

Redes Neurais Aplicadas ao Diagnóstico de Falhas Incipientes em Transformadores de Potência Imersos em Óleos

Sanderson E.U. de Lima¹, Laurinda Lucia N. Reis¹, Juliana Carvalho Oliveira¹, Leandro dos Santos Coelho², Otacilio M. Almeida¹

¹*Universidade Federal do Ceará*

*Laboratório de Automação e Robótica, Departamento de Engenharia Elétrica,
Caixa Postal 6001 - Campus do Pici- Bloco 705 - 60.455-760 Fortaleza - CE - Brasil
(sanderson; laurinda; Juliana; otacilio)@dee.ufc.br*

²*Pontifícia Universidade Católica do Paraná*

*Laboratório de Automação e Sistemas, LAS/PPGEPS/CCET/PUCPR
Rua Imaculada Conceição, 1155, CEP 80215-901, Curitiba, PR, Brasil
E-mail: leandro.coelho@pucpr.br*

Abstract - Incipient fault in power transformer is closely related to isolation condition assessment. Dissolved gas analysis in transformer insulating oil is a well-known diagnostic technique. Produced gases serve as indicators of the type and severity of electrical stress. This paper presents a comparative study of multilayer perceptron and radial basis function neural networks for fault identification in oil-immersed power transformers. The neural networks are trained by using improved learning algorithms. The used data in experiments are taken out chromatographic analysis of COELCE (Electrical Company of Ceara) transformers. The simulation results indicate the potentialities of the tested neural networks for diagnostic in oil-immersed power transformers.

Keywords: Neural Network, Artificial Intelligence, Fault Diagnosis, Power Transformer, Dissolved Gas-in-oil Analysis (DGA).

I. INTRODUÇÃO

O diagnóstico de falhas incipientes em transformadores de potência baseado em normas, tal como a IEC 599 [1, 2], tornou-se popular por causa de sua simplicidade, no entanto, a aplicação desses padrões requer experiências com os procedimentos e métodos e com o histórico do transformador em questão. Atualmente as falhas em equipamentos do sistema elétrico são de aspectos negativos e de grande preocupação para o setor, principalmente aquelas que ocorrem sem nenhum conhecimento prévio e que, por causarem interrupções abruptas no fornecimento de energia, têm conotações econômicas diretas [3, 4].

Desde que todos os fenômenos relacionados ao diagnóstico de falhas em transformadores são caracterizados por imprecisões, incertezas nas medidas e não-linearidades não modeladas, métodos convencionais, combinados com

métodos baseados em inteligência computacional e em especial, os sistemas nebulosos e redes neurais artificiais podem ser empregados de forma eficiente para diagnósticos automáticos de falhas. Na literatura, têm sido apresentados diversos estudos relativos ao uso de técnicas inteligentes para o diagnóstico de falhas em transformadores de potência imersos em óleo. Neste contexto, os modelos matemáticos concebidos baseados em algumas abordagens de redes neurais têm recebido atenção recentemente, tanto da comunidade científica quanto dos profissionais que atuam no meio elétrico-industrial, por tratarem-se de ferramenta de projeto que oferecem soluções promissoras para problemas de identificação e diagnóstico de falhas em transformadores [5]-[13]. Entre as características relevantes das redes neurais têm-se: processamento paralelo, aprendizado, memória associativa e distribuída. Estas características são inspiradas nas redes neurais biológicas, mesmo que rudimentarmente.

Neste trabalho é realizada uma comparação entre duas topologias de sistemas inteligentes baseado em redes neurais para realizar o diagnóstico de falhas incipientes em transformadores de potência através dos dados cromatográficos dos gases dissolvidos no óleo. Na realização dos testes foram utilizados dados reais de transformadores da COELCE (Companhia Energética do Ceará), provenientes de análises cromatográficas que foram realizadas nos anos de 2001 a 2003. As abordagens inteligentes empregadas para este propósito foram uma rede neural perceptron de múltiplas camadas (RN-MLP) e uma rede neural função de base radial (RN-RBF).

