

Classificação de Distúrbios Elétricos Utilizando Wavelets, Componentes Principais e Redes Neurais Artificiais

Milde M. S. Lira^{*}, Manoel A. C. Júnior^{*}, Ronaldo R. B. Aquino^{*}, Mêuser J. S. Valença⁺, José Júlio A. L. Leitão⁺

^{*}*Universidade Federal de Pernambuco*
CEP: 50.740-530, Recife-PE

⁺*Companhia Hidrelétrica do São Francisco*
CEP: 50761-901, Recife/PE

E-mails: milde@ufpe.br, macj@ufpe.br, rrba@ufpe.br, meuser@chesf.gov.br,
julio@chesf.gov.br

Resumo

Neste trabalho é apresentado um método para classificação automática de distúrbios elétricos baseado em redes neurais artificiais, utilizando no treinamento sinais reais de tensão coletados através de registradores digitais de perturbação existentes no sistema de monitoração da CHESF. O sinal de tensão perturbado é coletado a uma taxa de amostragem de 128 amostras/ciclo na frequência de 60Hz durante 14 ciclos. O sinal de tensão real coletado é processado em duas etapas: i) inicialmente é decomposto através da transformada wavelet até o quinto nível de resolução, extraindo suas principais características; ii) em seguida os coeficientes wavelets são processados via Análise de Componentes Principais que opera projetando os dados linearmente em um subespaço de menor dimensão. A classificação realizada pela combinação de três redes MLP com diferentes arquiteturas é baseada nas características extraídas a partir do pré-processamento do sinal de tensão. O algoritmo *Resilient Backpropagation* foi utilizado no treinamento das redes. Na combinação das redes, em cada um dos seis nós de saída, aplicou-se a média entre as três saídas das redes individuais. A decisão final do classificador corresponde à saída combinada de maior valor. Os resultados são bastante promissores para os cinco tipos de distúrbios testados – Afundamentos e Elevação de tensão, Harmônicos, Transitórios oscilatórios, Interrupção; e para ausência de distúrbio.

1. Introdução

Nas últimas décadas, a qualidade da energia elétrica - QEE tornou-se um assunto de grande interesse da indústria de energia elétrica, dos fabricantes de equipamentos e consumidores. Este interesse advém da grande proliferação de dispositivos e microprocessadores utilizados em vários tipos de equipamentos eletrônicos nos complexos processos industriais, incluindo computadores, controladores de velocidade e outros dispositivos não lineares.

Visando fornecer energia elétrica de alta qualidade, diversas empresas de energia vêm investindo de forma significativa no monitoramento em tempo real do seu sistema de potência. Através dos dados obtidos pela rede de monitoração é possível aos engenheiros responsáveis pela análise do sistema, além de avaliar a QEE fornecida pela empresa, diagnosticar o problema e recomendar soluções apropriadas.

A qualidade da energia engloba a análise, o diagnóstico, a solução e o impacto econômico de qualquer anomalia do sistema elétrico. Neste contexto, um dos mais importantes métodos de análise é a classificação de distúrbios. Isto ocorre porque a partir do tipo de distúrbio é possível identificar suas causas e tomar as providências cabíveis para mitigar ou até mesmo eliminar seus efeitos. Face ao elevado número de registros de distúrbios, é impraticável classificá-los manualmente por meio da visualização do sinal perturbado. Além disto, a inspeção visual requer profissionais altamente especializados, cuja experiência fora adquirida ao longo de anos de trabalho na área, capazes de extrair informações a partir destes

registros. Desta forma, é necessário um procedimento automático que seja capaz de analisar o sinal e disponibilizar as informações indispensáveis para que medidas preventivas adequadas sejam tomadas.

Neste trabalho é proposto um método automático de classificação da forma de onda da tensão através das *Redes Neurais Artificiais* - RNA. Esta metodologia analisa o sinal da forma de onda da tensão através da *Transformada Wavelet* - TW, posteriormente aplica *Análise de Componentes Principais* - PCA a todos os coeficientes wavelets resultantes da decomposição do sinal e finalmente classifica as diferentes perturbações utilizando a combinação das saídas de três redes MLP.

