

ANÁLISE DE SENSIBILIDADE NO USO DAS REDES NEURAIS AUTOASSOCIATIVAS E DA TEORIA DA INFORMAÇÃO MEANSHIFT NO DIAGNÓSTICOS DE FALHAS INCIPIENTES EM TRANSFORMADORES DE POTÊNCIA

SHIGEAKI L. LIMA* OSVALDO R. SAAVEDRA*, VLADIMIRO MIRANDA†

**Av. dos Portugueses, 1966, Bacanga - CEP 65080-805
Cidade Universitária
São Luís, Maranhão, Brasil*

†*Rua Dr. Roberto Frias, 378 - 4200 - 465
INESC Porto, Campus da FEUP
Porto, Portugal*

Emails: sllima@ieee.org, o.saavedra@ieee.org, vmiranda@inescporto.pt

Abstract— Power transformers are important equipments in a substation, from transmission to distribution of electricity to end users. The major importance of this equipment is the ability to transform voltage, which allows the transmission of power over long distances. Costs associated with the acquisition of a new transformer are high, on the order of a few million reais for power transformers in voltage class of 500 kV (100 MVA). The fines imposed by ANEEL are also significant when there is breach of contract, ie, in case of unavailability of the service because the equipment is out of operation for any defects or contingency. Therefore, energy companies have focused on predictive and preventive maintenance in order to keep the equipment in a effective maintenance program, avoiding the occurrence of failures. There are various techniques that are used for diagnosis and failure analysis of power transformers, but little has been discussed about the robustness of these techniques, if the sample being contaminated. This paper presents a sensitivity analysis to test the reliability and robustness of two techniques applied to fault diagnosis, one based on auto-associative neural network and the other through the algorithm Information Theoretic Mean shift. It is shown through the results of the systems robustness of the recognition failure, which in practice represents the impact on minimizing cost related to a wrong evaluation.

Keywords— Sensitivity analysis, Power Transformers, Autoencoder, Information theoretic mean shift.

Resumo— Transformadores de potência são equipamentos importantes em uma subestação, da transmissão, distribuição de energia elétrica aos usuários finais. A principal importância está na capacidade de transformação de tensão (elevadores e abaixadores) o que permite a transmissão de energia em grandes distâncias. Os custos associados a aquisição de um transformador de potência na classe de tensão dos 500 kV (100 MVA) é de alguns milhões. As multas aplicadas pela ANEEL são significativas quando há indisponibilidade do serviço devido o equipamento está fora de operação por algum defeito ou contingência. Logo, as empresas de energia tem se esforçado na manutenção preventiva e preditiva, a fim de manter esses equipamentos dentro de um programa de manutenção efetivo, evitando a ocorrência de falhas. Existem várias técnicas inteligentes utilizadas para o diagnóstico e análise de falhas em transformadores, mas pouco tem se discutido sobre a robustez dessas técnicas, caso a amostra esteja contaminada. Neste trabalho é apresentada uma análise de sensibilidade para testar a confiabilidade e robustez de duas técnicas aplicadas ao diagnóstico de falhas, uma baseada em redes auto-associativas e a outra através do algoritmo de teoria da informação *mean shift*. É mostrado através dos resultados a robustez dos sistemas no reconhecimento da falha, o qual na prática representa o impacto na minimização de custo referentes a uma avaliação errada.

Palavras-chave— Análise de sensibilidade, Transformadores de potência, Redes auto-associativas, Teoria da Informação *Mean shift*.

1 Introdução

O Brasil possui um dos maiores sistemas interligados do Mundo, com milhões de subestações desde a geração à distribuição. O Sistema Interligado Nacional (SIN) é formado pelas empresas do setor elétrico das regiões Sul, Sudeste, Centro-Oeste, Nordeste e parte da região Norte do Brasil, sendo previsto para julho de 2013 a interligação do sistema Norte (Tucuruí-Macapá-Manaus). A interligação apresenta vantagens no que se refere a confiabilidade, flexibilidade, continuidade, segurança e economia. Porém, para o subsistema Norte, poderá ocasionar temporariamente um significativo aumento no índice de perdas elétricas (EPE, 2011).