O restante do artigo está organizado da seguinte forma. Na seção 2 é apresentada uma metodologia de diagnóstico baseada em redes neurais, além de uma breve descrição das RN-MLP e RN-RBF utilizadas. Na seção 3 e 4 são

apresentados os resultados obtidos e a conclusão do artigo, respectivamente.

II. ANÁLISE DE GASES DISSOLVIDOS EM ÓLEO (DGA)

O mecanismo de formação de gases no interior de transformadores imersos em óleo mineral devido solicitações elétricas e térmicas obedece a um modelo termodinâmico que relaciona a taxa de formação dos gases e a temperatura na vizinhança do ponto onde a falta ocorre. Portanto, uma dada configuração de distribuição de concentrações de gases dissolvidos pode ser associada a uma determinada faixa de temperatura e conseqüentemente a um determinado tipo de falha.

Os principais gases formados a partir da decomposição do óleo mineral de transformadores submetidos à falhas são: hidrogênio (H_2), metano (CH_4), etano (C_2H_6), etileno (C_2H_4) e acetileno (C_2H_2). Todos os métodos de análise desenvolvidos utilizam as razões entre as concentrações dos gases definidas na tabela 1.

TABELA 1. RAZÕES ENTRE AS CONCENTRAÇÕES DOS GASES

razão	$\frac{CH_4}{H_2}$	$\frac{C_2H_2}{C_2H_4}$	$\frac{C_2H_2}{CH_4}$	$\frac{C_2H_6}{C_2H_2}$	$\frac{C_2H_4}{C_2H_6}$
notação	R1	R2	R3	R4	R5

A seguir são detalhadas as abordagens de RN-MLP e RN-RBF utilizadas, neste artigo, nos experimentos.

III. MECANISMO DO DIAGNÓSTICO BASEADO EM REDES NEURAIS

As RNAs utilizadas (MLP e RBF) são redes com propagação adiante (*feedforward*) e totalmente conectadas, como ilustra a figura 1 onde os círculos representam os neurônios e os retângulos denotam as unidades de entradas (x) e saídas (y), enquanto as setas correspondem à função de propagação direta dos sinais. Tais redes realizam um mapeamento não linear entre as entradas e saídas. Tendo em vista que o problema envolve a classificação de vetores n -dimensionais (espaço vetorial de entrada e a polarização) em diferentes classes, neste trabalho é utilizada uma unidade de saída para cada classe, portanto o espaço vetorial de saída é m -dimensional.

Todo projeto de RNAs apresenta duas fases: A primeira fase que corresponde ao processo de treinamento, durante o qual um conjunto de amostras de dados são fornecidas à rede para o ajuste iterativo dos pesos neuronais. A segunda fase corresponde ao processo de testes ou validações segundo algum critério pré-estabelecido.

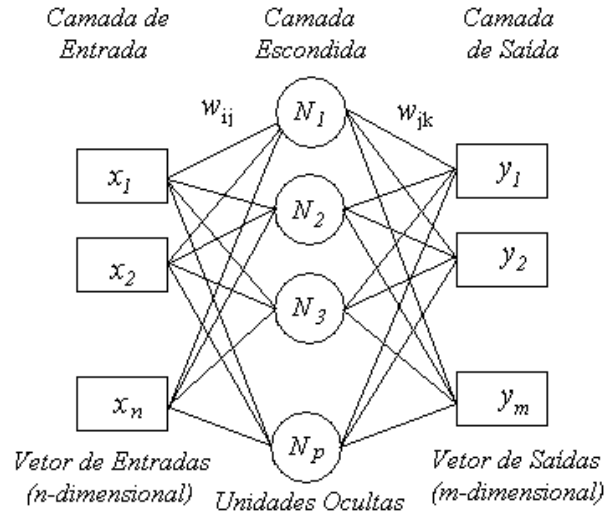


Figura 1 - RNA com propagação adiante (*feedforward*)