Para treinamento foi empregado o algoritmo *Resilient Backpropagation*, e para combinação das saídas das redes com diferentes arquiteturas foi utilizada a média das suas saídas para cada um dos seis nós de saída. A decisão final do classificador corresponde à saída combinada de maior valor.

A principal contribuição do trabalho é a utilização de informações relativas aos coeficientes wavelets de todos os níveis da decomposição do sinal de distúrbio através da introdução da Análise de Componentes Principais. Esta técnica de redução de dimensionalidade permite uma diminuição considerável na dimensão do vetor de dados formado pelos coeficientes wavelets sem perda de suas principais características. Este vetor de dados, agora efetivamente reduzido e resumindo características relevantes do sinal de distúrbios, é a entrada do classificador neural. Este classificador neural apresenta melhores resultados na classificação do sinal de distúrbio em relação aos métodos de classificação modernos que utilizam como entrada da rede neural apenas os coeficientes wavelets dos primeiros níveis ou do nível de decomposição de maior energia, visto que os coeficientes dos níveis desprezados contêm informações relevantes do sinal de distúrbio.

Este trabalho está organizado da seguinte maneira: inicialmente faz-se uma descrição da transformada wavelet e análise de componentes principais nas seções 2 e 3 respectivamente. Em seguida, na seção 4 é feita uma apresentação do pré-processamento dos dados. A seguir, a seção 5 mostra como foi realizado o treinamento das redes neurais. Na seção 6 são apresentados os resultados obtidos. Finalmente, na seção 7 as conclusões são apresentadas.

2. Transformada Wavelet Discreta

Um dos objetivos da análise de sinais elétricos é extrair informações relevantes sobre o processo ao qual o sinal está associado. A teoria wavelet, que se constitui um método para decomposição e reconstrução de sinais, emprega funções de base que são locais nos domínios do tempo e da frequência ([1]), características estas que a torna superior à transformada de Fourier na análise de sinais transitórios ([2]).

A transformada wavelet representa um sinal como uma soma de “ondinhas” em diferentes escalas e localizações, cujos coeficientes maximizam as contribuições destas “ondinhas” em cada uma destas escalas e localizações. A teoria wavelet, de aplicação relativamente recente na análise de sinais em sistemas de potência, vem sendo freqüentemente considerada como o método mais indicado para a análise de transitórios. A Transformada Wavelets Contínua essencialmente mapeia um sinal unidimensional (no tempo) em uma representação bidimensional (tempo, escala) que é altamente redundante. As Transformadas Wavelets Discretas (DWT) são mais atraentes do ponto de vista de implementação computacional. Elas não são transladadas nem escalonadas continuamente, mas sim em intervalos discretos ([3]). Além da discretização do plano escala-translação, a variável independente do sinal pode também ser discretizada. Isto é possível através de uma pequena modificação na wavelet contínua. Neste caso, tem-se:

$$DWT(m, n) = \frac{1}{\sqrt{a_0^m}} \sum_{k=0}^{N-1} f[k] \psi^* \left(\frac{k - nb_0 a_0^m}{a_0^m} \right) \quad (1)$$

em que $\psi(k)$ é a wavelet mãe, e sua versão deslocada e escalonada são as wavelets filhas.

Normalmente assumem-se a_0 e b_0 como inteiros. O menor passo inteiro para a escala, $a_0 = 2$, é conhecido como escalonamento diádico. E o menor passo inteiro de translação temporal é $b_0 = 1$. Assim as wavelets diádicas tornam-se:

