

## REDES NEURAIS ARTIFICIAIS NA ENGENHARIA NUCLEAR (PARTE 2)

Benedito Dias Baptista Filho

Instituto de Pesquisas Energéticas e Nucleares - IPEN-CNEN/SP  
Av. Prof. Lineu Prestes, 2242 Cidade Universitária  
05508-000, São Paulo, SP, Brasil

## RESUMO

O campo das *Redes Neurais Artificiais* (RNA) tem despertado muito interesse na Engenharia Nuclear (EN). As RNA se aplicam a problemas complexos de difícil modelagem, em situações em que os dados são falhos ou incompletos, e em problemas complexos de controle. A primeira parte deste trabalho iniciou uma discussão sobre RNA na EN iniciando pelas redes Alimentadas Adiante. Nesta parte são trabalhados exemplos de utilização das Redes de Contatos Sinápticos Múltiplos (MULSY) aplicadas a controle e dos Mapas Auto-Organizáveis (SOM) na seleção e classificação de padrões. O principal propósito do trabalho é mostrar que as RNA podem ser utilizadas com sucesso desde que seja feita a escolha do tipo adequado à aplicação.

**Keywords:** neural networks, nuclear engineering, self-organized maps.

## I. INTRODUÇÃO

Na primeira parte deste trabalho [1] foi avaliado o conceito de *Redes Alimentadas Adiante* treinadas com o método de retro-propagação (RAA). Foram vistos exemplos de controle de um manipulador bi-articulado, de previsão do comportamento de um sistema termo-hidráulico, e aproximação de funções, no caso, o tempo de queima do combustível de um reator como função da distribuição de temperatura no núcleo. Os resultados indicaram que as RAA podem não se apresentar como solução ótima para problemas de controle, mas são eficientes na previsão do comportamento e na aproximação de funções.

Neste trabalho são discutidas redes mais especializadas: as redes de contatos sinápticos múltiplos (MULSY), desenvolvidas especialmente para controle, e os mapas auto-organizáveis (SOM), incomparáveis na seleção e classificação de padrões e características. São apresentados dois exemplos de controle com as redes MULSY: o manipulador bi-articulado e um processo termo-hidráulico. É apresentado o método e um exemplo de uso de SOM na EN. O artigo é dividido em 4 seções, a primeira é esta introdução, a segunda aborda as aplicações das redes MULSY, a terceira os SOM e a quarta as conclusões.

## II. REDES MULSY

O conceito de redes neurais de contatos sinápticos múltiplos foi desenvolvido no Centro de Engenharia Nucle-

ar do IPEN para aplicações em controle. Utiliza o conceito de contatos sinápticos múltiplos, novas funções de transferência, e um método de treinamento não-supervisionado que se assemelha ao aprendizado por reforço, tendo por base um modelo de plasticidade sináptica semelhante aos biológicos e a arquitetura dos aglomerados de moto-neurônios dos reflexos. Maiores detalhes estão nas referências [2] a [4]. A Fig. 1 ilustra uma unidade básica da rede MULSY. O caminho de entrada de um nível de controle hierarquicamente superior, "*o desejo*" -  $x_D$ , e o caminho que vem do sistema sensorio, "*a condição atual*" -  $x$ , convergem para a unidade responsável pela percepção do erro atual,  $\varepsilon$ , e para a unidade de saída.

**Controle do Manipulador.** O manipulador bi-articulado no plano é acionado por dois motores, necessitando pelo menos duas unidades de controle motor. O modelo para controle de posição tem como sinais de entrada os ângulos desejados e atuais dos dois segmentos. As velocidades e acelerações angulares são estimadas por unidades internas. A Fig. 2 apresenta o esquema simplificado do acoplamento das unidades de controle motor com o manipulador. Os caminhos do sistema sensorio estão representados por linhas pontilhadas e os caminhos dos sinais de comando do nível hierárquico superior e os intermediários estão representados por linhas sólidas. Unidades de entrada fazem a soma das entradas.