Estas perdas elétricas representam custos adicionais, perda de faturação e conseqüentemente sobrecarregamento da rede base. Logo, as empresas de energia tem que investir em bons programas de manutenção, supervisão e monitoramento das condições de operação, a fim de mitigar as perdas e aliviar as subestações de potência. Na prática, implica em monitorar amiúde todos os equipamentos que compõem as subestações, dando um maior foco aos transformadores de potência, devido ao alto custo e a indisponibilidade de compra imediata, no caso de falha que leve a perda do equipamento.

De acordo com as estatísticas de falhas apresentados em (Suwanasri et al., 2008), as partes mais frequentemente afetadas são: o OLTC (*on-*

load tap-changer), o tanque (vazamento de óleo), as buchas, o enrolamento e o núcleo, sendo que o OLTC e o núcleo tem uma taxa de falha 40% e 35%, respectivamente. Porém, uma vez que as falhas se manifestam de forma precoce (nos primeiros anos de operação), resultante de estresse elétrico, dielétrico, eletromagnético, térmico ou químico, há a necessidade constante de um bom programa de manutenção que permita evitar falhas de maior gravidade, detectadas pelo surgimento de gases combustíveis (estresse químico).

As técnicas mais utilizadas para a compreensão dessas dinâmicas que ocorrem no interior dos transformadores são a análise físico-química e a cromatografia gasosa. A cromatografia gasosa permite a identificação de conjuntos separados de gases que estão relacionadas com falhas térmicas e elétricas. O diagnóstico através da interpretação dos gases dissolvidos no óleo é feita por normas padronizadas, tais como a IEEE, IEC e NBR, utilizando a técnica de análise de gás dissolvido (AGD).

Embora a AGD seja uma técnica mundialmente aplicada no diagnóstico de falhas em transformadores, à forma retangular descrita nas normas, com uma especificação simplista de intervalos para cada concentração de gás ou razão, usada na fase de interpretação, torna o método incapaz de mapear o espaço topológico de gases. Geralmente, a precisão alcançada no diagnóstico utilizando a norma IEC, com as bases de dados disponíveis é em média de 85%.

Isto motivou uma nova abordagem e a construção de um sistema de diagnóstico preciso com o uso de redes neurais auto-associativas (*autoencoders*), relatado em (Miranda et al., 2012). Ainda em (Miranda et al., 2012) devido ao pequeno número de amostras de falhas para validação das redes, obtidas através de normas específicas e bases de dados disponibilizadas em alguns artigos na área, foi proposto um novo método chamado de truque de densificação, baseado na teoria da informação *mean shift*, com a finalidade de compensar a existência de dados escassos.

Através do algoritmo *mean shift* é possível gerar dados virtuais e assim preencher conjuntos que possuem um pequeno número de elementos representativos. Más além, a capacidade do algoritmo *mean shift* em capturar a estrutura dominante de cada um dos conjunto de amostras de falhas, possibilitou, assim como no caso das redes auto-associativas em um arranjo competitivo, ser uma abordagem bem sucedida para o diagnóstico de falhas, atingindo um nível de 100% de precisão (Lima et al., 2013).

No entanto, apesar deste sucesso com as redes auto-associativas e com o *mean shift*, o trabalho está longe de estar completo, porque a base de dados é formada por um conjunto padrão de amostras pontuais, dando apenas uma imagem estática

do transformador. Logo, com o intuito de avaliar a robustez dos métodos propostos em (Miranda et al., 2012) e (Lima et al., 2013), foi introduzido um erro de distribuição uniforme (que representa uma contaminação do óleo) nas amostras de teste para pontuar a precisão dos sistemas de diagnósticos.

Neste trabalho pretende-se avaliar através de uma análise de sensibilidade, o percentual de acerto obtidos pelas redes auto-associativas e pelo *mean shift* quando da ocorrência de perturbações que modifiquem as amostras de gases, o qual em um sistema convencional para a análise de falhas estaria suscetível ao erro, mostrando a robustez dos métodos.

O artigo é organizado como segue: no tópico 2 e 3 os resultados com a aplicação das redes neurais auto-associativas e o algoritmo *mean shift* no diagnóstico de falhas. No tópico 4 a análise de sensibilidade aplicada aos dois sistemas e no tópico 5 as conclusões acerca do trabalho.