$$DWT(m, n) = 2^{-m/2} \sum_{k=0}^{N-1} f[k] \psi^*(2^{-m} k - n) \quad (2)$$

A wavelet diádica é implementada através da técnica multirresolução piramidal de decomposição. Em princípio, o sinal digitalizado $c_0[n]$ é decomposto em suas versões de detalhe $d_1[n]$ e de aproximação $c_1[n]$, utilizando-se os filtros $g[n]$ e $h[n]$, respectivamente. O filtro digital $g[n]$ é um filtro passa faixa. Logo, o sinal filtrado $d_1[n]$ é uma versão detalhada do sinal $c_0[n]$ e possui componentes de altas frequências (e.g. variações súbitas do valor instantâneo da tensão nos sinais de distúrbios de potência) em comparação ao sinal de aproximação $c_1[n]$, pois o filtro $h[n]$ é um filtro passa baixa. A decomposição de $c_0[n]$ em $c_1[n]$ e $d_1[n]$ é a decomposição de 1º nível. Matematicamente, elas são definidas por:

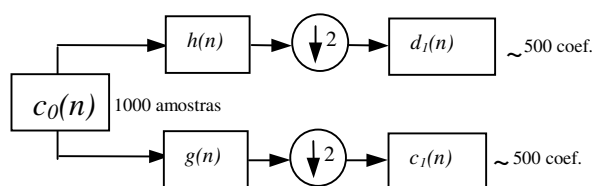


Figura 1- Processo de decomposição do 1º nível

$$c_1[n] = \sum_k h[k-2n]c_0[k], \quad (3)$$

$$d_1[n] = \sum_k g[k-2n]c_0[k]. \quad (4)$$

As equações 3 e 4 estão esquematicamente representadas na Figura 1. A cada nível de decomposição, o tamanho dos sinais decompostos (e.g. $c_1[n]$ e $d_1[n]$) é metade do tamanho do sinal no estágio anterior (e.g. $c_0[n]$). As decomposições nos níveis mais elevados se processam de forma similar.

3. Análise de Componentes Principais

Um dos problemas básicos na análise de dados, especialmente dados obtidos de um processo estocástico, consiste na extração de características (feature selection). A idéia é construir um processo (ou transformação) que leva o espaço de dados a um espaço de características, usualmente de mesma dimensionalidade do espaço de dados original.

Além da extração de características, um outro interesse é extrair apenas aquelas mais relevantes dos dados, ou seja, projetar uma transformada que reduza o número efetivo de características, retendo ainda a maior parte do conteúdo de informação intrínseca dos dados. Este processo é conhecido como redução de dimensionalidade. Esta simplificação é extremamente atrativa do ponto de vista de implementação e de análise dos dados. Uma grande variedade de métodos tem sido proposta na literatura para a redução do número de dimensões, dentre elas, a análise de componentes principais. A idéia central é remover (ou combinar) características altamente correlacionadas.

Hotelling ([4]) foi o primeiro a deduzir e publicar a transformação que mapeia variáveis discretas em coeficientes não correlacionados. Este método foi referenciado como o método das componentes principais. Uma transformada análoga, porém no caso contínuo, transformando sinais contínuos em um conjunto de coeficientes não correlacionados foi introduzida por Karhunen-Loève (a chamada expansão de Karhunen-Loève para Processos estocásticos).

Para realizar a redução de dimensionalidade em determinados dados de entrada, calculam-se os autovalores e autovetores da matriz correlação do vetor de dados. Em seguida, projetam-se esses dados de entrada ortogonalmente sobre o subespaço abrangido pelos autovetores pertencentes aos autovalores dominantes. Pode-se mostrar ([5]) que as variâncias correspondem aos autovalores da matriz correlação do vetor de dados de entrada e são arrumadas em ordem decrescentes. Logo, as primeiras componentes principais estão associadas às maiores variâncias.

4. Pré-processamento do sinal de tensão

O sinal de tensão capturado através dos *Registadores Digitais de Perturbações* – RDP durante o distúrbio a uma taxa de amostragem de 128 amostras/ciclo na frequência de 60Hz é apresentado em uma janela de 14 ciclos perfazendo um total de 1792 amostras. Este sinal é pré-processado em duas etapas:

1ª Etapa – O principal objetivo é extrair o máximo de informações no domínio das wavelets presentes nos diversos níveis de resolução. Neste sentido, o sinal é decomposto através da wavelet db6 ([6]) até o 5º nível de resolução produzindo um total de 1844 coeficientes wavelets.