IV. A REDE NEURAL PERCEPTRON MULTI-CAMADA (MULTI-LAYER PERCEPTRON - MLP)

A MLP é talvez a rede neural mais popular em aplicações de reconhecimento padrão [9]. As memórias são os pesos (w_{ij} e w_{jk}) entre as camadas, ilustrados na figura 1, o w_{ij} representa o peso que conecta o neurônio i ao j , a relação de entrada-saída de um neurônio j está descrita na equação (1).

$$y_j^{(l)} = \Phi(v_j^{(l)}) = \Phi\left(\sum_{i=0}^p w_{ij}^{(l)} x_{ij}^{(l)}\right) \quad (1)$$

Onde l indica a camada ($l > 0$, a camada de saída é a terceira camada), y_j é a saída do j -ésimo neurônio, v_j é a soma de pesos dos neurônios de entrada, x_{ij} é a i -ésima entrada do neurônio (p entradas da camada anterior e a polarização), w_{ij} é o peso da i -ésima entrada para o j -ésimo neurônio, e $\Phi(\cdot)$ denota a função de ativação do neurônio.

A função de ativação $\Phi(v)$ é uma função não-linear suave e pode ter várias formas, tais como a função logística da equação (2) e a função tangente hiperbólica da equação (3).

$$\Phi(v) = \frac{1}{1 + \exp(-av)} \quad a > 0 \quad \text{e} \quad -\infty < v < \infty \quad (2)$$

$$\Phi(v) = a \tan(bv) \quad (a, b) > 0 \quad (3)$$

O treinamento de uma RN-MLP geralmente usa um algoritmo de retropropagação do erro (*back-propagation*), que consiste de duas etapas – a etapa de alimentação direta (cálculo das saídas dos neurônios e determinação do erro da saída da RN-MLP) e a etapa de retropropagação do erro (*backward*). Na etapa direta, os pesos sinápticos da RN-MLP são fixos e a equação (1) é usada para obtenção das saídas dos neurônios de todas as camadas da RN-MLP. Durante a etapa de *backward*, todos os pesos sinápticos são ajustados de acordo com as seguintes equações de correção de erro:

$$e_j(n) = d_j(n) - y_j(n) \quad (4)$$

$$\delta_j^{(l)} = \begin{cases} e_j^{(l)} \Phi'(v_j^{(l)}(n)) \rightarrow \text{camada de saída} \\ \Phi' \left(v_j^{(l)}(n) \sum_k \delta_k^{(l+1)}(n) w_{kj}^{(l+1)}(n) \right) \rightarrow \text{camada escondida} \end{cases} \quad (5)$$

$$w_{ji}^{(l)}(n+1) = w_{ji}^{(l)}(n) + \alpha w_{ji}^{(l)}(n-1) + \eta \delta_j^{(l)}(n) y_i^{(l-1)}(n) \quad (6)$$

Onde n é o número da iteração de treinamento, e representa o sinal de erro entre a saída da RN-MLP e a saída desejada, d é a saída desejada, $\Phi'(\cdot)$ denota a derivada da função de ativação, η é o parâmetro da taxa de aprendizagem e α constante momento.

Durante o processo de treinamento, as amostras de dados devem ser apresentadas à rede randomicamente. A apresentação de todas amostras de dados à rede uma vez é denominada de “época”. Muitas épocas são geralmente necessárias para treinar uma rede. O treinamento termina quando o quadrado dos erros individuais e/ou a o erro quadrático médio da(s) saída(s) da rede neural é menor do que os valores pré-estabelecidos. Estes erros são definidos como:

$$e_{\text{quadrado}} = \frac{1}{2} [e_j(n)]^2 \quad (7)$$

$$e_{\text{medio}} = \frac{1}{2N} \sum_{j=1}^N [e_j(n)]^2 \quad (8)$$

Para a seleção de uma estrutura de MLP neste estudo foram considerados alguns fatores. Primeiro, o diagnóstico de falta é um problema de mapeamento não-linear complexo, devido ao fato das entradas e saídas serem ambas multivariáveis e não existe nenhuma relação linear obtida até o presente momento. Segundo, uma MLP de três camadas (com apenas uma camada escondida) tem a capacidade de aproximar qualquer função, com determinada precisão [15, 16], independente de sua complexidade. Em geral, as MLPs com mais do que uma camada escondida são mais eficientes. Terceiro, o algoritmo de treinamento retropropagação do erro tem sido aplicado a diversos problemas com sucesso. O treinamento é do tipo supervisionado, pois a cada iteração é conhecida a saída desejada e pode-se calcular o erro. Desta forma, espera-se que uma MLP seja capaz de modelar com suficiente precisão o problema estudado neste artigo.

V. IDENTIFICAÇÃO DA TOPOLOGIA MLP ÓTIMA

Para determinação do vetor ótimo de entradas foram testadas várias combinações de entradas. O desempenho utilizando diretamente as concentrações dos gases foi inferior ao desempenho quando são utilizadas as razões entre as concentrações dos gases (R1, R2 e R5). É adicionado ainda

uma polarização (*bias*) em todos os neurônios. A cada neurônio de saída é associado um padrão de falta (falta térmica de baixa temperatura, falta térmica de alta temperatura, descargas de baixa energia, descargas de alta energia e degradação da celulose).

O desempenho da rede MLP é função do número de camadas escondidas e da quantidade de neurônios presentes nestas camadas. Neste trabalho para determinar a topologia ótima realiza-se uma computação dos números de acertos de cada rede com um conjunto de dados de validação. Foram realizadas diversas simulações de modo a possibilitar a escolha do valor ótimo para o número de camadas escondidas e o número de neurônios destas camadas. Uma MLP com uma única camada escondida foi identificada como a melhor escolha para solução do problema. Portanto, o problema de otimização está parcialmente resolvido, sendo utilizadas MLPs de três entradas, cinco saídas e uma camada escondida.

A função de ativação tangente hiperbólica foi utilizada em todos os neurônios. O critério de parada do treinamento é realizado através do monitoramento do índice de acerto no conjunto de dados de validação para cada época de treinamento. O número máximo de épocas possíveis, neste caso, está limitado em 30000 épocas.

O algoritmo *Back-Propagation* convencional é susceptível aos problemas de velocidade de convergência e aprisionamento em mínimos locais. Para minimizar estes problemas foram utilizadas algumas modificações, tais como: utilização de constante de momento e adaptação da taxa de aprendizagem. Tal algoritmo é conhecido como *Adaptive Back-Propagation*. A seleção da taxa de aprendizagem (η) é importante para determinar a velocidade de convergência do treinamento. Para acelerar o treinamento deve-se manter o passo de aprendizagem o maior valor possível sem comprometer a estabilidade do processo de aprendizagem.

A taxa de aprendizagem inicial adotada foi de 0.08 e a constante de momento de 0.1. A cada época de treinamento a taxa de aprendizagem é atualizada. Se o erro quadrático médio da época atual é superior a 105% do anterior então a taxa de aprendizagem é multiplicada por 0,8 e ainda descartamos os novos valores dos pesos. Se o erro quadrático médio da época atual é inferior ao da anterior então a taxa de aprendizagem é multiplicada por 1,1. Desta forma conseguiu-se uma estratégia de redução da taxa de aprendizagem adequada para o problema em estudo.

VI. REDES DE FUNÇÃO DE BASE RADIAL (RADIAL-BASIS FUNCTION - RBF)

Uma rede RBF consiste de um conjunto de unidades locais (são utilizadas funções Gaussianas) posicionadas no espaço vetorial de entrada. As unidades locais são completamente conectadas às unidades de saída através de conexões ponderadas. As unidades de saída e utilizam o somatório ponderado de suas entradas para aplicar uma

função de ativação e determinar a saída. Esta camada é em geral linear e fornece a resposta da rede ao padrão de entrada.