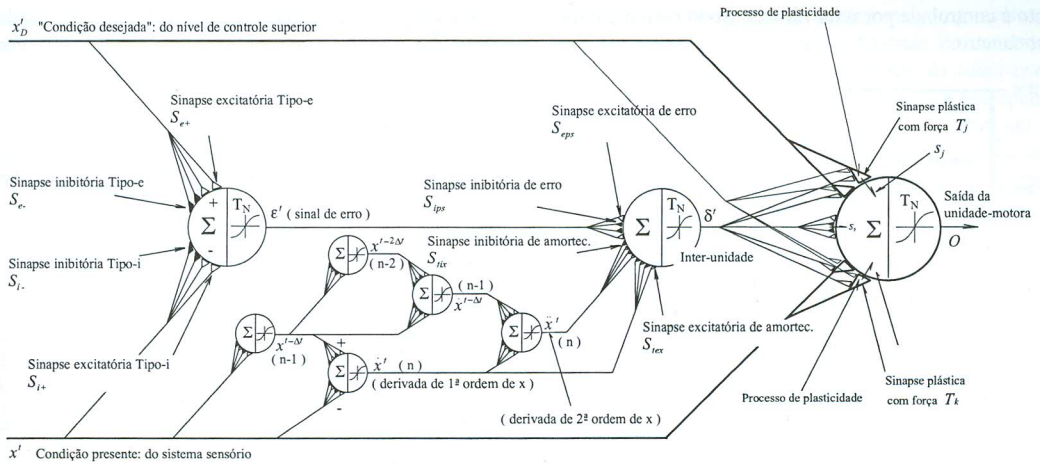


Figura 1. Unidade de Controle Motor Básica.

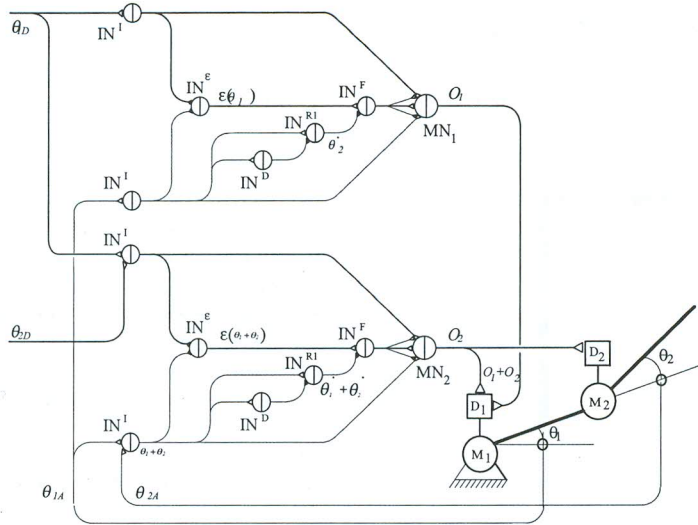


Figura 2. Acoplamento de Duas Unidades de Controle Motor com o Manipulador.

O treinamento da rede da Fig. 2 é não-supervisionado, com a execução de comandos de ação. É assemelhado ao aprendizado por reforço. Os pontos de treinamento foram os mesmos considerados na fase inicial do treinamento da RAA apresentada na parte 1 [1]. A seqüência de posições foi apresentada seis vezes variando-se o tempo de manutenção de cada comando (desejo). Na primeira e segunda apresentação da seqüência de treinamento o tempo foi de 5 segundos, na terceira e quarta foi de 10 segundos, e na quinta e sexta vezes foi de 20 segundos, totalizando 70 segundos de treinamento por ponto. Após as seis apresentações dos comandos de treinamento, a rede foi capaz de atender todos os comandos de posicionamento com excelente precisão. O treinamento correspondeu a um tempo simulado de 1960 segundos, contra as 7 horas gastas pela RAA. A comparação da precisão de posicionamento com os resultados da RAA está na Fig. 3. Não se observaram erros superiores a 5 mm. Na RAA treinada com o método de retro-propagação

os erros atingiram 300 mm. O aprendizado da rede MULSY foi aproximadamente 150 vezes mais rápido do que o da RAA. A capacidade de generalização da rede MULSY também foi muito superior, justamente por ser uma rede projetada e dedicada à tarefa específica de controle do manipulador. Para essa rede não foi necessário nenhum treinamento adicional para serem atingidos precisamente alvos fora do universo de treinamento.