O número de sub-bandas (nível de resolução), a ser usado na decomposição de um sinal, é escolhido de forma que a frequência fundamental do sinal se situe no centro da sub-banda das frequências mais baixas, a fim de limitar os efeitos do conteúdo espectral da fundamental nas outras sub-bandas. Baseado nessa hipótese, e sabendo-se que a taxa de amostragem do RDP utilizado para capturar o sinal de distúrbio na frequência fundamental (60Hz) é de 7.680Hz encontra-se que este sinal deve ser decomposto através da TW até o 5º nível de resolução.

2ª Etapa - Os coeficientes wavelets são submetidos a uma transformação para reduzir o número efetivo de características através da técnica PCA, porém retendo ainda a maior parte do conteúdo de informação ([7]). Nesta etapa foram mantidas as primeiras componentes, onde a soma das variâncias associadas foi de aproximadamente 85%, garantindo a minimização da perda da informação e maximização da redução dimensional dos dados de entrada original (coeficientes wavelets).

Todo o processo sofrido pelo banco constituído pelos sinais de tensão proporciona à rede neural um aprendizado mais rápido e eficiente, já que o mesmo após decomposição pela TW, fora normalizado e decorrelacionado.

4.1 Elaboração do banco de dados de entrada da rede neural

Neste trabalho é realizada a identificação de cinco tipos de distúrbios: Afundamentos e Elevação de tensão, Harmônicos, Transitórios, Interrupção; e para ausência de distúrbio. Assim, é necessário que a rede realize a classificação de seis classes, incluindo a ausência de distúrbio.

O banco de dados a ser pré-processado é constituído de 730 exemplos de cada classe, e o algoritmo para realizar este processo foi desenvolvido no MatLab®. Este programa prepara o banco de dados de treinamento da rede neural e armazena todos os parâmetros (média, desvio padrão e matriz transformação - PCA) necessários para posterior pré-

processamento do sinal de tensão, caso o sistema venha a operar on-line.

4.2 Divisão do conjunto de dados

De acordo com o método *10-fold cross-validation* com estratificação, os padrões devem ser divididos em 10 partições estratificadas. A Figura 2 ilustra a criação das 10 partições.

Dos 4.380 exemplos do banco de dados 20% foram escolhidos para validação, 20% para teste e 60% para treinamento.

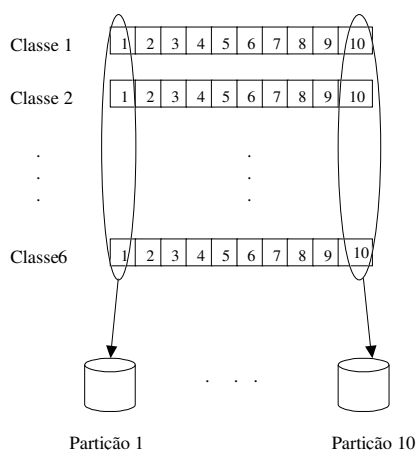


Figura 2- Formação das 10 partições

4.3 Codificação da resposta desejada

Na classificação de padrões selecionou-se a codificação do tipo 1 de M , onde o k -ésimo elemento do vetor resposta desejada é igual a um, se o vetor de entrada x pertence à classe k , e zero caso contrário, como mostra a Tabela 1.

Tabela 1. Saídas desejadas da rede neural

DISTÚRBIOS	SAÍDAS DESEJADAS					
Afundamentos	1	0	0	0	0	0
Tensão Normal	0	1	0	0	0	0
Harmônicos	0	0	1	0	0	0
Elevação	0	0	0	1	0	0
Transitórios	0	0	0	0	1	0
Interrupção	0	0	0	0	0	1

5. Treinamento das redes MLP com o algoritmo *Rprop*

Uma diferença fundamental entre o algoritmo *Rprop* ([8]) e o *Backpropagation* está no processo de ajuste dos pesos. No algoritmo *Rprop*, a atualização dos pesos é influenciada pelo sinal da derivada parcial, e não pelo seu valor, como ocorre no *Backpropagation*. Isto resulta em uma melhoria no processo de aprendizagem tornando a convergência nesta fase mais rápida.