Desta forma uma rede RBF, em sua forma mais básica, envolve três camadas com papéis totalmente diferentes. A camada de entrada é constituída por nós de fonte (unidades sensoriais) que conectam a rede ao meio ambiente. A segunda camada, a única oculta da rede, aplica uma transformação não-linear do espaço de entrada para o espaço oculto, geralmente o espaço oculto é de alta dimensionalidade. No contexto de uma rede neural, as unidades ocultas fornecem um conjunto de “funções” que constituem uma “base” arbitrária para os padrões de entrada, quando estes são expandidos sobre o espaço oculto, por isso estas funções são chamadas de funções de base radial.

Os centros das funções de base radial são os pesos dos neurônios escondidos. O conjunto de funções de base $G(\cdot)$ são funções não lineares aplicadas à distancia do padrão de entrada aos neurônios da camada escondida. A saída y é uma combinação ponderada das funções de base radial.

$$G(\mathbf{x}, \mathbf{x}_c) = \exp\left(-\frac{|\mathbf{x} - \mathbf{x}_c|^2}{2\sigma^2}\right) \quad (9)$$

$$y_j = \sum_{i=1}^N [w_{i,j} \cdot G(\mathbf{x} - \mathbf{x}_c)] \quad (10)$$

Onde $G(\mathbf{x}, \mathbf{x}_c)$ é a função de Green Gaussiana, $|\mathbf{x} - \mathbf{x}_c|$ é a norma euclidiana do vetor de diferença $(\mathbf{x} - \mathbf{x}_c)$ e \mathbf{x} é o vetor de entrada centrado em \mathbf{x}_c .

VII. IDENTIFICAÇÃO DA TOPOLOGIA RBF ÓTIMA

O projeto de uma RBF constitui em última análise um problema de ajuste de curva em um espaço de alta dimensionalidade. Quando uma rede RBF é usada para realizar uma tarefa complexa de classificação de padrões, o problema é resolvido basicamente pela sua transformação para um espaço de alta dimensionalidade, de forma não-linear. A justificativa por trás deste processo é encontrada no teorema de Cover [14] sobre a separabilidade de padrões.

O treinamento é realizado através da apresentação de um dado vetor de entrada e a correspondente saída desejada. Cada vetor de entrada ativa todas as funções Gaussianas em algum nível. Esses valores de ativação das funções locais são propagados através de conexões ponderadas com as unidades de saída, que, por sua vez, realiza o somatório destes valores ponderados. A comparação entre a saída atual e a saída desejada fornece o erro apresentado pela rede para este padrão em um dado instante do treinamento. Desta forma o treinamento de uma rede RBF é dividido em duas etapas. Uma etapa não supervisionada para a camada escondida das unidades locais. E uma etapa supervisionada para as unidades de saída.

Redes deste tipo oferecem uma interessante alternativa ao uso de MLPs, tendo em vista que podem ser treinadas

muito mais rapidamente. O treinamento envolve a determinação da localização das unidades locais dentro do espaço vetorial de entrada (utilizando um método de *clusterização*) e a determinação dos pesos das unidades de saída (como uma RBF apresenta apenas uma camada ponderada a regra delta é suficiente para determinar os pesos). A parte crítica neste desenvolvimento é a localização e parametrização das unidades locais, bem como, a quantidade necessária para otimizar uma determinada função de custo. Em busca da determinação automática destes valores foram utilizados diversos algoritmos, sendo, que o que apresentou melhores resultados com menor tempo de treinamento foi o *Incremental RBF* [10].