2 Diagnóstico com redes auto-associativas

O método desenvolvido para o diagnóstico de falhas incipientes baseado na AGD com as redes auto-associativas, tem por definição que todas as redes devem ser treinadas para armazenar na matriz de pesos as características de uma falha em particular. Foram consideradas sete redes, sendo duas para condição normal com e sem OLTC e cinco para os outros tipos de falhas. Todas as sete redes são concorrentes (as entradas em paralelo) competindo entre si, conforme a Figura 1.

Em uma arquitetura concorrentes, quando as redes são ativadas por uma nova amostra (não classificada), apenas a rede que estiver melhor sintonizada apresentará um menor erro entre entrada e saída em relação as demais. Isto permite identificar a qual conjunto (falha ou normal) pertence o novo vetor e, assim, fazer a classificação de forma correta.

A arquitetura do sistema proposto utiliza três neurônios nas camadas de entrada e saída e 15 neurônios na camada intermediária. A função de custo para o cálculo do erro e atualização dos pesos é realizada através do método clássico (*Mean squared error - MSE*). A função de ativação entre a camada de entrada e a camada escondida é sigmóide e linear para a camada de saída. O algoritmo de treinamento é o *Levenberg-Marquardt* implementado em Matlab, com o cuidado de evitar cair em soluções triviais.

Na Tabela 1 estão os resultados da validação com os dados reais e a comparação com a aplicação da norma IEC 60559. Pode-se observar que todos os dados reais foram classificados de forma correta, pois foram obtidos 100 % de acerto com todos os dados, incluindo os casos para transfor-

madores em estado normal. Não houve erros ou má classificação dos dados (352 acertos em 352 casos), sendo que o uso do truque de densificação permitiu que todos os casos reais fossem absolutamente testados.

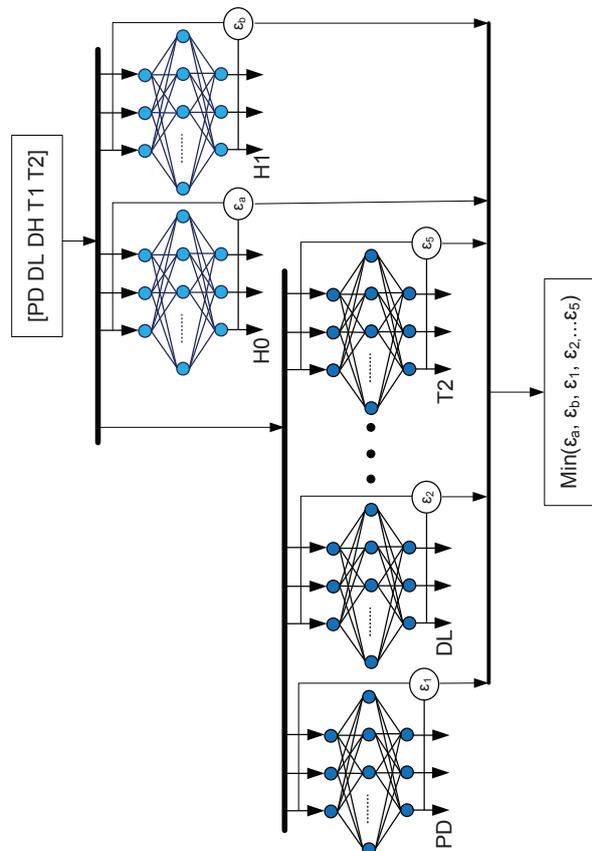


Figura 1: Arquitetura completa do sistema de diagnóstico com competição entre os *autoencoder*

Tabela 1: comparação da precisão do diagnóstico na discriminação do tipo de falha

| Modelo | Treino (virtual) | Teste (real) | Não classificado |
|--------------------|------------------|--------------|------------------|
| Miranda et al,2012 | 100% | 100% | 0 |
| IEC | 93 % | 95.28 % | 85 |

O resultado comparativo da aplicação da IEC 60559 indica que o conjunto de validação não era fácil de diagnosticar e a forma retangular (como é induzido pela aplicação da norma) não é conveniente para representar os conjuntos de falhas.