No sentido de encontrar a melhor arquitetura de rede MLP com uma única camada intermediária, um programa foi desenvolvido no MatLab® para gerar a partição dos conjuntos e variar o número de nodos intermediários. Neste processo de busca, a rede que apresenta menor erro médio quadrático - MSE no conjunto de validação é substituído pela rede anterior. O programa lê o banco de dados já pré-processados e funciona do seguinte modo:

- 1- Gera as 10 partições.
- 2- Combina essas partições de forma aleatória para compor o conjunto de treinamento, validação e teste.
- 3- A partir do número de nodos intermediários inicial N_i e final N_f estabelecidos pelo usuário, toma-se o número de nodos intermediários N_i .
- 4- Inicia os pesos aleatoriamente.
- 5- Começa o treinamento até parar seja pelo erro MSE no conjunto de treinamento, seja pelo número de épocas ou validação.
- 6- Armazena o erro MSE no conjunto de validação e os parâmetros da rede; e com a mesma arquitetura, faz-se uma nova reinicialização dos pesos iniciais e recomeça o treinamento.
- 7- Compara o erro de validação atual com a anterior, e armazena os parâmetros da rede associada à rede de menor erro MSE no conjunto de validação.
- 8- Acrescenta mais um nodo na camada intermediária e vai para o passo 4 até que N_f seja alcançado.
- 9- Quando N_f é alcançado, toma-se mais uma combinação das partições e vai para o passo 4 até que tenham sido realizadas 10 combinações das partições.
- 10- Caso o número de combinações seja 10, os parâmetros da rede de melhor arquitetura são armazenados e o programa é finalizado.

6. Resultados

Foram geradas três bases de conhecimento em função dos coeficientes wavelets nos 5 níveis de resolução resultantes da decomposição do sinal (Tabela 2).

Tabela 2. Composição das bases de conhecimento

Largura dos filtros (Hz)	BASE 1	BASE 2	BASE 3
0 – 120 (escala)	cA_5	cA_5	-
120 – 240 (wavelet)	cD_5	-	cD_5
240 – 480 (wavelet)	cD_4	-	cD_4
480 – 960 (wavelet)	cD_3	-	cD_3
960 – 1.920 (wavelet)	cD_2	-	cD_2
1.920 – 3.840 (wavelet)	cD_1	-	cD_1

Com cada uma das bases de conhecimento foram treinadas três redes com diferentes arquiteturas segundo o número de nodos intermediários.

Caso 1: Redes treinadas com a Base 1

A) Redes Individuais

A seleção das três melhores arquiteturas utilizando a base 1 foi executada em três faixas de nodos intermediários: i) 10-40; ii) 41-70; iii) 71-100.

As melhores arquiteturas encontradas foram: Rede-1) 188-40-6; Rede-2) 188-44-6 e Rede-3) 188-78-6. Para avaliar o desempenho das redes foram coletados novos dados do Sistema de Transmissão da Chesf, denominado de conjunto de teste extra. A Tabela 3 mostra o erro percentual de acerto por classe das três redes para este novo conjunto.

Tabela 3. Resposta das redes para o novo conjunto de teste

	Exemplos No.	Rede-1	Rede-2	Rede-3
Afundamentos	52	92,3	84,6	88,5
Tensão Normal	88	96,6	94,3	97,7
Harmônicos	27	100,0	100,0	100,0
Elevação	42	100,0	100,0	100,0
Transitórios	50	92,0	92,0	96,0
Interrupção	47	100,0	97,9	97,9
Total	306	96,4	94,1	96,4

B) Redes Combinadas

A combinação das saídas das redes com diferentes arquiteturas tem um grande potencial para assegurar um melhor desempenho no sistema de classificação. Há diferentes modos de combinar resultados de redes individuais ([9, 10]). Dois métodos comuns utilizados são a votação e a média. Neste trabalho, na combinação das redes, a cada um dos seis nós de saída, aplicou-se a média entre as três saídas das redes individuais. A decisão final do classificador corresponde a saída combinada de maior valor. A Tabela 4 apresenta os resultados das redes combinadas.

