



**XXII SNTPEE
SEMINÁRIO NACIONAL
DE PRODUÇÃO E
TRANSMISSÃO DE
ENERGIA ELÉTRICA**

BR/GTM/18
13 a 16 de Outubro de 2013
Brasília - DF

GRUPO -XIII

**GRUPO DE ESTUDO DE TRANSFORMADORES, REATORES, MATERIAIS E TECNOLOGIAS EMERGENTES -
GTM**

**ALGORITMO MEAN SHIFT APLICADO A CARACTERIZAÇÃO DE CLUSTERS PARA O DIAGNÓSTICO DE
FALHAS EM TRANSFORMADORES DE POTÊNCIA**

**Shigeaki L. Lima(*)Vladimiro MirandaOsvaldo R. Saavedra
UFMAINESCPORTO UFMA**

RESUMO

O diagnóstico através da interpretação dos gases chave é feito a partir de Normas padronizadas, como a IEEE, IEC, NBR que usam a técnica da Análise dos Gases Dissolvidos (AGD) e conseguem relacionar causa-efeito. A desvantagem dessa técnica é devida delimitação de fronteiras, quando um gás passa a pertencer a mais de um conjunto de falha, gerando o resultado de erro ou de não classificado. A incapacidade dos métodos descritos nas Normas em mapear o espaço topológico dos gases, se deve a forma retangular como os conjuntos de falhas são interpretados. Normalmente, a utilização das relações entre gases permite um acerto em média de 80 %. Neste trabalho pretende-se apresentar uma possibilidade para definição dos contornos e das fronteiras entre gases de falhas representados por *clusters*. Para isso é utilizado o algoritmo *mean shift* baseado em teoria da informação, como um método que permite mapear o espaço topológico dos conjuntos de falhas, por meio de certos critérios de entropia e do kernel gaussiano. O algoritmo *mean shift* gera a partir de uma base de dados (IEC TC10 entre outras), uma nuvem de pontos (dados virtuais) com as mesmas características dos conjuntos de falha originais, e isto é possível, pois minimiza entropia do conjunto enquanto mantém a similaridades entre os dados originais e os virtuais (comparando as funções densidades de probabilidade estimadas). A proposta pode ser aplicada as normas IEE, IEC e NBR possibilitando uma taxa de acerto bem superior à média dos 80 %, sendo os resultados validados com a comparação entre o modelo proposto e a norma IEC 60599.

PALAVRAS-CHAVE

Transformadores de potência, *Information Teoretic Mean Shift*, AGD, Classificação, Entropia.

1.0 - INTRODUÇÃO

Transformadores de potência são os principais equipamentos em um sistema de transmissão e distribuição de energia. Estes equipamentos são caros e não estão disponíveis para compra imediata, logo, as empresas de energia tem investido em bons programas de manutenção, supervisão e monitoramento das condições de operação, a fim de identificar falhas incipientes e evitar a ocorrência de falhas.

Em muitos casos, as falhas em um transformadores se manifestam de forma precoce (nos primeiros anos de operação), resultante de em algum tipo de estresse elétrico, dielétrico, eletromagnético, térmico ou químico. De acordo com as estatísticas de falhas, as partes do equipamento que são 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 (1).

No entanto, na maioria dos casos, as causas de falhas que levam à explosão do equipamento estão ligados à presença de gases combustíveis (estresse químico), formadas pela degradação do óleo a partir de descargas parciais de baixa energia e de alta intensidade.

(*) Cidade Universitária, Av. dos Portugueses, n° 1966, Bacanga – CEP 65.080-805 São Luís, MA, – Brasil
Tel: (+55 98) 3272-9202 – Fax: (+55 98) 3272-8243 – Email: sllima@ieee.org

As técnicas mais utilizadas para a compreensão dessas dinâmicas 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), o que pode, em certa medida, relacionar causa e efeito. As desvantagens desta técnica estão relacionadas com os percentuais de gases usados para indicar o tipo de falha e a delimitação de fronteiras entre os conjuntos de falhas que não são claras e as normas existentes não preenchem todo o espaço de possíveis relações entre gases, gerando ambiguidade ou erro ou mesmo casos não classificados.