O método proposto apresenta superioridade em relação à IEC, uma vez que o modelo do *autoencoder* foi capaz de identificar corretamente todos os casos considerados indecisos. O resultado ótimo de 100 % pode ser explicado pela capacidade do *autoencoder* em aprender um específico *manifold* de um conjunto de dados.

3 Diagnóstico com *mean shift*

A teoria da informação *mean shift* pode ser usado na forma de um algoritmo de classificação de dados, separando a base de dado em subconjuntos de falhas. Para isso utiliza duas metodologias: o GBMS (*Gaussian Blurring Mean Shift*) e o GMS (*Gaussian Mean Shift*).

Essas metodologias estão compiladas em (Miranda et al., 2012) e (Lima et al., 2013), sendo implementadas com a variação do parâmetro λ , que faz o *trade-off* entre as duas metodologias. Em síntese, quando $\lambda = 1$ o algoritmo gera as modas da função densidade de probabilidade (fdp), uma vez que a força de informação faz com que as partículas (conjunto de dados) sejam atraídas para a região de alta densidade.

Baseado na geração das modas que representam o conjunto de dados, o *mean shift* é aplicado a cada um dos conjuntos de falhas, sendo gerada uma moda para cada conjunto. Desta forma, quando uma nova informação é adicionada ao conjunto, este novo ponto deve migrar em direção a moda mais similar. Logo, dessa forma, pode-se fazer o diagnóstico de falhas.

O diagnóstico é realizado através da comparação de distâncias (erro médio absoluto) entre a nova informação e a nuvem de pontos gerados pelo *mean shift*. Esta nuvem de pontos é obtida com a variação do σ no intervalo entre [0.01 0.1] a cada iteração (Figura 2).

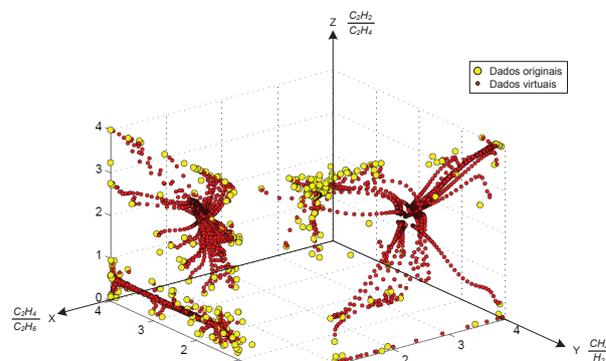


Figura 2: Geração de dados virtuais descobrindo os conjuntos fundamentais para cada modo de falha.

Matematicamente, o primeiro passo é o cálculo do menor valor médio absoluto de X_i em (1), para determinar a menor distância entre um ponto e um dos conjuntos. Em seguida, é calculado o mínimo dos mínimos da distância absoluta em (2), para determinar qual dos conjuntos está mais próximo do ponto.

$$K_i = \frac{\sum_{j=1}^n |x_j - y_i|}{n} \quad (1)$$

$$X_j = \min(\min(K_i)) \quad (2)$$

As equações supracitadas (1) e (2) são usadas para determinar a distância dos cinco pontos em relação à nova informação (representando cada conjunto de falha) e então é selecionada a menor dos cinco. O conjunto que apresentar o menor erro (menor distância) será considerado como aquele ao qual o ponto pertence.

Na Figura 3 é ilustrado um caso onde a distância entre o ponto em amarelo (nova informação) e os outros cinco pontos (o mais perto de cada conjunto) é 0.0011, o qual representa uma falha térmica do tipo ($T < 300^\circ \text{C}$).

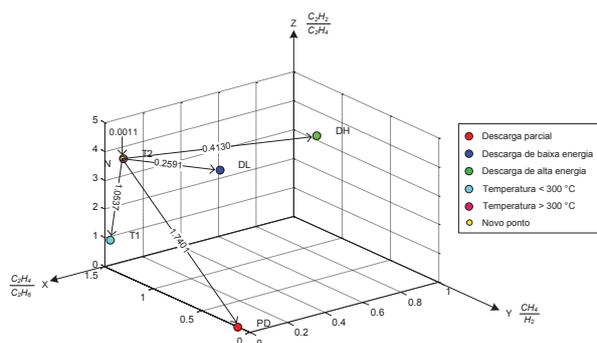


Figura 3: Distância entre o novo ponto de falha e um conjunto representado pela moda da fdp.