**Controle de Processo Termo-Hidráulico.** A Fig. 4 mostra um esquema da Bancada de Circulação Natural (BCN), que simula a sistema de resfriamento de emergência de um reator a água pressurizada avançado. Um aquecedor elétrico é a fonte quente do sistema, e um trocador de calor a fonte fria. O trocador de calor é constituído por dois tubulões horizontais conectados por um feixe de tubos vertical imerso em um tanque de água. Um medidor de vazão magnético está instalado na linha principal do circuito. A água de res-

friamento é controlada por uma válvula globo com o auxílio de um rotâmetro.

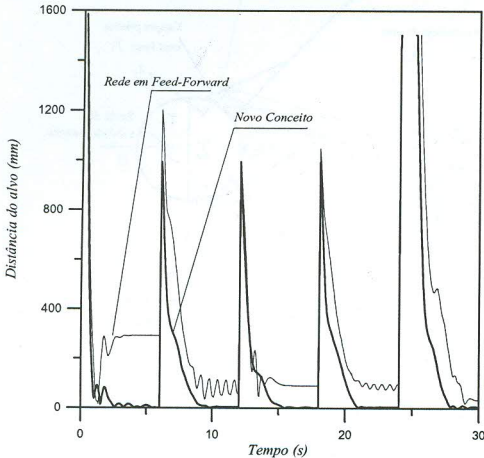


Figura 3. Comparação na Precisão de Posicionamento.

dades ligadas aos sinais de perturbação são usadas para modular o ganho da unidade motora de saída ( $O$ ), gerando o sinal de controle:

$$Sc = (O) (O_1) (O_2) \tag{1}$$

onde  $Sc$  é o sinal de controle,  $O$ ,  $O_1$  e  $O_2$  as saídas das unidades motoras.

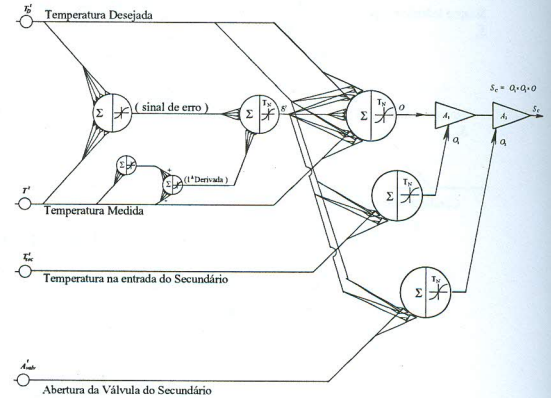


Figura 5. Unidade de Controle Modificada.

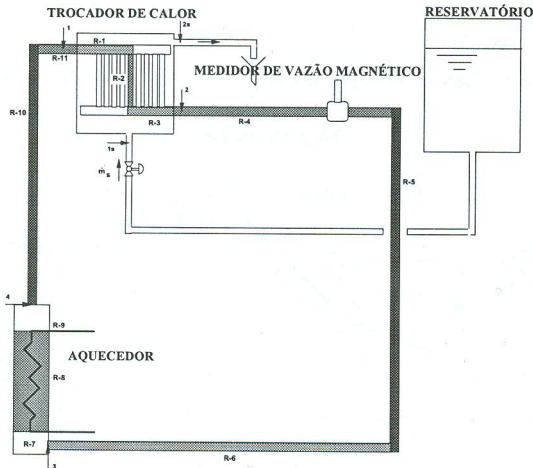


Figura 4. Esquema Simplificado da BCN.

O problema da BCN consiste em controlar a temperatura da água no lado primário do circuito, em qualquer ponto especificado, atuando somente na potência do aquecedor. A temperatura e a vazão da água no lado secundário são assumidas como perturbações, monitoradas mas fora do controle. Para a solução deste problema, a unidade básica de controle é modificada para receber, além dos sinais de *condição desejada* e *condição atual*, os sinais das *perturbações*. A Fig. 5 mostra a rede com essas modificações. A unidade de controle motor é ligada a dois ramos paralelos que processam os sinais das perturbações. Cada um dos ramos possui um conjunto de "sinapses plásticas." O sinal  $\delta$ , que combina o erro e as taxas de variação da temperatura controlada é utilizado para a adaptação das forças de ligação de todas as "sinapses plásticas." O sinal  $x_D$ , provê as características seletivas para a adaptação. A saída das uni-