A. Incremental RBF

A rede inicial consiste de duas unidades locais posicionadas de forma aleatória dentro do espaço de entrada. Ambas unidades estão conectadas através de uma “conexão de vizinhança” (*edge*). Esta conexão revela simplesmente o fato de que estas unidades estão próximas uma da outra, portanto, não constitui de modo algum um peso. Cada unidade local está associada a uma função Gaussiana. O valor do desvio padrão da função Gaussiana é calculado através da média aritmética dos comprimentos das conexões de vizinhança de cada unidade. A idéia por trás disto é dar a cada unidade um “campo receptivo” dentro do espaço vetorial de entrada.

Os pesos dos neurônios da camada oculta (“centros” das funções Gaussianas) são atualizados de forma semelhante ao utilizado em uma rede auto-organizada de Kohonen [11]. De modo que deslocamos o vencedor e também os seus vizinhos em direção ao vetor de entrada. Logicamente, os vizinhos são deslocados em um passo muito menor que o vencedor. Os pesos dos neurônios de saída são atualizados através da regra delta.

O erro quadrático médio da saída em cada iteração é atribuído ao neurônio vencedor, a acumulação deste valor é realizada para que se tenha uma variável de atribuição de crédito para os neurônios ocultos. O erro acumulado de cada neurônio oculto é o parâmetro utilizado para decidir a localização de novas unidades dentro da camada oculta.

A inserção de novas unidades é realizada sempre que o erro de validação se estabiliza, mostrando que para a atual configuração da rede a regra delta e a adaptação dos centros não são capazes de melhorar a desempenho da rede. Desta forma uma unidade oculta é inserida entre o neurônio que apresenta o maior erro acumulado e o vizinho deste mais distante. A localização da nova unidade dentro do espaço de entradas é tomada como a média dos pesos das unidades entre as quais ela é inserida. Para evitar inserções sucessivas na mesma região, após a inserção de um neurônio o erro acumulado dos vizinhos ao novo neurônio é reduzido em 50%. O erro acumulado atribuído ao novo neurônio é a média aritmética dos erros acumulados de seus dois vizinhos. Para

ênfátizar a importância dos erros mais recentes todos os erros acumulados são multiplicados por um fator α ($0 < \alpha < 1$) a cada iteração.

A formação da vizinhança é inspirada no método Hebbiano, desde que é realizada uma ligação de vizinhança (*edge*) entre neurônios que apresentam atividades correlatas (proximidade dentro do espaço vetorial de entrada) e são removidas ligações de vizinhança em caso contrário. A cada ligação de vizinhança é atribuída uma variável de idade que assume inicialmente o valor zero quando a ligação é criada e é incrementada sempre que esta ligação está conectada ao neurônio vencedor de uma dada iteração. A cada iteração os neurônios ocultos realizam pequenas movimentações dentro do espaço vetorial de entrada fazendo com que seja necessário atualizar a informação de proximidade e conseqüentemente a vizinhança. A cada iteração são localizados os dois neurônios mais próximos (menores distâncias Euclidianas) do atual vetor de entrada e é formada uma ligação de vizinhança entre estes. Caso a ligação já exista sua variável de idade é colocada em zero. Uma ligação de vizinhança é removida sempre que a variável de idade associada a esta atinge uma idade limite.

VIII. ANÁLISE DOS RESULTADOS

Para realização da etapa de testes e avaliação dos resultados foram coletados dados de transformadores da COELCE (Companhia Energética do Ceará), provenientes de 135 análises cromatográficas que foram realizadas nos anos de 2001, 2002 e 2003. Como exemplo dos dados utilizados são apresentados alguns deles na tabela 2.

TABELA 2. EXEMPLOS DE AMOSTRAS CROMATOGRÁFICAS.

amostra	H ₂	CH ₄	C ₂ H ₄	C ₂ H ₆	C ₂ H ₂
1	2	28	63	0,2	0
2	21	5	13	2	0
3	58	2	0	0	0
4	1	3	11	2	0
5	4	2	8	2	0
6	2	4	6	3	0
7	3	6	11	4	0
8	36	4	3	3	0

Dentre as 135 amostras são encontradas situações de funcionamento normais e falhas incipientes como pode ser observado na tabela 3. Os resultados apresentados no laudo técnico do especialista responsável são a saída desejada das redes neurais consideradas.