Tabela 4. Resposta das redes 1, 2 e 3 combinadas

	Acertos	Erro	Acerto Percentual (%)
Afundamentos	46	6	88,5
Tensão Normal	86	2	97,7
Harmônicos	27	0	100,0
Elevação	42	0	100,0
Transitórios	47	3	94,0
Interrupção	47	0	100,0
Total	295	11	96,4

Comparando os resultados das Tabelas 3 e 4, observa-se uma significativa melhora no desempenho do sistema quando as redes são combinadas, com exceção das classes *Afundamentos* e *Transitórios*, onde um das redes apresentou um índice de acerto ligeiramente superior ao da combinação. É importante ressaltar que a resposta da combinação, em geral, é

superior ou igual a resposta individual da rede que menos acertou.

Caso 2: Redes treinadas com a Base 2

As melhores arquitetura encontradas com a Base 2 foram: Rede-4) 10-38-6; Rede-5) 10-63-6; Rede-6) 10-100-6.

O número de nós de entrada das redes treinadas com essa base de conhecimento, que equivale ao número de componentes principais retidas, foi 10. Este resultado é interessante, pois difere bastante do número de nós de entrada nas redes treinadas com a base 1. Visto que o critério utilizado para retenção das componentes foi o mesmo, isto ocorre devido ao alto grau de correlação existente entre os coeficientes wavelets de aproximação do nível 5. Na Tabela 5 são apresentados os resultados da combinação dessas três redes para o mesmo conjunto de teste do caso 1.

Tabela 5. Resposta das redes 4, 5 e 6 combinadas

	Acertos	Erro	Acerto Percentual (%)
Afundamentos	51	1	98,1
Tensão Normal	85	3	96,6
Harmônicos	27	0	100,0
Elevação	42	0	100,0
Transitórios	32	18	64,0
Interrupção	47	0	100,0
Total	284	22	92,8

O baixo rendimento na classe *Transitório* é causado provavelmente pela exclusão dos coeficientes wavelets de detalhes desta base de conhecimento, responsáveis pelas componentes de altas frequências contidas nesta classe de sinais.

Caso 3: Redes treinadas com a Base 3

As melhores arquitetura encontradas com a Base 3 foram: Rede-7) 197-26-6; Rede-8) 197-62-6; Rede-9) 197-78-6.

Na Tabela 6 são apresentados os resultados da combinação dessas redes para o mesmo conjunto de teste do caso 1.

Tabela 6. Resposta das redes 7, 8 e 9 combinadas

	Acertos	Erro	Acerto Percentual (%)
Afundamentos	41	11	78,8
Tensão Normal	86	2	97,7
Harmônicos	26	1	96,3
Elevação	42	0	100,0
Transitórios	45	5	90,0
Interrupção	44	3	93,6
Total	284	22	92,8

Na Tabela 6, observa-se o baixo desempenho da classe *Afundamento* causado pela ausência dos coeficientes wavelets de aproximação nesta base de co-

nhecimento, responsáveis pelas componentes de baixas frequências contidas nesta classe de sinais.

Nos três casos apresentados, as redes combinadas foram treinadas com a mesma base de conhecimento. No próximo caso serão analisadas as respostas da combinação das redes treinadas apenas com os coeficientes de aproximação com as redes treinadas apenas com os coeficientes de detalhes.

Caso 4: Combinação das redes 4, 5 e 6 com as redes 7, 8 e 9

As saídas das seis redes formam as colunas de uma matriz, onde para cada linha dessa matriz é calculada a média e posteriormente, aplicado o critério *o vencedor leva tudo*. Isto garante que apenas uma classe seja indicada como a classe à qual pertence o sinal de tensão na entrada do sistema.

