

Reconhecimento Automático de Modulação Digital e Demodulação de sinais baseado em Redes Neurais Artificiais

Yona Lopes, David L. Acosta, Antonio J. C. R. Coutinho, Fidel E. Souza,

João M. M. Silva, José S. G. Panaro

yona@gteccom.uff.br, david.leonardo.acosta@gmail.com, ajc Coutinho@uol.com.br, fidel Souza30@yahoo.com.br

jmarcos@telecom.uff.br, panaro@telecom.uff.br

Departamento de Engenharia de Telecomunicações. Universidade Federal Fluminense.

Niterói, RJ - Brasil. Tel.: +55 21 2621-3488

Resumo—Redes Neurais Artificiais estão entre as mais bem-sucedidas ferramentas computacionais para tratamento de problemas não-lineares em processamento de sinal. Esse método é diretamente aplicado às amostras do sinal de rádio, e tem complexidade computacional relativamente baixa.

Neste artigo, são utilizadas Redes Neurais Artificiais ART (Teoria da Ressonância Adaptativa - *Adaptive Resonance Theory*) para reconhecer automaticamente o tipo de modulação digital, e, com isso, demodular um sinal de rádio gerado. Três tipos de modulações são investigadas: BPSK, 2FSK e 4PSK, sendo averiguada a robustez da rede Neural em relação a cada modulação gerada. São testadas duas métricas: a distância Euclidiana e a Norma de Manhattan e é proposta uma comparação do desempenho da Rede Neural ART simplificada em cada situação. Esse desempenho é avaliado com diversas simulações, levando-se em consideração canais com ruído branco aditivo gaussiano (*Additive White Gaussian Noise* - AWGN). É mostrado que no reconhecimento de um sinal modulado (não codificado), usando uma relação sinal ruído(SNR) superior a 10dB a porcentagem de acerto chega até 100%, e para valores inferiores a probabilidade de acerto é suficientemente alta.

Palavras-Chave—Redes Neurais Artificiais, Redes ART, Reconhecimento Automático de Modulação, Rádio Cognitivo, 2FSK, BPSK, 4PSK.

I. INTRODUÇÃO

Com o avanço da ciência da computação e com o surgimento das Redes Neurais e dos métodos de Inteligência Computacional, tornou-se possível modelar em *software* sistemas que antes eram processados por *hardware* dedicado. Portanto, maiores níveis de automação, eficiência de processamento, flexibilidade, integração e portabilidade são alcançados. O reconhecimento e a classificação automática de sinais utilizando recursos computacionais, é um assunto de grande interesse, tendo impacto direto nos Rádios Definidos por Software e nos Rádios Cognitivos.

O reconhecimento automático de modulação (*Automatic Modulation Recognition* - AMR) permite que o receptor reconheça automaticamente a modulação utilizada no transmissor. Esta técnica, associada a outras técnicas, de processamento de sinais possibilita: a utilização de mais de um tipo de modulação no canal; a escolha pelo receptor do sinal menos degradado pelo ambiente de transmissão; e a utilização de uma mesma faixa de frequências por vários serviços distintos,

aumentando a eficiência na utilização do espectro. O reconhecimento automático de modulação e demodulação são dois componentes essenciais em receptores de rádio cognitivos. Com isso, trabalhos envolvendo Redes Neurais Artificiais com o objetivo de desenvolver métodos que classifiquem automaticamente a modulação de sinais, e/ou propiciem a sua demodulação, vem sendo desenvolvidos [1], [2], [3], [4].

Os métodos de classificação automática de modulação digital podem ser divididos em duas categorias principais: Decisão Teórica e Reconhecimento de Padrões [5]. Um conhecido método de decisão teórica é o *Likelihood Function-LF*. Os métodos de reconhecimento de padrões podem ser baseados, ou não, em Redes Neurais Artificiais. Estes métodos possuem basicamente três estágios [6]: Pré-processamento, treinamento e validação, e Testes.