A incapacidade dos métodos descritos nas normas no mapeamento do espaço topológico de gases é também devido à forma retangular, induzida pela especificação simplista de intervalos para cada concentração de gás ou razão, usada na fase de interpretação. Geralmente, a precisão alcançada no diagnóstico utilizando a norma IEC, com as bases de dados disponíveis é em média de 85%. Isto motivou para uma nova abordagem e a construção de um sistema de diagnóstico preciso, em parte relatado em (2).

Em (2) um novo método foi proposto pela primeira vez, com a finalidade de compensar a existência de dados escassos, sendo chamado de truque de densificação e foi desenvolvido através da teoria da informação *mean shift* (3). Através do algoritmo *mean shift* é possível gerar dados virtuais e assim preencher conjuntos que possuem um pequeno número de elementos representativos.

Também em (2), a aplicação de um conjunto de redes neurais autoassociativas (*autoencoders*) em um arranjo competitivo, provou ser uma abordagem bem sucedida, atingindo um nível de 100% de precisão. No entanto, apesar deste sucesso, 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. Isto deriva, do modo tradicional de coleta de amostras de óleo. Assim, o envelhecimento natural do transformador não é levado em consideração.

Além disso, uma hipótese comum na construção de sistemas de diagnóstico de falhas incipientes é que apenas uma falha está presente. Logo, quando mais de um incidente causa a contaminação do óleo, os métodos existentes estão propensos a errar no diagnóstico.

Este trabalho apresenta uma exploração mais profunda das propriedades do algoritmo *meanshift*, ou seja, na determinação dos contornos e limites dos conjuntos de falhas, bem com outras propriedades dos clusters, tais como as modas. Isto permite pesquisar as propriedades de um novo algoritmo para o diagnóstico de falhas, o qual será descritos e apresentados.

As redes autoassociativas competitivas em (2) e o novo método, obtiveram a mesma precisão. Mas, como será apresentado, existem vantagens a favor do novo diagnóstico baseado no *mean shift*, e isto será discutido nas seções seguintes.

2.0 - METODOLOGIA

A aplicação do algoritmo *mean shift* é dividido em duas fases: a primeira determinar o contorno e as fronteiras dos conjuntos de falhas e a segunda aplicar o truque de densificação para gerar dados suficientes para treinar a rede neural autoassociativa e usar os dados reais para validar os resultados.

A determinação dos contornos, fronteiras e o truque de densificação é obtido através do algoritmo *mean shift* através de um processo de clusterização baseado em um método não paramétrico para a estimação da função densidade de probabilidade (unimodal ou multimodal) a partir dos dados existentes.

A função de densidade de probabilidade de uma variável aleatória, pode ser estimada usando a técnica de janelas e Parzen um *kernel* gaussiano. O algoritmo *mean shift* foi originalmente desenvolvido para encontrar as modas de uma função de densidade induzida por uma amostra discreta. A primeira conceituação demonstrou ser incorreta e instável, mas o modelo foi aperfeiçoado através da teoria da informação. Estas modificações permitiram um ajuste preciso entre a entropia dos novos pontos virtuais gerados e a partir da medida de separação baseada na distância de *Cauchy-Schwarz*, outras propriedades dos clusters podem ser encontradas.

Uma aplicação do algoritmo *mean shift* com a base dados de amostras de AGD (incluindo IEC TC10 (3)) para gerar uma nuvem de pontos (dados virtuais) que compartilham algumas propriedades estatísticas do conjunto original, o qual representa uma falha específica. Durante o processo de convergência para a moda, um conjunto de pontos intermediários é gerado ao longo de R^3 . A entropia do conjunto é minimizada enquanto mantém a similaridade entre o conjunto original e os novos pontos (comparação da estimação da função densidade de probabilidade). Esta nuvem de pontos permite observar as características de cada conjunto e a forma do *manifold* de dados (Figura 1).