TABELA 3. CLASSIFICAÇÃO DOS DADOS UTILIZADOS QUANTO AO TIPO DE FALTA.

tipos de falhas / normal	quantidade
sobreaquecimento no óleo	78
sobreaquecimento na celulose	17

descargas de baixa energia	25
descargas de alta energia	11
normal	4
total	135

O conjunto total de dados com as concentrações de gases dissolvidos no óleo é dividido em dois grupos: conjunto de treinamento e conjunto de validação. O conjunto de treinamento possui 100 amostras e o conjunto de validação tem 35 amostras. A cada simulação as 135 amostras foram divididas nestes dois grupos de forma aleatória com distribuição de probabilidade uniforme, no final foi tomada a média dos resultados de cada simulação.

Na tabela 4 são apresentados os valores médios do percentual de acerto para a RN-MLP e a RN-RBF em função do número de neurônios utilizados na camada escondida.

TABELA 4. PERCENTUAL MÉDIO DE ACERTO DAS REDES NEURAIS.

RN	neurônios na camada escondida	% médio de acerto do diagnóstico
MLP	5	89,64%
RBF	5	70,79%
MLP	25	94,04%
RBF	25	85,36%
MLP	50	94,54%
RBF	50	94,17%

Os melhores resultados apresentados conduzem a níveis de 95% de acerto no diagnóstico. É importante ressaltar que é praticamente impossível atingir o percentual de 100% de acerto, tendo em vista que a seleção dos dados de entrada não garante que o conjunto de treinamento é realmente uma amostra representativa dos cinco padrões de falhas analisadas e ainda que o diagnóstico do especialista (baseado na IEC 599 e na ABNT NBR-7070) está correto para todos os casos.

Ao analisar o desempenho das redes em função do número de neurônios da camada escondida percebe-se que as redes MLPs produzem resultados sensivelmente melhores do que as RBFs, principalmente para um número pequeno de unidades escondidas. É importante observar que o índice de acerto em função do número neurônios da camada escondida é muito mais sensível para as redes RBFs do que para MLPs. Portanto com uma MLP é possível utilizar uma rede de menor complexidade, apenas 10 neurônios na camada escondida, e ainda assim manter a taxa de acerto em torno 94%.

IX. CONCLUSÃO

Os resultados apresentados na utilização de redes neurais aplicadas ao diagnóstico de falhas internas de transformadores apresentaram um percentual elevado de diagnóstico correto. Neste trabalho foram considerados cinco tipos de falhas: térmica de baixa temperatura, térmica de alta temperatura, descargas de baixa energia e descargas de alta energia e

degradação da celulose. A condição de normalidade é indicada pela ausência de falhas.

A implementação de duas topologias de redes para e a comparação de desempenho foi realizada. Tendo em vista que os MLPs constroem aproximações globais de um mapeamento de entrada-saída não linear e as redes RBF, por outro lado, constroem aproximações locais utilizando não-linearidades localizadas com decaimento exponencial (neste trabalho, funções gaussianas) é natural esperar que o MLP requiera um número menor de parâmetros que a rede RBF para o mesmo grau de precisão. Como o critério de seleção de topologia adotado envolve apenas a questão da complexidade da estrutura, concluímos que a RN-MLP com uma camada escondida foi a melhor solução para o problema de classificação de padrões de falhas em transformadores de potência estudado neste artigo.

Foram testados algoritmos de treinamento que apresentam vantagens em relação aos convencionais, tais como: otimização da complexidade da rede (menor número de neurônios), maior capacidade de generalização, busca eficiente do mínimo global e diminuição do tempo de treinamento.

O algoritmo *Adaptive Back-Propagation* consegue reduzir significativamente o número de épocas necessárias para atingir o nível de acerto exigido em comparação ao *Back-Propagation*, tendo em vista que possui um método mais eficiente de atualização da taxa de aprendizagem.