O presente trabalho se propõe a demonstrar que é possível realizar a identificação do tipo de modulação de um sinal, e a sua correta demodulação, utilizando para isso, uma rede neural artificial do tipo ART, e se realizando uma normalização prévia dos sinais apresentados à rede. A escolha da rede ART teve o propósito de alcançar uma maior simplicidade computacional, que é fundamental para implementação de soluções em tempo real, sem atraso de processamento, o que inviabiliza a identificação da modulação em receptores de rádio. Para ter viabilidade prática, a rede neural do tipo ART deve possuir, além da baixa complexidade computacional, a capacidade de operar em um canal de propagação que possua níveis baixos de SNR. É indispensável que a rede tenha desempenho robusto mesmo em ambientes com níveis diferentes de SNR. Na seção II são apresentadas algumas considerações teóricas sobre o classificador em estudo, a Rede Neural ART, e as métricas utilizadas. A seção III apresenta o pré-processamento. Na Seção IV são descritas as fases de treinamento e classificação. Na seção V os resultados de desempenho e testes feitos com o classificador são apresentados para sinais modulados em 2FSK, BPSK e 4PSK. Por fim, na seção VI, são feitas as conclusões acerca deste trabalho.

II. REDE ART

Uma Rede Neural Artificial (RNA) é um modelo computacional abstrato do cérebro humano. É um método para

solução de problemas através da simulação deste, inclusive em seu comportamento, ou seja, aprendendo, errando e fazendo descobertas. São técnicas computacionais que apresentam um modelo inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento através da experiência. Essa rede possui inclusive a capacidade de generalização, ou seja, consegue produzir saídas razoáveis para as entradas desconhecidas, que não foram usadas no processo de treinamento [7].

As técnicas de redes neurais têm sido cada vez mais aplicadas em problemas de difícil modelagem computacional ou em áreas em que um modelo matemático seria complexo demais para ser útil. Em uma RNA os neurônios devem estar conectados entre si e dispostos em camadas. A disposição dos neurônios nas camadas, e o padrão de conexão entre estas, definem a arquitetura da RNA. Nas redes neurais conhecidas como competitivas, as células vizinhas em uma rede neural competem em suas atividades se desenvolvendo e se adaptando. Sua principal característica é seu aprendizado não-supervisionado, situação esta onde normalmente não se conhece “*a priori*” a que classe pertence cada padrão apresentado à rede.

A arquitetura ART modificada baseia-se essencialmente na camada de Kohonen com aprendizagem competitiva [8]. Cada neurônio da camada tem um vetor de pesos sinápticos associado.

A classificação de um vetor de entrada (x) representando um determinado padrão é realizada pelo neurônio que apresentar a menor distância para entre o centro da classe e este vetor, de acordo com uma métrica, desde que esta distância seja menor do que um raio pré-definido u_0 [9]. Este neurônio torna-se o vencedor, e é o único a ser treinado. A figura 1 exemplifica uma rede ART modificada.

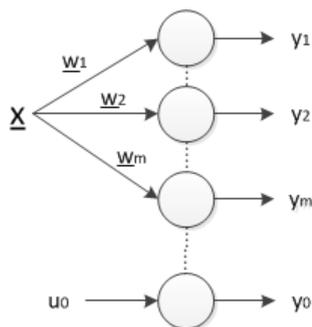


Fig. 1. Arquitetura rede ART

Tal operação, corresponde a um procedimento de segmentação (*clustering*), onde cada neurônio é atribuído a um padrão (*cluster*). Após algumas épocas de treinamento, em geral, o vetor de pesos w_i associado ao “*i-ésimo*” neurônio deve convergir para um ponto onde ele irá representar melhor o seu respectivo padrão - ponto este conhecido como centro de classe [9].

O propósito de uma rede ART é incorporar um modelo de aprendizagem competitiva dentro de uma estrutura de

controle auto-organizável, cujo reconhecimento e aprendizado continuam estáveis em resposta a uma sequência arbitrária de padrões de entrada. Esta teoria foi desenvolvida principalmente para resolver um dilema conhecido como estabilidade/plasticidade. A estabilidade é a garantia de que todos os elementos serão agrupados, a medida que os pesos são atualizados, os novos valores tendem a diminuir até estabilizar. A plasticidade é a característica que a rede possui de adaptar-se a mudanças no padrão originalmente aprendido.