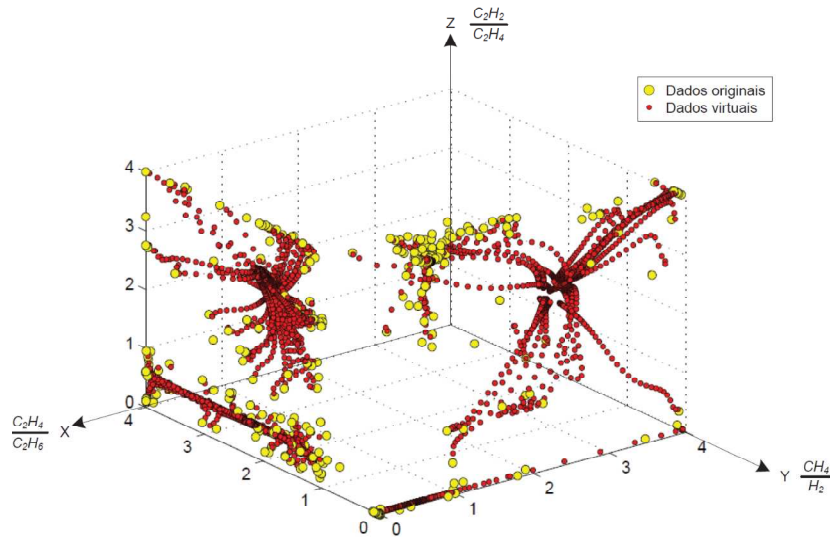


FIGURA 1 – Geração de dados virtuais descobrindo os *clusters* fundamentais para cada modo de falha. Os eixos representam proporções de gases dissolvidos de acordo com a norma IEC 60599

3.0 - TEORIA DA INFORMAÇÃO *MEAN SHIFT* E DENSIFICAÇÃO DE *CLUSTERS*

O algoritmo baseado na teoria da informação *mean shift* (5)(6) foi proposto como uma forma de capturar a estrutura dominante de um conjunto de dados, a partir da estimação da função densidade de probabilidade (fdp). Para explicar como a estrutura é capturada, alguns conceitos de entropia e distância devem ser apresentados.

A entropia quadrática de Renyis (7) para uma fdp é definida como:

$$H(X) = -\log \int_{-\infty}^{+\infty} p^2(x) dx \quad 1$$

e a fdp $p(X)$ pode ser estimada pela técnica de Janela de Parzen (8)

$$\hat{p}(X) = \frac{1}{N} \sum_{i=1}^N G_{\sigma}(x - x_i) \quad 2$$

no qual $G_{\sigma}(t) = e^{-\frac{t^2}{2\sigma^2}}$ é um kernel Gaussiano com $\sigma > 0$. Substituindo 2 em 1 tem-se:

$$H(X) = -\log V(X) \quad 3$$

$$V(X) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N G_{\sigma'}(x_i - x_j) \quad 4$$

com

tendo $\sigma' = \sqrt{2\sigma}$. $V(X)$ é chamado de informação potencia da fdp $p(x)$. A derivada desta expressão com relação a um único ponto x_i fornece a força de informação exercida por todas as partículas em x_i (9)(10)(11).

A entropia cruzada entre duas fdp é definida da forma:

$$H(X, X_0) = -\log V(X, X_0) \quad 5$$

$$V(X, X_0) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N G_{\sigma'}(x_i - x_{0j})$$

com

6

A distância de Cauchy-Schwartz medida entre duas fdp definida por d_{CS} é:

$$D_{CS}(X, X_0) = \log \left(\frac{\left(\int p^2(x) dx \right) \left(\int q^2(x) dx \right)}{\left(\int p(x)q(x) dx \right)} \right) \quad 7$$

$$D_{CS}(X, X_0) = -[H(X) + H(X_0) - 2H(X, X_0)] \quad 8$$

O objetivo do algoritmo baseado na teoria da informação *mean shift* é encontrar um conjunto de dados X que capture a informação estrutural de um conjunto X_0 . Isto é obtido pelo duplo critério de otimização, minimizando a entropia de X enquanto mantém a distância de *Cauchy-Schwartz* de algum valor k . Esta formulação de otimização sem restrições, é apresentada a seguir, no qual o parâmetro λ representa o *trade-off* entre os dois objetivos.