O algoritmo de formação e treinamento da rede RBF (*Incremental RBF*) conduz a redes relativamente pequenas e com boa capacidade de generalização e ainda com um tempo de treinamento menor que os algoritmos convencionais.

Os trabalhos futuros envolverão experimentos relacionados à análise de aspectos de projeto, custo computacional e velocidade de convergência de diversas abordagens de redes neurais e sistemas nebulosos para diagnóstico de falhas em transformadores de potência.

AGRADECIMENTOS

Os autores deste artigo agradecem o suporte financeiro da Companhia Energética do Ceará (COELCE) para este projeto.

REFERÊNCIAS

- [1] IEC 599, *Interpretation of the analysis of gases in transformers and other oil-filled electrical equipment in service*, International Electrotechnical Commission, Geneva, Switzerland, 1978.
- [2] IEC 599, *Revision of IEC 599 – Working Draft 2*, International Electrotechnical Commission, Geneva, Switzerland, 1996.
- [3] S. E. U. L. Sanderson, *Diagnóstico Inteligente de Falhas Incipientes em Transformadores de Potência*, Dissertação de Mestrado, UFC/CT/DEE, Fortaleza-Ce, 2005.
- [4] PROQUIP, *Relatório de divulgação transformadores de distribuição*, Publicação editada pela Eletrobrás, Rio de Janeiro, 1997.
- [5] Q. Su, C. Mi, L. L. Lai e P. Austin, *A fuzzy dissolved gas analysis method for the diagnosis of multiple incipient faults in a transformer*, IEEE Transactions on Power Systems, vol. 15, no. 2, pp. 593-598, 2000.
- [6] H. -C. Kuo, H. -K. Chang e Y. -Z. Wang, *Symbolic evolution-based design of fuzzy-neural transformer diagnostic system*, Electric Power Systems Research, vol. 27, pp. 235-244, 2004.
- [7] L. V. Ganyun, C. Haozhong, Z. Haibao e D. Lixin, *Fault diagnosis of power transformer based on multilayer SVM classifier*, Electric Power Systems Research, vol. 74, pp. 1-7, 2005.
- [8] J. L. Guardado, J. L. Naredo e P. Moreno, *A comparative study of neural network efficiency in power transformers diagnosis using dissolved gas analysis*, IEEE Transactions on Power Delivery, vol. 16, no. 4, pp. 643-647, 2001.
- [9] Y. -C. Huang e C. -M. Huang, *Evolving wavelet network for power transformer condition monitoring*, IEEE Transactions on Power Delivery, vol. 17, no. 2, pp. 412-416, 2002.
- [10] M. -H. Wang, *Extension neural network for power transformer incipient fault diagnosis*, IEE Proc.- Generation, Transmission and Distribution, vol. 150, no. 6, pp. 679-685, 2003.
- [11] M. A. Wang, *Artificial intelligence applications in the diagnosis of power transformer incipient faults*, Doctorad thesis, Blacksburg, Virginia, USA, 2000.
- [12] L. D. S. Araújo, S. F. Gomes e S. R. A. Tenius, *Transformadores autoprotégidos: análise de desempenho*, Enershow, São Paulo-SP, 1998.
- [13] M. Tsuchie, *Recent diagnostic technology on oil-immersed power transformer in Japan*, Mitsubishi Electric Corporation, 2002, pp. 1496-1499.
- [14] S. Haykin, *Neural networks: a comprehensive foundation*, McMillan Inc., Englewood Cliffs, NJ, USA, 2000.
- [15] G. Cybenko, *Approximation by superpositions of a sigmoidal function*, Mathematics of Control Signals and Systems, vol. 2, pp. 303-314, 1989.
- [16] J. T. -H. Lo, *Multilayer perceptrons and radial basis functions are universal robust approximators*, Proceedings of International Joint Conference. on Neural Networks, Anchorage, AK, USA, vol. 2, pp. 1311-1314, 1998.