Dentre as famílias da rede ART estão a ART-1, com entradas e saídas binárias, a ART-2, com entradas e saídas analógicas e a ART 3 que inclui Neuro transmissores. Esse artigo utiliza a ART 2-A que tem um sistema computacional mais simples, preciso e rápido. ART 2-A é executada aproximadamente duas a três ordens de magnitude mais rápida do que a ART-2, tornando mais fácil a solução de problemas de grandes dimensões [10].

A. Métricas de Dissimilaridade

Métricas de dissimilaridade são utilizadas em redes neurais para medir a distância entre o vetor de entrada (x) e os vetores de pesos dos neurônios (w_i). A métrica de dissimilaridade usualmente empregada é a norma Euclidiana (Eq. 2), porém outras normas de dissimilaridade também podem ser empregadas conforme as características do problema [11], como a norma de Manhattan (Eq. 1).

$$L(p=1) \longrightarrow d_1(x, y) = \sum_{i=1}^l w_i |x_i - y_i| \quad (1)$$

$$L(p=2) \longrightarrow d_2(x, y) = \sum_{i=1}^l \|x_i - y_i\|^2 \quad (2)$$

Estas são duas formas de normas ponderadas L_p . Foram testadas essas duas métricas, pois geralmente essas distâncias fornecem *clusters* mais homogêneos e penalizam, de maneira mais consistente, os dados chamados de *outliers* que, por motivos de erro grosseiro ou falha, estão muito distantes da média.

III. PRÉ-PROCESSAMENTO

Nenhuma técnica de codificação de canal é empregada neste trabalho. Os vetores do sinal, puramente modulados, são pré-processados antes das etapas de treinamento e de classificação, sofrendo uma normalização. Técnicas como Transformada Rápida de Fourier ou Wavelets são geralmente empregadas nesse processo, mas apresentam custo elevado de processamento.

A normalização permite que valores diferentes possam ser comparados. Os valores amostrados são colocados em uma escala comum e em uma mesma faixa de valores para que eles possam ser vistos de modo comparativo. Normalizam-se as amostras para evitar que uma dimensão se sobreponha as outras e para evitar que o aprendizado fique estagnado. O pré-processamento pode otimizar o treinamento da rede neural e melhorar a capacidade de aprendizagem e a identificação de

novos padrões. A equação 3 mostra a normalização utilizada nesse artigo na etapa prévia ao treinamento e à classificação. Para cada sinal são normalizados cada símbolo (vetor) de maneira independentemente.

$$R[k] = \frac{N * r[k]}{\sqrt{\sum_{n=1}^N r[n]^2}} \quad (3)$$

Onde N é o número de amostras do vetor a ser normalizado.

IV. AMR E DEMODULAÇÃO USANDO REDES ART

Dependendo dos vetores apresentados a uma rede ART, esta vai ativando ou modificando seus neurônios durante a evolução do seu treinamento. Como se busca classificar símbolos que podem ser semelhantes entre algumas modulações, é criada uma rede ART própria para cada tipo de modulação, permitindo assim diferenciar individualmente cada símbolo em cada tipo de modulação.

No presente trabalho, são criadas 3 redes ART que são treinadas independentemente com seu respectivos símbolos. Estas redes são usadas posteriormente no processo de classificação, apresentando os símbolos do sinal a ser identificado. Cada uma delas tem um neurônio vencedor, este é comparado com os vencedores das outras duas redes, para finalmente decidir qual é o vencedor definitivo, que representará o tipo de modulação desse símbolo.

Para preparar o processo de demodulação, uma espécie de treinamento supervisionado é realizado. Como os valores dos símbolos "0" ou "1", ou "2" e "3" quando usado 4PSK, são conhecidos no treinamento, estes são usados para definir os pesos iniciais dos neurônios e para estimar o valor do parâmetro de vigilância que permita a menor ativação de neurônios possível em cada rede. Desta forma, cada rede tem rotulados seus neurônios dependendo do símbolo. Durante a etapa de classificação, o neurônio ativado leva a etiqueta do símbolo ao qual corresponde, permitindo assim que depois de identificado o tipo de modulação, se obtenha o símbolo.