$$J(X) = \min H(X) + \lambda [D_{CS}(X, X_0) - k] \quad 9$$

Diferenciando $J(X)$ com relação a cada x é dada uma regra que permite transformar o conjunto X_0 dentro de outro conjunto em cada iteração $t+1$, fazendo uso da informação contida na fdp de X na iteração t , estimada por 2:

$$x_i^{t+1} = \frac{c_1 \sum_{j=1}^N G_{\sigma'}(\|x_i^t - x_j^t\|) x_j^t + c_2 \sum_{j=1}^N G_{\sigma'}(\|x_i^t - x_{0j}\|) x_{0j}}{c_1 \sum_{j=1}^N G_{\sigma'}(\|x_i^t - x_j^t\|) + c_2 \sum_{j=1}^N G_{\sigma'}(\|x_i^t - x_{0j}\|)} \quad 10$$

$$\text{no qual } c_1 = \frac{(1-\lambda)}{V(X)} \text{ e } c_2 = \frac{2\lambda}{V(X, X_0)}$$

Esta formulação permite que partículas de informação, em um campo de informação potencial, moverem-se sob a influência das forças de informação, tais como mencionado em 4.

Fazendo $\lambda = 1$ em 10, o algoritmo converge para as modas da fdp $p(X)$. Quando λ é maior que 1 e menor que 2, o algoritmo tende a fazer que os pontos x converjam para curva principal do conjunto de dados, e para valores de λ maiores que 2, os pontos são concentrados em torno da região de maior densidade da fdp. Além disso, em cada iteração os pontos x_i^t descrevem uma fdp $p(X^t)$ que retêm a informação de $p(X_0)$. Logo, cada ponto x_i^t ao longo da iteração t descreve um caminho de x_{i0} em direção a moda da fdp $p(X_0)$, ou para a curva principal do cluster de dados, ou para a região de alta densidade, dependendo do valor de λ adotado. Desta forma, uma sucessão de pontos $X_0, X^1, \dots, X^t, \dots$ podem se aproximar ou se afastar da moda, ou seja, seguir em direção da força de informação – $\partial V / \partial X$ como em 4 – ou em direção contrária.

Esta propriedade é usada neste trabalho como um novo caminho para densificar os *clusters* X_0 com a descoberta em cada iteração, de pontos intermediários, que podem ser usados como novos dados (pontos virtuais) compatíveis com a fdp original.

Com a teoria da informação *mean shift*, o conjunto de treino pode ser composto somente de dados virtuais, mantendo a totalidade dos dados reais para a fase de teste. Isto aumenta a robustez do procedimento de teste e a precisão dos resultados, como foi relatado em (2).

4.0 - DIAGNÓSTICO ATRAVÉS DA TEORIA DA INFORMAÇÃO MEAN SHIFT

O algoritmo *mean shift* pode ser usado como ferramenta para a clusterização de dados. Isto é realizado baseado nos modelos existentes na literatura, como o GBMS (*Gaussian Blurring Mean Shift*) e o GMS (*Gaussian Mean Shift*). Ambos modelos, estão compilados dentro da equação 10, implementados com a mudança do parâmetro λ (*trade-off*). Logo, quando λ é "1" o algoritmo gera as modas da fdp, uma vez que a força de informação faz com que as partículas (conjunto de dados) sejam atraídas para a região de alta densidade.

Com base neste conceito, o *mean shift* é aplicado a cada um dos conjuntos de falhas, sendo uma moda gerada para cada conjunto. Desta forma, quando uma nova informação é apresentada, o novo ponto deve migrar em direção a moda mais similar.