Mais detalhes desse problema podem ser encontrados na referência [3]. O treinamento da rede é não-supervisionado. Os dados são as condições desejadas e as perturbações. A saída, ou seja, o sinal de controle necessário é descoberto pela rede. O treinamento foi realizado em três etapas para não sobrepor efeitos. Enquanto se treina sobre uma das variáveis em mudança, a plastificação das sinapses de saída das outras unidades é "congelada" fazendo  $T_c = 0$ . Cada uma das etapas do treinamento é realizada em sessões, observando-se a evolução gradual do aprendizado. Todo o treinamento representou um total de 615600 segundos de processo simulado, ou seja, 171 horas. Após o treinamento, a rede foi capaz de atender comandos de temperatura desejada sob diversas condições de perturbação de abertura da válvula e de variação na temperatura de resfriamento. Os testes apresentados verificam a capacidade de generalização pela execução de comandos não presentes no treinamento. Nos testes são bloqueados os mecanismos de plasticidade anulando-se a força de ligação de todas as sinapses facilitadoras,  $T_c = 0$ .

A Fig. 6 mostra resultados de um teste com vários transitórios em um intervalo de 26 horas. O teste se inicia com temperatura ambiente de 25°C e temperatura de resfriamento de 20°C. O primeiro comando solicita uma temperatura de 50°C na saída do aquecedor, com a vazão de resfriamento determinada pela abertura da válvula em 25%. Alguns valores do teste estão fora das faixas consideradas no treinamento, outros dentro dessas faixas porém não coincidentes com nenhum. A duração da primeira condição permite à BCN aproximar o regime permanente. Para isso o processo requer 22.000 segundos, determinados experimentalmente. O tempo de 16.000 segundos é suficiente apenas para a aproximação do regime permanente no lado primário

do trocador de calor. Todos os outros estágios duraram 7.200 segundos, o que representa aproximadamente 1/3 do tempo necessário para estabilização e leva a uma aproximação em torno de 95% da temperatura de regime. Essa condição adiciona algum erro à aproximação mas é suficiente para avaliação da rede.

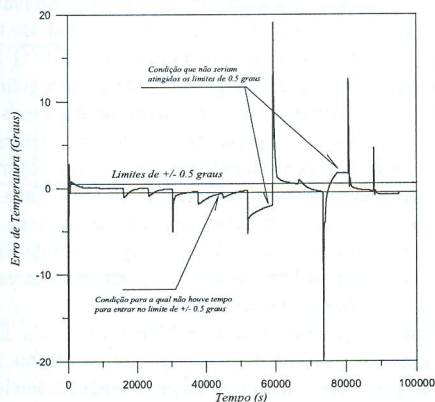


Figura 6. Evolução do Erro de Temperatura.

O erro na temperatura desejada, definido pela diferença entre a temperatura observada e a temperatura desejada, ficou fora de uma faixa de  $\pm 0,5^{\circ}\text{C}$  somente em três condições, sendo: uma por falta de tempo para acomodação da perturbação; outra somente justificada por se combinarem as duas maiores perturbações: *Temperatura de resfriamento de  $15^{\circ}\text{C}$* , e *Abertura da válvula de 0,08*; e a terceira onde se combinou a perturbação de *Abertura da válvula 0,08* com a demanda de uma temperatura desejada  $5^{\circ}\text{C}$  acima da maior temperatura considerada no treinamento, ou seja  $T_{des}=60^{\circ}\text{C}$ .

A opção de arquiteturas especializadas por tarefa e o uso de contatos sinápticos múltiplos aumenta a capacidade de integração da rede. Funções de transferência sinápticas mais complexas que pesos de ligação melhoram a relação entrada/saída e permitem a redução no número de unidades. O treinamento ocorre durante a execução de comandos, dispensando a necessidade de exemplos x/y: é um treinamento não-supervisionado. Os resultados obtidos no controle de temperatura do BCN mostraram que a rede apresenta boa capacidade de generalização e robustez.