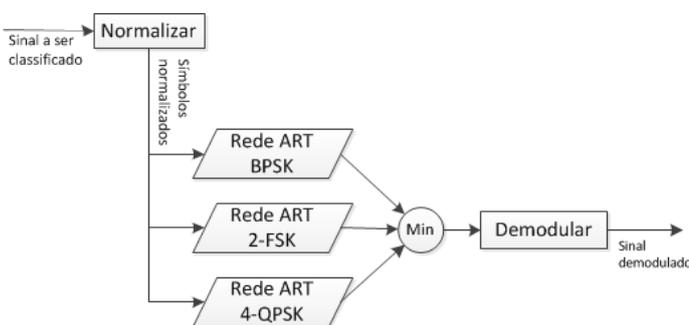


Fig. 2. Arquitetura da rede proposta

A. Fase de treinamento

A rede ART utiliza um algoritmo de aprendizado competitivo, onde os neurônios da camada de saída competem entre si para se tornarem ativos, ou seja, para ver quem gera a menor distância ao vetor de entrada. Assim, para cada padrão de

entrada apenas um neurônio de saída ou neurônio por grupo se torna ativo. Esta competição é chamada de *winner-takes-all*.

No processo de treinamento do sistema proposto, uma rede ART simplificada foi treinada independentemente para cada tipo de modulação, apresentando a cada rede unicamente os símbolos correspondentes à modulação. O processo de treinamento de cada rede ART segue os seguintes passos:

▷ Inicializar

Se definem os pesos iniciais com a média aritmética para cada tipo de símbolo. Se estabelece o parâmetro de vigilância.

▷ Apresentar entrada

Se apresenta um símbolo do banco de símbolos correspondente ao tipo de modulação. O símbolo é normalizado previamente.

▷ Calcular a distância

Se calcula a distância entre o vetor apresentado e os pesos existentes na rede. Depende da norma ponderada selecionada são usadas a equação 1 para Manhattan ou a equação 2 para Euclidiana.

▷ Selecionar o neurônio vencedor

É escolhido o neurônio vencedor de acordo com a menor distância encontrada. Se o "neurônio" vencedor é o de vigilância, será ativado um novo neurônio com os valores do vetor de entrada.

▷ Atualizar pesos

É usada equação 4 para atualizar os pesos do neurônio vencedor i . Os outros neurônios manterão seus pesos.

$$w_i(n+1) = w_i(n) + \alpha[x(n) - w_i(n)] \quad (4)$$

▷ Repetir

Continuar apresentando os vetores de entrada (símbolos) enquanto estejam pendentes.

▷ Finalizar

Repetir o processo para o número de iterações estabelecida.

B. Fase de classificação

Na fase de classificação, o sinal apresentado ao sistema é dividido em seus símbolos e estes são passados pelas três redes treinadas. Cada uma delas vai entregar o número do neurônio vencedor e sua distância calculada. É selecionada a menor das três distâncias a qual é usada para definir o tipo de modulação. Finalmente, depois de fazer o mesmo processo para todos os símbolos do sinal, é realizada a demodulação de acordo com o maior número de símbolos classificados e é usada a etiqueta do neurônio ativado para obter o tipo de símbolo correspondente à modulação.

C. Taxa de erro de bit

A transmissão não ideal de sinais ao longo de um canal AWGN provoca alterações na amplitude da onda de forma randômica, alterando assim a energia média de bit e de símbolo. Este fenômeno pode provocar a ocorrência de erros de detecção dos bits da informação. Assim pode ser definida a probabilidade de média de erro ou BER (*bit error rate*). A tabela I traz as equações que definem a BER para as

TABELA I
PROBABILIDADE MÉDIA DE ERRO [12]

Modulação	Probabilidade média de erro (BER)
FSK	$\frac{1}{2} \operatorname{erfc} \sqrt{\frac{E}{2N_0}}$
PSK	$\frac{1}{2} \operatorname{erfc} \sqrt{\frac{E}{N_0}}$
QPSK	$\frac{1}{2} \operatorname{erfc} \sqrt{\frac{E}{N_0}}$

modulações FSK, PSK e QPSK.