A contribuição deste trabalho é fazer o diagnostico de falhas com AGD baseado somente no algoritmo *mean shift*, sem que exista a necessidade de treinamento ou a construção de uma rede neural.

O diagnóstico é feito com a comparação da distância (erro médio absoluto) entre a nova informação e a nuvem de pontos gerados pelo algoritmo. Esta nuvem de pontos é obtida com a variação do σ no intervalo entre [0.01 0.1] em cada iteração (Figura 1).

Após a formação da nuvem para cada falha, é verificada a proximidade entre o novo ponto e cada um dos pontos dos conjuntos de falha. A proximidade entre um ponto e um dado conjunto, indicará o grau de pertinência àquele conjunto de falha (Figura. 2).

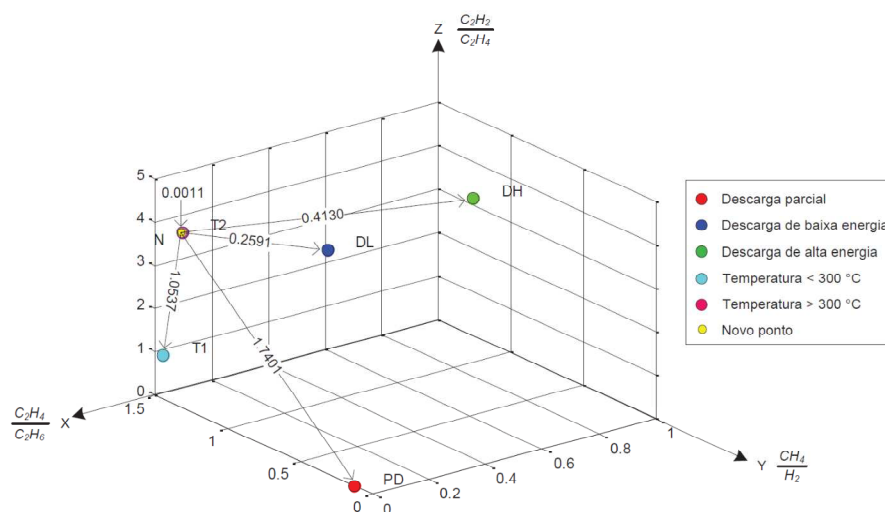


FIGURA 2 – Distância entre o novo ponto de falha e um conjunto representado pela moda da FDP

Para determinar a menor distância entre um ponto e um dos conjuntos, é calculado o menor valor médio absoluto de X_i através de :

$$K_i = \frac{\sum_{j=1}^n |x_j - y_i|}{n} \quad 11$$

e para determinar qual dos conjuntos está mais próximo do ponto, é calculado o mínimo dos mínimos da distância absoluta de X_i :

$$X_j = \min (\min(K_i)) \quad 12$$

Em outras palavras, é determinada 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 2 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 ($T < 300^\circ \text{C}$).

A base de dados usada foi composta por 352 amostras, com 318 amostras de falha e 34 para condição normal. Todas as amostras são comparadas com os 1400 pontos virtuais (nuvem de pontos), sendo 200 pontos por falha e por condição normal, como apresentado na Tabela 1.

Tabela 1: grupo de amostras para teste e validação

	Tipo de falha/sem falha	Conjunto de treino (virtual)	Conjunto de teste (real data)
T1	Falha térmica – $T < 300^\circ \text{C}$	200	77
T2	Falha térmica – $T > 300^\circ \text{C}$	200	71
PD	Descarga parcial	200	30
DL	Descarga de baixa energia	200	37
DH	Descarga de alta energia	200	103
H0	Estado normal (sem OLTC)	200	20
H1	Estado normal (com OLTC)	200	14

A nuvem de pontos foi gerada de duas etapas diferentes: na primeira, dados virtuais foram produzidos com $\lambda = 1$, em cada iteração, ou seja, cada nuvem foi composta por pontos que compartilham a mesma moda. Na segunda, foram adicionados novos pontos a nuvem com $\lambda > 2$, ou seja, isto significa que outras estruturas dos conjuntos foram compartilhadas pelo conjunto de pontos virtuais e o conjunto original.