### III. MAPAS AUTO-ORGANIZÁVEIS

Mapas Auto-organizáveis (Self-Organized Maps, SOM) são RNA's constituídas por uma única camada onde as unidades são colocadas em nós de uma grade, normalmente uni ou bidimensional (Fig. 7). Sua arquitetura reproduz uma característica do cérebro: *a organização, nos planos superiores, das entradas sensoriais que são representadas por mapas topográficos*. Nessa arquitetura as unidades se tornam sintonizadas a classes de padrões de entrada formando mapas, ou seja, as localizações espaciais das unidades se tornam indicativas das características estatísticas contidas nos padrões de entrada. O SOM representa uma

generalização não-linear da análise de componentes principais (PCA). Seu treinamento é baseado no conceito de aprendizagem competitiva: *as unidades da saída competem entre si para serem ativadas*. O resultado é que apenas uma unidade da saída, ou uma por grupo, está ativa em um determinado instante. O conceito de competição – *vencedor leva tudo* – é anterior ao SOM que porém representa o modelo mais geral e conhecido, desenvolvido na década de 80 por Teuvo Kohonen [6].

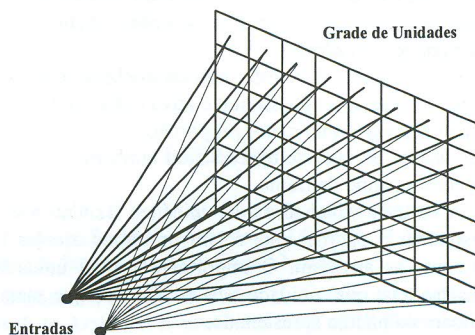


Figura 7. Rede SOM.

Nas redes SOM, as unidades possuem comunicação lateral que pode ser excitatória ou inibitória, obedecendo a uma função de "vizinhança," representada por uma Gauseana ou por uma função tipo "chapéu mexicano." O mais comum é o uso da seguinte função:

$$V_{j,j^0}(n) = \exp\left(-\frac{\|r_j - r_{j^0}\|^2}{2\sigma^2(n)}\right) \quad (2)$$

onde  $j^0$  representa a unidade vencedora,  $\|r_j - r_{j^0}\|^2$  é a distância entre a unidade  $j$  e a unidade vencedora e  $\sigma(n)$  varia em função do tempo, começando com um valor que faça a vizinhança inicial envolver todas as unidades, diminuindo com o número da época ( $n$ ) na forma:

$$\sigma(n) = \frac{\sigma_0}{1 + b_{\sigma} n} \quad (3)$$

A atualização dos pesos obedece a expressão:

$$w_j(n+1) = w_j(n) + \eta(n)[x(n) - w_j(n)] \quad (4)$$

onde a taxa de aprendizado  $\eta$  varia em função da distância entre a unidade  $j$  e a unidade vencedora  $j^0$ , conforme a função vizinhança e em função da época conforme:

$$\eta(n) = \frac{\eta_0}{1 + b_{\eta} n} \quad (5)$$

Não existem saídas desejadas ou conhecidas, o objetivo da rede é procurar a existência de padrões similares e classificar os dados de entrada em grupos de padrões. O aprendizado segue o seguinte roteiro:

1. Os pesos iniciais  $[w_{k,i,j}]$  são atribuídos aleatoriamente;
2. Os padrões são apresentados;

3. A cada apresentação de um padrão é determinada a unidade que possui os pesos "mais próximos" do vetor de entrada  $x$  (menor distância Euclidiana). Essa é eleita a *unidade vencedora*;
4. Os pesos da unidade vencedora são atualizados na direção de diminuir a *distância Euclidiana*;
5. Os pesos das unidades vizinhas são atualizados em uma proporção menor, em função de seu afastamento dado pela *função vizinhança*. A função de vizinhança varia com o número de épocas: engloba praticamente todas as unidades no início e se reduz até se limitar exclusivamente à unidade vencedora;
6. O padrão seguinte deve estar mais sintonizado com outra unidade, ou seja, sendo diferente ativa outra unidade, que será atualizada a uma maior proporção;
7. A constante da taxa de aprendizado também sofre redução com o tempo de treino.

A unidade vencedora e as unidades vizinhas a ela irão responder a padrões que sejam próximos àqueles que inicialmente as ativaram. O importante é qual unidade é ativada, ou seja, qual unidade tem as ligações que mais se aproximam do padrão apresentado.