É importante notar que a expressão da BER para a modulação QPSK é igual a da PSK. Para as modulações binárias FSK e BPSK a probabilidade de erro de bit é igual a probabilidade de erro de símbolo pois um bit é também um símbolo. Já para a modulação QPSK que possui dois bits por símbolo a probabilidade de erro de símbolo é dada pela seguinte equação [12]

$$P_{e_{sim}} = 1 - \left(1 - \frac{1}{2} \operatorname{erfc} \sqrt{\frac{E_b}{2N_0}} \right)^2 \quad (5)$$

Onde E é a energia do símbolo. Como $E = 2E_b$ para a modulação QPSK temos:

$$P_{e_{sim}} = 1 - \left(1 - \frac{1}{2} \operatorname{erfc} \sqrt{\frac{E_b}{N_0}} \right)^2 \quad (6)$$

V. TESTES E RESULTADOS

Neste artigo a complexidade computacional é reduzida, trabalhando apenas com a atualização dos neurônios vencedores, um por classe. Apesar da redução da complexidade deixando de usar a função de vizinhança, é comprovado que o resultado é satisfatório, como pode ser visto nas tabelas abaixo, e o ganho com a simplicidade e rapidez da rede é alto.

Os dados são pré-processados usando uma normalização simples segundo a equação 3, ao invés de cálculos complexos como a FFT ou a Transformada de Wavelet [4]. O banco de dados para o treinamento foi criado com sinais aleatórios gerados no Matlab, modulados em 2FSK, BPSK e 4PSK e com o mesmo valor de portadora. Posteriormente foram contaminados com ruído branco aditivo gaussiano (AWGN) com SNR de 0 a 20dB.

Para o processo de aprendizado foi utilizada uma taxa de aprendizado de 5% (0.05) para 200 iterações, e os pesos iniciais dos neurônios foram calculados com a média aritmética que representava cada símbolo do banco de dados.

Foram feitos testes usando as normas Euclidiana e Manhattan para a medição de dissimilaridade nas fases de treinamento e de classificação. Na fase de treinamento foi usado um banco de dados com 20 vetores de cada símbolo para cada modulação, cada um com 30 amostras e com o mesmo valor de portadora, 2.16Ghz. Para os testes de classificação foram usados sinais de 20 símbolos, os quais foram apresentados ao

sistema treinado. Para o pré-processamento a normalização de cada símbolo foi feita independentemente.

Para o processo de classificação, foram apresentados 1000 sinais gerados aleatoriamente contaminados com ruído branco aditivo gaussiano (AWGN), com valores de SNR de 1,10, 15 e 20 dB. Assim, foram gerados 1000 sinais de 1 dB e apresentados para o sistema de demodulação, 1000 sinais com 2 dB, e assim por diante.

As figuras 3 e 4 ilustram os resultados para cada grupo de 1000 sinais gerados. Nela, pode-se perceber que para 0dB de SNR usando a norma Euclidiana é alcançado 81% de acerto na identificação dos símbolos, enquanto que, com a utilização da norma Manhattan é alcançado 70.87%. Além disso, o comportamento é similar até 10 dB, sendo sempre a norma Euclidiana a que teve maior número de acertos na identificação de símbolos e correta demodulação.