Com esta configuração, a precisão foi de 98% com o uso da primeira etapa ($\lambda = 1$) com 318 amostras de casos de falha. Para o uso de toda a base de dados (352 amostras, incluindo a condição normal) a precisão foi de 99.42%. Para a segunda etapa ($\lambda > 2$) a precisão para ambos os casos (condição normal e de falha) foi de 100%.

Quando se obtém 100% de acerto, isto significa que o conjunto formado pela nuvem de pontos conseguiu mapear adequadamente a densidade de toda a região de cada subconjunto e o sistema está modelado e pronto para realizar o diagnóstico.

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 uma capacidade de discriminação, uma vez que não há respostas falsas e 100% de precisão foi alcançada. Isto é ainda mais notável, uma vez que o *mean shift* consegue discriminar entre transformadores com e sem OLTC apenas a partir das amostras de gases.

Tabela 2 – Precisão do diagnostic em discriminar a condição normal e de falha

Modelo	% correta identificação (condição norma e de falha)		Nº de casos com diagnóstico errado
	Conjunto treino	Conjunto teste	
Miranda-Castro-Lima[2]	100 %	100 %	0
Lima-Miranda-Saavedra Case 1 ($0 < \lambda < 1$)	98%	98%	2
Lima-Miranda-Saavedra Case 2 ($1 < \lambda < 2$)	100%	100%	0

5.0 - ANALISES FUTURA

Na Figura 3 é apresentado um gráfico de barras com a normalização da distância para a amostra nº 39 da base de dados com casos reais. O resultado indica uma falha do tipo descarga de baixa energia (corona). Contudo, também é possível observar a presença de uma descarga parcial (DP) com 13 % de importância, em relação às demais possibilidades. Embora este não seja um valor alto (em um banco de dados estático), esse percentual poderia ser confundido (ou dar um indicativo) com um caso de falhas múltiplas.

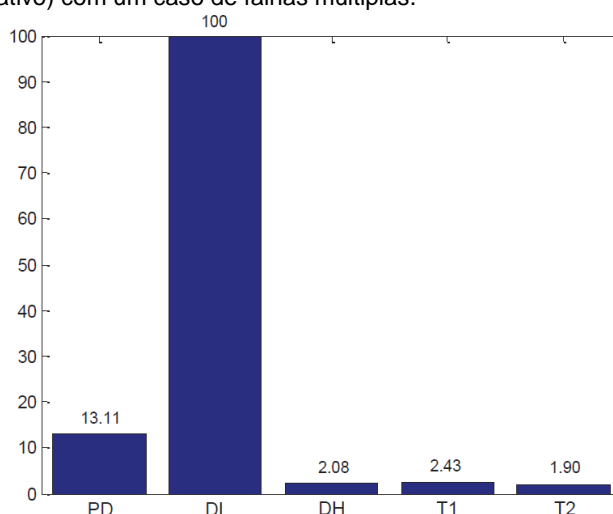


FIGURA 3 – Gráfico de barras indicando a amostra nº 39, descarga de baixa energia, mas também indicando um segundo candidato a falha, do tipo descarga parcial.

Uma análise simples desta possibilidade (falha DP) não seria correta, pois a análise de gases dissolvidos com base apenas em três relações de concentração de gás não seria capaz de distinguir falhas múltiplas de falhas simples. Além disso, vestígios de falhas múltiplas sobrepostas, não significam a ocorrência de diferentes falhas sobrepostas, pelo contrário, no caso normal poderia ser proveniente das ocorrências de falhas ao longo do tempo, com intensidades diferentes e com diferentes efeitos de contaminação.

Contudo, isto ajuda a explicar o porquê de uma diversidade de métodos baseados em AGD e relações utilizando de 3 a 5 gases, permitindo uma maior margem para o diagnóstico, mapeando a forma de desenvolvimento da falha incipiente (lenta, rápida, alta temperatura, etc.). A proposta deve ser de identificar indícios que apontem diferentes falhas e relacionar estas falhas, com a idade do equipamento, que gradativamente muda a composição do óleo, devido a uma lenta acumulação de resíduos.