O SOM é uma excelente ferramenta para a fase exploratória de dados. Projetando o espaço de entrada em protótipos de baixa dimensionalidade: grades regulares de uma ou duas dimensões; permite a visualização das propriedades dos dados. Esse é o primeiro passo na classificação de padrões e é realizada de um modo totalmente não-supervisionado. O SOM pode ser usado para se obter informações qualitativas das características dos dados de en-

trada. A "seleção" é a "extração" das características. O segundo passo seria a classificação propriamente dita, onde as características selecionadas dos dados de entrada são atribuídas a classes individuais. Hoje em dia, considera-se que o desempenho na classificação de padrões pode ser melhorado quando a "seleção" é seguida pela classificação acompanhada por um esquema de aprendizado supervisionado, ou seja, um "classificador adaptativo de padrões" de natureza híbrida. Uma boa definição de um classificador desse tipo é a do "quantizador vetorial por aprendizado"[7]. O SOM também pode ter sua utilidade ampliada com a rotulação dos dados, principalmente após a seleção de características, formando-se o que se chama de "mapas contextuais." A idéia é atribuir classes a unidades em uma grade bidimensional, dependendo de como cada padrão de teste" (não visto anteriormente) excita uma unidade particular na rede auto-organizada. Como resultado desse estágio "estimativo" as unidades da grade bidimensional são particionadas em um número de "regiões coerentes."

Este artigo aborda um problema típico da EN: a *classificação de eventos operacionais de um reator* como função de parâmetros medidos como a potência nuclear, a pressão no primário, etc. A Tabela 1 contém dados de 15 eventos utilizados no treinamento de uma rede SOM bidimensional. Esses dados, produzidos a partir de uma típica análise de transientes/acidentes de um PWR, representam as máximas variações relativas de 7 parâmetros do primário; do secundário e a máxima variação do nível de radiação e da pressão na contenção.

TABELA 1. Dados para Classificação

VARIÁVEIS		$P_{PR}(+)$	$P_{PR}(-)$	$L_{PZ}(+)$	$L_{PZ}(-)$	$T_R(+)$	$T_R(-)$	$Q_R(+)$	$Q_R(-)$	$W_S(+)$	$W_S(-)$	$P_S(+)$	$P_S(-)$	$W_{PR}(+)$	$W_{PR}(-)$	$R_C$	$P_C$	
NORMAIS	Variação Potência –degraus de +10%	0,000	0,020	0,013	0,040	0,006	0,003	0,144	0,000	0,111	0,000	0,000	0,059	0,000	0,000	0,000	0,000	
	Variação Potência –degraus de –10%	0,020	0,010	0,040	0,013	0,003	0,006	0,000	0,130	0,000	0,100	0,046	0,000	0,000	0,000	0,000	0,000	
	Rampa – 100 a 40% - 10%/min.	0,020	0,020	0,013	0,013	0,000	0,043	0,000	0,600	0,000	0,600	0,177	0,000	0,000	0,000	0,000	0,000	0,000
	Rampa – 40 a 100% - 10%/min.	0,020	0,020	0,013	0,013	0,045	0,000	1,500	0,000	1,500	0,000	0,000	0,200	0,000	0,000	0,000	0,000	0,000
	Desligamento do Reator ("Trip")	0,000	0,063	0,000	0,280	0,000	0,086	0,000	0,950	0,000	0,950	0,250	0,000	0,000	0,000	0,000	0,000	0,000
ANORMAIS	Desligamento da Turbina ("Trip")	0,025	0,032	0,013	0,120	0,000	0,049	0,000	0,700	0,000	0,700	0,300	0,000	0,000	0,000	0,000	0,000	0,000
	Fechamento inad. 1 válv. Vapor	0,013	0,050	0,013	0,400	0,000	0,091	0,000	0,950	0,020	1,000	0,250	0,000	0,000	0,000	0,000	0,000	0,000
	Fechamento inad. 4 válv. Vapor	0,063	0,100	0,093	0,410	0,000	0,101	0,000	0,500	0,000	0,450	0,170	0,000	0,000	0,340	0,000	0,000	0,000
	Perda Parcial Vazão (1 bomba)	0,025	0,038	0,000	0,160	0,006	0,028	0,000	0,500	0,000	0,100	0,980	0,375	0,000	0,000	0,000	0,000	0,000
	Falha de todas bombas	0,025	0,032	0,013	0,120	0,006	0,049	0,000	0,950	0,000	1,000	0,375	0,000	0,000	0,000	0,950	0,000	0,000
EMERGÊNCIA	"Trip" Turbina sem desvio de vapor	0,070	0,101	0,093	0,410	0,000	0,101	0,000	0,950	0,100	0,980	0,375	0,000	0,000	0,000	0,000	0,000	0,000
	Falha de tubos do GV	0,000	0,519	0,650	0,493	0,000	0,117	0,000	0,950	0,100	1,000	0,203	0,047	0,000	0,000	0,000	0,000	0,000
	Pequeno "LOCA"	0,000	0,524	0,000	0,900	0,000	0,210	0,000	0,950	0,000	1,000	0,000	0,250	0,000	0,000	0,950	0,100	0,100
	Pequena ruptura tubul. Vapor	0,000	0,175	0,000	0,450	0,000	0,160	0,000	0,950	0,000	1,000	0,000	0,280	0,000	0,000	0,000	0,000	0,100
F	Grande LOCA	0,000	0,974	0,000	1,000	0,000	0,828	0,000	0,950	0,000	1,000	0,000	0,747	0,000	1,000	1,000	4,000	0,000

