um desvio padrão de 1,14%. A pontuação suave obteve uma media de 86,57% e um desvio padrão de 1,48%. Na Figura 2 é mostrado o desempenho mínimo, médio e máximo, utilizando a pontuação rígida (esquerda) e suave (direita).

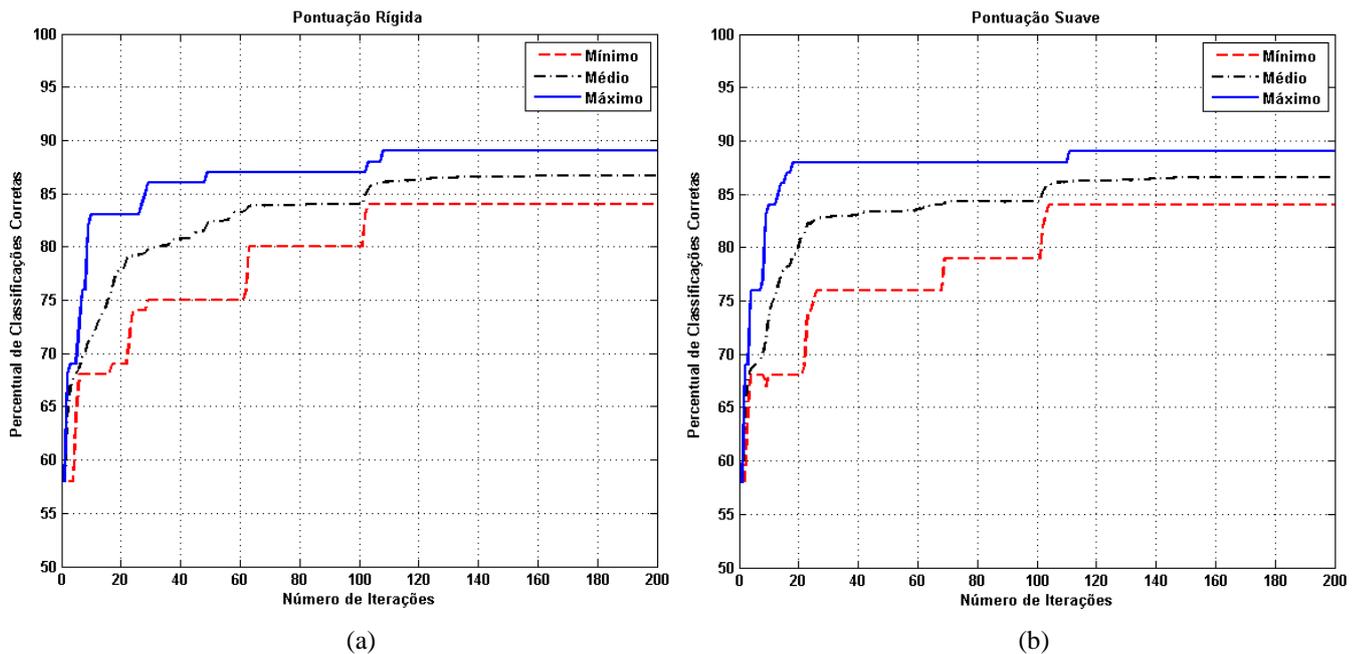


Figura 2 – Desempenho mínimo, médio e máximo dos otimizadores: (a) pontuação rígida e (b) pontuação suave.

Para o propósito de comparação, uma Rede Neural (RN) *Perceptron* de Múltiplas Camadas, [10] e [11], foi implementada. Tomaram-se os valores logarítmicos de base 10 de (Energia, Quantidade de DPs, Ruído) como seu vetor de entrada e a (Classificação), cujo valor numérico é um dentre {4 – Péssimo, 3 – Ruim, 2 – Bom, 1 – Ótimo}, como sua saída. O pré-processamento do vetor de entrada (aplicação de \log_{10} em seus valores) se mostrou eficaz para o aumento da precisão da RN, visto que a ordem de grandeza das entradas do conjunto de treinamento da RN varia consideravelmente. Treinou-se a RN com 100 exemplos, validou-se com mais 50 e testou-se com outros 50. Variou-se o número de neurônios sigmoidais da camada oculta de 1 a 10, executando-se 100 treinamentos independentes para cada configuração da RN, e aplicando-se o algoritmo Levenberg–Marquardt. Obteve-se um máximo de 83% de classificações corretas com 5 neurônios na camada oculta, sendo esta então inferior à abordagem por Rede Bayesiana proposta neste trabalho.

6 Conclusão

Neste artigo, apresentou-se uma metodologia capaz de criar uma Rede Bayesiana adequada à classificação de níveis de falhas incipientes em transformadores de potência. A Rede é capaz de indicar não somente o estado do transformador, mas também qual a porcentagem de crença no resultado, dadas as evidências apresentadas. Foi proposto um método de otimização dos limites dos parâmetros usados pela Rede Bayesiana, em que um exemplo ilustrou a capacidade do algoritmo em aumentar a taxa de classificações corretas da Rede. Foram dadas duas alternativas de pontuação a serem usadas no otimizador: pontuação suave e pontuação rígida. Com a pontuação suave se obteve um desempenho equivalente ao da pontuação rígida. A metodologia proposta foi comparada à uma abordagem por Rede Neural, tendo-se obtido um melhor desempenho por parte da primeira.

Novos ensaios em transformadores de potência estão sendo programados para serem realizados na concessionária de energia, os quais servirão para alimentar a Rede Bayesiana. É possível que, na medida em que forem feitos novos ensaios, a metodologia seja ampliada para, não somente indicar o estado do equipamento, mas também identificar o tipo de defeito no transformador.

7 Agradecimentos

Os autores agradecem à CELG Distribuição (CELG D) por todo apoio ao trabalho e à Capes/MEC pelo suporte financeiro.

8 Referências

- [1] C. H. Azevedo, C. J. Ribeiro, A. P. Marques, Methodology for the Detection of Partial Discharges in Power Transformers using the Acoustic Method, **Eurocon, San Petersburg**, (2009)
- [2] E. Mohammadi, M. Niroomand, M. Rezaeian, Z. Amini, Partial Discharge Localization and Classification Using Acoustic Emission Analysis in Power Transformer, **31st International Telecommunications Energy Conference - INTELEC 2009**, (2009).
- [3] W. H. Tang, Z. Lu, Q. H. Wu, A Bayesian Network Approach to Power System Asset Management for Transformer Dissolved Gas Analysis, **Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies**, (2008).
- [4] X. Hao, S. Cai-xin, Artificial Immune Network Classification Algorithm for Fault Diagnosis of Power Transformer, **IEEE TRANSACTIONS ON POWER DELIVERY**, Vol. 22, No. 2, (2007).
- [5] C. Kuo, Artificial recognition system for defective types of transformer by acoustic emission, **Expert Systems with Applications**, 36 (2009), 10304 - 10311.
- [6] F. V. Jensen, T. D. Nielsen, Bayesian Networks and Decision Graphs. 2nd edition, **Springer**, (2007).
- [7] T. Pansombut, Learning Bayesian Belief Networks. Department of Computer Science, **North Carolina State University**, (2006).
- [8] A. Darwiche, Modeling and Reasoning with Bayesian Networks, **Cambridge University Press**, (2009).
- [9] A. E. Eiben, J. E. Smith, Introduction to Evolutionary Computing, **Springer**, (2003).
- [10] Z. Michalewicz, D. B. Fogel, How To Solve It: Modern Heuristics, **Springer**, (2000).
- [11] S. Haykin, Redes Neurais: Princípios e Prática, **Bookman**, (1999).