No caso da amostra nº 39, à proximidade dos dois clusters de falha pode ser bem significativa. Logo, podem-se pensar duas possibilidades:

- a) Pode ter ocorrido contaminação dos dados em função de outra falha, no qual gerou um efeito de sobreposição, ou
- b) Uma pequena “possibilidade” de outra falha ter realmente ocorrido, caso o primeiro diagnóstico esteja errado.

Os testes realizados levaram a um resultado de 100% de precisão acerto, porém, não se pode descartar a possibilidade de eventualmente se encontrar um caso em particular com uma mistura de características que levem a um diagnóstico errado.

Por outro lado, um método que pode gerar uma segunda “opinião” com algum grau de probabilidade é vantajoso se comparado com um método que proporcione uma única resposta. Portanto, o diagnóstico baseado no *mean shift* é mais eficaz do que os outros métodos propostos.

Com o monitoramento em tempo real dos gases dissolvidos, pode-se facilmente acompanhar a evolução do gráfico como mostrado na Figura 3 e, desta forma, confirmar ou descartar a possibilidade de falhas múltiplas com efeito de sobreposição. Isto também pode ajudar a formar um histórico da vida do transformador e acompanhar através de um gráfico a evolução de transformadores com um diagnóstico de condição normal. Espera-se que as bases de dados compostas por esses registros em tempo real estejam disponíveis em um futuro próximo, permitindo progressos nos sistemas automáticos de diagnóstico.

6.0 - CONCLUSÃO

O problema da condição do diagnóstico em transformadores de potência, ou seja, quando sujeitas a um acompanhamento em tempo real de gases dissolvidos em óleo, tem tido um avanço apontando uma nova direção para o trabalho relatado neste artigo: um novo processo de diagnóstico, com base na geração de nuvens virtuais de dados, alcançou 100% de precisão no diagnóstico, tanto na detecção de condições normais/falha e em identificar de forma exata, em cada caso, o tipo de falha, bem como discriminar casos de transformadores com e sem OLTC.

As novidades neste trabalho estão resumidas da seguinte forma:

- A confirmação da possibilidade de utilizar a teoria da informação *mean shift* para a densificação dos *clusters* de dados, gerando as nuvens virtuais de pontos para compensar a escassez de dados reais disponíveis.
- A ideia de usar as nuvens de pontos virtuais produzidos pelo *mean shift* permitiu que a base de dados fosse aplicada diretamente para fazer o diagnóstico e obter a mesma precisão do sistema de redes neurais autoassociativas.
- A capacidade de fazer o reconhecimento de estados normais e com defeito, ambos no mesmo sistema de diagnóstico.

A precisão obtida com este método é comparável com a precisão alcançada pelo método proposto anteriormente, baseado no *mean shift* e no sistema de redes neurais autoassociativas [2]. A maior vantagem do método atual é que não há necessidade da fase de treinamento, um segundo ponto a favor é que permite mais facilmente e naturalmente um *ranking* de possibilidades para o diagnóstico.

O sucesso relatado neste artigo não significa que o *mean shift* é infalível, por exemplo, pode haver casos, na prática, onde os efeitos simultâneos de mais de uma falha podem alterar as relações de concentração de gases e tornar um caso ambíguo para o diagnóstico. No entanto, como o método ajuda a visualizar os *clusters* que estão associados com as diferentes modas de falha, mais informações podem ser analisadas pela operadora tomada de decisão.