As variáveis indicadas representam:  $P_{PR}$  – pressão no primário;  $L_{PZ}$  – nível no pressurizador;  $T_R$  – temperatura na saída do reator;  $Q_R$  – Potência no reator;  $W_S$  – vazão de vapor;  $P_S$  – pressão do vapor;  $W_{PR}$  – vazão no primário;  $R_C$  – nível de radiação na contenção; e,  $P_C$  – pressão na contenção. Cada valor indica a máxima variação relativa ao valor do início do transiente, p. ex.:  $|P_R^{\max} - P_R^0| / P_R^0$ . Os sinais (+) e (-) indicam, respectivamente, a variação máxima acima do valor inicial e a variação máxima abaixo dele, p. ex.:

no primeiro evento a pressão caiu um máximo de 2% e não subiu acima do valor inicial, desse modo  $P_R(+)=0$  e  $P_R(-)=0,020$ . Para identificar qual unidade é ativada por um dado padrão de entrada, esses foram "rotulados" acrescentando-se um vetor no qual cada coluna representa uma das entradas, p. ex., 0,001 seguido de 14 "zeros" representa o primeiro evento. Esse valor pequeno foi utilizado para assegurar que a rotulação não irá interferir no resultado.

Foram testadas redes bidimensionais em arranjo quadrado com diversas quantidades de unidades. Os resultados apresentados se referem a uma rede com uma camada de  $10 \times 10$  unidades. As constantes utilizadas foram:  $\eta_0 = 0,9$ ;  $b_1 = 0,02$ ;  $\sigma_0 = 18$ ;  $b_0 = 0,02$ . A Fig. 8 mostra os resultados obtidos. As linhas separando os padrões de entrada foram traçadas para ilustrar a separação produzida. A seqüência de 0 a 14 segue a ordem de eventos da Tabela 1.

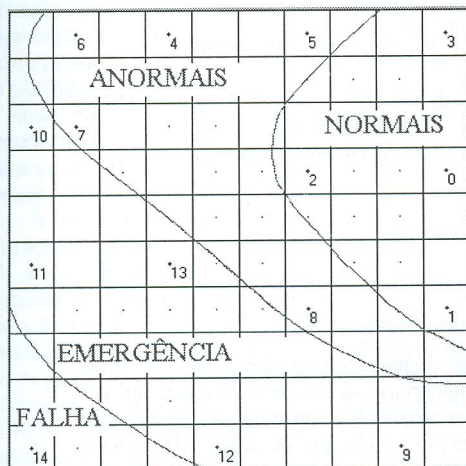


Figura 8. Classificação de Eventos pelo SOM.