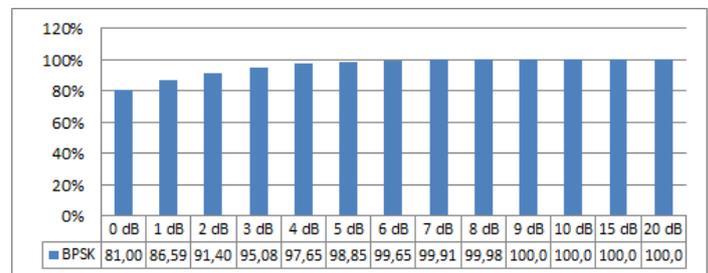


Fig. 3. Porcentagem para BPSK com norma Euclidiana

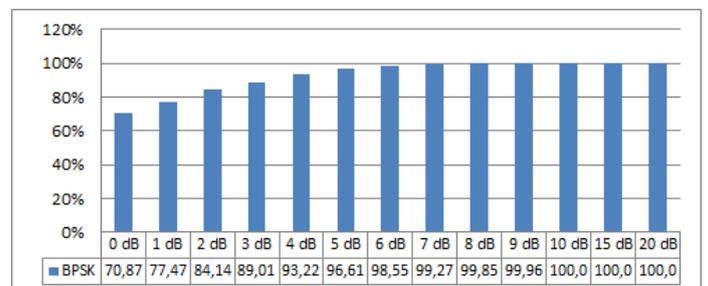


Fig. 4. Porcentagem para BPSK com norma Manhattan

Nas figuras 5 e 6 é mostrado que para a modulação 2FSK os resultados são similares, o maior número de acertos é alcançado com a norma Euclidiana. Para sinais de 0 dB de ruído branco, é mostrado que usando a norma Euclidiana obtem-se um resultado de 78.96% de acertos contra 69.16% da norma Manhattan, e que estabiliza em 100% de acertos para sinais com SNR superior a 9 dB em L_2 e depois de 11 dB em L_1 .

A situação é semelhante nos testes usando modulação 4PSK. As figuras 7 e 8 mostram que usando 1000 sinais com SNR=0dB, a identificação de seus símbolos alcançaram 80.93% de acertos com a Norma Euclidiana contra 75.31% com a norma Manhattan.

É importante ressaltar que está sendo estudada uma etapa de “pós-processamento”. Esta etapa envolve a leitura da saída classificada e a correção dos erros encontrados com

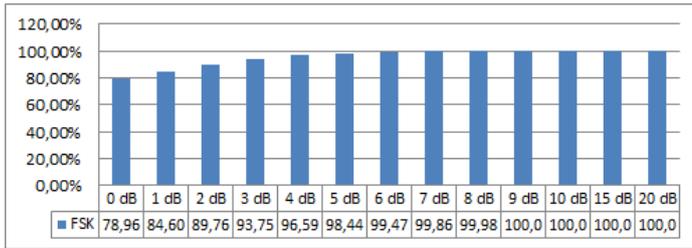


Fig. 5. Porcentagem para 2FSK com norma Euclidiana

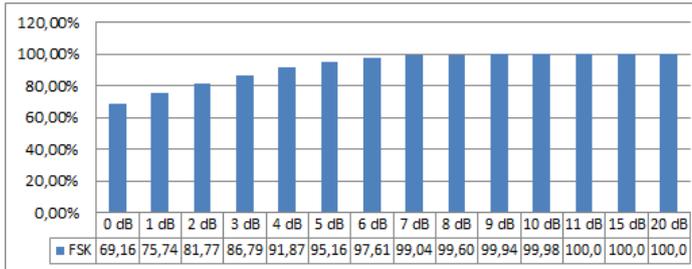


Fig. 6. Porcentagem para 2FSK com norma Manhattan

base na maioria dos símbolos classificados, fazendo uma rerepresentação da entrada à rede. Assim os dados são reclassificados obtendo uma porcentagem de acertos maior inclusive com SNR baixo. A complexidade computacional aumenta, porém, a assertividade é muito maior.

VI. CONCLUSÕES

O presente artigo mostra, através dos resultados obtidos na simulação, que é perfeitamente exequível a utilização de Redes Neurais do tipo ART para realizar as tarefas de reconhecimento automático dos tipos de modulação citadas, bem como, implementar a demodulação do sinal apresentado à entrada da rede neural. Foi verificado também que a rede ART obteve uma ótima capacidade de generalização alcançando um bom desempenho mesmo submetida a uma SNR baixa. Os resultados reforçam a ideia de que, quanto menor a SNR mais difícil é de se reconhecer automaticamente a modulação usada, bem como realizar a demodulação do sinal.

A utilização de uma normalização das amostras dos símbolos representa uma alternativa interessante no lugar de técnicas de pré-processamento com Fourier ou Wavelet, dentre outras. Isto permite que a implementação da solução em microprocessadores de baixo desempenho seja viável.

Foi visto que os resultados obtidos são um aporte importante e um ponto no rumo das redes neurais ART aplicadas a reconhecimento de modulação. Existem interessantes trabalhos usando outro tipo de redes neurais artificiais. Um exemplo é o estudo feito por [4] que trabalharam com mapas auto-organizáveis de 18x18 neurônios para o reconhecimento de modulações BPSK, QPSK, 2FSK e MSK.