7.0 - REFERÊNCIAS BIBLIOGRÁFICAS

- (1) Suwanasri, T., Chaidee, E. & 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.
- (2) V. Miranda, A. R. G. Castro and S. Lima, “Diagnosing Faults in Power Transformers With Autoassociative Neural Networks and Mean Shift”, *IEEE Transactions on Power Systems*, vol. 27, no. 3, July 2012.
- (3) R. Sudhir, A. d. M. Martins, and J. C. Principe, “Mean shift: An information theoretic perspective,” *Pattern Recogn. Lett.*, no. 30, pp. 222–230, 2009

- (4) M. Duval and A. Pablo, "Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases", *Electrical Insulation Magazine*, *IEEE* 17(2): 31-41, 2001.
- (5) Sudhir Rao, Allan de Medeiros Martins, Weifeng Liu, Jose C. Principe, "Information Theoretic Mean Shift Algorithm", Intl. Work. on Neural Networks for Signal Processing, Maynooth, Ireland, pp.1 – 9, 2006.
- (6) Sudhir Rao, Allan de Medeiros Martins, Jose C. Principe, "Mean shift: An information theoretic perspective", *Pattern Recognition Letters*, no. 30, pp. 222–230, 2009.
- (7) A. Renyi, "Some Fundamental Questions of Information Theory", *Selected Papers of Alfred Renyi*, vol 2, pp. 526-552, Akademia Kiado, Budapest, 1976.
- (8) E. Parzen, "On the estimation of a probability density function an the mode", *Annals Math Statistics*, vol 33, 1962.
- (9) J. C. Principe and D. Xu "Information-theoretic learning using Renyi's quadratic entropy", in J.-F. Cardoso, C. Jutten, and P. Loubaton, editors, *Proceedings of the First International Workshop on Independent Component Analysis and Signal Separation*, Aussois, France, pages 407- 412, 1999.
- (10) D. Erdogmus and J. C. Principe, "Generalized Information Potential Criterion for Adaptive System Training", *IEEE Transactions on Neural Networks*, vol. 13, no. 5, September 2002, pp. 1035-1044.
- (11) R.A. Morejon, J.C. Principe, "Advanced search algorithms for information-theoretic learning with kernel-based estimators", *IEEE Transactions on Neural Networks*, vol. 15(4), pp. 874 - 884, July 2004

8.0 - DADOS BIOGRÁFICOS

Shigeaki Leite Lima possui graduação em Engenharia Elétrica pela Universidade Federal do Maranhão (2005) e Mestrado em Engenharia Elétrica (2008). Atualmente é Doutorando do Programa de pós-graduação em Engenharia Elétrica (UFMA). Tem experiência na área de Engenharia Elétrica, com ênfase em Fontes Renováveis de Energia e Diagnóstico de Falhas em Transformadores de Potência, atuando principalmente nos seguintes temas: energia solar, energia eólica, sistemas híbridos, perda de vida e diagnóstico de falhas em transformadores de potência.

Vladimiro Miranda (M'90, SM'04, F'05) recebeu o título de Bacharel e Doutor em Engenharia Elétrica pela Faculdade de Engenharia da Universidade do Porto (FEUP) em 1977 e 1982 respectivamente. Em 1981 ingressou na FEUP e, atualmente, ocupa o cargo de Professor Titular. Também é investigador do INESC desde 1985 e atualmente Diretor do INESC Porto, um instituto de pesquisas avançadas em Portugal. É autor de diversos artigos e foi responsável por vários projetos em áreas relacionadas com a aplicação de Inteligência Computacional para Sistemas de Energia

Osvaldo R. Saavedra Mendez concluiu o doutorado em Engenharia Elétrica pela Universidade Estadual de Campinas em 1993. Atualmente é professor titular da Universidade Federal do Maranhão. É co-fundador do programa de Pós-Graduação de Engenharia de Eletricidade da UFMA, fundador do Núcleo de Energias Alternativas da UFMA. Coordenou a fundação do Instituto de Energia Elétrica na UFMA. Coordenou e/ou participou em mais de 20 projetos de pesquisa e de desenvolvimento, de colaboração nacional e internacional, e de pesquisa aplicada. Seus interesses abrangem operação de sistemas de potência em ambiente competitivo, serviços ancilares e energias renováveis. Desde meados de 2011 está a disposição do Governo do Estado do Maranhão, ocupando o cargo de Secretário Adjunto de Ciência e Tecnologia.