O objetivo deste artigo não é realizar uma análise detalhada desse problema de classificação em específico, mas somente ilustrar o uso de um SOM, por isso a discussão se limita a poucas observações. A primeira observação sobre os resultados de classificação apresentados no mapa de “unidades vencedoras” se refere à clara separação dos eventos de efeitos similares, ou seja, que geram variações equivalentes nos parâmetros medidos. Um fato notável se refere à “proximidade” topográfica entre alguns eventos de classes diferentes, como p. ex. o 7 e o 10. Esses se referem respectivamente a “Fechamento inadvertido de 4 válvulas de vapor” e “Trip de turbina sem desvio de vapor.” Observando-se a variação dos parâmetros fornecidos percebe-se claramente que os efeitos dos dois eventos são quase rigorosamente os mesmos, o que os diferenciou foi apenas a rotulação adicionada. Outro fato nítido é que o evento 5, “Desligamento da Turbina”, quanto a seus efeitos poderia ser classificado como um evento normal, o que fica evidente observando-se os valores da Tabela 1. Essas análises, limitadas pelos múltiplos objetivos e pelo espaço do artigo, apenas ilustram uma típica aplicação de SOM na EN. Outras análises poderão ser ilustradas em outros trabalhos, como por exemplo a entrada de padrões diferentes dos utilizados, ilustrando a capacidade de classificação de padrões novos. Essa habilidade de seleção de características do SOM pode ser utilizada, associando-o com uma RAA, produzindo um poderoso classificador híbrido.

#### IV. CONCLUSÕES

Neste artigo, dividido em duas partes, foram abordadas aplicações de três diferentes tipos de RNA na Engenharia Nuclear. Esta segunda parte focalizou as redes de contatos sinápticos múltiplos (MULSY) e os mapas auto-organizáveis (SOM). Os exemplos apresentados procuraram mostrar que a aplicação deve definir a escolha do tipo de rede, ou seja, problemas mais genéricos devem ser resolvidos por RAA, problemas de controle podem ter soluções mais eficientes com redes MULSY e problemas de classificação de padrões podem ser resolvidos por SOM, eventualmente associados a RAA.

#### REFERÊNCIAS

- [1] Baptista F., B.D., **Redes Neurais Artificiais na Engenharia Nuclear (Parte 1)**, Anais do XIII ENFIR, 11 a 16 de Agosto, Rio de Janeiro, Abstract Code 17, 2002.
- [2] Baptista F., B.D., Cabral, E.L.L., Soares, A.J., **A New Approach to Artificial neural Networks**, IEEE Transactions on neural Networks, Vol. 9, No. 6, pp. 1167-1179, 1998.
- [3] Baptista F., B.D., Cabral, E.L.L., **A New Neural Network Concept for Control of Nuclear Reactors Systems**, Eighth International Conference on Nuclear Engineering (ICONE-8), Track 7, Session 7.08, April 2-6, Baltimore, USA, 2000.
- [4] B.D. Baptista F., E.L.L. Cabral, **Planar two-link manipulator control with multiple-synapse neural network controller**, IEEE/ASME International Conference on Advanced Intelligent Mechatronics, pp. 647-652, Sept.19-23, Atlanta, USA, 1999.
- [5] Rosenblatt, F., **The Perceptron: A probabilistic model for information storage and organization in the brain**, Psychological Review, vol. 65, pp. 386-408, 1958.
- [6] Kohonen, T., **Self-Organized Formation of Topologically Correct Feature Maps**, Biological Cybernetics, v.43, p. 59-69, 1982.
- [7] Haykin, S., **Redes Neurais, Princípios e Prática**, trad. Paulo Martins Engel, 2. Ed., Porto Alegre: Bookman, ISBN 85-7307-718-2, 2001.

#### ABSTRACT

The field of *Artificial Neural Networks (ANN)*, one of the branches of *Artificial Intelligence (IA)* has been waking up a lot of interest in the Nuclear Engineering (NE). ANN can be used to solve problems of difficult modeling, when the data are fail or incomplete and in high complexity problems of control. The first part of this work began a discussion with Feed-Forward Neural Networks in back-propagation. In this part of the work, the Multi-synaptic neural networks (MULSY) is applied to control problems. Also, the self-organized maps (SOM) is presented in a typical pattern classification problem: *transients classification*. The main purpose of the work is to show that ANN can be successfully used in NE if a carefully choice of its type is done: *the application sets this choice*.