Na Tabela 2 é apresentado o resultado obtido após o treinamento, e a discriminação da condição normal e de falha. Pode ser observado que foi alcançada a capacidade de discriminação, uma vez que não há respostas falsas e 100 % de precisão foi alcançada.

Tabela 2: precisão do diagnóstico

| Modelo | Conjunto | | Não classificado |
|---------------------|----------|-------|------------------|
| | treino | teste | |
| Miranda et al, 2012 | 100% | 100% | 0 |
| Lima et al, 2013 | 98 % | 98 % | 2 |
| | 100 % | 100 % | 0 |

4 Análise de sensibilidade

Com os dois métodos de diagnóstico de falhas incipientes através de redes auto-associativas e com o *mean shift*, foi possível obter uma boa discriminação dos conjuntos com acerto de 100 %. Contudo, para avaliar a robustez dos métodos quanto ao diagnóstico correto, fez-se uma análise de sensibilidade aplicando uma perturbação no conjunto de amostras utilizado.

Para verificar a robustez de identificação e a capacidade de aprendizagem das redes, foi introduzido um ruído de distribuição uniforme em todo

o conjunto de dados, inicialmente variando entre $\pm 1\%$ e $\pm 5\%$ e, posteriormente, entre $\pm 5\%$ a $\pm 20\%$.

O objetivo foi provar que mesmo se houvesse uma alteração na amostra de dados, devido a erros de medição, contaminação das amostras, entre outros fatores, o sistema conseguiria ter uma boa taxa de acerto.

Na Figura 4 é apresentado o uso de um ruído uniforme com diferentes larguras, no qual foi replicado 20 vezes para cada uma das amostras, contaminando cada componente com o vetor de ruído. Observar-se que para o maior nível de ruído ($\pm 20\%$) a precisão se manteve acima de 90 %, um valor que está acima de muito dos resultados publicados na área e que foram comentados em (Miranda et al., 2012).

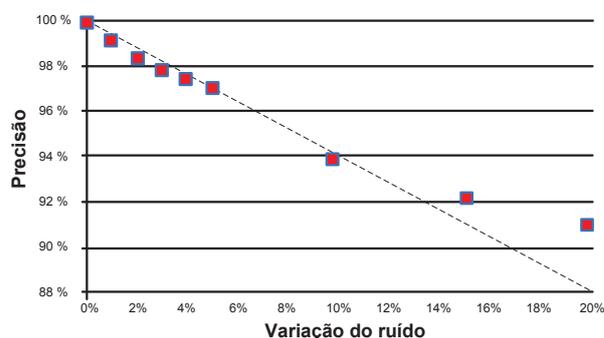


Figura 4: Sensibilidade da AANN na presença de ruído.

O mesmo procedimento aplicado as redes auto-associativas foi replicado no *mean shift*, considerando a base de dados com um ruído variando entre $\pm 1\%$ a $\pm 20\%$.

Os resultados mostrados através da Figura 5 indicam uma similaridade com as redes auto-associativas, no que diz respeito a capacidade de classificar a falha para um ruído entre $\pm 1\%$ e $\pm 5\%$, porém a precisão de acerto se mantém acima dos 98 % em oposição as redes que foi de 96 %.

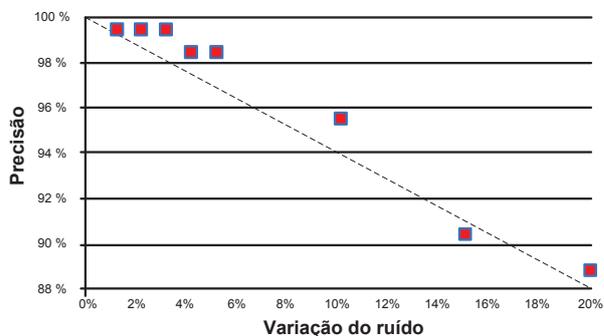


Figura 5: Sensibilidade do *mean shift* na presença de ruído.

Contudo, a medida que o ruído cresce acima dos 15 %, o *mean shift* não consegue manter a precisão compatível com as redes, possivelmente por

que os limites de fronteiras que foram definidos na determinação da moda de cada conjunto de falha, já ultrapassam as fronteiras dos outros conjuntos.