A Tabela 7 apresenta a resposta da combinação dessas redes para o mesmo conjunto de teste anterior.

Tabela 7. Resposta das redes 4, 5, 6, 7, 8 e 9 combinadas

	Acertos	Erro	Acerto Percentual (%)
Afundamentos	52	0	100,0
Tensão Normal	88	0	100,0
Harmônicos	27	0	100,0
Elevação	42	0	100,0
Transitórios	48	2	96,0
Interrupção	47	0	100,0
Total	304	2	99,3

Este resultado era esperado, visto que as redes treinadas com apenas os coeficientes de detalhes ou com apenas os coeficientes de aproximação tornaram-se mais especialistas no reconhecimento de determinadas classes. A combinação delas superou a deficiência que tinham no reconhecimento de outras classes de distúrbios, resultando em um sistema mais eficiente e robusto.

7. Conclusões

Neste trabalho, abordou-se o emprego das TW para extrair as principais características da forma de onda do sinal de tensão no domínio das wavelets até o quinto nível de resolução; e a PCA para reduzir o número efetivo destas características. Estas técnicas de transformação foram utilizadas para pré-processar o sinal de tensão antes de entrar no sistema classificador.

Para selecionar as três melhores arquiteturas de redes MLP com uma única camada intermediária utilizando o treinamento *Rprop* foram realizados diversos experimentos através da adição de nodos na camada intermediária. Esta busca foi realizada em três faixas de nodos intermediários e os treinamentos foram realizados simultaneamente em três computadores, com o objetivo de reduzir o tempo de treinamento das redes.

Os resultados da combinação das saídas das redes neurais na classificação dos distúrbios de tensão mostraram-se superiores aqueles apresentados pelas saídas individuais das redes. Além disso, a combinação de redes especializadas no reconhecimento de características específicas do sinal apresentou melhores resultados que redes treinadas para reconhecer todas as características presentes no mesmo.

Referências Bibliográficas

- [1] Kim, C. H. and Aggarwal, R. *Wavelet Transforms in Power Systems. Part1: General Introduction to the Wavelet Transforms*. IEEE Power Engineering Journal, pp.81-87.
- [2] Gu, YuHua and Bollen, M. H. J. *Time-Frequency and Time-Scale Domain Analysis of Voltages Disturbances*. IEEE Transactions on Power Delivery, 15 v., pp. 1279-1284.
- [3] Mallat, S. *A Theory for Multiresolution Signal Decomposition: The Wavelet Representation*. IEEE Trans. Pattern Analysis and Machine Intelligence, 11 v., pp.674-693.
- [4] Hotelling, H. *Analysis of a Complex of Statistical Variables into Principal Components*. J. Educ. Psychol 24 v., pp.417-441, 498-520.
- [5] Jolliffe, I. T. *Principal Component Analysis*. Springer-Verlag, New York.
- [6] Arruda, Elcio F. et al. *Um Estudo das Famílias Wavelets Aplicadas à Qualidade da Energia Elétrica*. XVI - CBA - Congresso Brasileiro de Automática.
- [7] Lam, S. L. Y. and Lee, Dik Lun. *Feature Reduction for Neural Network Based Text Categorization*. 6th International Conference on Database Systems for Advanced Applications, pp.195-202.
- [8] Riedmiller, M. and Braun, H. *A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm*. In Proceedings of the IEEE International Conference on Neural Networks. 1 v., pp.586-591.
- [9] Mao, J. *A Case Study on Bagging, Boosting and Basic Ensemble of Neural Networks for OCR*. IEEE International Joint Conference on World Congress on Computational Intelligence. 3 v., pp.1828-1833.
- [10] Wang, Z, Liu, Y. and Griffin, P. J. *A Combined ANN and Expert System Tool for Transformer Fault Diagnosis*. IEEE Transaction on Power Delivery, 13 v., pp.1224-1229.