As métricas usadas, Manhattan e Euclidiana, atenderam as expectativas do estudo. A Norma Euclidiana, apesar da pouca diferença, apresentou maior eficiência, formando *clusters* mais homogêneos e assim, um melhor resultado. Atualmente está

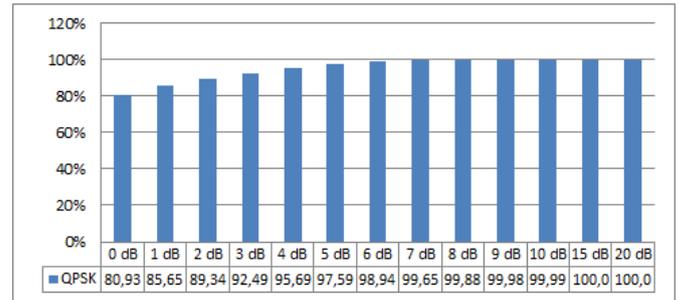


Fig. 7. Porcentagem para 4PSK com norma Euclidiana

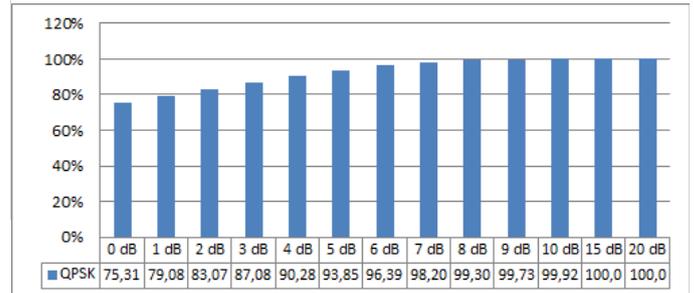


Fig. 8. Porcentagem para 4PSK com norma Manhattan

sendo estudado o uso de outras normas de dissimilaridade e similaridade para a obtenção do neurônio vencedor, as características das redes ART para se aplicar a modulações mais complexas sem requerer maiores requisitos computacionais e uma etapa de pós-processamento. E pensando em aplicações em tempo real, onde o tempo de processamento é um fator crítico, o emprego de uma rede com complexidade reduzida, menor número de neurônios, como a proposta por este trabalho, se torna uma alternativa válida.

REFERÊNCIAS

- [1] C. Louis e P. Schier, "Automatic modulation recognition with a hierarchical neural network", em *Military Communications Conference, 1994. MILCOM '94. Conference Record, 1994 IEEE*, p. 713 -717 vol.3, oct 1994.
- [2] Y. Xu, L. Ge e B. Wang, "Digital modulation recognition method based on self-organizing map neural networks", em *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference on*, p. 1 -4, oct. 2008.
- [3] Z. Yaqin, R. Guanghui, W. Xuexia, W. Zhilu e G. Xuemai, "Automatic digital modulation recognition using artificial neural networks", em *Neural Networks and Signal Processing, 2003. Proceedings of the 2003 International Conference on*, vol. 1, p. 257 - 260 Vol.1, dec. 2003.
- [4] L. Zhou, Q. Cai, F. He e H. Man, "Msom based automatic modulation recognition and demodulation", em *Sarnoff Symposium, 2011 34th IEEE*, p. 1 -5, may 2011.
- [5] A. Ebrahimzadeh e S. Seyedin, "Automatic digital modulation identification in dispersive channels", 2006.
- [6] M. Horewicz, C. Nascimento e W. Perrella, "Reconhecimento automatico de modulação digital de sinais de comunicações", 2007.
- [7] S. Haykin, *Neural Networks*. 2001.
- [8] L. Fausett, "Fundamentals of neural networks - architectures, algorithms and applications", 1994.
- [9] M. Vassali, J. Seixas e L. Caloba, "A neural particle discriminator based on a modified art architecture", em *Circuits and Systems, 2002. ISCAS 2002. IEEE International Symposium on*, vol. 2, p. II-121 - II-124 vol.2, 2002.
- [10] G. Carpenter, S. Grossberg e D. Rosen, "Art 2-a: An adaptive resonance algorithm for rapid category learning and recognition", *Neural Networks*, vol. 4, no. 4, p. 493-504, october 1991.

- [11] S. T. K. Koutrombas, "Pattern recognition", 1994.
- [12] S. H. e M. Moher, *Modern Wireless Communication*. Prentice Hall, 2004.