A taxa de acerto na maioria dos trabalhos citados em (Miranda et al., 2012) está acima de 90 %, porém, recentemente, outros trabalhos que tratam do diagnóstico com AGD também obtiveram 100 % de acerto (Isa and Mamat, 2011). Contudo, não foram consideradas análises de sensibilidade para testar a robustez ao ruído.

Os resultados apresentados neste trabalho adicionam confiabilidade e robustez aos sistemas desenvolvidos, tanto com o uso das redes auto-associativas quanto com o *mean shift*, de modo que, mesmo em caso de contaminação das amostras, ambos sistemas ainda conseguem identificar a falha com um bom nível de acerto.

A precisão de acerto comentados em (Miranda et al., 2012) é relativa, pois a única base de dados padronizada é a IEC TC 10, a qual torna-se pequena para o treinamento com redes neurais, dando liberdade para que sejam usadas outras bases de dados de uso restrito. Logo, os resultados (precisão) apresentados em um trabalho não podem ser diretamente comparados a outro.

O uso da análise de sensibilidade para indicar o nível de robustez, torna-se interessante, pois permite que o modelo em análise possa ser indiretamente comparado com os demais trabalhos com precisão acima de 90 %. Considerando que mesmo em caso de variações de ± 20 % os resultados obtidos com o uso das redes e com o *mean shift* tem sido satisfatórios.

5 Conclusão

O uso das redes auto-associativas e do algoritmo *mean shift* como ferramenta para o diagnóstico de falhas em transformadores de potência referenciados ao longo do trabalho, obtiveram um ótimo resultado com 100 % de precisão de acerto.

Contudo, a base de dados utilizada é formada em parte por dados obtidos de outros trabalhos existentes na área e pela base de dados normatizada IEC TC10. Isto, de certa forma, impede a comparação direta com outros trabalhos e, portanto, a confirmação da capacidade de generalização do método.

A análise de sensibilidade foi introduzida neste trabalho, como um método de avaliar a robustez dos sistemas propostos (rede e *mean shift*) na determinação da falha. Permitindo analisar até que grau os sistemas se mantêm confiáveis, mesmo com a contaminação das amostras.

Uma vez que não é possível comparar diretamente todos os trabalhos existentes na área, a análise de sensibilidade indica uma margem de erro do sistema de diagnóstico. Esta margem de erro, com a introdução de um ruído de até ± 20 % foi menor que 10 %, ou seja, a taxa de acerto (maior que 90

%) é superior a muitos dos trabalhos já publicados e que não foram submetidos ao teste de robustez.

Logo, a classificação e a identificação das falhas para ambos sistemas, mesmo na presença de um ruído é visto com boas perspectivas e demonstram a capacidade dos sistemas em capturar a informação predominante em um conjunto de dados.

Na prática, representa a economia de alguns milhões para a empresa, devido ao impacto de uma avaliação errada que leve a perda do equipamento e indisponibilidade do serviço.

Agradecimentos

Shigeaki L. Lima agradece ao CNPq pelo suporte financeiro.

Referências

- EPE (2011). Empresa de pesquisa energética, Projeção da demanda de energia elétrica para os próximos 10 anos.
- Isa, N. and Mamat, W. (2011). Clustered-hybrid multilayer perceptron network for pattern recognition application, *Applied Soft Computing Journal*, Vol. 11, pp. 1457–1466.
- Lima, S., Saavedra Mendez, O. and Miranda, V. (2013). Information theoretic mean shift compensating for scarce data in fault diagnosis in power transformers, *PowerTech, 2013 IEEE Grenoble*, pp. 1–6.
- Miranda, V., Castro, A. and Lima, S. (2012). Diagnosing faults in power transformers with autoassociative neural networks and mean shift, *Power Delivery, IEEE Transactions on* **27**(3): 1350–1357.
- Suwanasri, T., Chaidee, E. and Adsoongnoen, C. (2008). Failure statistics and power transformer condition evaluation by dissolved gas analysis technique, *Condition Monitoring and Diagnosis, 2008. CMD 2008. International Conference on*, pp. 492–496.